RAPPORT DE STAGE DE MASTER INFORMATIQUE DE L'UNIVERSITE PIERRE ET MARIE CURIE

Sécurité des infrastructures critiques.

DELAMARE Simon
Stage réalisé à l'Ecole Nationale Supérieure des Télécommunications.

30 août 2006

RESUME DU RAPPORT

Ce rapport présente les travaux réalisés autour de la protection des infrastructures critiques durant le stage de Master. Il est tout d'abord décrit ce que sont les infrastructures critiques, ainsi que les enjeux que représentent leurs protections. Le document s'oriente ensuite vers une recherche des mécanismes nécessaires pour assurer la disponibilité des services en cas de panne d'un ou plusieurs éléments du réseau. Nous nous intéresserons plus particulièrement à la connectivité au sein du réseau nécessaire à la délivrance des services. Un état de l'art sur ce sujet est présenté, qui englobe les bases du routage dynamique jusqu'aux récents travaux sur les réseaux de distribution de contenu et plus particulièrement les réseaux overlay (ou de recouvrement). Afin de pouvoir évaluer les besoins en terme de disponibilité des services des infrastructures critiques, un exemple détaillé d'une telle infrastructure, le réseau d'un fournisseur d'accès à Internet, est présenté. Ceci permet de bien comprendre quels sont les domaines de la sécurité à aborder, en fonction des objectifs recherchés. Le chapitre suivant est une évaluation des capacités des protocoles de routage à rétablir la connectivité après une panne. Plus précisément, une comparaison des performances de RON, un mécanisme de routage overlay, avec les protocoles de routage classiques est effectuée. Ceci permet de mettre en évidence les avantages et les inconvénients de ces protocoles pour la récupération rapide de la connectivité après une panne. Le dernier chapitre de ce document consiste en une explication du travail futur à réaliser autour de ce sujet.

Table des matières

Table des figures

Liste des tableaux

Chapitre 1

Introduction

1.1 Généralités sur la protection des infrastructures critiques

Les réseaux informatiques et de télécommunications prennent une place sans cesse grandissante dans la société moderne. En effet, ils prennent petit à petit la place des réseaux traditionnels pour la délivrance de services tels que la téléphonie ou la télévision, mais ils apportent aussi de nouveaux services aux usagers. De ce fait, ces services font maintenant partie de la vie quotidienne de nombreux usagers et il s'est créé une dépendance à ces réseaux.

Le terme d'infrastructure critique illustre cette situation. On considère qu'une infrastructure est critique si son non-fonctionnement porte préjudice à un grand nombre d'usagers. Cette définition, très large, regroupe toute sorte d'infrastructures appartenant à des domaines variés. Ce document se concentre sur les infrastructures d'informations critiques. Ces infrastructures englobe les réseaux de télécommunications délivrant des services quotidiens aux usagers, mais aussi les réseaux de télécommunications qui intéragissent avec d'autres infrastructures qui vont à leur tour fournir un service critique [1].

Il donc essentiel de mettre en place des mécanismes pour protéger ces infrastructures, afin que, dans la mesure du possible, les services puissent être délivrés en permanence, même lorsque l'infrastructure est endommagée, de manière provoquée ou non.

1.2 Contribution du document

La contribution de ce document s'articule en deux points. Tout d'abord, il est présenté un exemple concret d'infrastructure critique : le réseau d'un fournisseur d'accès à internet. Dans cette partie sont décris les services délivrés, l'architecture du réseau ainsi que les équipements nécessaires au déploiement des services. Ensuite, on étudiera les menaces pesant sur la disponibilité des services, quelles sont leurs incidences sur la délivrance des services aux usagers, ainsi que des pistes pour s'en protéger. Enfin, les besoins pour assurer la robustesse du réseau sont étudiés.

L'autre contribution du document est une mesure de performances de différents protocoles de routage. Plus particulièrement, il s'agit d'étudier les avantages apportés par un protocole de routage de niveau applicatif, déployé dans un réseau overlay, par rapport aux protocoles de routage classiques. Il sera montré quels sont les avantages et les inconvénients d'une telle solution, ce qui permettra de dégager des pistes pour la conception d'un protocole de récupération rapide en cas de panne.

Le document s'organise comme suit : Le chapitre deux est un aperçu de l'état de l'art sur divers domaines concernant la robustesse et la reprise sur panne. Les domaines abordés concernent le routage dynamique ainsi que les réseaux overlay pour apporter de la robustesse dans la délivrance de contenu. Le chapitre trois est consacré à la description d'un exemple d'infrastructure critique, le réseau d'un fournisseur d'accès à Internet (FAI). Le chapitre quatre présente la mesure de performance du protocole de routage overlay RON en comparaison avec un protocole de routage classique, OSPF. Après avoir décrit RON et l'architecture de test, on commentera les résultats obtenus. Le cinquième chapitre est un aperçu du travail futur, tel que l'on peut l'envisager après les travaux effectués durant le stage.

Chapitre 2

Etat de l'art

Dans cette partie, on abordera différents travaux qui sont en relation avec la disponibilité des services. Ces travaux visent à améliorer la robustesse de la disponibilité des services. On séparera cet état de l'art en trois parties :

- La description des principaux protocoles de routage dynamique et plus particulièrement, celle de OSPF ;
- Les travaux sur le re-routage rapide en cas de panne ;
- Les travaux sur la continuité des services, notamment les solutions overlay destinées à augmenter la fiabilité de délivrance des services. Une description de RON, un protocole de routage overlay, sera effectuée.

2.1 Les principaux protocoles pour la connectivité des nœuds

On va décrire un peu plus précisément les protocoles permettant d'assurer la connectivité entre les nœuds d'un réseau. La description plus poussée du protocole OSPF est motivée par le fait qu'il sera utilisé dans le test du chapitre 4.

Il existe de nombreux protocoles de routage dynamique permettant de renseigner automatiquement les tables de routage des routeurs d'un réseau. On peut noter qu'il existe deux catégories de protocole : les Internal Gateway Protocol (IGP) et les External Gateway Protocol (EGP). Les EGP sont dédiés au routage entre différents domaines administratifs, c'est-à-dire au routage entre différents Autonomous System (AS), et ne seront pas abordés ici. Les IGP sont dédiés au routage au sein d'un même AS. Voici les IGP les plus répandus :

- RIP [5] : Protocole à vecteur de distance (pour calculer sa table de routage, un nœud calcul le chemin de coût minimum pour arriver à chacun des autres nœuds). Fortement utilisé par le passé, il a été délaissé au profit de OSPF. La version 2 de RIP est conçue pour supporter les sous-réseaux ;
- OSPF [2] : Protocole à état de lien (chaque nœud connaît l'état de ses voisins, calcul la topologie complète du réseau avec les informations transmises par les autres nœuds et transmet ces informations en cas de changement de topologie) ;
- IS-IS [4] : Assez similaire à OSPF, ce protocole est plus particulièrement destiné aux grand réseaux. En effet, il sépare l'AS en sous-zones afin de répartir le travail de calcul des routes ;
- IGRP [6] : Protocole à vecteur de distance. Sa principale différence avec RIP est qu'il permet l'utilisation de différentes métriques. Il est la propriété de Cisco et n'est utilisé que sur leur matériel ;
- EIGRP [7] : Evolution de IGRP qui apporte un mécanisme pour empêcher les boucles de routage.

On va maintenant s'intéresser plus particulièrement à OSPF qui sera utilisé dans le test du chapitre suivant.

2.1.1 OSPF [2], [3]

Présentation

Open Shortest Path First (OSPF) est un IGP destiné à distribuer les informations de routage à l'intérieur d'un seul AS. Il est basé sur une technologie à état de lien, qui fonctionne de la façon suivante :

- A l'initialisation ou après une modification, le routeur envoie un message contenant l'ensemble de l'information sur l'état de ses liens ;
- Chaque routeur échange cette information par inondation. Un routeur conserve l'information qu'il reçoit des autres routeur ;
- Une fois l'ensemble des données collectées, le routeur calcule le chemin le plus court pour chaque destination selon l'algorithme de Dijkstra. C'est ainsi que sont construites les tables de routage avec le next-hop pour atteindre ces destinations ;
- Tant qu'aucun changement sur l'état des liens n'intervient, OSPF reste silencieux.

Caractéristiques

Voici quelques caractéristiques du protocole :

- La métrique utilisée pour évaluer le coût d'un chemin est la bande passante. Le coût peut aussi être spécifié explicitement ;
- Dans le cas où le réseau serait séparé en sous réseaux communiquant au travers de routeurs de bordure, les inondations peuvent être limitées aux sous-zones ;
- Il est possible d'utiliser un mécanisme d'authentification lors de l'échange des messages OSPF qui est basé sur un secret partagé ;
- Les routeurs prennent connaissance de leurs voisins (et donc de l'état de leur liens) par l'échange de messages HELLO. L'envoie de messages HELLO s'effectue tout les HELLO INTERVAL secondes (10 s par défaut) et un lien est considéré comme mort au bout de ROUTER DEAD INTERVAL secondes (40 s par défaut) pendant lesquelles aucun paquet HELLO n'a été reçu.

2.2 La récupération après panne dans le routage dynamique

2.2.1 Principes généraux

Le routage a pour but l'acheminement des paquets d'une source à sa destination, grâce aux routeurs, les nœuds intermédiaires. On dit que le but du routage est d'assurer la connectivité entre les différents nœuds du réseau. Ce qui nous intéresse plus particulièrement dans ce document concerne le routage dynamique, c'est-à-dire les protocoles permettant d'assurer la connectivité entre les machines du réseau automatiquement, et de façon dynamique en cas de changement de topologie du réseau, qui survient par exemple lors de la panne d'un lien ou d'un routeur.

On s'intéresse donc à la résistance de la connectivité en cas de pannes d'un élément du réseau. Nous étudierons les travaux dont le but est de fournir des mécanismes permettant, en cas de panne, de rétablir la connectivité entre les nœuds le plus rapidement possible.

Il existe trois phases pour le rétablissement de la connectivité par le protocole de routage

- La détection de la panne ;
- La dissémination aux autres nœuds de l'information concernant la présence d'une panne ;
- Le calcul par les nœuds de nouvelles routes en correspondance avec la nouvelle topologie.

L'essentiel des travaux présentés ici consiste à minimiser la durée d'une ou plusieurs de ces phases.

2.2.2 La détection des pannes

La détection des pannes joue un rôle essentiel dans le re-routage. En effet, lors de la panne d'un lien ou d'un équipement, avant de pouvoir démarrer une action rétablissement de la connectivité, il faut bien entendu détecter cette panne. Généralement, la détection s'effectue au moyen de sondes ou probes, souvent appelées messages HELLO, et consiste à envoyer un message à son voisin, qui y répond immédiatement. En l'absence de réponse, au bout d'un certain nombre de messages HELLO envoyés, le lien est considéré comme défectueux.

Il est indispensable que le temps de détection d'une panne soit le plus court possible pour pouvoir rétablir la connectivité dans les meilleurs délais. Quelques études ont été menés dans ce sens. En particulier, [11] considère le temps minimum que l'on peut attribuer aux valeur de HELLO INTERVAL et ROUTER DEAD INTERVAL du protocole OSPF afin d'effectuer une détection la plus rapide possible d'une panne, tout en minimisant le taux de « fausses alertes », c'est à dire de liens signalés défectueux alors qu'il ne le sont pas et qui sont provoqués par la perte de messages HELLO à cause de la congestion du réseau. Les résultats de ces études ont montré qu'en cas de forte congestion du réseau, un intervalle inférieur à une seconde provoquait de trop nombreuses « fausses alertes ».

Pour empêcher le risque de fausses alertes en cas perte du message HELLO lors d'une congestion du réseau, [13] stipule que les paquets HELLO doivent être traités en priorité par le routeur les réceptionnant. Ceci permet de diminuer le HELLO INTERVAL tout en gardant un faible taux de fausses alertes.

Dans [12], un nouvel algorithme pour la détection de pannes sur un lien est proposé. Il prend en compte l'impact des différents messages de contrôle du protocole de routage sur les performances du réseau. Sa particularité est de faire varier l'intervalle d'envoie de messages HELLO en fonction de la congestion du réseau. Les résultats montrent que l'algorithme peut détecter une panne en moins d'une seconde en ayant un très faible taux d'erreurs. De plus, si les paquets HELLO sont traités en priorité par le routeur, le temps de détection d'une panne peut être diminué à une demi-seconde.

2.2.3 Mécanismes de restauration

On peut distinguer deux familles de mécanismes pour la restauration de la connectivité après une panne :

- Les mécanismes proactifs, qui mettent en place des mesures de protection préventives, déclenchée une fois que la panne est apparue ;
- Les mécanismes réactifs, qui mettent en place des mesures de réparation une fois que est la panne détectée.

Les mécanismes proactifs

Les mécanismes proactifs consistent à anticiper la panne d'un lien et calculer une route de secours à emprunter en cas de panne. Cette méthode est efficace car elle permet de réduire à zéro le temps normalement dédié au calcul d'une nouvelle route lors de l'apparition d'une panne. Parmi les différents mécanismes existant, on distingue plusieurs méthodes selon la topologie de la route de secours pré calculée.

Routes de secours de bout en bout Cette méthode consiste à calculer, pour chaque routeur, deux routes possibles pour atteindre chaque destination. Ceci permet, en cas de panne sur un lien de la première route, d'emprunter la deuxième route pré calculée. Il existe néanmoins des contraintes à cette méthode. Il est en effet nécessaire que la route de secours soit valide, et donc qu'elle ne soit pas touchée par la panne. Il existe plusieurs algorithmes permettant le calcul d'une route de secours. On peut notamment citer TDSPF [10], une extension de l'algorithme à état de lien de Dijkstra qui permet de calculer, en plus du plus court chemin, un deuxième plus court chemin, totalement disjoint du premier, sous la condition qu'un tel chemin existe.

Routes de secours locales Une autre technique permettant le re-routage rapide est de prévoir des routes de secours locales, c'est-à-dire des routes qui contournent le routeur susceptible d'être affecté par une panne. Ces techniques, appliquées à MPLS et RSVP-TE, sont décrites dans [14].

Cycles A l'origine conçu pour les réseau WDM, la protection par *p-cycles* a été adaptée [15] pour la restauration rapide dans les réseaux IP conventionnels. Les *p-cycles* sont des boucles fermées, traversant un certain nombre de nœuds du réseau. Elles vont permettre la restauration en cas de panne d'un certain nombre de liens ou de routeurs, suivant la topologie de la boucle par rapport à celle du réseau.

Lors de la panne d'un lien, un routeur adjacent au lien en panne qui aurait dû acheminer des données via le lien défectueux encapsule ces données de façon à ce qu'elles soient acheminées par le cycle pré calculé. Ensuite, les données voyagent dans le cycle jusqu'à ce qu'elles arrivent dans un routeur où le

coût du chemin pour atteindre la destination initiale des données est moindre que le coût relevé lors de l'encapsulation. Ce routeur décapsule alors les données pour les acheminer jusqu'à la destination.

Les difficultés rencontrées pour la mise en place de cette protection est la façon d'organiser les cycles dans un réseau donné de façon à minimiser l'augmentation du temps d'acheminement induit par le passage des données dans le cycle, tout en se protégeant de l'ensemble des pannes possibles. Les auteurs donnent les conditions pour parvenir à cette optimalitée, mais le calcul de la position des cycles est une optimisation combinatoire NP difficile.

Arbres Le but de cette technique est de construire deux arbres recouvrant le graphe correspondant au réseau à protéger. La racine de ces arbres est la source d'émission des données. Les arbres doivent être construits de telle façon que si l'on supprime n'importe quel nœud autre que la source, tous les autres nœuds du graphes doivent être reliés à au moins un des deux arbres. Ceci n'est réalisable que sous l'hypothèse que le graphe est redondant, c'est-à-dire que chaque nœud peut être relié à la source par deux chemins disjoints. Les auteurs de [16] proposent deux algorithmes pour parvenir à ce but, en cas de panne d'un lien ou de panne d'un nœud. Malheureusement, l'algorithme ne calcule pas les arbres de coût minimum pour parvenir à une destination. Des améliorations de ces algorithmes ont cependant été développées pour construire des arbres qui amènent à une diminution notable du délai entre la source et la destination [17].

Resilient Routing Layers Cette technique [20] consiste à calculer à l'avance des tables de routage dans un réseau où un ou plusieurs liens sont défaillant. Ainsi, lorsqu'une panne sur un lien est détectée, le routeur utilise la table de routage correspondant au sous-réseau dont le lien défectueux à été omis. L'efficacité de ce protocole dépend grandement du nombre de sous-réseaux pré calculés, c'est-à-dire du nombre de liens omis dans chaque sous-réseau pré calculé. En effet, dans le cas où par exemple, on pré calcule deux sous-réseaux, et que dans chaque sous-réseau la moitié des liens ont été retirés, lors d'une panne, le sous-réseau utilisé pour le routage va être loin de l'optimal en terme de nombre de sauts. A l'inverse, si l'on calcul autant de sous réseaux qu'il y a de liens, le sous réseau utilisé sera optimal, mais le coût en calcul et en mémoire est trop important.

Les mécanismes réactifs

Parmi les méthodes réactives (c'est-à-dire qu'aucune anticipation de la panne n'est effectuée), on peut citer deux approches permettant l'amélioration du temps de re-routage.

Diminution du temps de calcul de l'algorithme à état de lien Les auteurs de [18] proposent un algorithme qui permet, en se basant sur les informations de routage déjà calculées, de diminuer la complexité du calcul des tables de routage et ainsi diminuer le temps pour rétablir la connectivité.

Limitation de la dissémination des information sur l'état des liens Une approche décrite dans [19] consiste à limiter la dissémination de l'état des liens lors d'une panne au strict nécessaire pour opérer un re-routage rapide. Pour cela, lorsqu'un nœud détecte une panne, s'il ne peut faire transiter le trafic directement vers un autre nœud et ainsi éviter la panne, il informe le ou les nœuds qu'il convient pour permettre l'évitement de la panne.

2.3 Les réseaux overlay

2.3.1 Principes généraux

Les réseaux overlay sont depuis quelques années un sujet en plein développement. De nombreux projets ont montré que, déployés au-dessus d'Internet, l'overlay pouvait permettre d'optimiser la distribution des données. En effet, dans le cadre des Content Delivery Network (CDN), il est apparu que l'utilisation de la couche application pour la conception d'un réseau overlay permettait d'apporter une gestion poussée de la distribution des données, par l'application du client.

Dans la plupart des cas étudiés ici, le réseau overlay correspond à un réseau pair à pair (P2P), où chaque nœud de l'overlay possède des caractéristiques similaires.

Les overlay vont permettre la mise en place d'un réseau optimisé pour la distribution des données. Ces optimisations peuvent concerner le délai ou la bande passante, mais surtout, et c'est le sujet qui nous intéresse le plus ici, la robustesse. En effet, la mise en place du réseau overlay va permettre d'accroître la disponibilité des services en cas de pannes dans le réseau sous-jacent.

Dans cette partie, nous décrirons tout d'abord Resilient Overlay Network (RON). Nous étudierons ensuite des concepts plus théoriques concernant les réseaux overlay : dans un premier temps, les architectures d'overlay existantes, puis les mécanismes de probing entre les nœuds de l'overlay. Pour terminer, nous aborderons les systèmes overlay existants pour l'amélioration de la robustesse de délivrance des services.

2.3.2 Routage Overlay : Resilient Overlay Network (RON)[8]

Nous allons nous intéresser à RON, qui est un protocole de routage déployé sur un réseau overlay. Il sera utilisé dans le test du chapitre 4.

Présentation

Internet présente aujourd'hui des problèmes de résistance aux pannes. Il faut en effet plusieurs minutes pour arriver à rétablir une route correcte après la défaillance d'un lien ou d'un routeur. Le but de Resilient Overlay Network (RON) est d'apporter une solution à ces problèmes. Pour cela, son principe de fonctionnement consiste à placer des nœuds RON dans le réseau, en particulier dans des AS différents. Ces nœuds vont coopérer entre eux afin d'acheminer les communications, par sauts successifs de nœuds RON. La valeur ajoutée de cette méthode par rapport aux protocoles de routage standards est que les nœuds RON coopèrent alors qu'ils sont présents dans différents AS. Ce n'est pas le cas dans Internet, où les AS sont gérés indépendamment les uns des autres et ne s'échangent pas d'information de routage.

Les nœuds RON s'échangent des informations sur la qualité des chemins les reliant et construit une table de routage en fonction de plusieurs métriques telles que le délai, le taux de perte, la bande passante disponible. Ces informations sont recueillies par probing actif du chemin entre les différents nœuds. L'utilisation d'un probing agressif permet de détecter rapidement un changement de topologie. Ceci se fait au détriment d'un surcoût d'utilisation de bande passante. De plus, l'utilisation de métriques variées permet de faire un choix plus fin des routes à emprunter, en fonction des besoins de l'application.

Conception

Chaque programme qui communique avec le programme RON présent sur un nœud est appelé client RON. Le premier nœud RON à recevoir des données d'un client est le nœud d'entrée. Le dernier, le nœud de sortie.

Pour propager l'information sur la topologie du RON, on utilise un algorithme à état de lien. Dans un réseau à N nœuds RON, chaque nœud à N-1 liens virtuels. Chaque nœud effectue périodiquement une requête sur les N-1 liens virtuels afin d'évaluer les performances des liens. Les messages de routage de RON sont eux-mêmes des clients RON et sont donc propagés selon les chemins définis par RON. Ainsi, si un nœud RON n'a aucune information sur un autre nœud RON, c'est que ce nœud n'a plus aucun lien virtuel avec l'ensemble des autres nœuds RON.

Les nœuds RON ont besoin d'un algorithme afin de déterminer la qualité d'un lien et choisir la route à emprunter. La détection d'une panne sur un lien s'effectue de la manière suivante :

- En tant normal, le probing s'effectue tous les PROBE_INTERVAL (12s par défaut) ;
- Si, au bout de PROBE_TIMEOUT (3s par défaut), un probing ne répond pas, on envoie une série de probe séparés par un temps plus court (3s, par défaut, tant que le probing ne répond pas) ;
- Au bout de OUTAGE_THRESH (3s par défaut) envois de probing sans réponse, le lien est considéré comme étant rompu.

Voici comment sont calculées les différentes métriques :

Calcul du délai : On mesure le délai d'un lien (entre deux nœuds RON) en calculant la moyenne pondérée entre le délai de réponse du dernier probing et l'ancien délai mésuré. Le délai d'un chemin, qui est une succession de liens, se calcule en ajoutant les délais des liens :

Version	Hop limit	Routing Flags
RON Source Address		
RON Dest Address		
Source Port		Dest Port
Flow ID		
Policy Tag		
Packet Type		

TAB. 2.1 – L'entête RON

$$lat(lien) \leftarrow A * lat(lien) + (1 - A) * new\_lat$$

avec $A = 0.9$

$$lat(chemin) \leftarrow \sum(lat(lien)), pour\_chaque\_lien\_de\_chemin$$

Calcul du taux de perte : Le calcul du taux de perte d'un lien est effectué sur les 100 derniers probes envoyés. Le taux de perte d'un chemin est le complémentaire du taux de succès de transmission sur ce chemin, qui est le produit des taux de succès de transmission sur l'ensemble de ses liens :

$$loss\_rate(lien) \leftarrow \frac{nombre\_probe\_perdu}{nombre\_probe\_envoye}, sur\_100\_derniers\_probes$$

$$loss\_rate(chemin) \leftarrow 1 - \prod(1 - loss\_rate(lien)), pour\_chaque\_lien\_de\_chemin$$

Evaluation de la bande passante : On calcule un score pour comparer les bandes passantes disponibles entre les liens. Ce score est basé sur une simplification de l'équation de l'évaluation de la bande passante disponible par TCP [9]. De façon à éviter de calculer une bande passante infinie, cette équation est calculée avec un taux de perte d'au moins 2%.

$$score(lien) \leftarrow \frac{\sqrt{1.5}}{(rtt * \sqrt{p})}$$

et

$$p = max(2\%, loss\_rate(lien)/2)$$

Base de données des performances : Chaque nœud RON utilise une base de données de performances pour stocker les résultats des différentes mesures réalisées auprès de chaque nœud RON du réseau.

Acheminement des données : RON encapsule les données qu'il transporte. L'entête RON est décrite dans la table 2.1. Quand un paquet arrive dans un nœud RON, soit c'est le dernier nœud RON et l'entête est décapsulée, soit il a besoin d'être acheminé. Dans ce cas, le nœud RON :

- examine le Policy Tag du paquet, et cherche une métrique compatible dans la table de préférence qui contient la liste des métriques utilisables associées avec le Policy Tag ;
- sélectionne la table de routage associée avec la métrique ;
- sélectionne le prochain saut selon la métrique ;
- envoie le paquet.

Le flow ID est utilisé pour court-circuiter ce processus lorsqu'il appartient à une communication déjà traitée.

Implémentation

RON est implémenté en C++ et disponible sous la forme de librairie pour FreeBSD. Malheuresement, les auteurs de RON ont limité leur implémentation aux seuls besoins de leurs tests. Ainsi, l'implémentation de RON impose en particulier deux contraintes qui vont amener à certaines restrictions lors du test des performances de RON de la partie 4. :

- La source et la destination du trafic « pris en charge »par RON doivent être des nœuds RON ;
- Un nœud RON ne doit posséder qu'une seule interface réseau pour pouvoir encapsuler ou décapsuler du trafic.

2.3.3 Architecture des overlays

L'architecture d'un réseau overlay englobe deux sujets distincts :

- La position des nœuds du réseau overlay dans la topologie du réseau initial ;
- La topologie du réseau overlay indépendamment du réseau sous-jacent.

Peu de documents sont consacrés à l'étude de l'architecture du réseau overlay par rapport au réseau initial. Cependant, quelques études théoriques ont essayé de montrer l'impact de la topologie de l'overlay par rapport à celle du réseau sous-jacent sur la bande passante optimale. Plus particulièrement, [21] a montré que, dans le cas particulier de l'utilisation d'un overlay dédié au multicast dont les nœuds sont en bordure de réseau et en utilisant la technique du « network coding »[22], la bande passante optimale délivré aux clients dans le réseau overlay est égale à 95% de celle pouvant être délivrée dans le réseau initial.

L'étude de la topologie propre au réseau overlay consiste principalement à mesurer les avantages de telle ou telle architecture d'overlay. On peut distinguer deux catégories d'architecture de réseau overlay :

- les réseaux mesh, où chaque nœud de l'overlay est en communication directe avec chacun des autres. Ainsi, chaque nœud est adjacent à chaque autre, il n'est pas nécessaire de déployer des mécanismes de routage au sein du réseau overlay. Le problème de ce système est qu'il ne passe pas à l'échelle dans le cadre de réseau overlay comprenant de nombreux pairs. Des systèmes tels que RON ont portés leur choix sur ce type d'architecture. En effet, RON n'est pas destiné à être un grand réseau overlay donc le problème de passage à l'échelle est négligé.
- les réseaux non mesh tels que CAN [23] ou CHORD [24], où les nœuds de l'overlay sont en communication avec un groupe restreints d'autres nœuds voisins. Il est ici nécessaire de déployer des mécanismes de routage pour que tous les nœuds puissent communiquer entre eux. Le problème de ce système est d'ordre topologique. En effet, si aucun mécanisme n'est mit en place, deux nœuds adjacents dans le réseau overlay peuvent être forts éloignés dans le réseau sous-jacent. Cela peut entraîner de forts délais et une dégradation des performances dans le réseau overlay. Il convient donc de mettre en place des mécanismes pour positionner judicieusement les nœuds dans le réseau overlay, de façon à ce que deux nœuds proches dans le réseau overlay soient proches dans le réseau sous-jacent, ainsi que des mécanismes pour modifier dynamiquement la topologie du réseau overlay, en cas de modification de topologie ou de charge du réseau sous-jacent. Des systèmes tels que [25] ont été proposés afin de permettre le positionnement intelligent des nœuds dans l'overlay. De plus, des systèmes tels que [29], qui proposent d'appliquer le positionnement intelligent et dynamique des nœuds à des protocoles similaires à Bittorent [27], permettent une amélioration des performances d'au moins 25% en terme de temps de téléchargement.

