

N° d'ordre : 4461

THÈSE

présentée à

L'UNIVERSITÉ BORDEAUX I

ÉCOLE DOCTORALE DE MATHÉMATIQUES ET D'INFORMATIQUE

Par **Jingxian LU**

POUR OBTENIR LE GRADE DE

DOCTEUR

SPÉCIALITÉ : **Informatique**

**L'auto-diagnostic dans les réseaux autonomes. Application à la
supervision de services multimédia sur réseau IP de nouvelle génération**

Soutenue le : 19 Décembre 2011

Après avis des rapporteurs :

Eric Fabre Directeur de recherche, INRIA
Pascal Lorenz Professeur, Université de Haute Alsace

Devant la commission d'examen composée de :

Eric Fabre Directeur de recherche, INRIA
Zoubir Mammeri Professeur, Université Paul Sabatier (*Président du Jury*)
Pascal Lorenz Professeur, Université de Haute Alsace
Toufik Ahmed Professeur, IPB, Université Bordeaux
Yacine Ghamri-Doudane ... Maître de conférences (HDR), Université Paris-Est
Christophe Dousson Expert sénior, France Télécom (*Responsable de thèse*)
Francine Krief Professeur, IPB, Université Bordeaux (*Directrice de thèse*)

Titre : *L'auto-diagnostic dans les réseaux autonomes. Application à la supervision de services multimédia sur réseau IP de nouvelle génération*

Résumé : Les réseaux autonomes représentent un intérêt certain pour les opérateurs de télécommunications. L'auto-diagnostic, pour la détection des pannes et des dysfonctionnements, est une fonction critique dans le cadre de ces réseaux.

Nous avons opté pour l'utilisation d'un diagnostic à base de modèles car il permet un diagnostic automatique, distribué et adapté à l'architecture des réseaux autonomes. Ce diagnostic est basé sur une modélisation explicite des comportements normaux ou anormaux du système. Nous utilisons ensuite un algorithme de diagnostic générique qui s'appuie sur cette modélisation pour réaliser l'auto-diagnostic. La modélisation utilisée est à base de graphe causal. Elle est une représentation intuitive et efficace des relations de causalités qui existent entre les observations et les pannes.

Notre algorithme d'auto-diagnostic qui s'appuie sur l'utilisation de graphes causaux, fonctionne sur le principe suivant : lorsqu'une alarme est déclenchée, l'algorithme est lancé et, grâce aux relations de causalité entre l'alarme et les causes, les causes primaires vont pouvoir être localisées. Puisque le graphe causal permet une modélisation modulaire et extensible, il est possible de le séparer ou de le fusionner pour répondre aux besoins des services et architectures de communication. Cette caractéristique nous permet de proposer un algorithme distribué qui s'adapte à l'architecture des réseaux autonomes. Nous avons, ainsi, proposé un algorithme d'auto-diagnostic qui permet de réaliser le diagnostic distribué correspondant à l'architecture du réseau autonome afin de réaliser un diagnostic global.

Nous avons implémenté cet algorithme sur une plateforme OpenIMS, et nous avons montré que notre algorithme d'auto-diagnostic pourrait être utilisé pour différents types de service. Les résultats obtenus correspondent bien à ce qui est attendu.

Mots clés : *auto-diagnostic, IMS, VoIP, algorithme de diagnostic distribué.*

Title : *Self-diagnosis in autonomic networks. Application to the supervision of multimedia services on next generation IP network*

Abstract : The autonomic networks show certain interest to manufacturers and operators of telecommunications. The self-diagnosis, the detection of failure and malfunction, is a critical issue in the context of these networks.

We choose based-model diagnosis because it allows an automatic diagnosis, and is suitable to distributed network architecture. This diagnosis is based on an explicit modeling of normal and abnormal behavior of the system. We then use a generic diagnostic algorithm that uses this modeling to perform self-diagnosis. The modeling used is based on causal graph. It is an intuitive and efficient representation of causal relationships between observations and failures.

The self-diagnosis algorithm we proposed based on the use of causal graphs. The principle is : when an alarm is triggered, the algorithm is run and, with the causal relationships between alarms and causes, the principal causes will be located. Since the causal graph modeling allows a modular and extensible model, it is possible to separate or merge according to the needs of services and communication architectures. This feature allows us to propose a distributed algorithm that adapts to autonomic network architecture. We have thus proposed a self-diagnosis algorithm that allows for the diagnosis corresponding to the autonomic network architecture to realize a global diagnosis.