2.3.4 Probing

Il est tout à fait pertinent de s'intéresser à la problématique du choix et de la mesure des métriques pour mesurer la qualité de la connexion entre deux nœuds du réseau overlay. Il existe principalement deux approches pour réaliser cette mesure :

- L'utilisation de sondes, qui sont de petits paquets échangés entre deux peers, et qui vont par exemple permettre de mesurer le RTT ou la bande passante disponible ;
- La mesure sur le trafic effectivement échangé de la qualité de la connexion.

Voici ce qu'on peut retenir de l'utilisation des sondes : Les résultats de [30] ont montré le manque de précision des sondes « conventionnelles », tel que échange de messages HELLO pour mesurer le RTT et envoie d'un paquet de 10 KO pour mesurer la bande passante disponible. En effet, lorsqu'il s'agit de

choisir le peer optimal par rapport à une métrique donnée, la qualité du peer effectivement choisit ne dépasse pas en moyenne 50% de celle du peer optimal. L'utilisation de probes plus élaborées (comme la combinaison de sondes conventionnelles) permet de sensiblement améliorer les résultats, sans toutefois atteindre des résultats très satisfaisant. Les auteurs expliquent ce phénomène par la fluctuation du trafic et des routes dans le réseau sous-jacent.

Il parait alors naturel de s'intéresser à l'autre méthode pour évaluer la qualité de la connexion entre deux nœuds du réseau overlay. Une autre approche, développée dans [28], consiste à mesurer en temps réel le trafic effectivement échangé entre deux nœuds, et ainsi évaluer la qualité de connexion entre ces deux nœuds. Les auteurs de ce papier développent un système assez similaire à Bittorrent basé sur cette technique. Ils constatent que ses performances en terme de temps de téléchargement sont améliorées de 33% environ par rapport à celles de Bittorrent.

Dans le cadre d'un réseau overlay dédié au re-routage, la mesure de la connexion entre deux pairs de l'overlay est destinée à détecter la panne d'une route sous-jacente. On peut donc se questionner sur les points suivant :

- Comment évaluer la connexion, faut-il utiliser des sondes ou mesurer directement le trafic transmis ?
- Dans le cas de l'utilisation de sondes, quels sont les métriques à utiliser pour s'assurer que la détection d'une panne est fiable (la route est effectivement coupée), tout en étant suffisamment rapide ?
- Dans le cas ou l'on mesure directement le trafic, quels sont les mesures pertinentes pour la détection d'une panne ? Comment calculer le RTT lors de l'envoi de trafic unidirectionnel, comme UDP ?

2.3.5 Systèmes déployés

Parmi les systèmes existant, seul RON [8], qui a été décrit précédemment, est un protocole de routage overlay destiné à la robustesse des communications. Il existe cependant d'autres systèmes dédiés à la diffusion vidéo qui permet, grâce à la réplication du contenu vidéo, d'augmenter la robustesse de la délivrance du contenu.

Le premier et le plus connu de ces systèmes est Splitstream [31]. Dans ce système, le contenu vidéo est divisé en sous-bloc. Pour chaque bloc, un arbre de recouvrement différent incluant l'ensemble des nœuds du réseau est crée. Ainsi, la répartition de la charge entre les différents nœuds est assez équitable. Spiltstream utilise le réseau overlay Pastry [32] et utilise Scribe [33] pour la gestion du multicast.

Chapitre 3

Un exemple d'infrastructure critique : Le réseau d'un FAI

On va dans ce chapitre, s'intéresser au réseau d'un FAI. Il convient en effet d'étudier un cas concret d'infrastructure critique afin de déterminer les risques pesant sur la bonne délivrance des services aux usagers et les besoins pour assurer une meilleure disponibilité de ces services en cas de panne dans le réseau. C'est pourquoi, après avoir identifié les services rendus par le système, ainsi que les moyens mis en œuvre pour implémenter ces services, nous nous intéresserons aux risques d'indisponibilité pesant sur ces services, puis aux techniques permettant de limiter les risques d'indisponibilité en cas de pannes ou d'attaques.

Il faut noter que dans un souci de concision, nous nous intéresserons principalement au service de diffusion de vidéo à la demande proposé par le fournisseur d'accès. Cependant, ce service est représentatif d'un service critique. En effet, le FAI doit s'assurer de sa bonne délivrance à l'usager car celui-ci paye pour y accéder. De plus, ce service est délivré en temps réel, ce qui impose des contraintes, notamment de délai et de synchronisation, pour sa livraison.

3.1 Le service de diffusion vidéo

Le service consiste à fournir aux clients du FAI, par l'intermédiaire de sa Set Top Box reliée à sa télévision, un service de Vidéo à la Demande (service VoD). Le client peut donc choisir parmi une liste de films proposés par le FAI, et, en échange d'une somme d'argent automatiquement prélevée sur la facture de son abonnement, le client peut visionner le film sur sa télévision.

Typiquement, les films proposés par le FAI sont présentés au client sur sa télévision, sur un canal spécial. Le client peut alors se renseigner sur les films proposés et en choisir un grâce à la télécommande. Le client a ainsi accès au film pendant H heures. Il peut réceptionner le film sur sa télévision et a aussi accès à des fonctions de contrôle (lecture, pause, etc.).

3.2 Architecture du réseau [34]

Nous allons présenter les différents éléments de l'infrastructure nécessaires au déploiement des services.

Réseau FAI - Client

Voici les différents éléments permettant de relier le réseau du FAI au client, dans le cas d'une connexion xDSL. L'architecture de cette partie du réseau est décrite dans la figure 3.1 :

- Le DSLAM : Le rôle du DSLAM est de relier chaque client, par l'intermédiaire du câble téléphonique, au réseau du FAI. Un DSLAM peut actuellement connecter environ 1000 clients au maximum. Le DSLAM ne s'occupe que de la conversion de la technologie xDSL en technologie Ethernet ou ATM ;

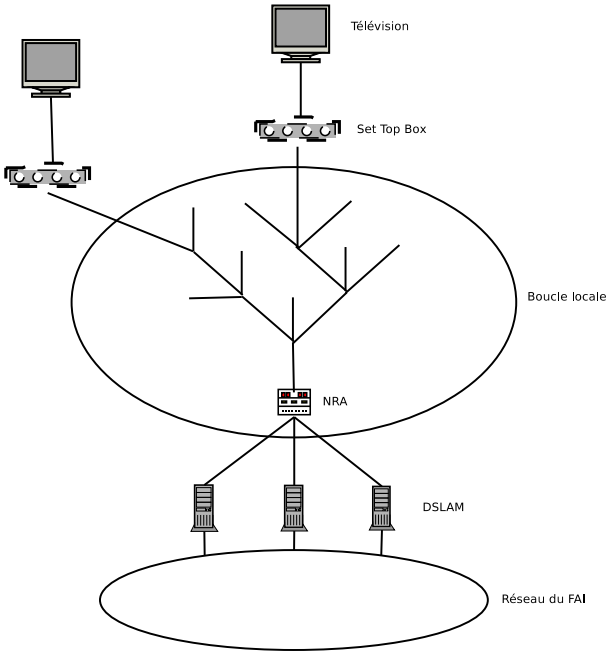


Fig. 3.1 – Architecture du réseau FAI vers client

- Le Nœud de Raccordement Abonné : Le NRA est le point de concentration de l'ensemble des câbles téléphoniques situés dans une même zone géographique. C'est le dernier équipement utilisant la technologie téléphonique utilisée avant le DSLAM. A un NRA, il est relié plusieurs milliers de clients donc un FAI peut avoir à y connecter plusieurs DSLAM ;
- La boucle locale : On appelle une boucle locale l'ensemble des câbles téléphoniques et équipement reliés au même NRA (le NRA étant compris dans la boucle locale). Le lien client NRA est loué par le FAI au propriétaire de la boucle locale. C'est le propriétaire de la boucle locale, qui n'est généralement pas le FAI, qui a la charge de son entretien ;
- La Set Top Box : C'est en général un modem xDSL qui apporte des fonctionnalités de routeur pour le réseau domestique, ainsi que des fonctions multimédia, comme la connexion à la télévision dans notre cas.

Réseau du FAI, partie centrale

Voici maintenant les différents éléments internes au réseau du FAI, tels que décris dans la figure 3.2. Tout d'abord, les éléments faisant partie de l'infrastructure centrale du réseau :

- La base de donnée (BDD) d'information sur les clients : L'annuaire contient la liste des clients du FAI, ainsi que des informations, notamment de facturation, sur ces derniers. Voici les informations dont le FAI pourrait avoir besoin :
  - L'identité du client ;
  - La facturation courante du client ;
  - Le DSLAM du client ;
  - Le port du DSLAM auquel le client est connecté ;
  - L'adresse MAC de la Set Top Box du client ;
  - Les chaînes auxquelles le client est abonné.
- Le serveur Authentication, Authorization, Accounting (AAA) : Le serveur AAA est interrogé par le client AAA pour authentifier un client lorsque celui-ci accède au réseau du FAI. Le rôle du serveur est d'authentifier le client, et de s'assurer de la continuation de l'authentification tout au long de la connexion du client. Le serveur AAA informe aussi l'annuaire de l'achat de l'accès à un film par un client. Le protocole AAA peut être, par exemple, RADIUS [35] ou DIAMETER [36] ;
- Le serveur de vidéos : Le serveur VoD stocke l'ensemble des vidéos proposées par le FAI. Il possède

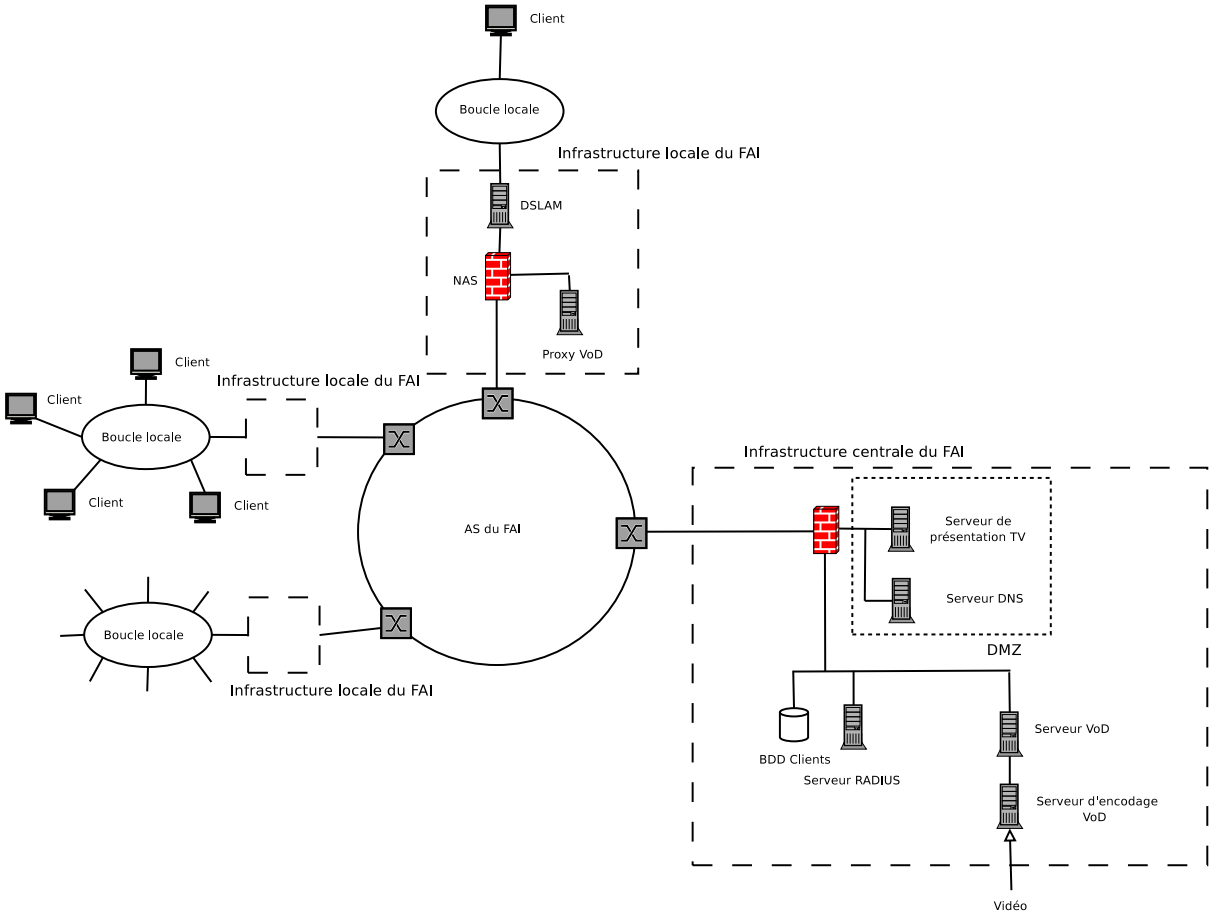


FIG. 3.2 – Architecture du réseau du FAI

des capacités de diffusion lorsqu'une requête appropriée lui est délivré ;
- Le serveur d'encodage vidéo : Le rôle de ce serveur est de récupérer la vidéo depuis son support de stockage original, de l'encoder dans le format de diffusion approprié, puis de transmettre la vidéo au serveur de VoD ;
- Le serveur de présentation TV : Ce serveur offre un service de présentation d'informations pour le client, consultables sur sa télévision. Pour le service VoD, il offre un service de présentation des films, sous forme de serveur Web, auquel la Set Top Box se connecte afin de présenter au client une interface de sélection et de renseignement sur les films ;
- Le serveur de nom ou serveur DNS : Les clients accèdent à ce serveur afin d'effectuer la correspondance entre nom de domaine et adresse IP.

Réseau du FAI, partie locale

Et voici les éléments déployés localement, pour chaque boucle locale connectée au réseau :
- Le Network Acces Server (NAS ou client AAA) : C'est le NAS que les clients interrogent lorsque ceux-ci veulent se connecter. Le NAS soumet alors les informations d'authentification au serveur AAA. En cas d'authentification réussie, le NAS relie le client au reste du réseau du FAI. Pour cela, il attribue une adresse IP à la Set Top Box et modifie les règles de contrôle d'accès afin que la Set Top Box puisse accéder au réseau ;
- Le proxy VoD : Le proxy VoD est interrogé directement par un client et lui diffuse la vidéo. Lors de la mise à disposition d'une nouvelle vidéo par le FAI, le proxy va télécharger cette vidéo auprès du serveur de vidéo.

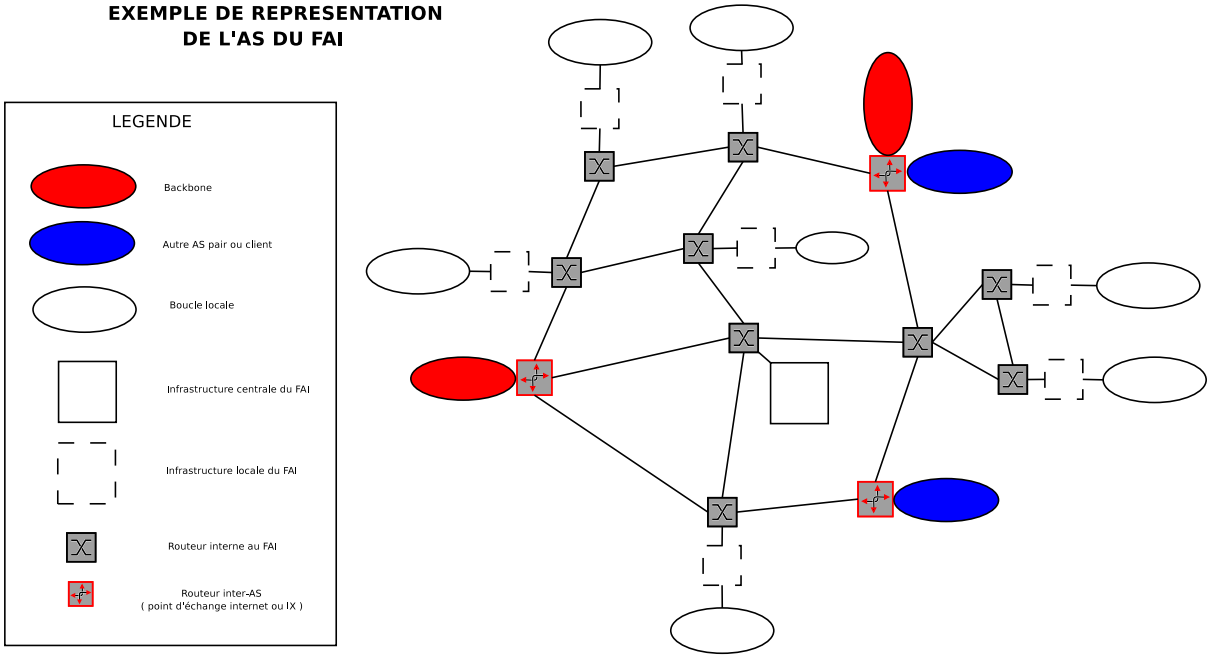


FIG. 3.3 – L'AS du FAI

Réseau du FAI, l'Autonomous System

Voici maintenant la description de l'architecture du réseau interne du FAI ou Autonomous System (AS) du FAI. Celui-ci assure la connectivité entre les différents DSLAM, les services fournis par le FAI et le réseau internet. Son architecture est décrite dans la figure 3.3. Il est composé de différents éléments :

- Les routeurs internes au FAI : Ce sont les routeurs destinés à assurer la connectivité entre les éléments appartenant au FAI, c'est-à-dire : Le réseau central du FAI et les clients reliés au DSLAM du FAI. On peut noter que les routeurs peuvent être connectés à la boucle locale, connectés au réseau central, mais aussi uniquement connectés à d'autres routeurs ;
- Les routeur inter-AS : Le routeur inter-AS est un routeur dont le rôle est de transmettre les communications entre différents AS. Il assure donc la connectivité du FAI avec le reste du réseau Internet ;
- Le routeur reliant les différents AS s'appelle IX (pour Internet eXchange point) et n'appartient pas en général au FAI. Par conséquent, on appellera ici routeur inter-AS le dernier routeur relié à l'IX et appartenant au FAI, ou l'IX lui-même si celui-ci appartient au FAI.

De plus, on appelle [39] :

- Backbone un réseau « de cœur »de l'Internet, c'est-à-dire un réseau dédié au communications longues distances et possédant des capacités de débits très élevés ;
- Autres AS, les AS appartenant à d'autres FAI ou organisations. Ceux-ci peuvent être clients (ils payent pour transiter par l'AS du FAI, en particulier pour accéder au backbone) ou pairs (l'AS du FAI et l'autre AS sont associés).

Les différentes interactions entre les éléments du réseau du FAI pour l'accès au réseau sont résumés par la figure 3.4, pour l'accès au film VoD par la figure 3.5. On notera qu'il a été choisi d'utiliser les protocoles RADIUS ACCONTING [37] et RTSP [38], qui sont des standards de l'IETF.

3.3 Pannes et attaques

On décrira dans cette section les pannes et attaques pouvant mettre en péril le bon fonctionnement du système et perturber la délivrance des services à l'utilisateur. Ainsi, dans le cas des attaques volontaires sur le système, nous nous concentrerons uniquement sur les attaques de déni de service. C'est en effet

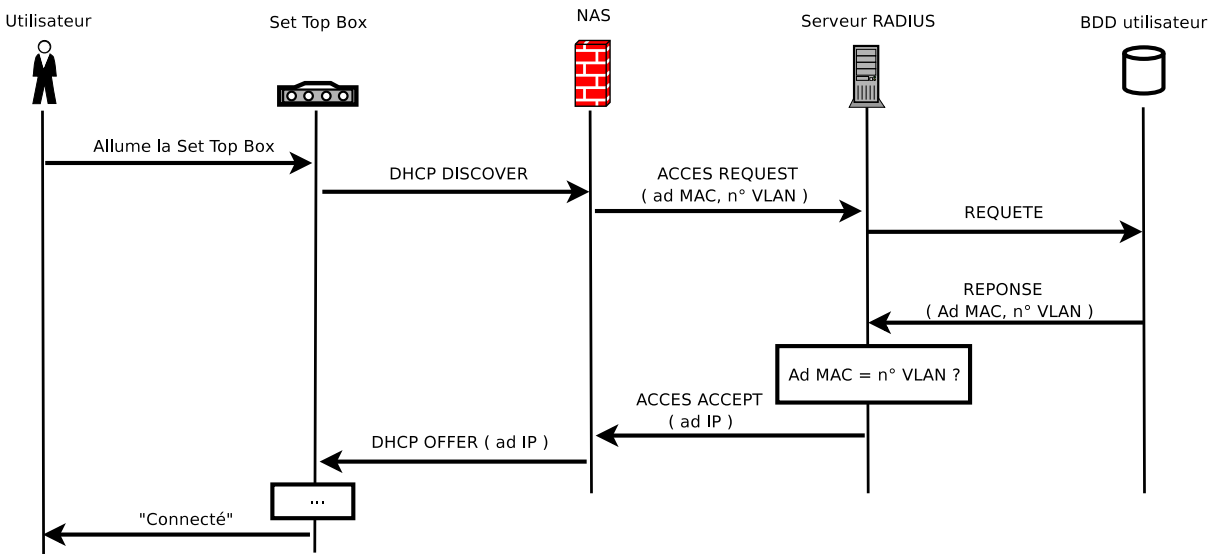


FIG. 3.4 – Scénario de connexion détaillé

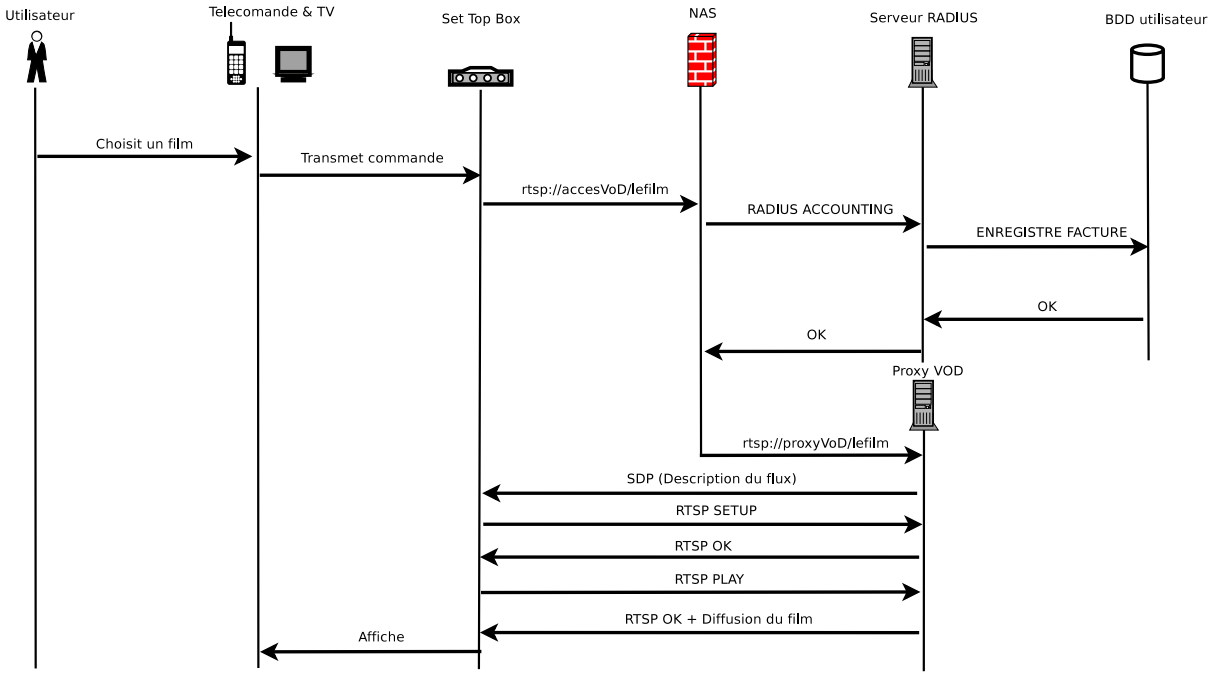


FIG. 3.5 – Scénario de l'accès VoD détaillé

Equipement	Accès au réseau	Service VoD
Set Top Box	Pour un client	
Boucle locale	Pour un ensemble clients	
DSLAM	Pour un ensemble de clients	
NAS	Pour un ensemble de clients	
Proxy VoD		Pour un ensemble de clients
Liens ou routeurs de l'AS	Pour un ensemble de clients	
Serveur VoD		Impossibilité de fournir de nouveaux films à tous les clients
Serveur RADIUS	Pour tous les clients	
Base de donnée client	Pour tous les clients	
Serveur de présentation TV		Pour tous les clients
Serveur DNS		Pour tous les clients

TAB. 3.1 – Nombre de clients concernés par l'indisponibilité des services en cas de panne d'un équipement

ce type d'attaque qui perturbe la disponibilité des services et peut par conséquent être assimilé à une panne. On ne traitera donc pas dans les attaques dont le but est, par exemple, l'espionnage ou la fraude.

3.3.1 Pourquoi a-t-on des besoins en sécurité ?

Les menaces :

On va étudier les menaces sur la disponibilité des services qui pèsent sur le système. On appelle menace un présage concernant la réalisation d'un fait qui va nuire au système. On appelle l'indisponibilité du service, l'impossibilité pour un client de se voir délivrer un des services normalement rendu par le système.

Source des menaces :

Les sources des menaces, c'est-à-dire les causes qui font que le système est menacé, sont multiples. On peut les séparer en deux catégories : Les sources liées à l'environnement du système et les sources liées aux attaquants potentiels du système.

Parmi les sources environnementales, on peut distinguer les pannes [40] (l'arrêt accidentel de fonctionnement d'une partie du système) dont les causes peuvent être multiples (dysfonctionnement interne au système, logiciel ou matériel, ou externe, panne de courant ou catastrophe naturelle) ou encore des erreurs dues au personnel et qui entraînent la compromission du système [41].

On peut de même distinguer plusieurs profils d'attaquants dont le but est de mettre en péril la disponibilité du système. Ils peuvent être des rivaux économiques, être motivés par le jeu, ou même par le vandalisme.

Cible des attaques :

Voici plus précisément la cible potentielle qui sera l'objet de la réalisation de la menace, c'est-à-dire la cible de l'attaque de déni de service :

- Sur l'infrastructure de récupération de la vidéo : Consiste à perturber le fonctionnement de l'équipement chargé de récupérer les sources vidéos et de les injecter dans le réseau du FAI ;
- Sur l'infrastructure de diffusion de la vidéo : Consiste à perturber le fonctionnement de l'équipement chargé de diffuser les flux vidéos à travers le réseau du FAI jusqu'aux clients ;
- Sur l'accès à la vidéo par les clients : Consiste à perturber les mécanismes et équipements permettant l'accès des clients aux flux vidéos.

On présente dans le tableau 3.1 les éléments critiques de l'infrastructure, c'est-à-dire les éléments dont dépend le fonctionnement du service de VoD. Nous y verrons quels sont les éléments dont le non-fonctionnement entraînerait le non-fonctionnement général du service ainsi que les éléments dont le non-fonctionnement entraînerait son indisponibilité locale, limité à un nombre restreint de clients.

Point de vue du FAI :

Voici le point de vue du FAI concernant la gravité des problèmes engendrés par la réalisation des menaces : Le FAI, en échange d'une certaine somme d'argent, apporte un ensemble de services aux clients. Par conséquent, il est essentiel que la disponibilité de ces services soit assurée. En effet, dans le cas contraire, le FAI s'expose, dans le meilleur des cas, à une forte perte de crédibilité, et au pire, à un non-respect de ses engagement pris dans le contrat. Notons que les engagements du FAI sont pris auprès des clients, mais aussi auprès des fournisseurs de vidéos, qui pourraient se sentir lésés si, par exemple, leur chaîne de télévision n'était plus visible par les clients.

Ainsi le FAI souhaite éviter, par ordre décroissant des dommages engendrés, l'indisponibilité généralisée du service ou son indisponibilité locale, limitée à une partie des clients.

3.3.2 Menaces environnementales

On s'attachera dans cette partie à décrire plus en détail les causes et les conséquences des menaces dont la source est l'environnement du réseau du FAI. Le tableau 3.2 récapitule les différents éléments exposés ici.

Les pannes [40] :

Dans cette partie, on s'intéressera aux différentes pannes pouvant survenir dans le réseau, leurs conséquences, et les moyens de s'en prévenir. On peut classifier les pannes selon l'équipement qu'elles frappent :

- La panne touche un élément du réseau dédié à l'acheminement des communication. Ces éléments correspondent aux liens du réseau et aux routeurs de l'AS ;
- La panne touche un élément délivrant un service. Ces éléments correspondent aux serveurs et proxy VoD, les serveurs DNS, etc.
- Le NAS est un élément à part car il fournit un service (l'accès au réseau) mais fait aussi partit des éléments nécessaires à l'acheminement.

Il est possible de prévoir certaines pannes et ainsi d'appliquer des méthodes pro-actives pour s'en prémunir. Il est ainsi envisageable de prédire l'apparition d'une panne en évaluant la durée pendant laquelle un élément du système est conçu pour fonctionner correctement. Par exemple, on peut considérer que le disque dur possède une durée de vie de 3 ans, et qu'au bout de ces 3 ans, une panne est susceptible de se produire. Donc la panne est prévisible. On peut ainsi prévoir les pannes de manière probabiliste en caractérisant les probabilités d'apparition d'une panne sur les différents éléments du système. De plus, il est parfois possible de prévoir une panne si celle-ci découle d'un autre incident qui est apparu dans le système. Il faut pour cela caractériser les relations entre les différents incidents pouvant survenir dans le système.

Cependant, certaines pannes ne sont pas prévisibles. Par conséquent, les méthodes pour garantir une continuité de service lors de l'apparition de ces pannes seront plutôt réactives, c'est-à-dire déclenchées après l'apparition de la panne.

Voici des pistes pour assurer la continuité du service en cas de panne

- d'un élément du réseau dédié à l'acheminement : Il faut dans ce cas faire transiter le trafic par un autre élément. Pour que cela soit possible, il faut s'assurer qu'un autre élément est disponible pour assurer la connectivité entre les différents points du réseau. Cela est particulièrement important pour l'infrastructure centrale, qui fournit les services les plus critiques et ne doit pas être reliée par un seul point avec le reste du réseau. La détection des pannes ainsi que la modification des chemins de transit est normalement assurée par le protocole de routage du réseau ;
- d'un élément du réseau délivrant un service : Dans ce cas, il faut qu'un autre élément du réseau puisse apporter ce service, et que les requêtes pour ce service ne soit plus acheminées à l'élément défectueux mais à l'élément de remplacement. Il faut par conséquent s'assurer que les éléments apportant les services soient dupliqués.

Lors d'une panne, le système doit être placé dans un état de secours et continuer, dans la mesure du possible, à délivrer les services. L'état de secours peut impliquer une détérioration des services fournis si cela est nécessaire pour continuer à assurer leur livraison. Par exemple, si la panne d'un élément de transit du réseau entraîne une diminution de la bande passante disponible pour l'acheminement des

vidéos, on peut envisager de détériorer la qualité des vidéos afin de réduire les besoins en bande passante et continuer à délivrer le service. Il est aussi possible de vouloir fermer temporairement l'accès au service afin de réparer la panne dans de meilleurs délais.

Personnel en erreur [41] :

Le personnel chargé de la conception, de l'implémentation, du déploiement et de la maintenance du système peut être une source de vulnérabilités si ceux-ci commettent des erreurs. Voici une liste des erreurs et problèmes potentiels qui peuvent être issus des employés du FAI.

Fautes humaines On listera ici les vulnérabilités dues à des erreurs humaines. On appelle erreurs humaines l'exploitation par un attaquant de la position priviligiée d'un employé du FAI. Cette exploitation va permettre la réalisation ultérieure d'une attaque. La plupart de ces problèmes peuvent être évités en informant et éduquant les employés sur les menaces possibles. Dans le cas du FAI, le nombre de personnes ayant réellement accès à la configuration du réseau est limité, il donc assez facile de sensibiliser ces personnes aux problèmes de sécurité.

On peut citer différents cas de fautes humaines. Par exemple, le fait d'amener un employé a divulger des informations secrêtes ou à configurer les équipements du réseau au profit d'un attaquant par la corruption ou la menace. Cette situation est par ailleur assez similaire au Social Engineering, où l'attaquant, pour arriver à ses fins, se fait passer pour une personne qu'il n'est pas auprès de l'employé. On peut encore citer les problèmes qui peuvent être liés aux mots de passe, lorsque ceux-ci sont faciles à deviner ou à récupérer sur le lieu de travail de l'employé par exemple.

Fautes techniques On listera ici les erreurs techniques pouvant être commises par un employé du FAI. Les erreurs techniques concernent un plus grand nombre d'employés que les erreurs humaines, car ces erreurs ne reposent pas sur la divulgation d'un secret. Ces erreurs peuvent intervenir tout au long du processus de conception, de développement, de déploiement et de maintenance et ne concernent donc pas uniquement les techniciens en charge du réseau. Une protection envisageable est la mise en place d'un processus de mise en production permettant la vérification des produits avant leur exploitation. Cette vérification peut être humaine, automatique, inclure des tests, etc. Cependant, il faut avoir conscience qu'une vérification ne s'opère que sur un élément particulier et dans un certain contexte. La vérification n'apporte donc pas de garantie absolue sur le fonctionnement du système.

Parmi les fautes techniques, on peut citer par exemple :

- Différence entre la spécification et l'implémentation : Une différence entre la spécification d'un produit (ce qu'il doit faire) et son implémentation (ce qu'il fait) peut amener à la présence de failles de sécurité exploitables par un attaquant. La programmation formelle, qui va permettre de vérifier l'implémentation mathématiquement, permet d'empêcher ce type de problème. Cependant, il est difficile de concevoir de tels programmes, et il est peu envisageable de concevoir de façon formelle un système entier si celui-ci est complexe et dynamique ;
- Fonctions non documentées : La présence de fonctionnalités non documentées (dans un programme, dans un protocole, etc.) peut entraîner la non prise en compte d'une menace de sécurité pesant sur le produit, donc l'absence de mécanismes de sécurité pour pallier à cette menace ;
- Erreur sur les droits d'accès : Une mauvaise attribution des droits d'accès au différents éléments du réseau (machine, périphérique, fichier ...) peut entraîner des problèmes de confidentialité (si un fichier confidentiel peut être lu par quelqu'un de l'extérieur du réseau, par exemple) et peut même amener à la compromission d'une machine si les droits d'accès à une ressource importante sont inadéquats (droit de lecture sur un fichier de mot de passe par exemple) ;
- Attaque lors de la maintenance : Lors de la maintenance d'un élément du réseau (ou lorsqu'une procédure d'urgence est lancée, en cas de panne par exemple), il est possible que le système fonctionne dans un état moins sécurisé que dans l'état normal. Cette phase de maintenance peut donc être exploitée pour lancer une attaque ;
- Mauvaises valeurs par défaut : Les valeurs par défaut d'un système d'exploitation ou d'un logiciel ne sont pas changées. Ces valeurs peuvent être connues d'un attaquant qui peut utiliser ces informations pour mener une attaque.

Source de vulnérabilité	Cas	Défense
Panne	Sur un élément dédié à l'acheminement	Modifier les routes d'acheminement
Panne	Sur un élément dédié à la livraison d'un service	Rediriger les requêtes à ce service vers un élément redondant qui fournit ce service
Personnel en erreur	Erreur humaine	Formation du personnel
Personnel en erreur	Erreur technique	Vérification avant mise en exploitation

Tab. 3.2 – Les menaces environnementales : Risques et protection

3.3.3 Les attaques par déni de service

Dans cette partie, on présentera les attaques de déni de service possibles, une évaluation de leurs faisabilité, et les défenses envisageables. Le tableau 3.3 récapitule les différents points abordés dans cette partie.

Description :

Un déni de service consiste à perturber le fonctionnement normal du système dans le but d'empêcher l'accès à un service. Le but d'une telle action peut être le vandalisme (c'est-à-dire que le but est uniquement d'empêcher l'accès au service), mais aussi de perturber le fonctionnement du système afin de mener une autre attaque. Enfin, un concurrent du FAI peut vouloir réaliser un déni de service afin de discréditer le FAI.

Les conséquences de la réalisation d'un déni de service sont donc l'impossibilité d'accès à l'ensemble ou une partie du service, pendant une durée variable.

Attaques possibles :

Un attaquant voulant perturber l'accès aux services vidéos du FAI peut envisager plusieurs approches pour parvenir à ses fins :

- Rendre indisponible l'infrastructure de récupération de la vidéo, c'est-à-dire perturber le fonctionnement de l'équipement chargé de transmettre les sources vidéos dans le réseau du FAI. Cette infrastructure n'est normalement pas sous la responsabilité du FAI mais à la charge des vendeurs de contenu vidéo, c'est-à-dire les chaînes de télévision ou les vendeurs de film du service VoD. Il est donc important que le FAI et les vendeurs de contenus mettent en place des mesures de protection conjointes pour assurer la robustesse de cette infrastructure ;
- Rendre indisponible l'infrastructure de diffusion de la vidéo, c'est-à-dire perturber le fonctionnement de l'équipement chargé de diffuser les flux vidéos à travers le réseau du FAI jusqu'aux clients. C'est éléments sont : le serveur VoD, le réseau du FAI qui achemine ces vidéos jusqu'aux infrastructures locales ;
- Rendre indisponible l'accès à la vidéo par les clients, c'est-à-dire perturber les mécanismes et équipements permettant l'accès des clients aux flux vidéo, c'est-à-dire la Set Top Box et les proxies et VoD qui distribuent ces flux jusqu'aux clients, ainsi que les protocoles utilisés pour la diffusion de ces flux.

Il existe plusieurs approches afin de réaliser un déni de service :

- Perturbation physique :
  Consiste à agir physiquement sur le système, de manière à empêcher son fonctionnement. Pour cela, il est par exemple possible de couper les câbles de communication, l'alimentation électrique, voir d'atteindre directement une machine du système.
- Surcharge du système :
  Consiste à saturer volontairement le système de manière à ce que les services ne puissent être rendus à tous les clients. C'est la façon la plus répandue de procéder à une attaque de déni de service. Les machines les plus menacées par ce type de d'attaque sont les machines publiques, accessible depuis

l'extérieur du réseau du FAI, c'est-à-dire le plus souvent les serveurs Web et DNS, mais aussi les machines temporairement accessibles par les clients, comme les proxy vidéo.
On peut envisager, par exemple, les attaques suivantes :
- non-terminaison d'une communication : On établit un début de communication avec un serveur de manière à ce que ce dernier réserve des ressources pour la communication mais on ne termine pas cette communication proprement. Les ressources réservées ne sont donc pas libérées. Si cette opération est répétée un grand nombre de fois, le serveur n'a plus de ressources disponibles. Un exemple connu de cette attaque est l'attaque par TCP SYN [42] ;
- Saturation du système [43] : On effectue un grand nombre de requête à un service de manière à saturer ses ressources. Les ressources peuvent être : La bande passante, la mémoire, le processeur, etc.
- De plus, l'attaque par surcharge du système peut être opérée de façon distribuée et coordonnée [44], c'est-à-dire que l'attaque utilise un ensemble de machines subsidiaires pour lancer l'attaque par surcharge en même temps. Cette méthode, bien plus efficace qu'une attaque unique, présente de plus l'avantage qu'il est difficile de remonter jusqu'à la source réelle de l'attaque.

- Exploitation des (inter)dépendances du système :
On reporte l'attaque sur un élément du système dont dépend le service que l'on veut atteindre. En effet, il peut arriver que certains éléments dont dépende un service soient moins bien protégés que les éléments attaqués. On peut envisager plusieurs scénarios pour ce type d'attaque : Une série d'attaque sur des éléments mineurs du système qui vont provoquer une série de pannes mineures mais dont va finalement découler une panne majeure du service. Le cas d'une attaque sur un point faible du système mais dont le service à atteindre (un point fort) dépend correspond bien à l'attaque sur l'infrastructure de récupération des vidéos qui peut être effectivement moins bien protégée que le réseau du FAI lui-même.
- Exploitation d'une faille de sécurité [45] :
Consiste à exploiter une faille de sécurité dans le but de perturber le bon fonctionnement du système. La faille de sécurité peut concerner un protocole, mais aussi les logiciels et matériels utilisés par une machine du système. Pour effectuer un déni de service on peut envisager les exploitations suivantes :
  - On fournit à un logiciel des données illégales mais acceptées par l'implémentation afin d'amener le logiciel dans un état de non-fonctionnement
  - On désynchronise les flux basés sur le temps. On manipule les flux de manière à les désynchroniser et les rendre illisibles pour le récepteur. Cette attaque est en particulier envisageable sur les protocoles de transport de flux vidéo temps réel. Par exemple, la modification des numéro de séquence ou des estampilles temporelles des paquets RTP [47] entraînerait l'impossibilité pour le client de lire la vidéo.

Pour se protéger des attaques de déni de service, plusieurs défenses sont envisageables :
- Protéger physiquement l'infrastructure : Il s'agit de protéger l'accès aux câbles, aux machines, ainsi que prévoir des protection en cas de perturbation environnementales (coupures électriques, etc.) ;
- Introduire des mécanismes de détection d'intrusion et d'attaques distribuées : Il existe aujourd'hui plusieurs systèmes de détection d'intrusion (IDS) qui permettent de détecter certainnes attaques de déni de service [48]. Il faut cependant envisager qu'un attaquant puisse contourner ce système s'il connaît son fonctionnement. Ces systèmes peuvent utiliser les logs et le trafic pour détecter une attaque. On peut cependant se demander s'il est possible et comment détecter l'intrusion une fois l'attaque réussie ;
- Analyser les dépendances des services de façon à déterminer leur vulnérabilité et les protéger en conséquence ;
- Se tenir au courant des failles de sécurité des protocoles et des logiciels et procéder à leur mise à jour si une menace est découverte [46].

Il faut aussi envisager les conséquences d'une réussite d'une telle attaque. Pour cela, puisque la réussite d'une attaque de déni de service revient à l'apparition d'une panne sur le réseau, les mécanismes de protection utilisés seront les mêmes.

Attaque	Méthode	Défense
Déni de service	Perturbation physique	Protection de l'accès physique au réseau
Déni de service	Surcharge provoquée	Surveiller le bon comportement du système (IDS ...)
Déni de service	Exploitation des (inter)dépendances	Analyse et protection des éléments dont dépendent les services
Déni de service	Exploitation d'une faille de sécurité	Surveiller les failles potentielles, mise à jour si besoin, Surveiller le bon comportement du système (IDS ...)

TAB. 3.3 – Les attaques : Risques et défenses

3.4 Assurer la disponibilité des services en cas de panne

3.4.1 Les besoins

Après avoir étudié les éléments du réseau nécessaires à la délivrance des services ainsi que les menaces, pannes et attaques, qui pèsent sur la délivrance de ces services, nous allons maintenant mettre en évidence les principaux besoins pour assurer leur disponibilité en cas de réalisation de ces pannes ou attaques.

Nous l'avons vu, le non-fonctionnement de certains éléments du réseau entraîne l'impossibilité de délivrer le service aux clients. Ces éléments sont typiquement les serveurs, comme le serveur VoD. Il est donc nécessaire de répliquer en plusieurs points du réseau ces serveurs, de manière à s'assurer qu'en cas de panne de l'un de ces éléments, il soit encore possible de délivrer le service en un point du réseau.

3.4.2 Réplication des éléments sensibles de l'infrastructure

Voici quelques propositions pour assurer une meilleure disponibilité des services en cas de panne des différents éléments de l'infrastructure du FAI. En cas d'indisponibilité du réseau central, l'ensemble des accès aux services sont compromis. Pour pallier cela, on va répliquer dans les infrastructures locales les éléments indispensables au bon fonctionnement du service, sous forme de serveurs secondaires appelés ici proxy. Cette nouvelle infrastructure est représentée dans le schéma 3.6 :

- Un proxy RADIUS : Il sert d'intermédiaire entre le NAS et le serveur RADIUS. Le proxy fait parvenir l'ensemble des requêtes au serveur RADIUS ;
- Un proxy BDD client : Il contient la réplication des informations de la BDD centrale qui concerne les clients reliés à l'infrastructure locale ;
- Un proxy de présentation : C'est un proxy HTTP standard, avec mise en cache des pages Web ;
- Un proxy DNS : C'est un serveur DNS standard dont la position dans la hiérarchie des serveurs DNS est en dessous du serveur DNS central.

3.4.3 Comportement en cas de panne

Il est maintenant nécessaire d'envisager quel serait le comportement des nouveaux éléments du système en cas d'apparition d'une panne dans le réseau. Nous allons tout d'abord étudier le comportement normal du système, puis son comportement en cas de non-fonctionnement d'une partie du réseau.

Voici le comportement des proxy en cas de bon fonctionnement de l'infrastructure centrale :

- Le proxy RADIUS intercepte les requêtes RADIUS. Il peut normalement répondre à la requête en consultant et/ou modifiant le proxy BDD client. Il peut éventuellement faire parvenir une requête à laquelle il ne peut répondre au serveur RADIUS ;
- Le proxy BDD client : La BDD locale contenue dans le proxy BDD est synchronisée avec la portion de la BDD centrale qui lui correspond. Chaque modification effectuée sur l'une des BDD est donc immédiatement répercutée sur l'autre ;
- Le proxy de présentation : Le proxy de présentation est interrogé directement par les clients. S'il ne possède pas la page Web demandée en cache, il interroge le serveur de présentation TV et place

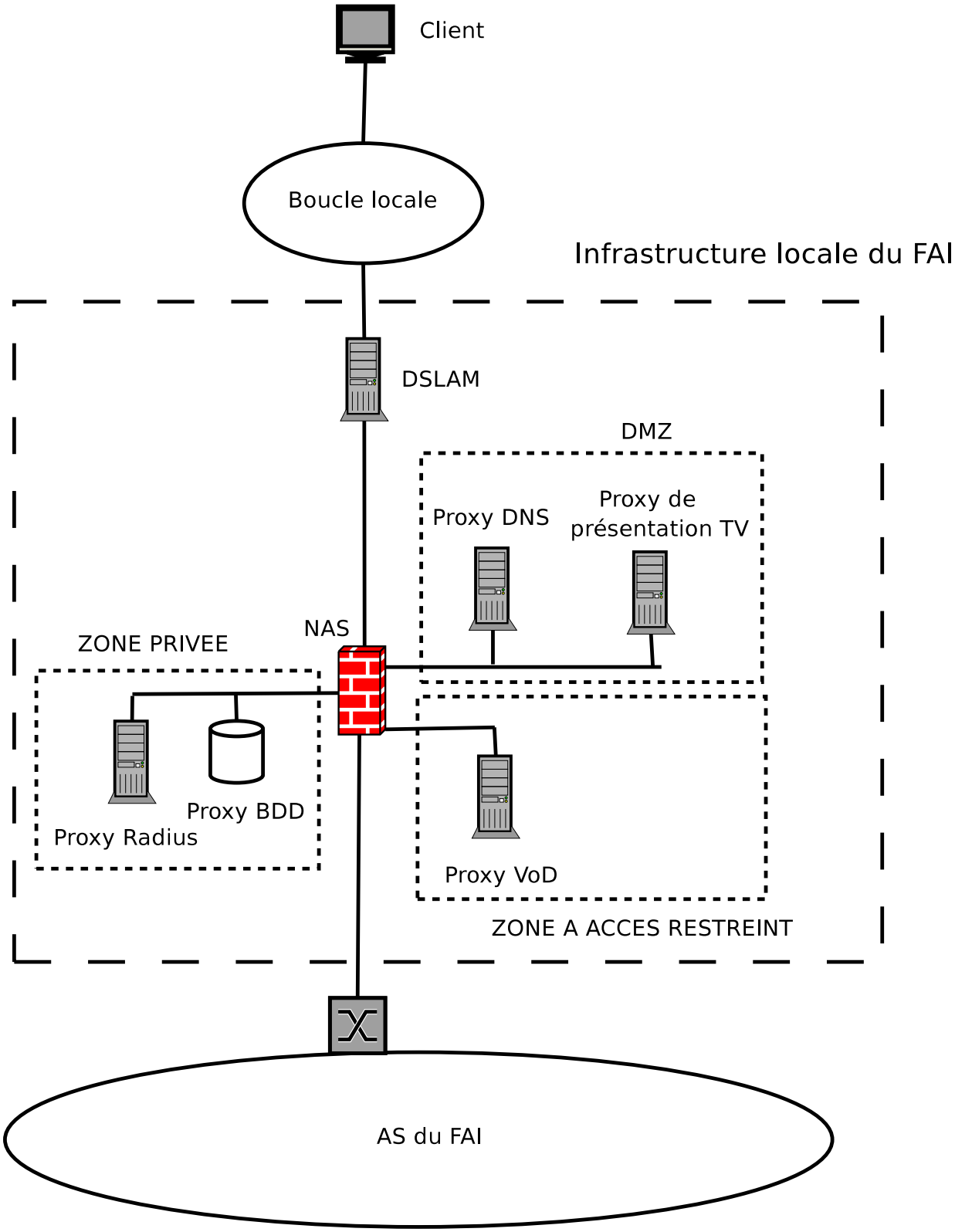


FIG. 3.6 – Nouvelle infrastructure locale

cette page en cache. De plus, il interroge régulièrement le serveur de présentation TV pour vérifier si des pages qu'il possède en cache ont été modifiées ;
- Le proxy DNS : Il répond directement aux requêtes des clients s'il connaît l'adresse associée au nom de domaine demandée. Sinon il interroge le serveur DNS central et met en cache sa réponse.

Voici le comportement des proxies en cas de non-fonctionnement de l'infrastructure centrale :
- Le proxy RADIUS, s'il ne peut répondre à une requête à l'aide des informations de la BDD locale, peut par exemple directement accepter la requête. Ainsi, on assure la disponibilité du service au détriment de la facturation ;
- Le proxy BDD client : La BDD locale attend que la BDD centrale soit à nouveau disponible pour se synchroniser avec elle et ainsi lui transmettre les modifications ayant eu lieu pendant son indisponibilité ;
- Le proxy de présentation : S'il ne possède pas la page demandée, il interroge un ou plusieurs autres proxies de présentation présents dans une ou plusieurs autres infrastructures locales ;
- Le proxy DNS : S'il ne connaît pas l'adresse associée au nom de domaine demandée, il peut :
  - Interroger un ou plusieurs autres proxy DNS présents dans une ou plusieurs autres infrastructures locales ;
  - Interroger un serveur DNS « public », accessible sur Internet, mais uniquement pour un nom de domaine n'appartenant pas au réseau du FAI.

Il faut maintenant envisager le comportement du réseau en cas d'indisponibilité des éléments de l'infrastructure locale, afin de garantir la disponibilité des services localement. Voici donc le comportement du système en cas d'indisponibilité de :
- La Set Top Box : Une panne de la Set Top Box entraîne la non-disponibilité des services pour un client. Les solutions envisageables sont :
  - La possibilité de connecter un autre modem pour accéder à Internet ;
  - Le remplacement de la Set Top Box.
- La boucle locale : La boucle locale n'étant en général pas la propriété du FAI, ce n'est pas à lui d'assurer son bon fonctionnement. On peut cependant envisager la mise à disposition au client d'un accès au réseau autre que xDSL, WiMAX (boucle locale radio) par exemple ;
- Le DSLAM : Le DSLAM est l'élément qui relie « physiquement »le réseau xDSL au réseau IP. C'est pourquoi en cas de non-fonctionnement, on ne peut éviter la perte de l'accès au réseau pour l'ensemble des clients reliés au DSLAM. La seule solution envisageable est la duplication du DSLAM ;
- Le NAS : On l'a vu, le NAS est un élément crucial dans la sécurité du réseau. Ainsi, en cas de panne de ce dernier, il faut analyser la politique de contrôle d'accès. On peut envisager de refuser tout le trafic, de refuser uniquement l'accès aux proxy de l'infrastructure locale ou encore de ne refuser aucun trafic. Quoiqu'il en soit, le non-fonctionnement du NAS entraînerait une forte perturbation des services du FAI et il convient donc d'assurer le plus possible la continuité de son fonctionnement, par exemple en le dupliquant ;
- Le proxy VoD : En cas de non-fonctionnement du proxy VoD, le NAS peut rediriger les requêtes qui lui étaient initialement destinées vers d'autres proxy VoD présents dans d'autres infrastructures locales. Le NAS doit essayer de répartir ces requêtes sur plusieurs proxy VoD afin de ne pas en surcharger un seul ;
- Le lien infrastructure locale - AS du FAI : La perte de ce lien entraînerait l'indisponibilité de l'ensemble des services, excepté la VoD, pour l'ensemble des clients connectés à l'infrastructure locale. Pour diminuer ce risque, il faut que l'infrastructure locale soit reliée à l'AS par au moins deux liens ;
- Le proxy RADIUS : En cas de non-fonctionnement du proxy RADIUS, le NAS redirige les requêtes RADIUS vers le serveur RADIUS de l'infrastructure centrale ;
- Le proxy BDD client : En cas de non-fonctionnement du proxy BDD client, les modifications et/ou consultations effectuées sur la base de donnée locale sont redirigées vers la BDD centrale. Cette dernière attend que la BDD locale soit à nouveau disponible pour se synchroniser avec elle et ainsi lui transmettre les modifications ayant eu lieu pendant son indisponibilité ;
- Le proxy de présentation TV : En cas de non-fonctionnement du proxy de présentation, les requêtes sont redirigées vers le serveur de présentation ou un autre proxy de présentation d'une autre infrastructure locale ;
- Le proxy DNS : En cas de non-fonctionnement du proxy DNS, les requêtes sont redirigées vers le

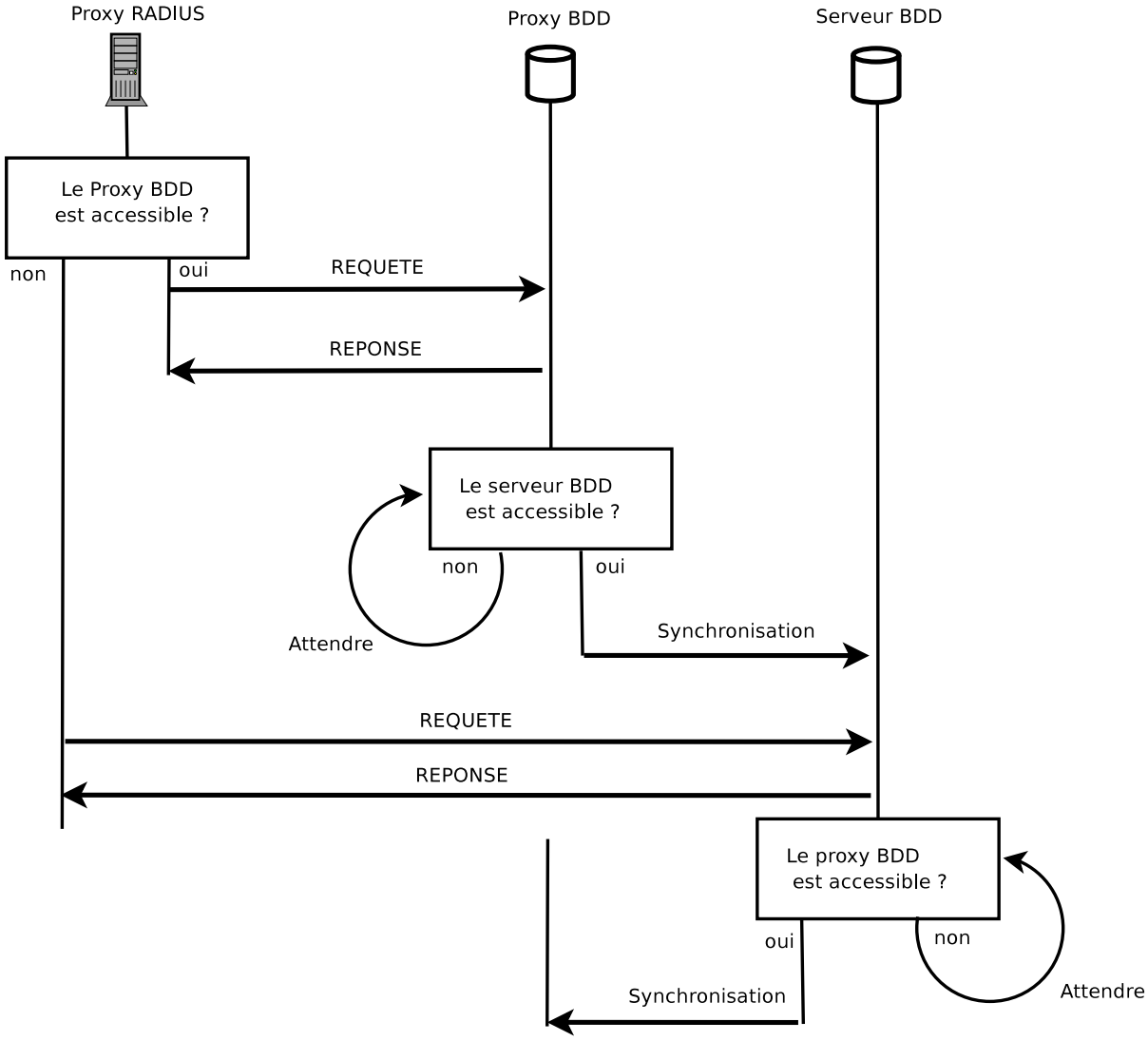


FIG. 3.7 – Comportement du proxy BDD

serveur DNS ou un autre proxy DNS d'une autre infrastructure locale.

Les schémas 3.7, 3.8, 3.9 et 3.10 récapitulent les comportements des différents éléments de l'infrastructure tel qu'ils ont été décris dans cette partie.

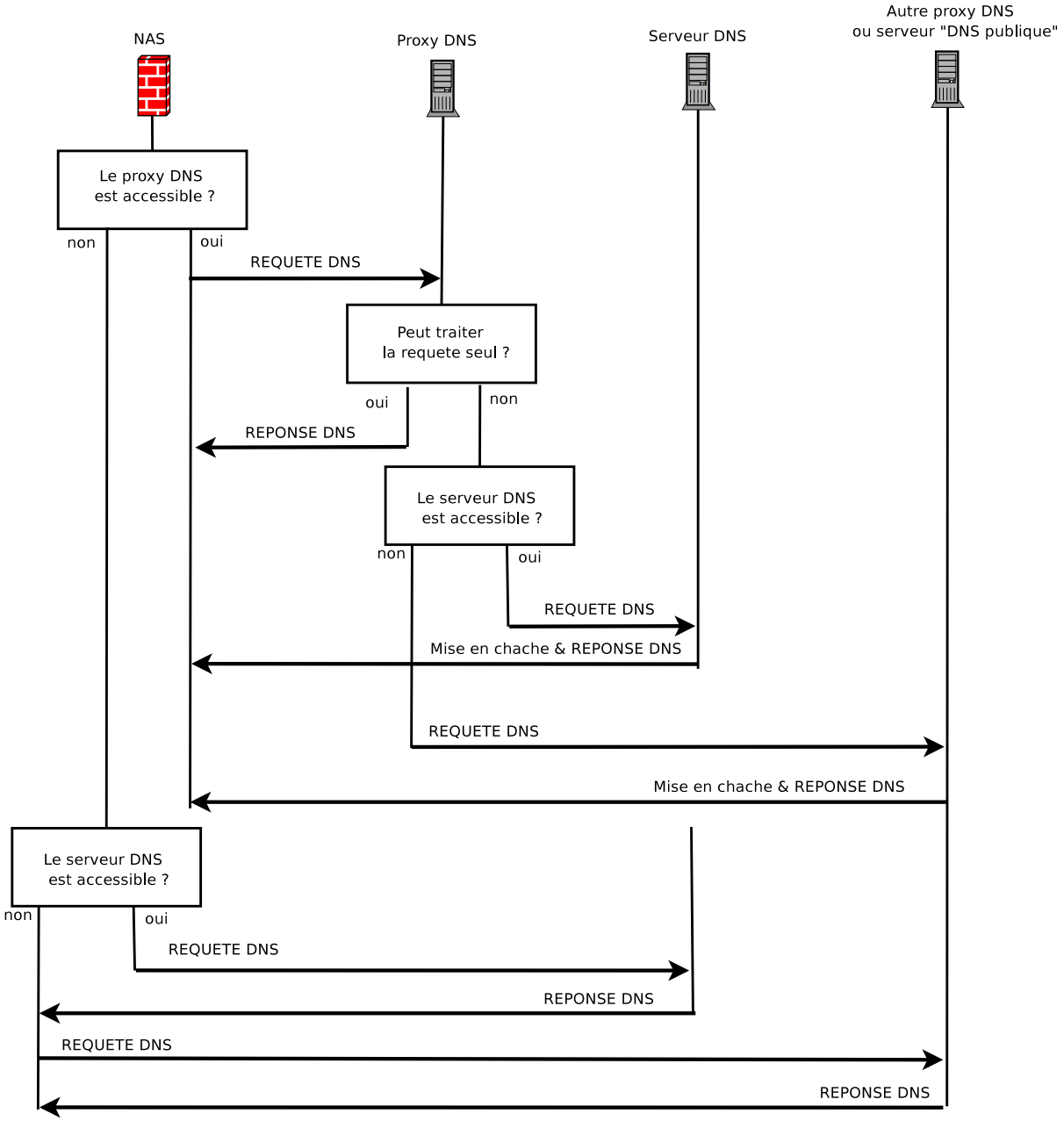


FIG. 3.8 – Comportement du proxy DNS

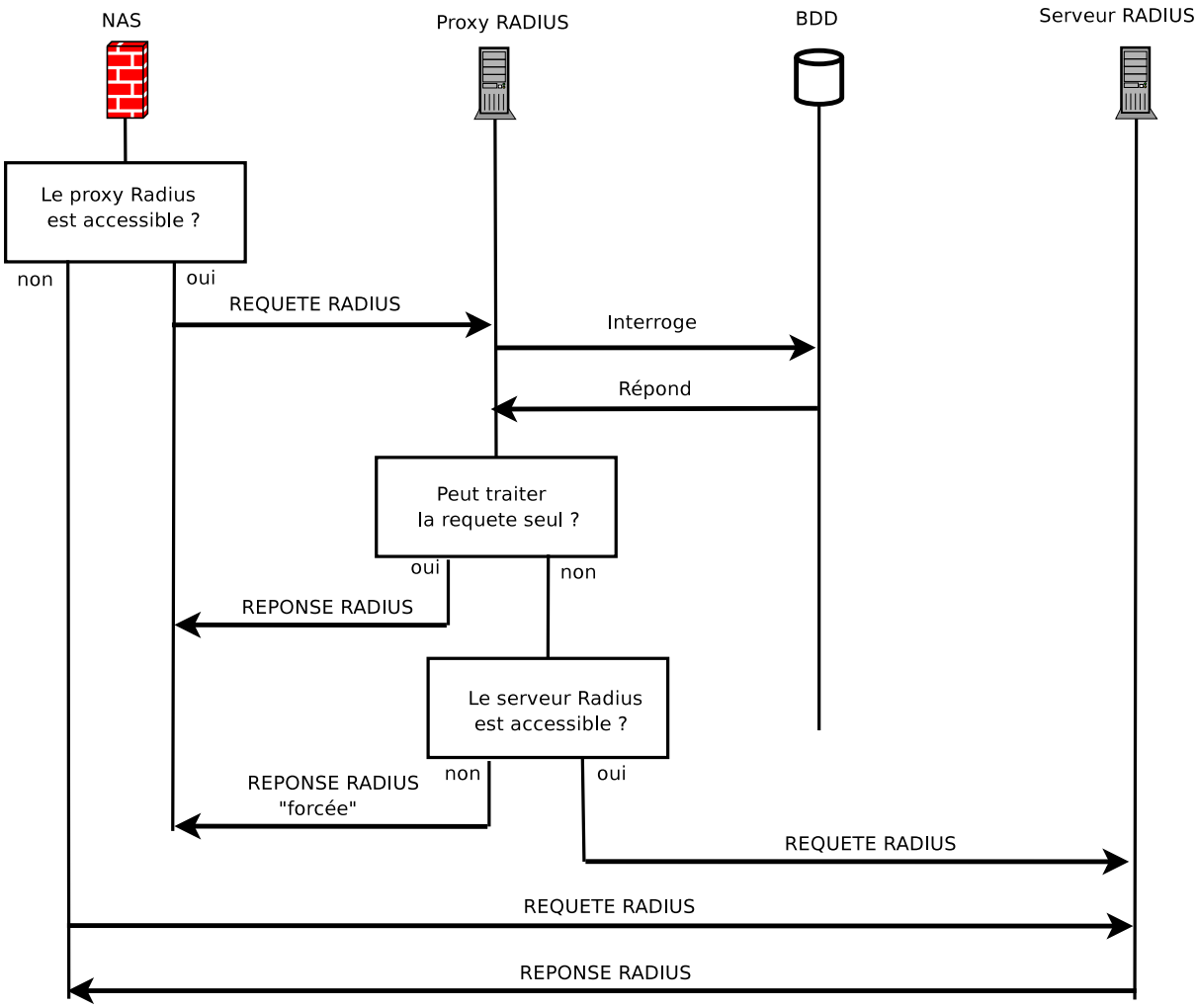


FIG. 3.9 – Comportement du proxy AAA

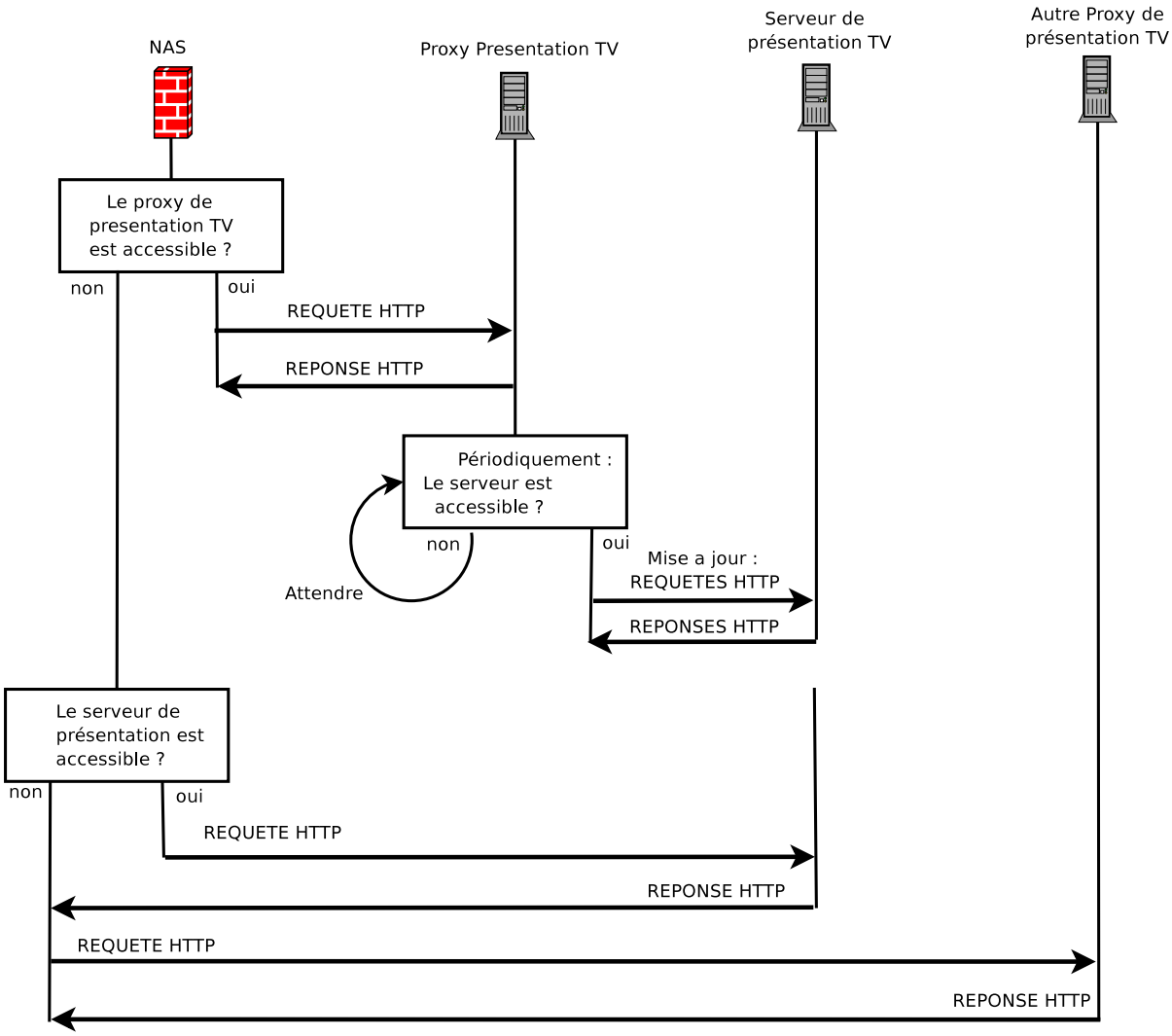


FIG. 3.10 – Comportement du proxy de présentation TV

Chapitre 4

Evaluation des capacités de récupération sur panne des protocoles de routage

4.1 Présentation

Le but de cette étude est d'analyser des protocoles destinés à conserver la connectivité entre les nœuds d'un réseau lors de l'apparition d'une panne. Ainsi, le but à atteindre est d'assurer la délivrance des services à l'utilisateur en cas de panne d'un élément de l'infrastructure du réseau, routeur ou lien. Nous nous concentrerons dans cette étude à évaluer les forces et les faiblesses de deux protocoles de routage à apporter de la robustesse.

Ainsi, il faudra étudier la rapidité avec laquelle les routes seront reconstruites pour apporter la connectivité des nœuds Il faudra de plus étudier la façon dont la surcharge de bande passante induite par le protocole de routage influe, et en quelles proportion, sur la vitesse de récupération sur panne.

On étudiera le protocole OSPF. Nous essayerons d'en évaluer ses atouts, ses inconvénient, ainsi que ses possibilités de paramétrage afin d'apporter une meilleur disponibilité. Nous soulignerons le fait qu'OSPF est, dans la plupart des cas, inadapté à fournir une haute disponibilité des services. Il a été choisis d'étudier OSPF car c'est l'un des protocoles les plus couramment utilisé. De plus c'est un protocole standard de l'IETF et son implémentation est très accessible, aussi bien sur les équipements de l'industrie que sur les plates-formes open source. OSPF a donc été préféré à RIP, qui est un protocole ancien et quelque peu dépassé et à IS-IS, qui est de fonctionnement proche de celui d'OSPF, mais qui présente l'inconvénient de ne pas être largement déployé.

Ensuite, nous étudierons RON, un protocole de routage overlay. Nous évaluerons quels sont les avantages à utiliser un tel système par rapport aux protocoles de routage classiques. Nous essayerons de comprendre si RON, initialement conçu pour être déployé sur Internet, donc sur différents AS, apporte de par sa conception originale des atouts transposables à une infrastructure unique. En effet RON a été initialement conçu pour apporter une sélection de la meilleure route en fonction d'une métrique choisie entre différent AS. De cette façon, il permet de « court-circuiter »les routes imposées par BGP dans le routage Internet standard, qui peuvent être soumises à des contraintes administratives et ne pas être optimales. Nous essayerons d'éclaircir les points suivant :

- Le mécanisme de détection des pannes de RON est plus agressif que celui d'OSPF, cependant, le probing s'effectue au dessus de plusieurs liens, puisque l'architecture de RON est un overlay. Nous verrons comment cela influe sur les performances de re-routage dans un cas concret de délivrance de services à un client ;
- Puisque le réseau RON est un réseau overlay, RON n'est pas déployé sur tous les nœuds Par conséquent, il est intéressant de se demander si le temps de convergence du réseau RON en cas de re-routage est inférieur au temps de convergence d'OSPF, déployé sur tous les nœuds ;
- Finalement, nous verrons si un IGP paramétré dans un but de haute disponibilité après panne peut se comporter de manière satisfaisante et si, par conséquent, RON apporte réellement quelque chose à la robustesse du réseau.

L'étude de ces protocoles va ensuite nous permettre de mettre en évidence les besoins et les difficultés

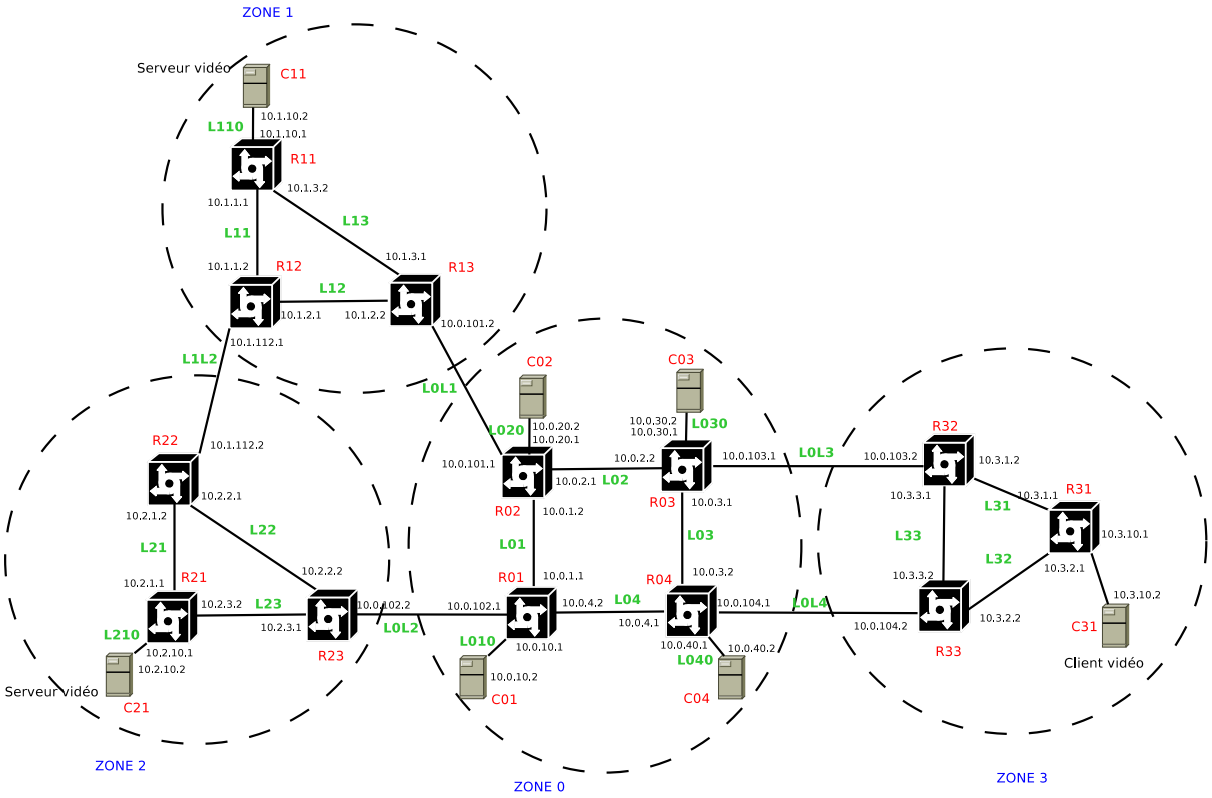


FIG. 4.1 – L'architecture du réseau de test

pressenties pour la conception d'un protocole permettant d'assurer la connectivité entre les nœuds d'un réseau qui soit résistant aux pannes et qui permette le re-routage du trafic en un temps court.

4.2 Test

4.2.1 Architecture de test

On déploie un réseau composés de 20 machines, réparties dans 4 zones, numérotées de 0 à 3. Parmi ces machines, certaines sont appelées « clientes », en bout de réseau. Leur dénomination est de la forme CXY, où X est la zone ou est présente la machine et Y le numéro de la machine dans la zone. Les autres machines sont appelées « routeurs », au cœur du réseau. Leur dénomination est de la forme RXY, où X est la zone où est présente le routeur et Y le numéro du routeur dans la zone. Chaque lien entre les machines est dénommé :

- LXY, si le lien relie 2 routeurs d'une même zone, avec X le numéro de zone et Y le numéro du lien dans la zone ;
- LXY0, si le lien relie un routeur et une machine cliente d'une même zone, avec X le numéro de zone et Y le numéro de la machine cliente dans la zone ;
- LXLY, si le lien est entre 2 routeurs de zones différentes, avec X le numéro de zone du premier routeur, et Y le numéro du deuxième routeur, et $X < Y$.

Chaque machine possède une ou plusieurs interfaces réseau. L'adresse IP de ces interfaces est de la forme 10.X.Y.Z, où X est le numéro de zone de la machine, Y le numéro du lien dans la zone de la machine et Z est 1 ou 2 suivant la position de la machine dans le réseau. (Pour les interfaces inter-zones reliées par un lien LXLY, les adresses IP de ces interfaces sont de la forme 10.X.1XY.Z)

Le schéma 4.1 représente l'architecture du réseau de test.

4.2.2 Mode opératoire

Emulation du réseau :

Toutes les machines du réseau étudié sont des machines émulées par le logiciel *QEMU* [49], lancé depuis plusieurs machines hôtes. A chaque zone correspond une machine hôte sur laquelle est émulée l'ensemble des machines de la zone. *QEMU* permet l'émulation d'un ordinateur complet avec son processeur et ses interfaces réseau. Une machine émulée par *QEMU* se comporte ainsi de façon identique à une machine normale. Il est ainsi possible d'y installer un système d'exploitation, ainsi que tous les logiciels souhaités sans y apporter de modifications.

Les interfaces réseau sont donc elles aussi émulées par *QEMU*, et tout le trafic émis vers une interface d'une machine émulée ressort vers une pseudo interface TAP de la machine hôte. Afin d'assurer la connectivité entre les machines émulées, on relie à un pont réseau les interfaces TAP qui correspondent à des interfaces émulées connectées entre elles dans le réseau émulé. Pour assurer la connectivité entre machines hôtes et donc entre les différentes zones du réseau émulé, on relie à un pont réseau la pseudo interface TAP inter-zones à l'interface réseau de la machine hôte. Cette interface est reliée à une autre machine hôte par un câble Ethernet.

Afin d'introduire du délai dans les liens émulés, on utilise le programme *linux tc* (traffic control) appliqué aux pseudos interfaces TAP qui ne sont pas des interfaces inter-zones.

Les avantages à utiliser une telle architecture, plutôt que de réaliser les tests par simulation sont multiples :

- Il n'est pas nécessaire de développer une extension implémentant RON pour un simulateur ;
- Les mesures sont effectuées en condition réelle, sur du trafic existant dans le réseau. Par exemple, bien que cela ne soit pas nécessaire pour les tests réalisés ici, il serait possible de visualiser la vidéo perçue par le client, afin d'en évaluer la qualité ;
- L'utilisation de *QEMU* permet de déployer un grand réseau, tout en utilisant un faible nombre de machines physiques.

Configuration des machines clientes et des routeurs :

Le système d'exploitation de ces machines est FreeBSD 5.4. Ce choix est motivé par le fait que c'est le seul système sur lequel RON fonctionne.

Les interfaces réseaux sont configurées dans le fichiers */etc/rc.conf* conformément à ce qu'il est décrit dans la partie architecture du réseau, avec un masque de sous réseau /24. Les machines clientes ont pour passerelle par défaut le routeur directement relié à elle.

4.2.3 Cas étudiés

Les différents cas d'étude sont les suivants :

- OSPF ;
- OSPF avec probing agressif ;
- RON sur un nombre différent de machines, associé au routage statique ;
- RON sur 5 machines, associé à OSPF.

Déroulement du test :

On va exécuter trois scénarios de test sur le réseau afin de mesurer différents paramètres :

Scénario 1 : Ce test consiste à diffuser une vidéo depuis la machine C11 jusqu'à la machine C31. La vidéo, encodée en MPEG-2 pour la partie vidéo et MPEG-1 Audio pour la partie audio, a un débit moyen de 1 Mbit/s et est encapsulée par le conteneur MPEG-TS (qui permet la diffusion). Sa longueur est de 10 minutes. La diffusion s'effectue par UDP.

Au bout de 2 min, le lien L0L1 est coupé. Il est rétablit à la 7e minute.

Lors de ce scénario, on réalise les mesures suivantes :

- Le délai de réception entre la machine C11 et C31 en fonction du temps ;
- Le délai de récupération de la connectivité entre C11 et C31 après la panne du lien L0L1 ;
- La répartition des paquets reçus en fonction de leur délai d'acheminement ;

– L'évolution des tailles des files d'attentes des machines R11 et R31, lors de la diffusion vidéo ;
– Le nombre de paquets induit par les messages de contrôle des différents protocoles de routage sur le lien L04, en fonction du temps, ainsi que la taille moyenne de ces messages de contrôle ;
– Le débit induit par les messages de contrôle du protocole de routage sur le lien L04, en fonction du temps et en moyenne.

De plus, ces mesures spécifiques à RON sont réalisées :
– L'overhead engendré par l'encapsulation des paquets ;
– Le temps de traitement d'un paquet par u n routeur RON comparé à un routeur non-RON ;

Scénario 2 : Ce scénario est identique au scénario précédent, mais on provoque la panne du routeur R03, afin d'évaluer le comportement des protocoles en cas de panne d'un routeur en comparaison avec la panne d'un lien.

Les mesures réalisées sont les mêmes que pour le scénario précédent

Scénario 3 : Le test consiste à transférer 100 Mo de données depuis la machine C11 jusqu'à la machine C31 via le protocole TCP. On effectue ce transfert avec les liens fonctionnant normalement (scénario 3.1), puis en situation de re-routage, avec le lien L0L1 coupé et en laissant le temps aux protocoles de routage de rétablir la connectivité (scénario 3.2).

On mesure ensuite le débit reçu par la machine C31 en fonction du temps, en situation normale et en situation de re-routage.

Mesure des performances :

Pour les mesures du scénario 1 et 2 : La vidéo est diffusée par le logiciel *VLC*. Toutes les trames transitant dans le réseau sont capturé sur chaque pseudo interface par le logiciel *TCPDUMP*. Toutes les commandes de capture *TCPDUMP* sont exécutées depuis les machines hôtes. En effet, l'enregistrement du trafic d'une interface réseau d'une machine émulée est effectuée via la pseudo interface de la machine hôte correspondante.

Pour les mesures du scénario 3 : On mesure le trafic reçu par le client C31 lors de l'envoi d'un fichier par TCP de 100 Mo par le serveur C11.

Cette mesure est réalisée à l'aide du logiciel *NETPERF*. Le trafic est enregistrer par *TCPDUMP* et ensuite analysé par des scripts. On va ainsi pouvoir mesurer le temps nécessaire à la transmission des 100 Mo de données mais aussi l'évolution du débit au cour du temps.

Scripts pour le traitement des fichiers *TCPDUMP* : Des scripts dédiés au traitement des fichiers TCPDUMP ont été crées afin de produire des traces du trafic, et par la suite des graphiques.

4.2.4 Résultats et commentaires

Une sélection des résultats des mesures sont présentes dans l'annexe A. Cette partie est consacrée à l'interprétation de ces résultats. Les différents résultats présentés dans les tables de cette partie sont les moyennes de l'ensemble des mesures effectuées et ne correspondent par conséquent pas nécessairement aux résultats présentés en annexe, qui sont les résultats d'une unique mesure.

Premières constatations

Voici les premières constations que l'on peut faire après étude des graphiques de l'annexe A.

Il apparait tout d'abord que les performances de OSPF en terme de temps de récupération après panne sont moins bonne que celles de RON (graphiques A.1.1 et A.3.1, A.4.1, A.5.1, A.6.1, A.7.1), y compris lorsque OSPF est comnfiguré avec des paramètres de probing agressif (grahique A.2.1). Ce point important sera détaillé plus loin.

Ensuite, l'étude du délai et de sa répartition (graphiques A.1.2 et A.3.2, A.4.2, A.5.2, A.6.2, A.7.2, A.8.2) se montre à l'avantage d'OSPF. En effet, une des première chose qui apparaît lors de l'observation des figures sur la variation du délai de bout en bout au cour du temps est qu'en situation de re-routage, RON augmente le délai à 100 ms, alors que ce délai est de 75 ms avec OSPF. Ceci est dû à la nécessité

Protocole	Temps de tratiement en ms
OSPF	0.999
RON	1.452

TAB. 4.1 – Temps de traitement d'un paquet

Protocole et configuration	Temps de récupération en seconde
OSPF	
paramètres par défaut	100.5
probing agressif	66
RON avec routage statique	
sur 3 nœuds	52
sur 5 nœuds	53.5
sur 7 nœuds	56.5
sur 10 nœuds	57
sur 14 nœuds	60
sur 20 nœuds	117
sur 7 nœuds, panne d'un routeur	121
RON sur 5 nœud avec OSPF	
paramètres par défaut	116

TAB. 4.2 – Temps de récupération après panne

pour le trafic de passer par le nœud C21, même si d'autres nœud RON sont présents sur le chemin du trafic. Ce phénomène est dû aux insuffisances de l'implémentation de RON, qui ont été expliquées dans la partie 2.3.2. De plus, la variation du délai dans le cas de l'utilisation de nombreux nœuds RON est importante. Ce phénomène, expliqué par les nombreux messages de contrôle transitant dans le réseau, sera analysé dans la suite.

On constate aussi que les performances de RON dans le cas de la panne d'un routeur (graphique A.9.1) sont bien moindres que dans le cas de la panne d'un unique lien. Il semble que la panne du routeur R03, qui correspond ici à la panne de quatre liens en simultané soit très préjudiciable au fonctionnement de RON.

L'étude des performances de TCP (graphiques A.11.1, A.11.1, A.11.2, A.11.2)montre que le débit mesuré est premièrement, plus important en situation normale qu'en situation de re-routage, et, deuxièment, plus important pour OSPF en situation de re-routage que pour RON en situation de re-routage. Ces résultats étaient attendus. En effet, puisque le débit de TCP est indexé sur le RTT et que le délai est plus important pour RON en situation de re-routage que pour OSPF en situation de re-routage, comme il l'a été expliqué plus haut, il est normal d'observer ces différences de débit.

Enfin, le tableau 4.1 montre les temps nécessaires à un routeur RON ou OSPF pour acheminer un paquet, dans le cas ou ses files d'attentes sont vides. On constate une augmentation de 0.5 ms pour le traitement d'un paquet par RON.

Etude du temps de récupération

OSPF : L'analyse des temps de récupération du protocole OSPF montre les sérieuses insuffisances de ce protocole à rétablir la connectivité après une panne. En effet, comme le montre le tableau 4.2, OSPF, configuré avec les paramètres par défaut, met plus de 100 secondes pour rétablir la connectivité entre le serveur et le client. Ce temps est évidement trop long pour ne pas perturber les applications temps réel.

L'utilisation d'un probing agressif diminue sensiblement le temps nécessaire à la récupération de la connectivité. En effet, même si le temps de convergence reste le même, la détection de la panne est effectuée plus rapidement, ce qui permet l'amélioration du temps de récupération global.

Il apparaît donc que l'utilisation du protocole OSPF n'est pas adapté à la récupération rapide après panne. Cependant, on s'aperçoit que ses piètres performances sont en grande partie dues à la lenteur de la détection des pannes. Une étude plus approfondie de l'incidence du mécanisme de probing sur les

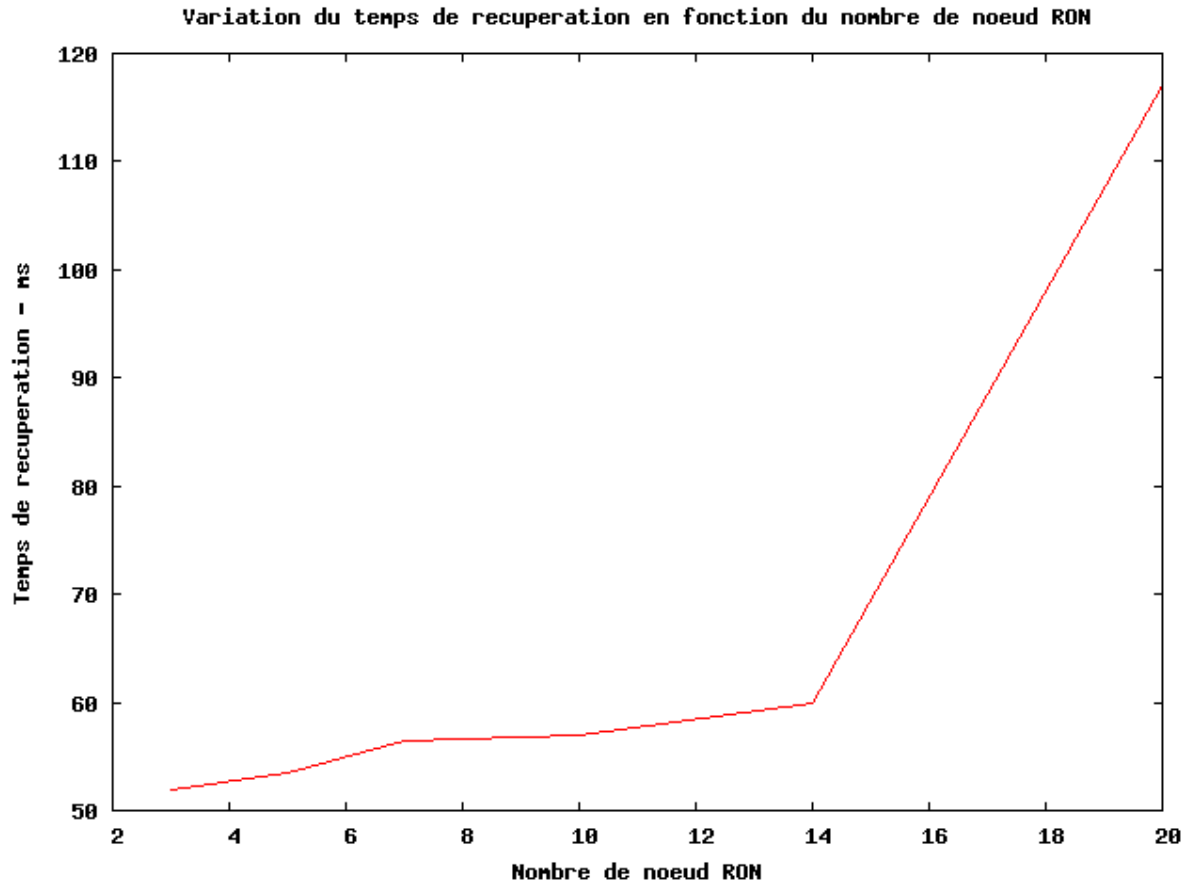


FIG. 4.2 – Temps de convergence et nombre de nœuds

performances de récupération après panne sera effectuée plus bas.

RON : L'analyse des temps de récupération montre que RON se comporte globalement mieux qu'OSPF pour la récupération après panne. En effet, si le nombre de nœuds RON dans le réseau est compris entre 3 et 14, le temps de récupération est compris entre 46 et 60 secondes, ce qui généralement plus faible qu'avec OSPF. On peut expliquer ce phénomène par deux principales raisons, le probing plus agressif de RON par rapport à OSPF avec les paramètres par défauts, ce qui permet de détecter plus rapidement les pannes, ainsi que le temps de convergence plus faible dû à l'architecture en overlay. Nous allons expliciter ce dernier point plus en détail.

On peut en effet considérer que la diminution du nombre de nœud dans le protocole de routage entraîne une baisse du temps de convergence, qui est le temps nécessaire, une fois la panne détectée, pris par chaque nœud pour recalculer les routes et ainsi rétablir la connectivité. Le graphique 4.2 illustre bien cette idée en montrant clairement que dans des circonstances identiques, le temps de convergence augmente avec le nombre de nœuds dans le réseau.

Le fait de diminuer le nombre de nœuds dans le réseau overlay permet donc une diminution du temps de récupération global dans le réseau. Cependant, il faut prendre en considération que le nombre et la position des nœuds de l'overlay par rapport à la topologie du réseau initial influe sur la possibilité de recouvrement du protocole overlay en cas de panne. En effet, par exemple, dans le cas étudié ici de l'overlay composé de trois nœuds RON, l'overlay couvre bien la plupart des pannes de lien possible. Ceci est dû au fait que les nœuds en communication sont situés en extrémité de réseau. Cependant, si les nœuds en communication comprenaient un nœud du centre du réseau, l'overlay serait parfaitement inutile. On peut donc conclure sur ce point en disant que diminuer le nombre de nœuds dans le réseau overlay permet une amélioration du temps de convergence après une panne, et donc du temps de récupération global,

Protocole et configuration	Débit moyen dû aux messages de contrôle, en kbit/s
OSPF	
paramètres par défaut	0.035
séparation des zones 0, 1, 2 et 3	0.041
probing agressif	0.061
séparation des zones et probing agressif	0.070
RON avec routage statique	
sur 3 nœuds	0.080
sur 5 nœuds	0.257
sur 7 nœuds	0.612
sur 10 nœuds	1.68
sur 14 nœuds	3.50
sur 20 nœuds	8.42
sur 7 nœuds, panne d'un routeur	0.537
RON sur 5 nœud avec OSPF	
paramètres par défaut	0.445

Tab. 4.3 – Débit des messages de contrôle

mais que cette diminution entraîne une moindre capacité de l'overlay à agir sur l'ensemble des pannes possibles. Par conséquent, le choix des nœuds composant l'overlay doit prendre en compte quels sont les nœuds entre lesquels on souhaite établir la communication ainsi que l'architecture du réseau sous-jacent, tout en essayant de minimiser le nombre de ces nœuds.

Influence du probing et trafic induit

L'observation des graphiques concernant la bande passante induite par les messages de contrôles nécessaires au fonctionnement des protocoles (graphiques A.1.3, A.1.4, A.2.3, A.2.4, A.3.3, A.3.4, A.4.3, A.4.4, A.5.3, A.5.4, A.6.3, A.6.4, A.7.3, A.7.4, A.8.3, A.8.4) va nous permettre de mieux comprendre l'influence de ces messages sur l'efficacité du protocole à rétablir rapidement la connectivité après une panne ainsi que l'efficacité de ce protocole indépendamment du mécanisme de probing.

Nous allons tout d'abord comparer la quantité de bande passante nécessaire à l'échange des messages de contrôle, telle qu'elle l'a été mesurée sur le lien L04. Le tableau 4.3 montre le débit moyen utilisé par les messages de contrôle.

La forte augmentation de la bande passante nécessaire lors de l'augmentation du nombre de nœud RON est expliqué par le fait que le réseau overlay RON est un réseau mesh ; chaque nœud RON est ainsi en communication avec l'ensemble des autres nœuds RON. Par conséquent, à la différence d'OSPF, où sur le lien L04 ne transite que les messages de contrôles échangés entre les machines R01 et R04, avec RON, il y transite les messages de contrôle de l'ensemble des nœuds de overlay dont la communication entre eux passe par ce lien. Par exemple, lors de l'utilisation de 3 nœuds RON, il transite par L04 les messages de contrôle allant de C21 à C31 et de C31 à C21. Lors de l'utilisation de 5 nœuds RON, il y transite en plus les messages allant de C04 à C21, de C21 à C04, de C04 à C02 et de C04 à C11. Il y a donc, lors du passage de 3 à 5 nœuds, 3 fois plus de messages qui transitent par ce lien. Ceci concorde bien avec les débits mesurés et indiqués dans le tableau 4.3

La nature mesh du réseau overlay de RON fait qu'il n'est pas performant lorsqu'il est déployé sur cette topologie de réseau, où un grand nombre de nœuds sont présents dans l'overlay et où les communications entre ces nœuds passe par le même lien. On constate effectivement que lors de l'utilisation de nombreux nœuds RON, une forte variation du délai est observée dans les graphiques représentant l'évolution du délai au cour du temps. De plus, on a constaté un plus fort remplissage des files d'attentes dans ces conditions. Ceci permet de penser que RON ne passerai pas à l'échelle dans le cas particulier de l'utilisation d'un grand nombre de nœuds RON et d'une utilisation intensive du réseau. Il faut de plus considérer la diminution de la charge utile lors de l'utilisation de RON due à l'encapsulation des données dans l'entête RON, comme il est montré dans le graphique 4.3. Ce dernier point, ainsi que l'important nombre de messages de contrôle nécessaire à son fonctionnement rendent problématique la surcharge dans le réseau

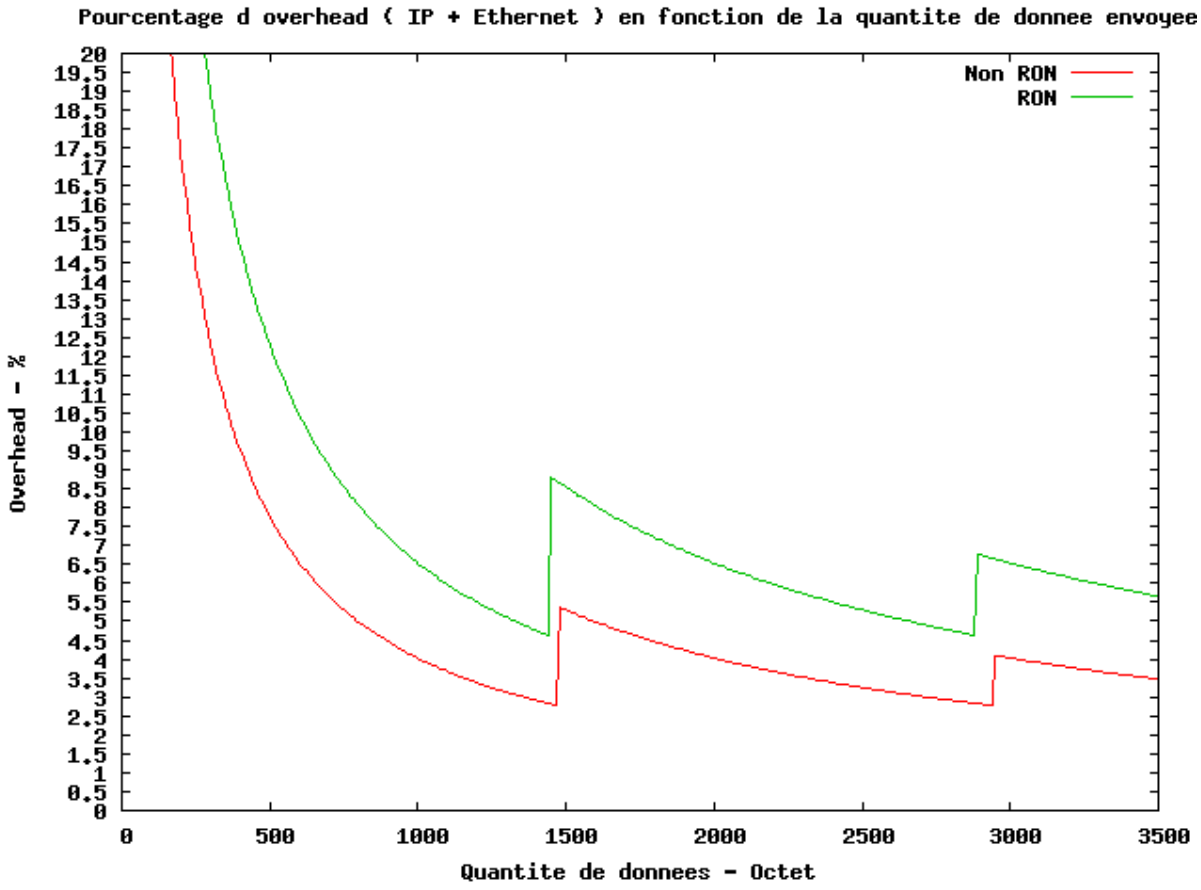


FIG. 4.3 – Overhead du à l'encapsulation RON

provoquée par RON.

Nous allons maintenant procéder à une évaluation de la performance des protocoles en mettant en perspective le temps de récupération sur panne en fonction du débit requis pour faire fonctionner le protocole. La figure 4.4 montre les résultats obtenus. Il n'a été pris en compte que le cas du réseau RON à trois nœuds pour les raisons données ci-dessus. On constate ainsi que le ratio débit requis / temps de récupération n'est pas nécessairement à l'avantage de RON. En effet, dans cette mesure, OSPF avec les paramètres de probing agressifs se révèle plus performant que RON. Cependant, cette constatation est dûe à certains choix de conception et d'implémentation de RON, qui utilise plusieurs métriques et par conséquent, les messages de contrôle doivent inclure les informations pour ces différentes métriques et sont donc de plus grande taille.

Concentrons nous maintenant sur le temps de détection des pannes. Comme il l'a été expliqué plus haut, ce temps a une grande incidence sur le temps total de récupération car on peut approximer la formule

$$temps\_de\_recuperation = temps\_detection\_panne + temps\_convergence$$

et

$$temps\_convergence = temps\_propagation\_infos\_etat\_des\_liens + temps\_calcul\_nouvelles\_routes$$

Ainsi, nous allons nous intéresser à l'espérance du temps de détection. On le calcule en fonction des intervalles de probing ainsi que de l'intervalle au bout duquel le routeur est considéré comme injoignable. Les documents [8] et [11] nous permettent de renseigner le tableau 4.4.

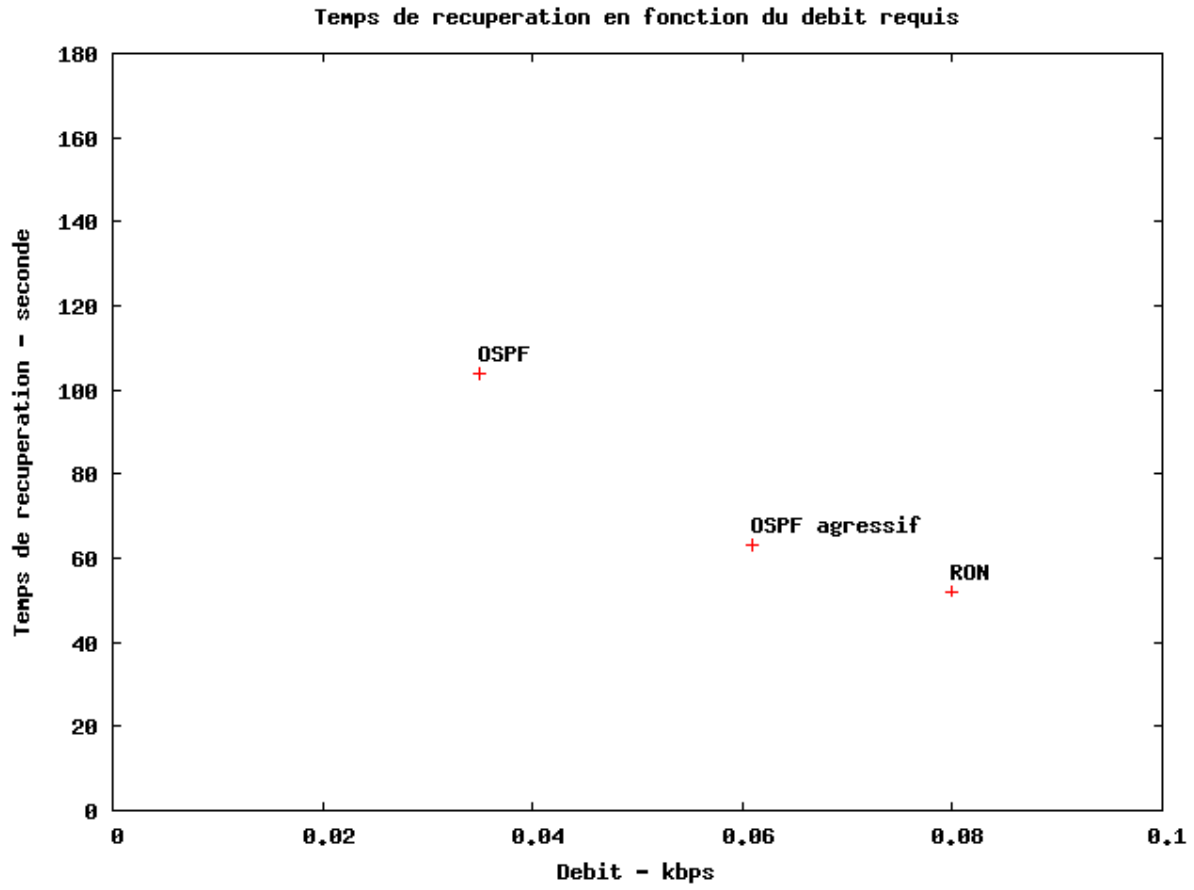


FIG. 4.4 – Temps de récupération après panne en fonction du débit nécessaire aux messages de contrôle

Protocole et configuration	Temps de détection en secondes
OSPF, paramètres par défaut	35
OSPF, probing agressif	7.5
RON	19

TAB. 4.4 – Espérance du temps de détection d'une panne

Protocole et configuration	Temps de convergence
OSPF, paramètres par défaut	65.5
OSPF, probing agressif	63.5
RON sur 3 nœuds	33
RON sur 5 nœuds	34.5
RON sur 7 nœuds	37.5
RON sur 10 nœuds	38
RON sur 14 nœuds	41
RON sur 20 nœuds	101

TAB. 4.5 – Temps de convergence des protocoles

Temps de convergence

Il est maintenant possible de comparer les temps de convergence des différents protocoles en soustrayant les résultats obtenus pour le temps de détection de panne au temps de récupération. Les résultats sont présentés dans le tableau 4.5. On constate, tout d'abord, que les temps de convergence d'OSPF, configuré avec les paramètres par défaut ou avec le probing agressif, sont équivalent. Ce résultat était attendu car le probing agressif n'influe que sur le temps de détection de la panne. Par ailleurs, on constate que RON, s'il est déployé sur 3, 5, 7, 10 ou 14 nœuds a un temps de convergence inférieur à celui d'OSPF. Il faut rappeler que OSPF est déployé sur 13 nœuds et donc qu'à priori, à complexité de calcul des routes équivalente, RON a un temps de convergence plus faible qu'OSPF.

Explicitons plus précisément les différentes étapes composant le temps de convergence :

$$temps\_convergence$$
$$=$$
$$temps\_de\_propagation\_infos\_etat\_des\_liens$$
$$+$$
$$temps\_de\_mis\_a\_jour\_routes$$

OSPF, lors de la réception d'une nouvelle information sur l'état d'un lien l'amenant à recalculer de nouvelles routes, attend un certain temps avant d'effectuer ce calcul (c'est un paramètre appelé *spfDelay*, qui vaut par défaut 5 secondes). Cette attente a pour but de laisser le temps à d'autres paquets d'information sur l'état des lien de parvenir au routeur, afin de s'assurer de posséder l'ensemble des informations sur la nouvelle topologie du réseau avant de démarrer le calcul de nouvelles routes. Ceci est particulièrement utile lors de la panne d'un routeur, où plusieurs liens sont concernés, car OSPF laisse le temps au paquets d'information sur l'état des liens de parvenir aux routeurs avant de lancer le calcul des nouvelles routes.

Un nœud RON envoie les mises à jour des l'état de ces liens tous les ROUTING_INTERVAL secondes, qui vaut 14 secondes par défaut. Si l'on néglige le temps d'acheminement effectif des informations sur l'état des lien aux différents nœuds RON, ce qui est acceptable dans des conditions de faible délai dans le réseau, l'espérance du temps de propagation de l'état des liens après détection d'une panne est donc de 7 secondes. A la différence d'OSPF, RON envoie périodiquement l'information sur l'état de ses liens aux autres nœuds RON. En effet, même en l'absence de modifications topologiques, les métriques utilisées par RON sont de nature à varier en permanence et il est donc nécessaire d'informer les autres nœuds de leurs évolution régulièrement. RON présente l'inconvénient, à l'inverse de OSPF, de ne pas temporiser le calcul des nouvelles routes après une panne. Ainsi, un nœud RON ne s'assure pas de bien avoir reçu l'ensemble des modifications topologiques avant de débuter le calcul des nouvelles routes. Par conséquent, si un ensemble de modifications ont lieu dans un cour laps de temps le nœud peut être amené à refaire ce calcul plusieurs fois de suite au fur et à mesure que les nouvelles informations sur l'état des liens lui parviennent. Ce phénomène peut expliquer la forte dégradation des performances de récupération de RON observée lors de la panne d'un routeur. Ce point de conception, négligé dans RON, révèle ainsi toute son importance dans la conception d'un protocole de récupération sur panne efficace.

Conclusions

L'analyse des résultats nous permet donc dégager quelques informations sur les performances de re-routage.

Tout d'abord, évaluons quels sont les avantages et inconvénients de RON et OSPF dans le contexte de la récupération rapide après panne.

OSPF a été conçu pour assurer la connectivité entre les différents nœuds d'un réseau, mais n'est pas spécialement dédié à la récupération rapide de la connectivité après une panne. Ceci est particulièrement vérifié lorsque les paramètres du protocole sont laissés à leur valeurs par défaut, car on observe alors qu'il faut plus de 100 secondes pour arriver à rétablir la communication entre deux nœuds aux extrémités d'un réseau de taille moyenne.

OSPF présente tout de même l'avantage d'établir la connectivité entre tous les nœuds du réseau, ce qui n'est pas le cas avec RON, qui dans certains cas peut être inefficace contre certaines pannes affectant le réseau, en fonction de la topologie de l'overlay par rapport à celle du réseau sous-jacent.

De plus OSPF est nettement moins consommateur de bande passante que RON. Ceci est dû à certains choix de conception de RON qui introduit plusieurs métriques dans la mesure de la qualité d'une connexion, ainsi qu'une communication régulière de ces informations à ses voisins, tandis que OSPF se contente d'une seule métrique, et ne communique ces informations que s'il observe un changement de topologie.

Enfin, OSPF temporise le calcul des routes lorsqu'il réceptionne une information indiquant un changement de la topologie. Ceci permet de laisser le temps à toutes les informations de changement de topologie de parvenir au nœud dans le cas, par exemple, de la panne d'un routeur. Ceci permet de ne pas calculer de nouvelles routes inutilement.

On l'a vu, RON possède quelques concepts intéressants pour permettre une récupération après panne rapide. Tout d'abord, son mécanisme de détection des pannes est plus agressif que celui d'OSPF par défaut. Un tel mécanisme est un élément important si l'on souhaite parvenir à un rétablir la connectivité dans des délais raisonnables.

De plus, le temps de convergence de RON est inférieur à celui d'OSPF. Nous l'avons vu, RON bénéficie du fait que le nombre de nœuds dans l'overlay est en général inférieur au nombre de routeur OSPF. Par conséquent, le travail de dissémination des informations sur l'état des liens ainsi que le calcul des routes est plus faible.

Son architecture en overlay mesh entraîne cependant des problèmes dus à la charge engendrée par les messages de contrôle. Ceci est particulièrement flagrant dans un réseau tel que celui utilisé dans les tests, où les nœuds RON sont rapprochés les uns des autres, et où par conséquent il arrive régulièrement qu'un grand nombre de communications entre les nœuds RON empruntent le même lien physique, ce qui entraîne une surcharge non négligeable lors du déploiement d'un overlay comportant beaucoup de nœud. On peut donc s'interroger si une autre architecture, où un nœud RON ne serait en communication qu'avec un nombre limité d'autres nœuds RON. Il faut cependant s'interroger sur le mécanisme à utiliser pour le choix de ces voisins, car pour être efficace celui-ci doit choisir les voisins du nœud RON comme étant proches dans l'overlay.

Pour terminer, on peut dire que l'utilisation de RON ou d'un système similaire peut s'avérer particulièrement adaptée pour protéger une partie du trafic du réseau, entre certains nœuds. En effet, il est alors possible d'utiliser un nombre minimum de nœuds RON de façon à ce que l'overlay recouvre bien le plus grand nombre de pannes possible. De cette façon, lors d'une panne, le trafic sensible sera « pris en charge »par RON, jusqu'à ce que OSPF rétablisse la connectivité et puisse acheminer l'ensemble du trafic. Comme le montrent les résultats des mesures effectuées en utilisant conjointement RON et OSPF (graphiques A.10.1 et A.10.3) cette utilisation s'avère efficace en terme de temps de récupération et permet de limiter la surcharge de bande passante dûe aux messages de contrôle de RON.

Chapitre 5

Travail futur

Nous avons vu que RON est une approche intéressante pour améliorer la robustesse des réseaux. Plusieurs travaux peuvent être envisagés. Tout d'abord, nous l'avons dit dans la partie 2.3.2, l'implémentation de RON présentant des limitations majeures par rapport à son concept initial, une re-écriture de RON pourrait s'avérer intéressante afin d'exploiter au mieux ses atouts.

De plus, ce travail a permis d'identifier un certain nombre de points intéressant développés dans RON et permettant d'apporter plus de robustesse dans le réseau d'une infrastructure critique. On peut, par exemple, citer l'importance du mécanisme de détection des pannes, l'architecture overlay qui permet un temps de récupération après panne plus faible. De la même façon, nous avons identifié un certain nombre de ses points faibles : l'architecture mesh de l'overlay, qui entraîne une forte consomation de bande passante, ainsi que l'impossibilité de détecter certaines pannes en fonction du choix des nœuds RON par rapport à la topologie du réeau sous-jacent. Cette limite ést imposée par la nature des réseaux overlay.

Les constatations faites dans ce document ainsi que les travaux décris dans l'état de l'art permettent de mettre en évidence un certain nombre de problèmes à étudier afin de progresser dans la recherche des mécanismes nécessaires à la conception d'un protocole apportant une bonne résistance aux pannes dans un réseau. Par exemple on peut citer les sujets suivants :

- **Déterminer une architecture de réseau l'overlay qui soit performante en terme de consommation de bande passante tout en étant adaptée à la détection de panne et au re-routage.** En effet, nous avons vu que l'architecture mesh de RON n'était pas performante si elle était déployée dans des réseaux de taille moyenne, comme l'AS d'un FAI. Il faut donc étudier les autres possibilités d'architecture qui peuvent être utilisés, où le nombre de voisins de chaque nœud dans l'overlay est réduit. Il faut de plus déterminer quelles sont les mécanismes apropriés pour la sélection de ces voisins de façon à ce qu'ils soient choisis parmi les nœuds proches dans la topologie du réseau sous-jacent.
- **Etudier l'impact de la topologie du réseau overlay par rapport au réseau sous-jacent sur l'efficacité et les performances du re-routage.** Peu d'études ont été consacrées à ce sujet, pourtant, l'influance de la topologie du réseau sous-jacent est grande sur le comportement de l'overlay. Il est donc essentiel de caractériser cette influence et éventuellement de mettre en place des mécanismes qui permettront à l'overlay de s'adapter au mieux à la topologie du réseau qu'il recouvre.
- **Déterminer un mécanisme incluant sondes et métriques performantes pour détecter les pannes.** On l'a vu, le mécanisme de détection des pannes est un élément déterminant dans l'efficacité du protocole à rétablir la connectivité dans un temps court. Il est donc essentiel de mettre en place un mécanisme de détection des pannes qui soit à la fois rapide et fiable. Il faut de plus s'interoger sur les sondes et métriques à utiliser pour concevoir un tel mécanisme.
- **Déterminer quels sont les mécanismes proactifs et réactifs permettant le re-routage rapide après détection d'une panne, éventuellement au sein d'un réseau overlay.** Comme on a pu le constater dans le chapitre consacré à l'état de l'art, de nombreuses solutions ont été proposées afin d'apporter une solution au problème du re-routage rapide, et ainsi essayer de diminuer

le temps de convergence. Ces différents mécanismes méritent d'être étudiés plus en détail, de façon à les comparer et à identifier leurs forces et faiblesses. De plus, on peut s'interroger sur les bénéfices ou problèmes éventuels qu'amènerai leur intégration dans un protocole overlay tel que RON.

– **Envisager d'autres approches que le routage afin d'assurer la livraison des services aux clients en cas de panne.**

Ce dernier point nous amène à souligner les travaux réalisés dans le cadre des CDN, qui ont été brièvement abordés dans l'état de l'art. Ces travaux présentent des approches concentrées sur la délivrance du service aux clients, à la différence des protocoles de routage qui ont une approches centrée sur la connectivité entre le serveur et le client. L'approche développée dans le cadre des CDN (par rapport à celle des protocoles de routage classiques) est pertinente pour plusieurs raisons :

– Assurer la connectivité entre le client et le serveur ne sert à rien si c'est le serveur qui en panne ;
– Si le service peut être rendu par plusieurs serveurs en différents endroits du réseau, il n'est pas forcément nécessaire de rétablir la connectivité dans le réseau le plus rapidement possible, mais plutôt de diriger le client vers un serveur encore accessible après la panne.

On peut présumer que certains des travaux présentés ici, même s'il ne concernent pas directement les CDN et ont une approche connectivité plutôt que délivrance des services, pourront être réutilisés dans le cadre des CDN. On peut notamment penser aux mécanismes de détection des pannes, de choix d'une route optimale, d'architecture de l'overlay, etc.

Chapitre 6

Conclusion

La robustesse des réseaux est un sujet essentiel notamment depuis l'essor des grands réseaux de télécommunications, qui a entraîné l'apparition d'infrastructures d'information critiques, dont le non-fonctionnement s'avère préjudiciable aux utilisateurs. C'est le cas des réseaux informatiques qui, avec certaines applications temps réel telles que la diffusion de vidéo en direct ou la voix sur IP, amène des besoins de robustesse face aux pannes. Nous avons vu que les protocoles de routage classiques très répandus, tels que OSPF, ne permettent pas le rétablissement de la connectivité en cas de panne dans un temps satisfaisant. Il convient donc de mettre au point et de déployer de nouveaux systèmes plus adaptés à la récupération rapide en cas de panne.

L'étude d'une infrastructure critique nous a permis de dégager les menaces pesant sur la disponibilité des services. Ceci a permis de mettre en évidence un certain nombre de besoins pour améliorer la robustesse. Parmi ces besoins, nous nous sommes particulièrement intéressés au rétablissement de la connectivité entre les différentes machines du réseau en cas de panne d'un équipement. A ce sujet, nous avons étudié dans l'état de l'art le fonctionnement des protocoles de routage les plus répandus, nous nous sommes intéressé aux principales avancées dans le domaine du re-routage rapide, dont le but est de parvenir à un rétablissement de la connectivité dans les meilleurs délais. Nous avons aussi étudié les travaux récents concernant les réseaux d'overlay P2P, qui sont une approche innovante pour aborder ce problème, comme nous le montre RON.

L'étude de RON en comparaison avec un protocole de routage classique, OSPF, a permis de mettre en évidence les avantages à utiliser un système overlay pour le re-routage, ainsi qu'un certains nombre d'autres points ayant un rôle important pour parvenir à un re-routage rapide, tout particulièrement le mécanisme de détection des pannes.

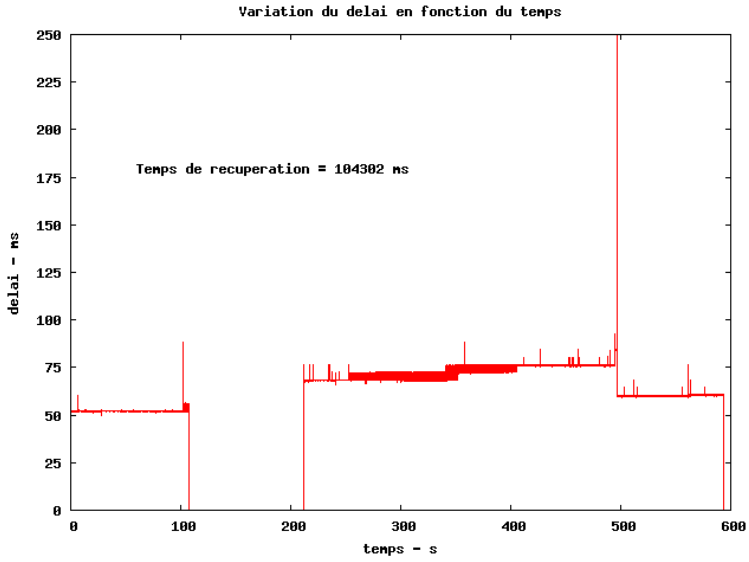
Cependant, le but principal à atteindre reste la disponibilité des services du point de vue de l'utilisateur. Les récents travaux sur les réseaux CDN et les architectures P2P offrent des perspectives intéressantes pour la robustesse dans la fourniture de ces services. Ces solutions, qui agissent sur un plan différent que celui des protocoles de routage, sont un axe de recherche tout à fait intéressant pour améliorer la fiabilité des infrastructures de télécommunications actuelles.

Annexe A

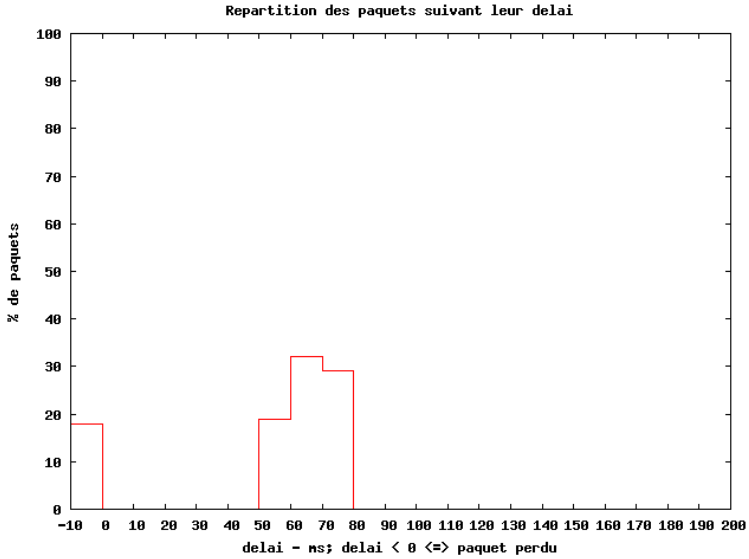
Résultats des tests

A.1 OSPF

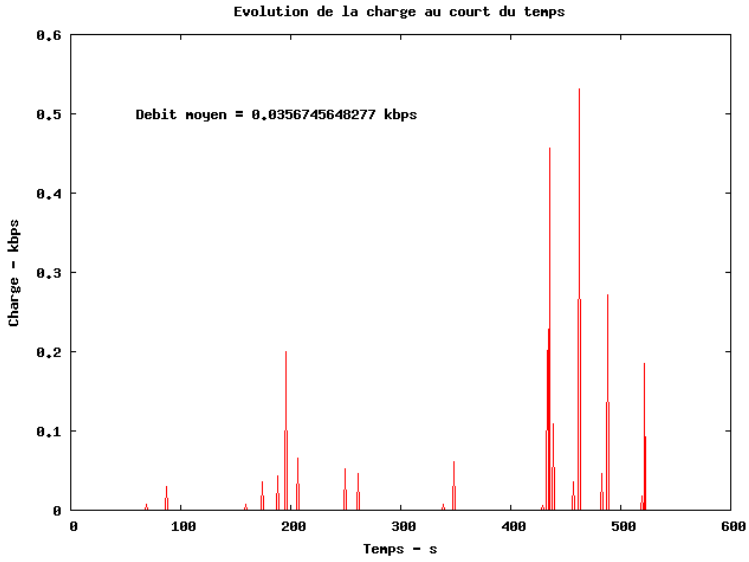
A.1.1 Variation du délai



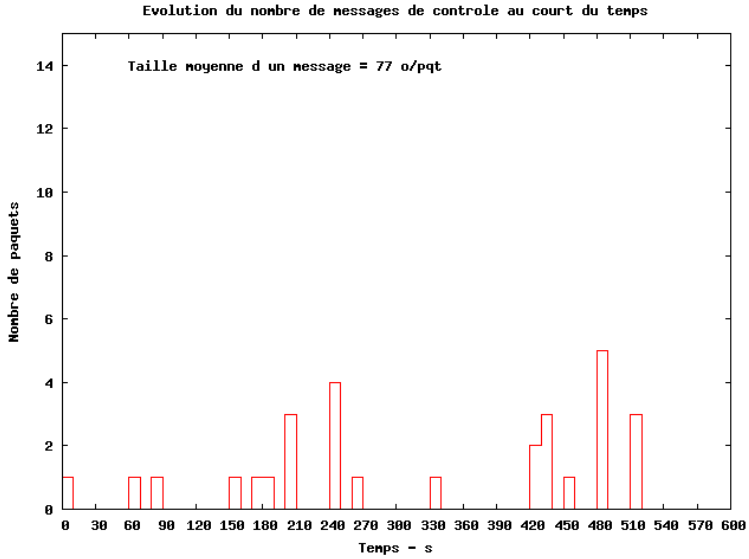
A.1.2 Répartition du délai



A.1.3 Variation du débit des messages de contrôle

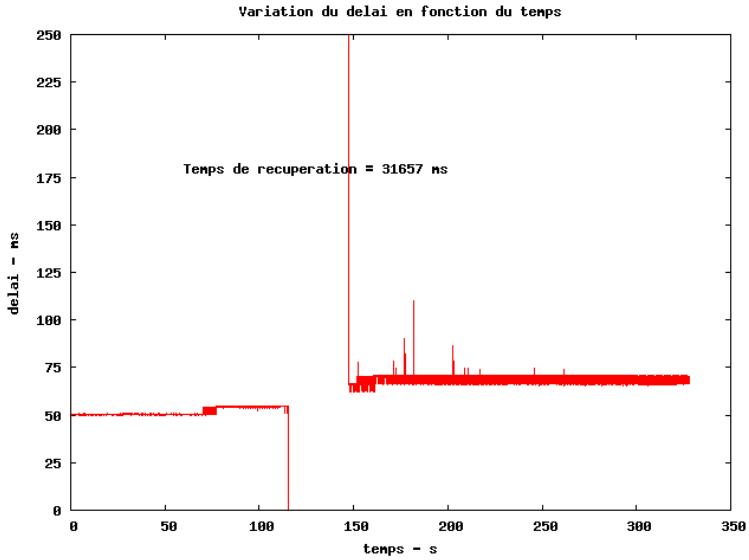


A.1.4 Evolution du nombre de messages de contrôle

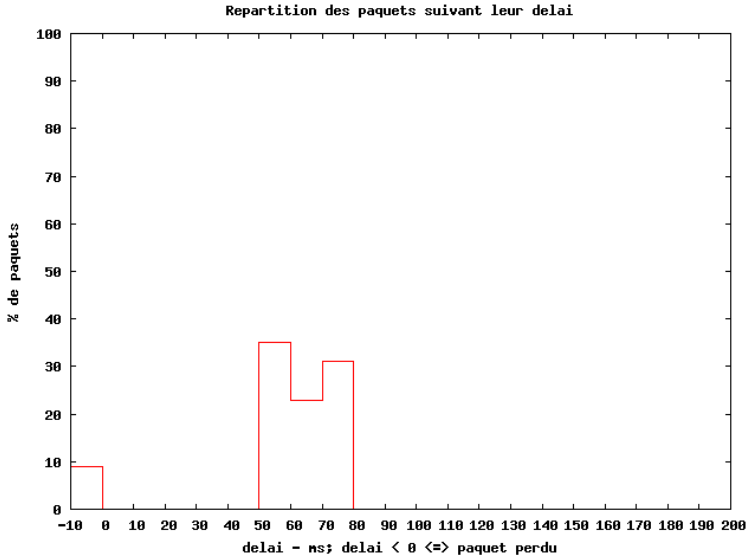


A.2 OSPF et probing agressif

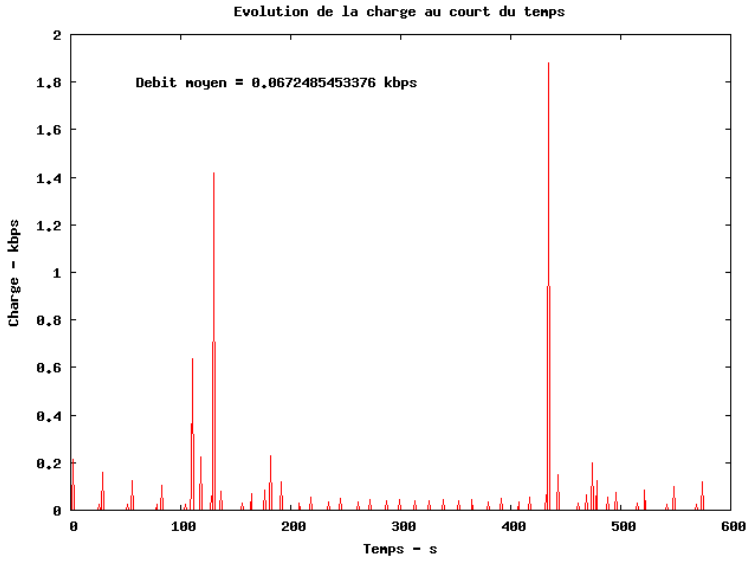
A.2.1 Variation du délai



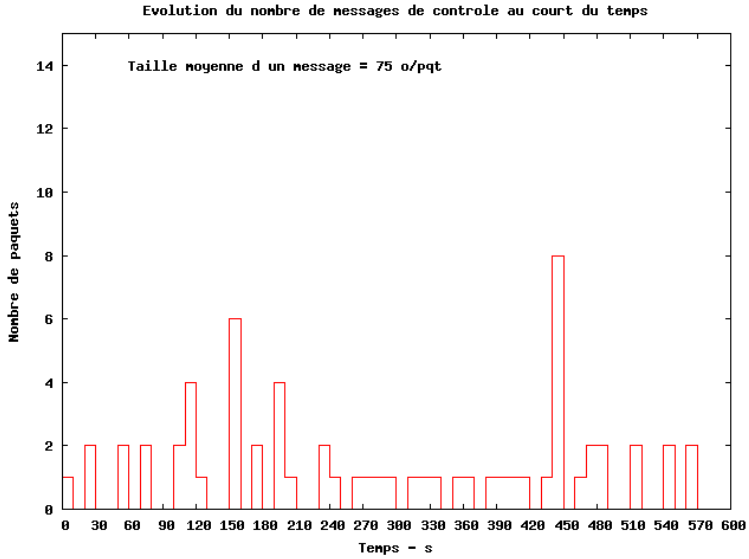
A.2.2 Répartition du délai



A.2.3 Variation du débit des messages de contrôle

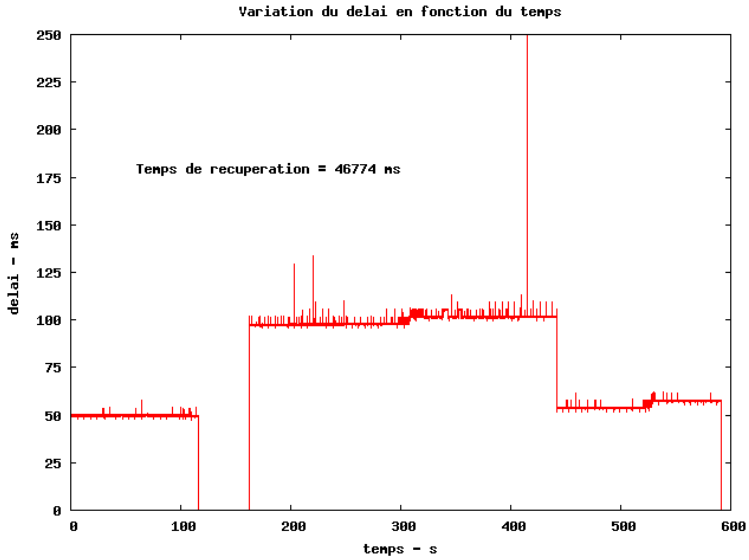


A.2.4 Evolution du nombre de messages de contrôle

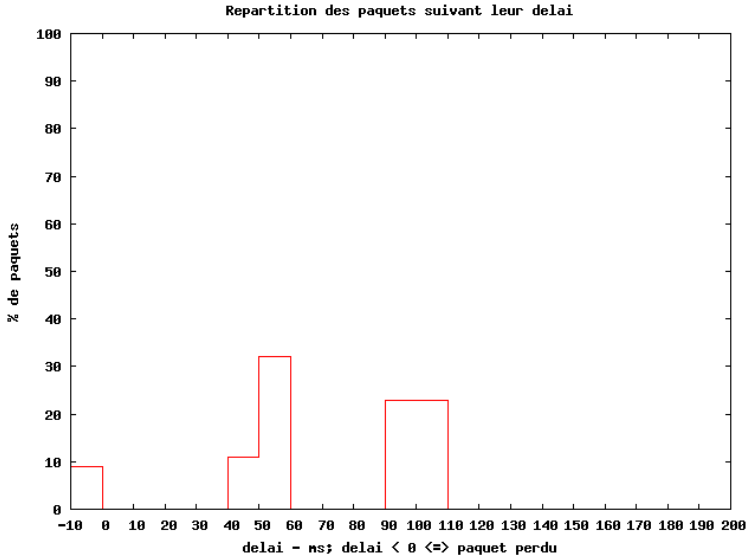


A.3 RON sur 3 nœuds

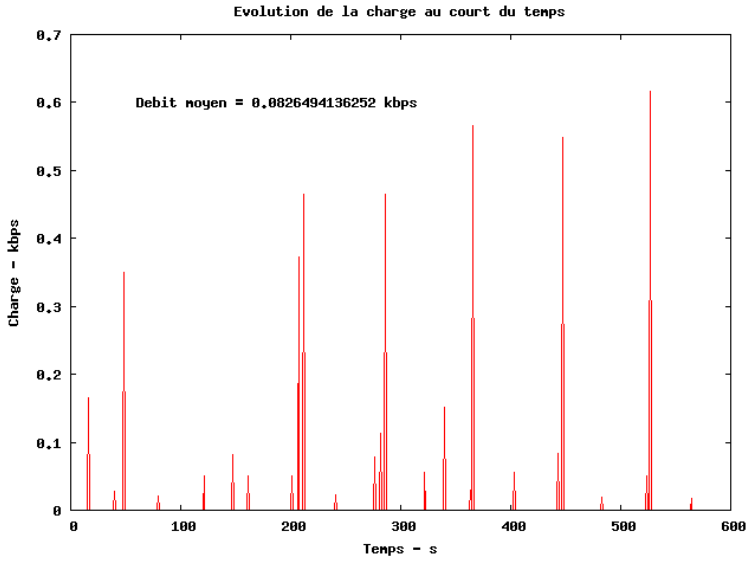
A.3.1 Variation du délai



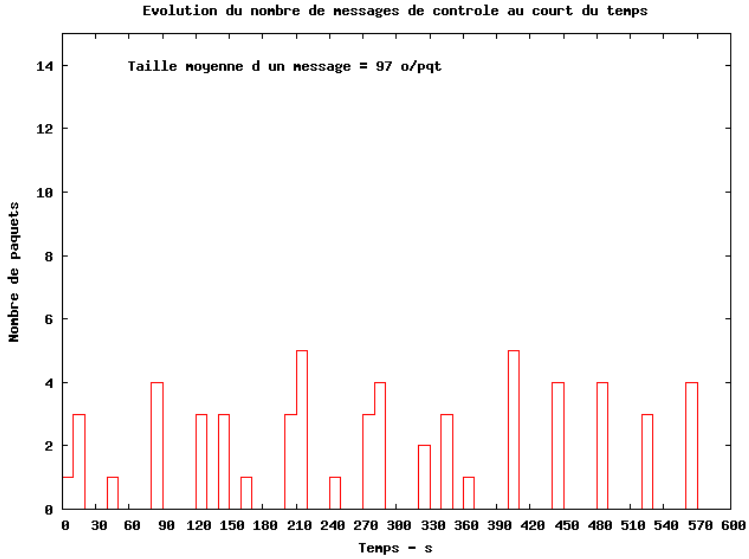
A.3.2 Répartition du délai



A.3.3 Variation du débit des messages de contrôle

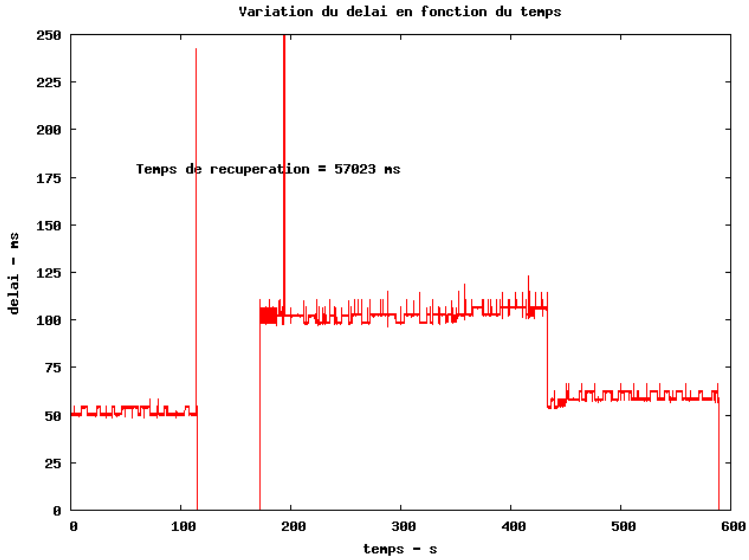


A.3.4 Evolution du nombre de messages de contrôle

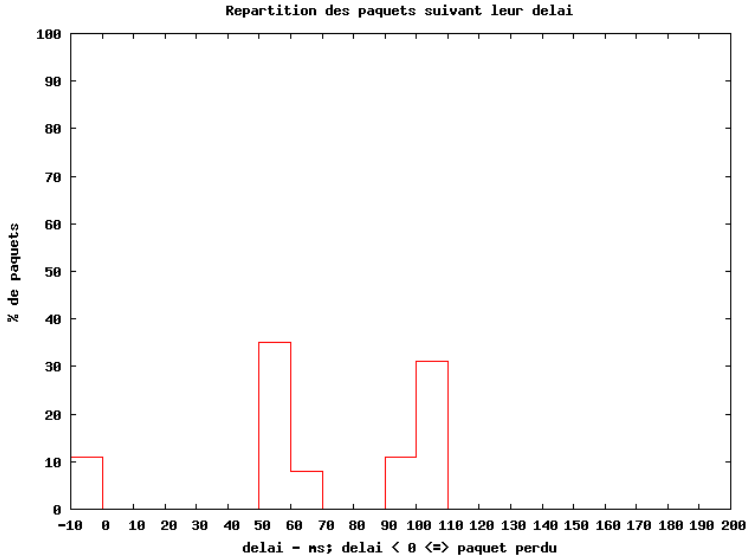


A.4 RON sur 5 nœuds

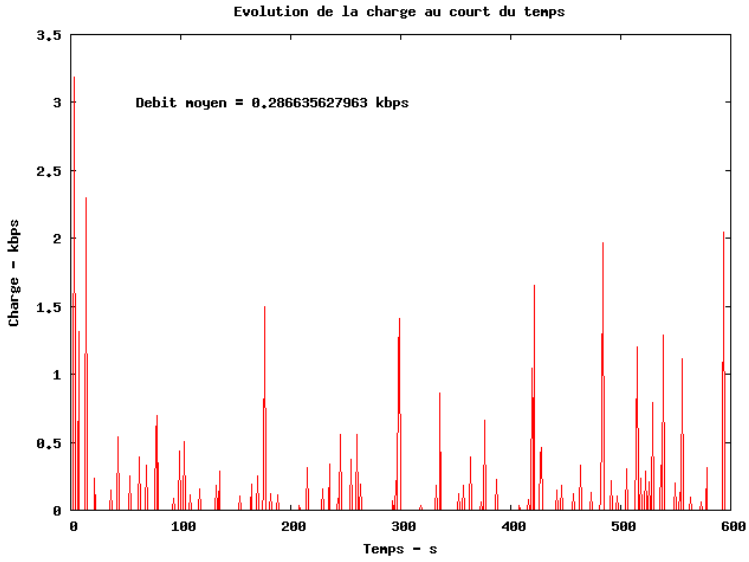
A.4.1 Variation du délai



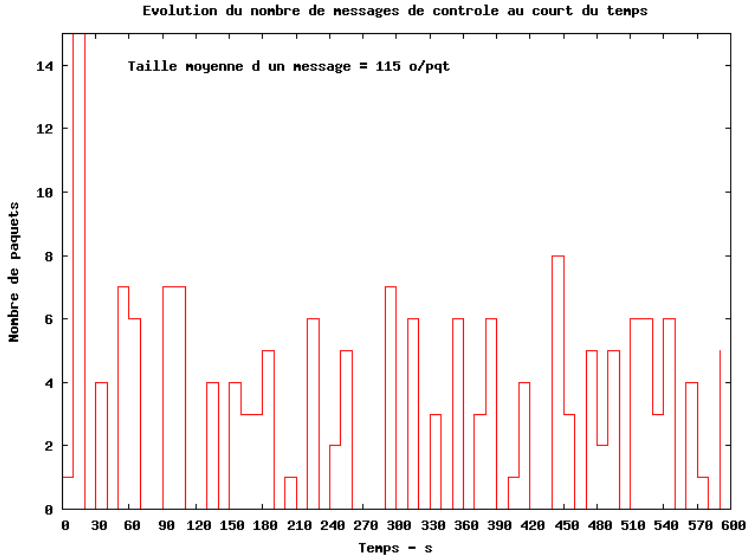
A.4.2 Répartition du délai



A.4.3 Variation du débit des messages de contrôle

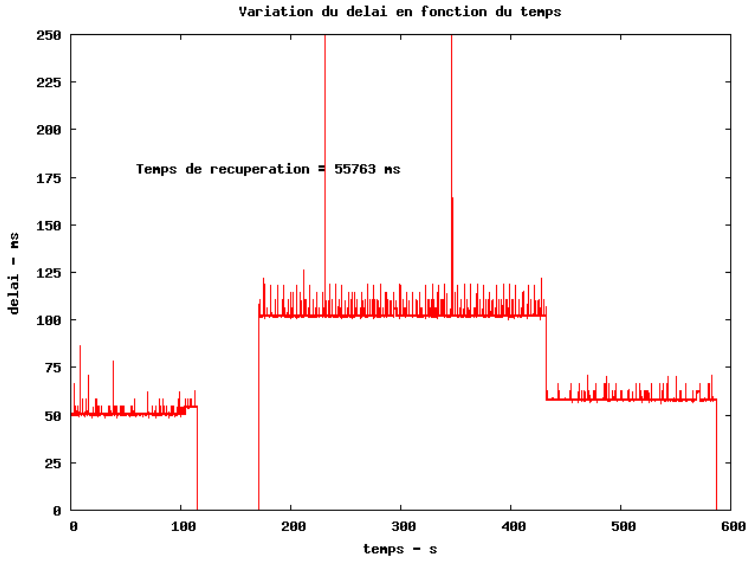


A.4.4 Evolution du nombre de messages de contrôle

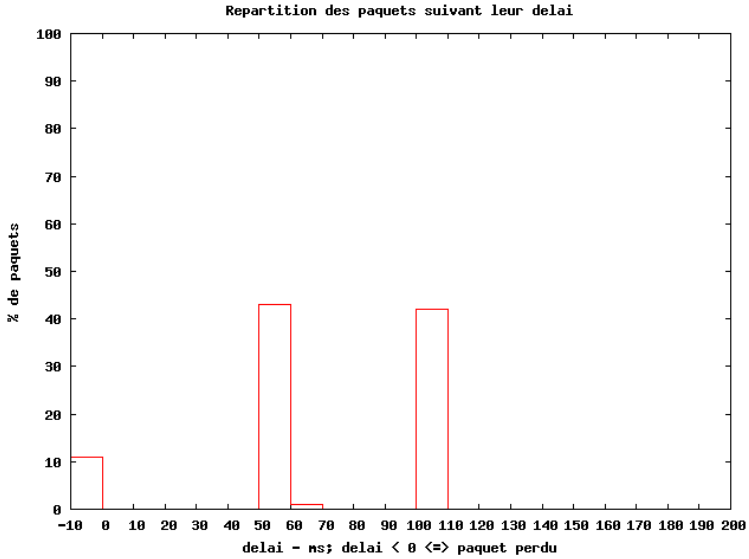


A.5 RON sur 7 nœuds

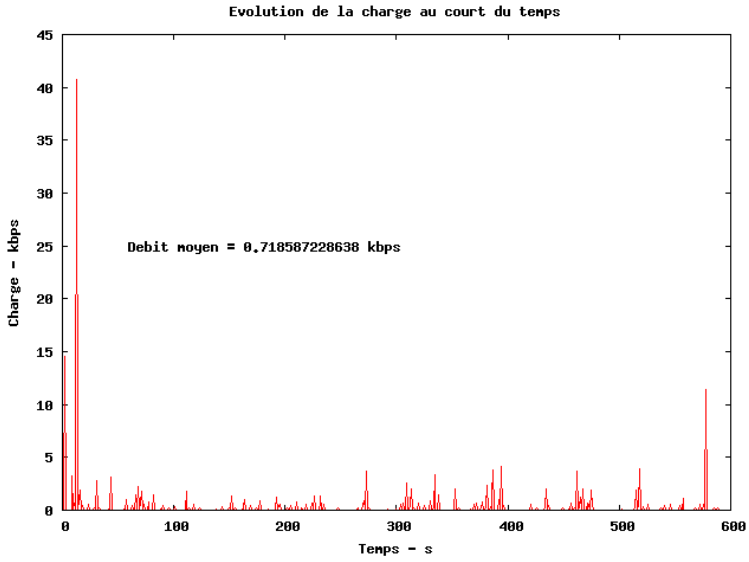
A.5.1 Variation du délai



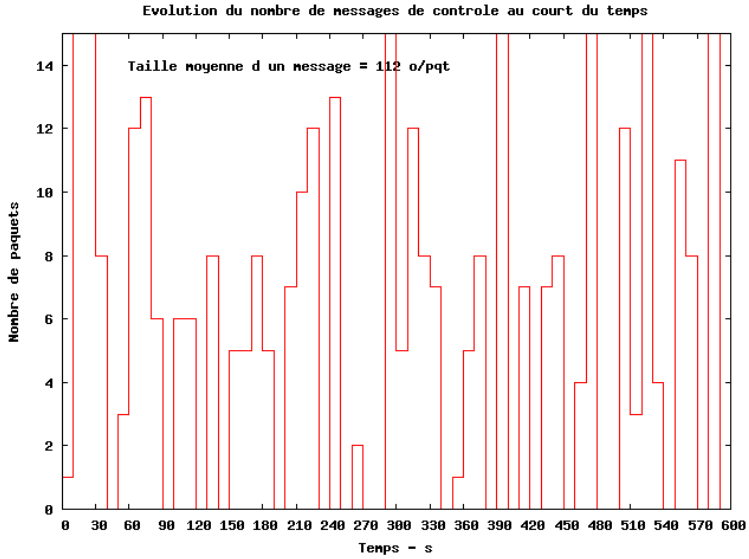
A.5.2 Répartition du délai



A.5.3 Variation du débit des messages de contrôle

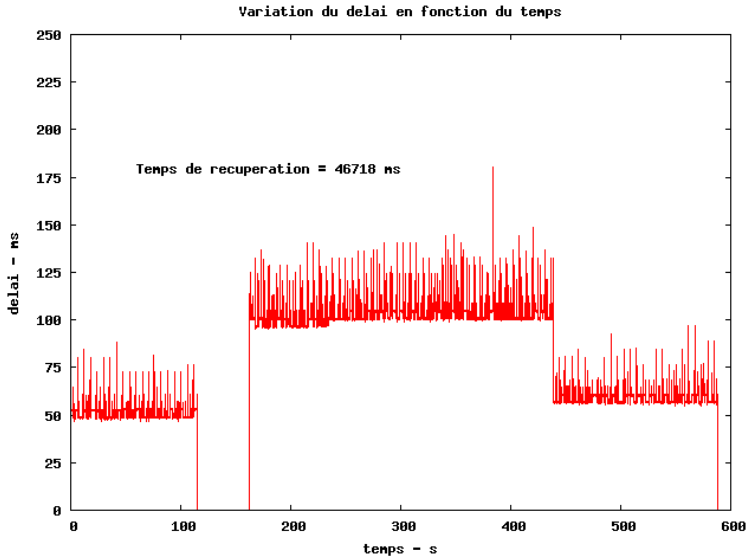


A.5.4 Evolution du nombre de messages de contrôle

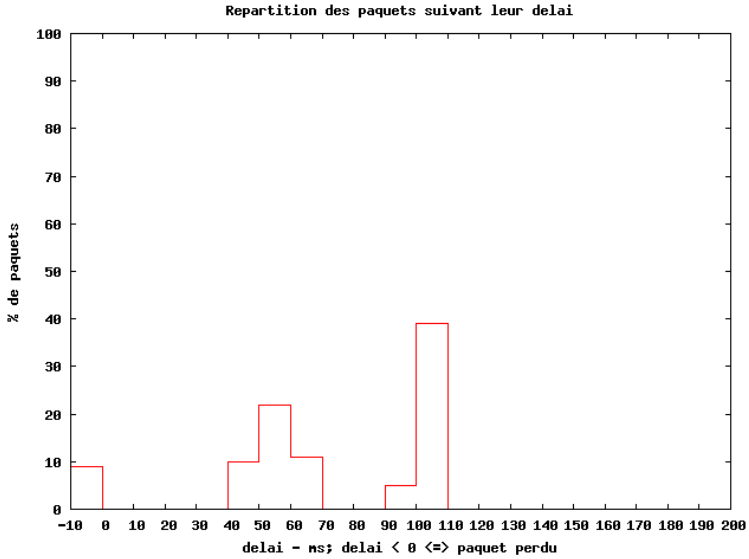


A.6 RON sur 10 nœuds

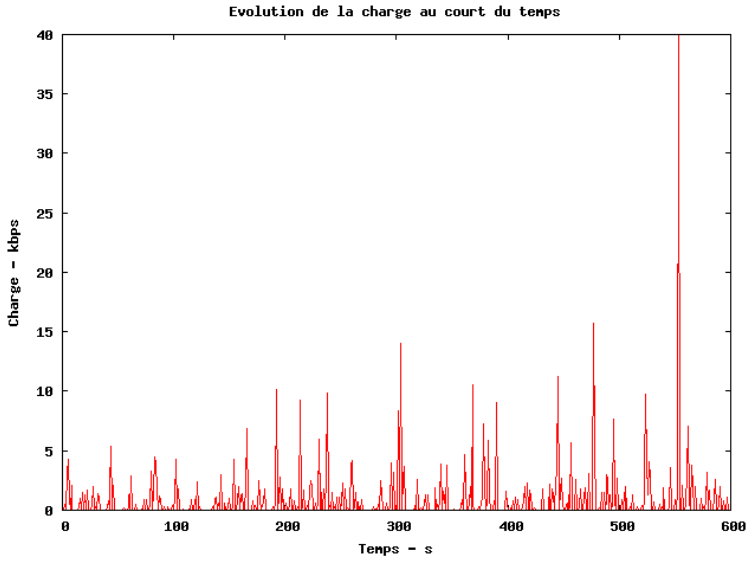
A.6.1 Variation du délai



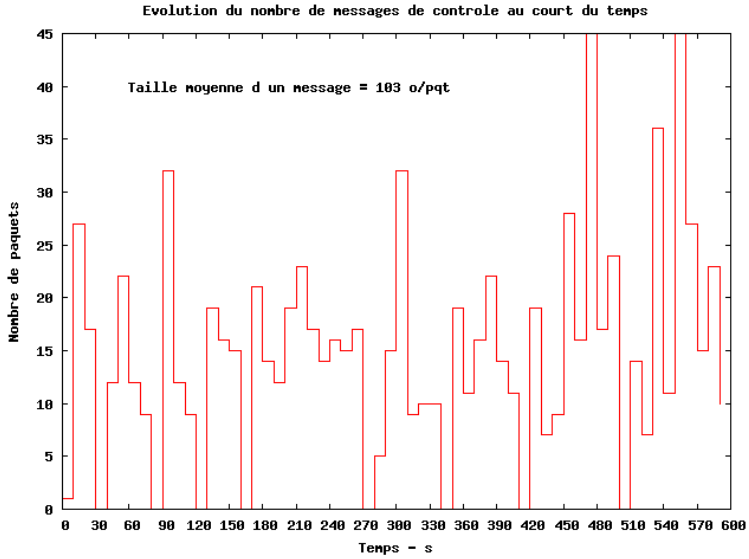
A.6.2 Répartition du délai



A.6.3 Variation du débit des messages de contrôle

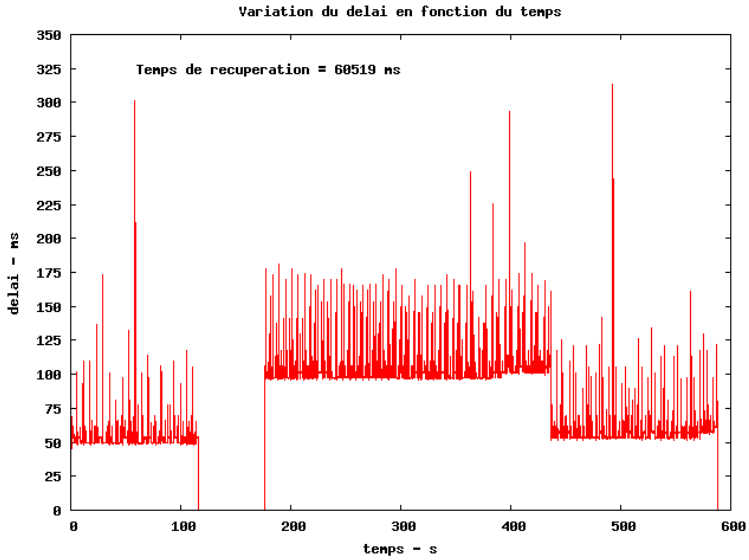


A.6.4 Evolution du nombre de messages de contrôle

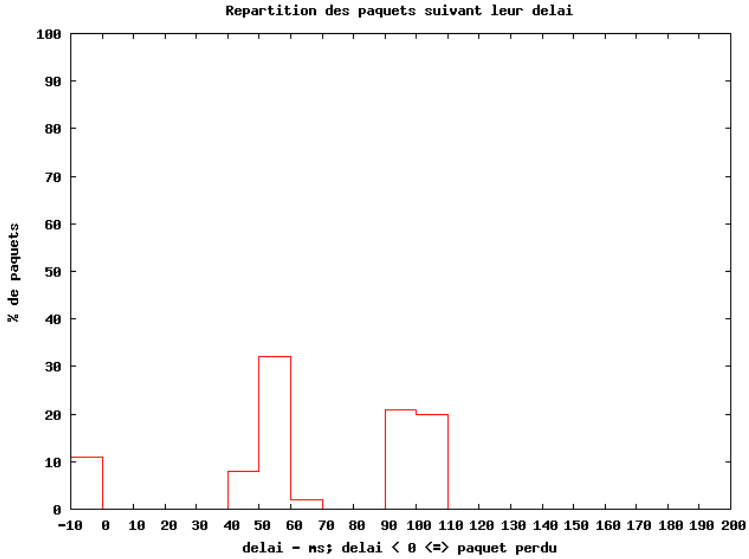


A.7 RON sur 14 nœuds

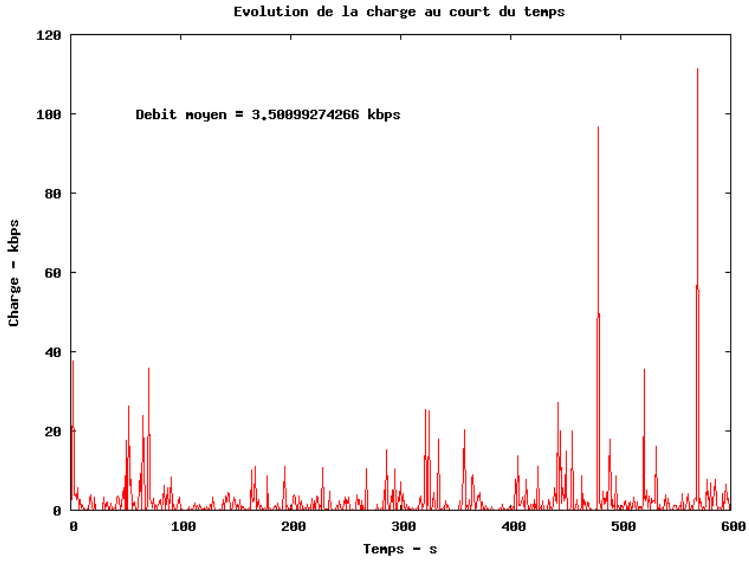
A.7.1 Variation du délai



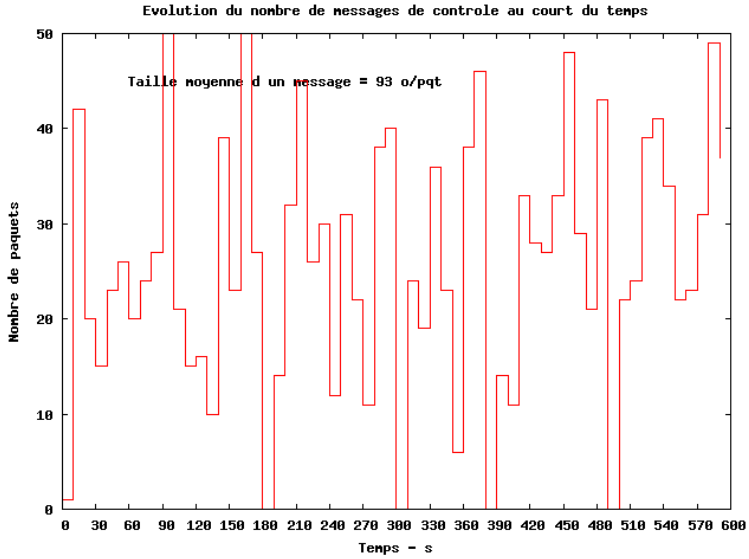
A.7.2 Répartition du délai



A.7.3 Variation du débit des messages de contrôle

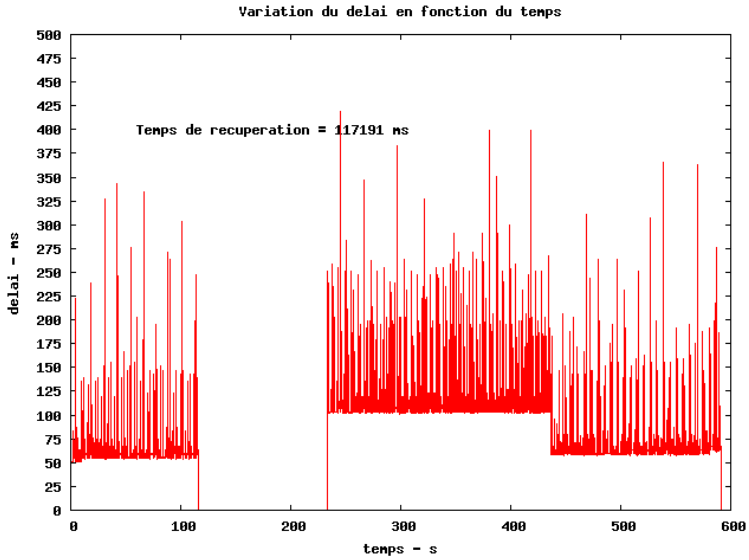


A.7.4 Evolution du nombre de messages de contrôle

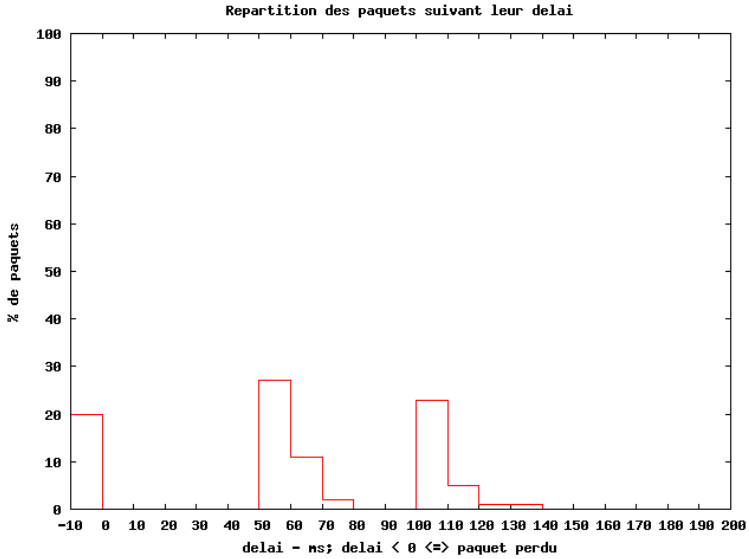


A.8 RON sur 20 nœuds

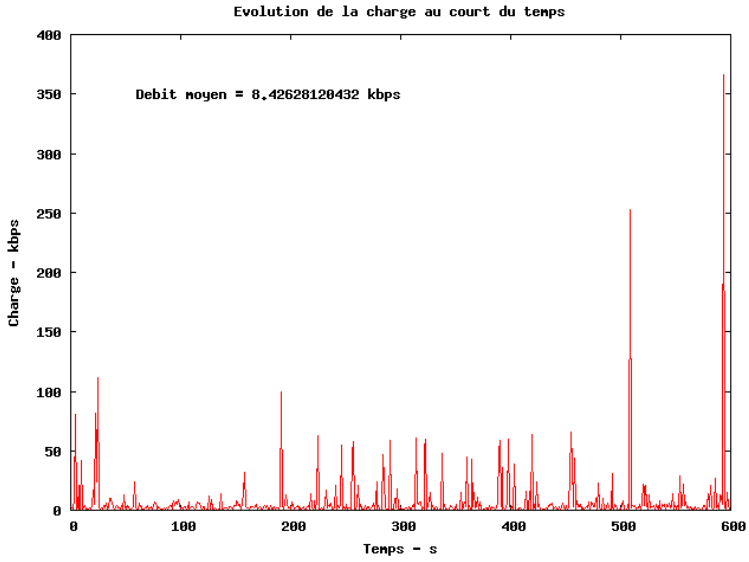
A.8.1 Variation du délai



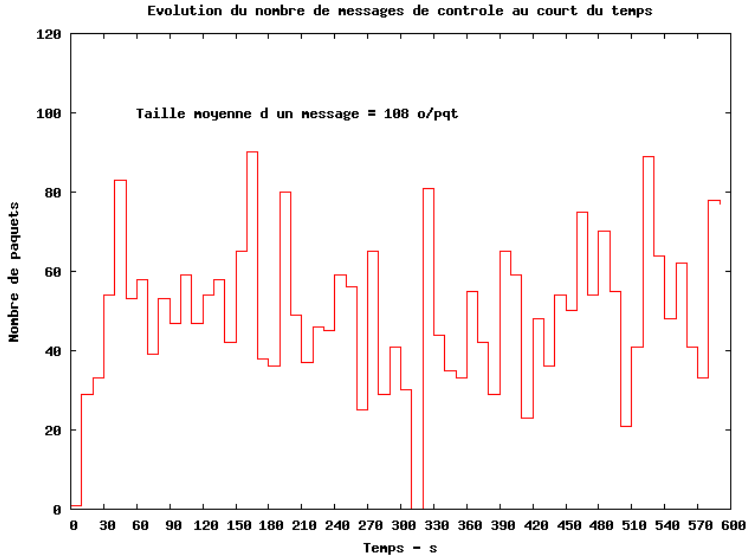
A.8.2 Répartition du délai



A.8.3 Variation du débit des messages de contrôle

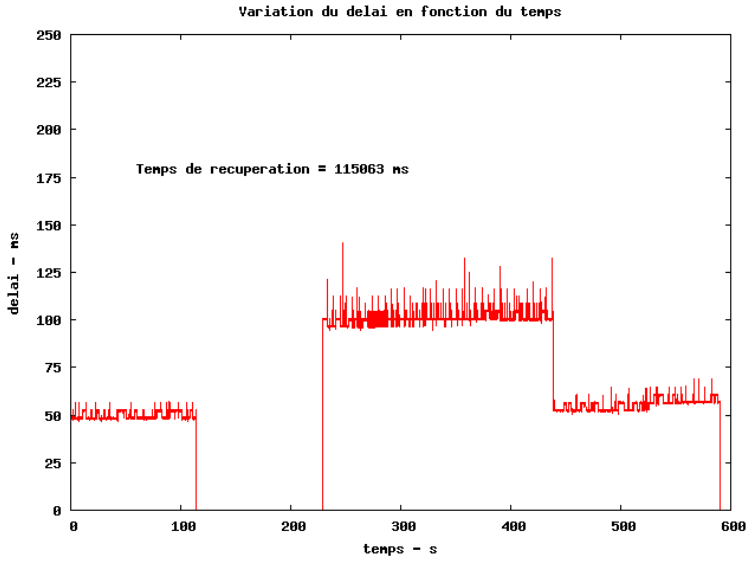


A.8.4 Evolution du nombre de messages de contrôle

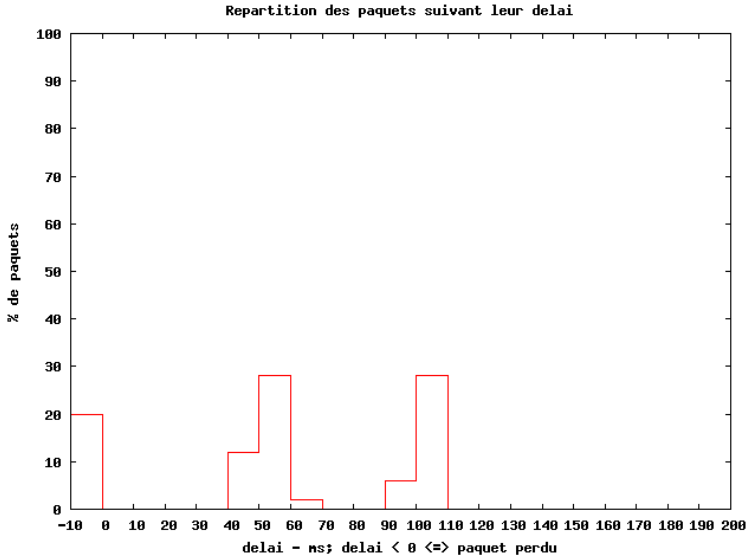


A.9 RON sur 7 nœuds et panne d'un routeur

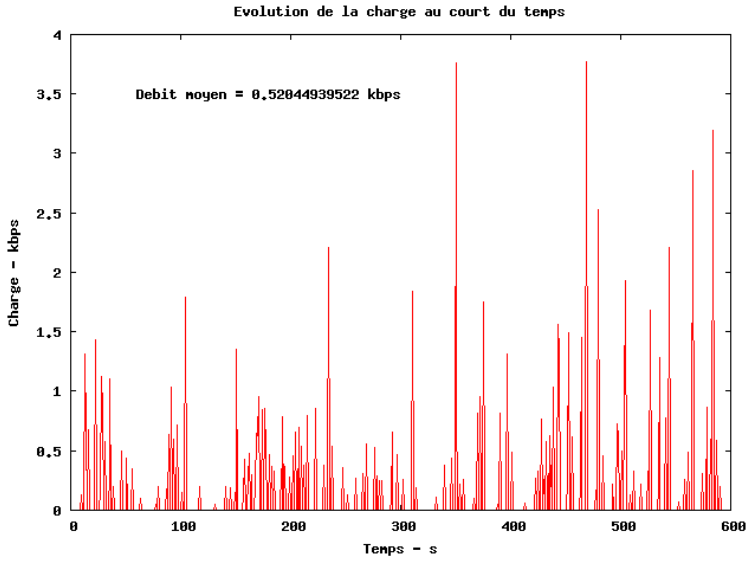
A.9.1 Variation du délai



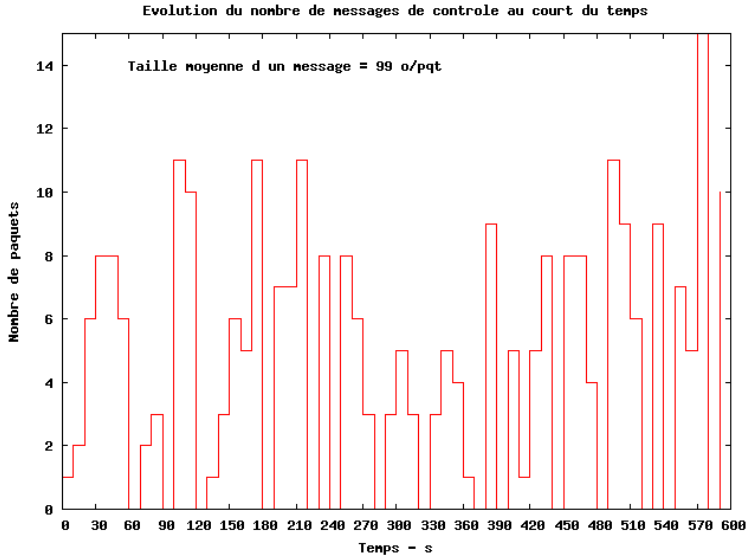
A.9.2 Répartition du délai



A.9.3 Variation du débit des messages de contrôle

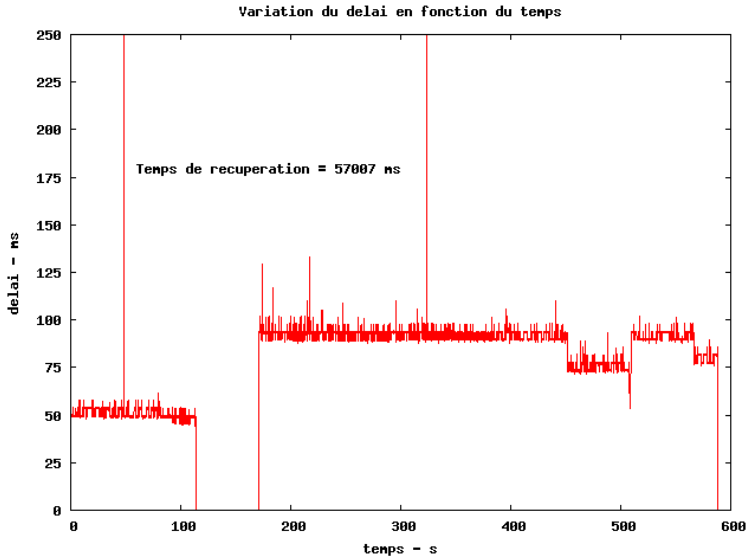


A.9.4 Evolution du nombre de messages de contrôle

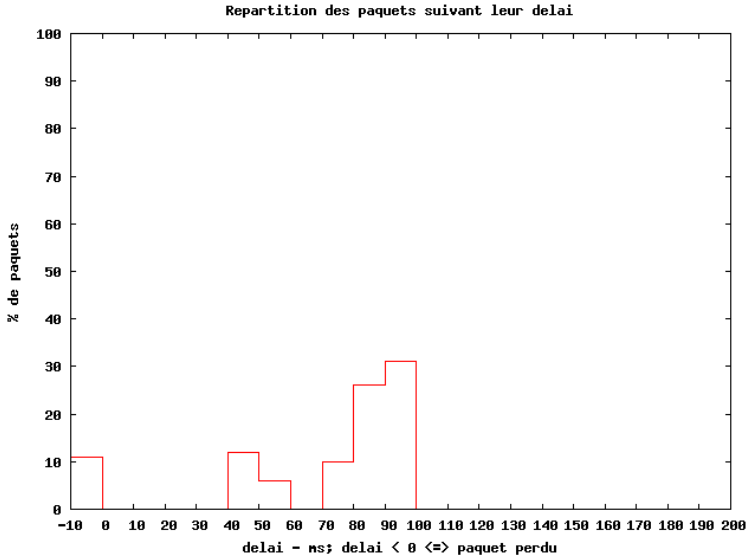


A.10 RON sur 5 nœuds avec OSPF

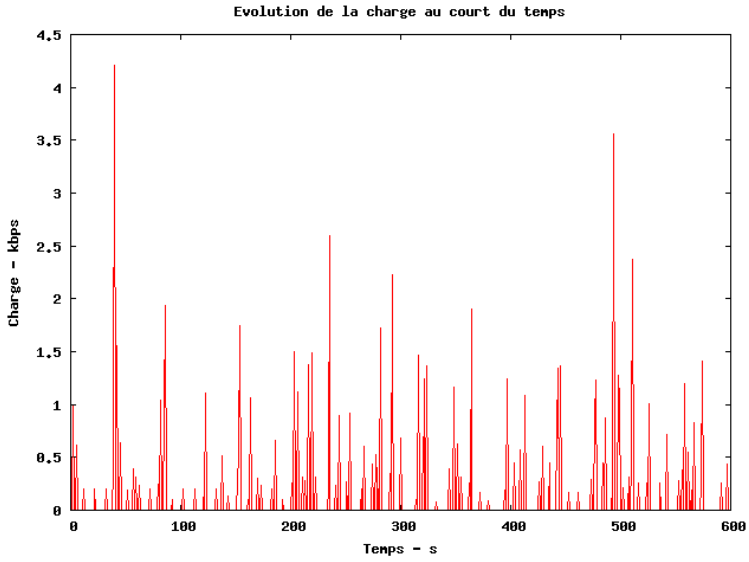
A.10.1 Variation du délai



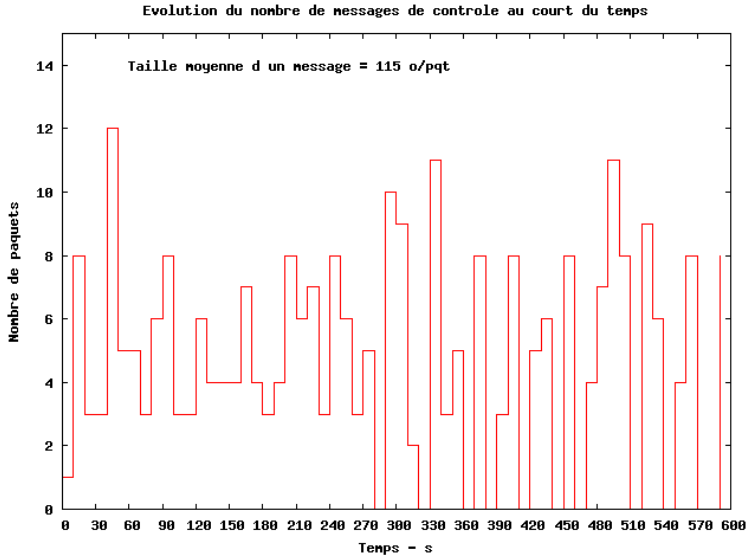
A.10.2 Répartition du délai



A.10.3 Variation du débit des messages de contrôle



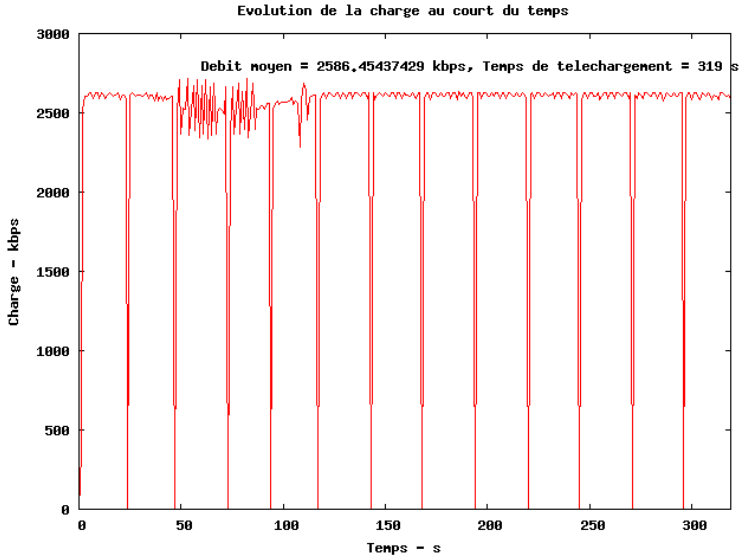
A.10.4 Evolution du nombre de messages de contrôle



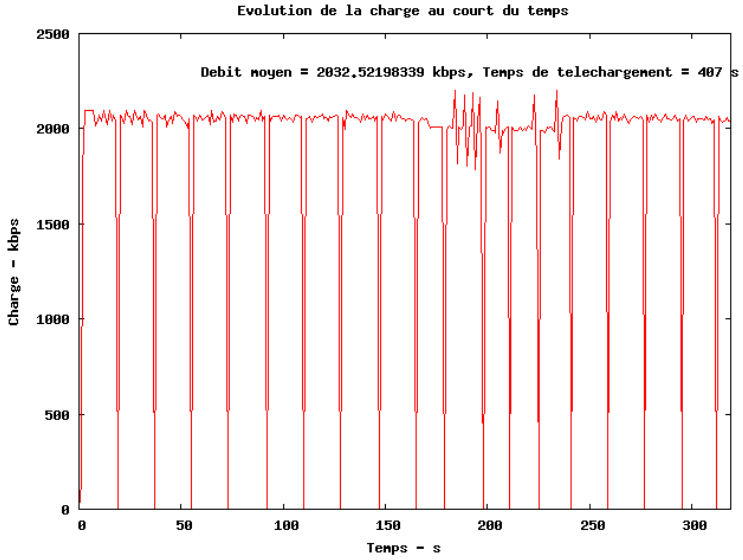
A.11 Mesure de performance de TCP

A.11.1 OSPF

Débit TCP en situation normale

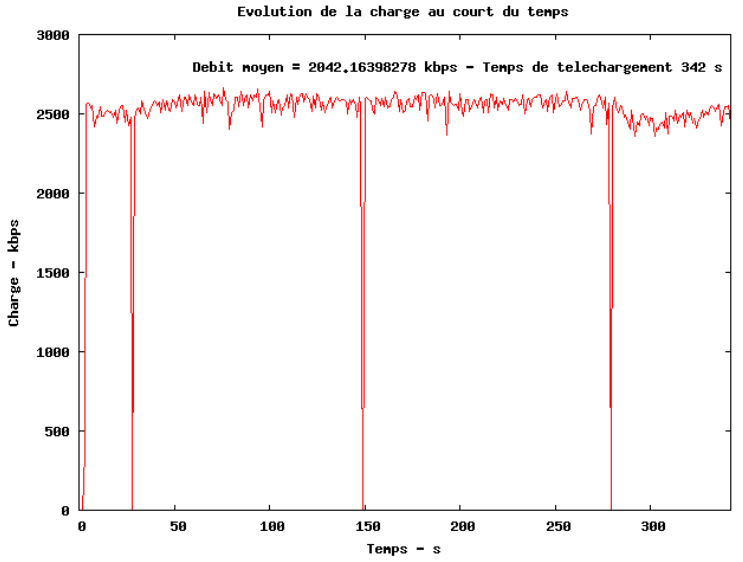


Débit TCP en situation de re-routage

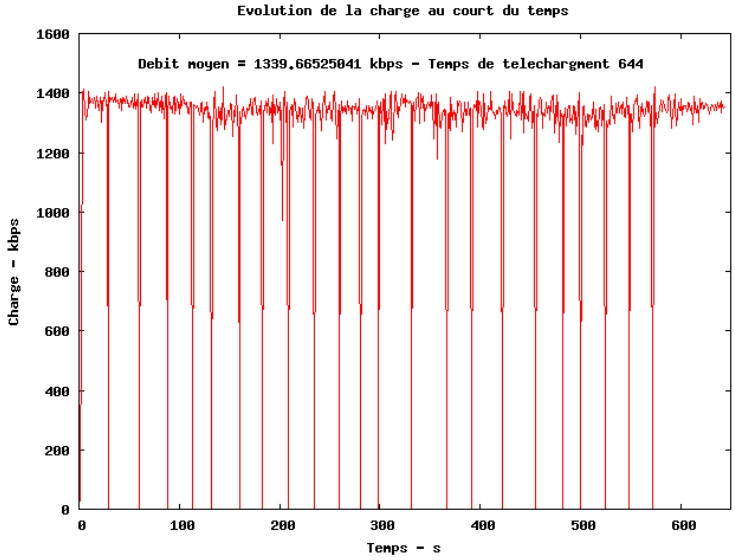


A.11.2 RON

Débit TCP en situation normale



Débit TCP en situation de re-routage



Annexe B

Signification des acronymes utilisés

- AAA : Authentication Authorization Accounting
- AS : Autonous System
- ATM : Asynchronous Transfer Mode
- BDD : Base de données
- CAN : Content Addressable Network
- CERT : Computer Emergency Response Team
- CDN : Content Delivery Network
- DNS : Domain name system
- DSL, xDSL : Digital Subscriber Line
- DSLAM : Digital Subscriber Line Access Multiplexor
- EGP : Exterior Gateway Protocol
- EIGRP : Enhanced Interior Gateway Routing Protocol
- FAI : Fournisseur d'Accès à Internet
- HTTP : Hypertext Transfer Protocol
- IDS : Intrusion Detection System
- IETF : Internet Engineering Task Force
- IGP : Interior Gateway Protocol
- IGRP : Interior Gateway Routing Protocol
- IP : Internet Protocol
- IS-IS : Intermediate System to Intermediate System
- IX : Internet Exchange Point
- OSPF : Open Shortest Path First
- MAC : Medium Access Controller
- MPEG : Moving Picture Experts Group
- MPEG-TS : Moving Picture Experts Group - Transport Stream
- MPLS : Multiprotocol Label Switching
- NAS : Network Access Server
- NP : Non-deterministic Polynomial time
- NRA : Nœud de Raccordement Abonné
- P2P : Peer To Peer
- RIP : Routing Information Protocol
- RON : Resilient Overlay Network
- RSVP-TE : Resource Reservation Protocol - Traffic Engineering
- RTP : Real-time Transport Protocol
- RTSP : Real Time Streaming Protocol
- RTT : Round Trip Time
- TCP : Transmission Control Protocol
- TV : Television
- UDP : User Datagram Protocol

– VoD : Video on Demand
– WDM : Wavelength Division Multiplexing
– WiMAX : Worldwide Interoperability for Microwave Access

Bibliographie

[1] *Critical Information Infrastructure Research Co-ordination Project*. http ://www.ci2rco.org/.

[2] J. Moy. *OSPF version 2*. Internet Engineering Task Force, 1998. RFC 2328.

[3] Cisco. *OSPF Design Guide*. http ://www.cisco.com/warp/public/104/1.html.

[4] D. Oran. *OSI IS-IS intra-domain routing protocol*. Internet Engineering Task Force,1990. RFC 1142.

[5] G. Malkin. *RIP Version 2*. Internet Engineering Task Force, 1998. RFC 2453.

[6] Cisco. *Interior Gateway Routing Protocol*. http ://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/igrp.htm.

[7] Cisco. *Enhanced IGRP*. http ://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/en_igrp.htm.

[8] D. Andersen, H. Balakrishnan, M. Frans Kaashoek, R. Morris. *Resilient Overlay Networks*. Proc. 18th ACM SOSP, Banff, Canada, October 2001.

[9] J. Adhye, V. Firoiu, D. Towsley, J. Kurose. *Modeling TCP Throughput : A Simple Model and its Empirical Validation*. In Proc. ACM SIGCOMM, Vancouver, Canada, September 1998.

[10] M. Molnár, M. Tezeghdanti. *Reroutage dans OSPF avec des chemins de secours*. Projet ARMOR, Rapport de recherche n°4340, Décembre 2001.

[11] M. Goyal, K. Ramakrishnan, W. Feng. *Achieving faster failure detection in OSPF networks*. Proc. IEEE International Conference on Communications 2003, vol. 1, May 2003, pp. 296-300.

[12] D. Gao, Z. Zhou, H. Zhang. *A Novel Algorithm for Fast Detection of Network Failure*. Photonic Network Communications, Volume 9, Issue 1, Jan 2005, Pages 113 - 120.

[13] G. Choudhury. *Prioritized Treatment of Specific OSPF Version 2 Packets and Congestion Avoidance*. Internet Engineering Task Force, 2005. RFC4222.

[14] P. Pan et al. *Fast Reroute Extensions to RSVP-TE for LSP Tunnels*. Internet Engineering Task Force, 2005. RFC4090.

[15] D. Stamatelakis, W. Grover. *IP Layer Restoration and Network Planning Based on Virtual Protection Cycles*. IEEE Journal on Selected Areas in Communications (JSAC) 18(10), October 2000.

[16] M. Medard, S. Finn, R. Barry, R. Gallager. *Redundant trees for preplanned recovery in arbitrary vertex-redundant or edge-redundant graphs*. IEEE/ACM Transactions on Networking, 7(5) :641–652, Oct. 1999.

[17] G. Xue, L. Chen, K. Thulasiraman. *Delay reduction in redundant trees for pre-planned protection against singlelink/node failure in 2-connected graphs*. in Proc. IEEE GLOBECOM, 2002, pp. 2691-2695.

[18] P. Narvaez, K.Y. Siu, H.Y. Tzeng. *New dynamic SPT algorithm based on a ball-and-string model*. In INFOCOM, pages 973–981, 1999.

[19] Y. Liu, A.L.N. Reddy. *A fast rerouting scheme for OSPF/ISIS Networks*. http ://www.ece.tamu.edu/ reddy/papers/yong_icccn04.pdf.

[20] A. Kvalbein, A.F. Hansen, T. Cicic, S. Gjessing, O. Lysne. *Fast recovery from link failures using resilient routing layers*. Computers and Communications, 2005. ISCC 2005. Proceedings. 10th IEEE Symposium on , vol. 27-30, no.pp. 554- 560, June 2005.

[21] Z. Li, B. Li, D. Jiang, L.C. Lau. *On Achieving Optimal End-to-End Throughput in Data Networks : Theoretical and Empirical Studies*. ECE Technical Report, University of Toronto, February 2004

[22] C. Gkantsidis, P.R. Rodriguez, *Network coding for large scale content distribution*. INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE , vol.4, no.pp. 2235- 2245 vol. 4, March 2005.

[23] S. Ratnasamy, P. Francis, M. Handley, R. Karp, S. Shenker. *A Scalable Content-Adressable Network*. SIGCOMM01, August 2001.

[24] I. Stoica, R. Morris, D. Karger, M.F. Kaashoek, H. Balakrishnan. *Chord : A scalable peer-to-peer lookup service for internet applications*. In Proceedings of the 2001 Conference on Applications, Technologies, Architectures, and Protocols For Computer Communications. SIGCOMM '01.

[25] F. Dabek, R. Cox, F. Kaashoek, R. Morris. *Vivaldi : A Decentralized Network Coordinate System*. In Proc. of SIGCOMM 2004.

[26] P. A. Felber, E. W. Biersack. *Self-scaling Networks for Content Distribution*. In Self-Star : International Workshop on Self Properties in Complex Information Systems, 2003.

[27] B. Cohen. *Incentives to Build Robustness in BitTorrent*. http ://bitconjurer.org/BitTorrent/bittorrentecon.pdf, 2003.

[28] D. Bickson, D. Malkhi. *The Julia Content Distribution Network*. In the 2nd Usenix

[29] D. Kosti, R. Braud, C. Killian, E. Vandekieft, J.W. Anderson, A.C. Snoeren, A. Vahdat. *Maintaining High-bandwidth under Dynamic Network Conditions*. Proceedings of 2005 USENIX Annual Technical Conference, April 2005.

[30] N.G. Tse, Y.H. Chu, S.G. Rao, K. Sripanidkulchai, H. Zhang. *Measurement based optimization techniques for bandwidth demanding peer to peer systems*. Proceedings of the USENIX/ACM Symposium on Networked Systems Design and Implementation, 2004.

[31] M. Castro, P. Druschel, A.M. Kermarrec, A. Nandi, A. Rowstron, A. Singh. *SplitStream : High-Bandwidth Multicast in Cooperative Environments*, Proc. of the 19th ACM Symposium on Operating Systems Principles (SOSP 2003), October 2003.

[32] A. Rowstron, P. Druschel. *Pastry : Scalable, decentralized object location and routing for large-scale peer-to-peer systems*. IFIP/ACM International Conference on Distributed Systems Platforms (Middleware), Heidelberg, Germany : 329-350, Nov 2001.

[33] M. Castro, P. Druschel, A.M. Kermarrec, A. Rowstron. *Scribe : A large-scale and decentralized application-level multicast infrastructure*. IEEE Journal on Selected Areas in Communications, 20 :1489-1499, 2002.

[34] P. Maigron. *Le tour du net en questions*. http ://www-public.int-evry.fr/ maigron/Internet/. Juin 2006

[35] C. Rigney et al. *Remote Authentication Dial In User Service (RADIUS)*. Internet Engineering Task Force, 2000. RFC 2865.

[36] P. Calhoun et al. *Diameter Base Protocol*. Internet Engineering Task Force, 2003. RFC 3588.

[37] C. Rigney. *RADIUS Accounting*. Internet Engineering Task Force, 2000. RFC 2866.

[38] H. Schulzrinne et al. *Real Time Streaming Protocol (RTSP)*. Internet Engineering Task Force, 1998. RFC 2326.

[39] B. Halabi. *Internet Routing Architectures*. Cisco Press. 1997.

[40] N. Feamster, D.G. Andersen, H. Balakrishnan, M.F. Kaashoek. *Measuring the effects of internet path faults on reactive routing*. SIGMETRICS '03. ACM Press, New York, NY, 126-137. 2003.

[41] J. Post, S. Jerrold, E. Shaw, K. Ruby. *Managing the threat from within*. Information Security Magazine. http ://infosecuritymag.techtarget.com/articles/july00/features2.shtml. July 2000.

[42] C. Schuba,I. Krsul,M. Kuhn,E. Spafford, A. Sundaram, D. Zamboni. *Analysis of a denial of service attack on TCP*. PROC IEEE COMPUT SOC SYMP RES SECUR PRIVACY , pp. 208-223. 1997

[43] L. Garber. *Denial-of-service attacks rip the internet*. Computer , vol.33, no.4pp.12-17, Apr 2000

[44] F. Lau, S.H. Rubin, M.H. Smith, L. Trajkovic. *Distributed denial of service attacks*. Systems, Man, and Cybernetics, 2000 IEEE International Conference on , vol.3, no.pp.2275-2280 vol.3, 2000

[45] Z. Wilson. *Hacking : The Basics*. SANS Institue. http ://rr.sans.org/toppapers/hack_basics.php. 2003.

[46] *CERT : Computer Emergency Response Team*. http ://www.cert.org.

[47] H. Schulzrinne et al. *RTP : A Transport Protocol for Real-Time Applications*. Internet Engineering Task Force, 2003. RFC 3550.

[48] L. Wenke, S. Salvatore, M. Kui. *A Data Mining Framework for Building Intrusion Detection Models*. sp , p. 0120, 1999.

[49] *QEMU Open Source Processor Emulator*. http ://fabrice.bellard.free.fr/qemu/.