We have implemented this algorithm on a platform OpenIMS, and we showed that our self-diagnostic algorithm could be used for different types of services. The results of implement correspond to what is expected.

Keywords : *self-diagnosis, IMS, VoIP, distributed diagnosis algorithm.*

Remerciements

Je commencerai naturellement par remercier le professeur Francine Krief et Monsieur Christophe Dousson, pour leur encadrement. Ils n'ont ménagé ni leur temps, ni leur expérience pour que ce travail aboutisse. J'ai pris beaucoup de plaisir à travailler avec deux personnes aussi sympathiques.

J'adresse également mes plus sincères remerciements aux professeurs Eric Fabre et Pascal Lorenz qui m'ont fait l'honneur d'évaluer les travaux de cette thèse.

Je remercie aussi l'ensemble des membres du jury pour leur présence lors de ma soutenance.

Je voudrais remercier aussi tous les membres de l'équipe SID et tout le personnel administratif d'Orange Labs et du LaBRI.

Sur un plan plus personnel, je tiens à remercier tout particulièrement mes parents, mon Qidong et tous les membres de ma famille. Je les remercie pour leur soutien de tous les jours, leur compréhension dans les moments difficiles. Sans eux, certains jours auraient été bien plus noirs.

Encore un grand **MERCI** à toutes et à tous !

Liste des publications

Conférences internationales avec actes

- “*Towards an Autonomic Network Architecture for Self-healing in Telecommunications Networks*”, **Jingxian Lu**, Christophe Dousson, Benoit Radier, Francine Krief, In proceeding of 4th International Conference on Autonomous Infrastructure, Management and Security (AIMS’10), 21 to 25 June 2010 in Zurich, Switzerland.
- “*A Self-diagnosis Algorithm based on Causal graphs*”, **Jingxian Lu**, Christophe Dousson, Francine Krief, In proceeding of 7th International Conference on Autonomic and Autonomous Systems (ICAS 2011), 22 to 27 2011 in Venice, Italy.
- “*Un algorithme distribué à base de graphe causal pour la fonction d’auto-diagnostic des réseaux autonomes*”, **Jingxian Lu**, Christophe Dousson, Francine Krief. 25ème Congrès DNAC : De Nouvelles Architectures pour les Communications, 16 au 18 Novembre 2011.
- “*Distributed self-diagnosis algorithm for VoIP service over IMS network*”, **Jingxian Lu**, Christophe Dousson, Francine Krief. In 9th IEEE - RIVF International Conference on Computing and Communication Technologies, in Ho Chi Minh City, Vietnam, 2012. (*Submitted*)
- “*Self-diagnosis architecture for QoS and QoE management in IPTV services over IMS network*”, **Jingxian Lu**, Christophe Dousson, Francine Krief, In Third International Conference on Communications and Networking, in Hammamet, Tunisia, 2012. (*Submitted*)

Table des figures

2.1	Les propriétés de la gestion autonome	15
2.2	Les 4 objectifs principaux d'un système autonome	16
2.3	Le processus de l'auto-rétablissement	18
2.4	L'architecture d'une gestion autonome (Boucle autonome)	20
2.5	Les défis clés du projet <i>UniverSelf</i>	28
3.1	La gestion des pannes	34
3.2	Principe du diagnostic à base de modèles	40
4.1	L'architecture NGN	49
4.2	L'architecture IMS	51
4.3	La structure interne du cœur IMS et les interfaces	52
4.4	La signalisation du processus de VoIP	58
4.5	Architecture IMS-IPTV	60
5.1	L'architecture du gestionnaire autonome chargé de l'auto-diagnostic	67
5.2	Représentation graphique des cinq types de nœuds du graphe causal	69
5.3	Représentation graphique des liens "Cause sûre" et "Cause possible"	70
5.4	Graphe causal pour IMS (extrait)	71
5.5	Illustration des définitions S , $\hat{S}$, S^{-1} et $\hat{S}^{-1}$	75
6.1	Exemple d'un diagnostic insuffisant	85
7.1	Différentes architectures d'un système de diagnostic	93
7.2	Frontières des effets et des causes	96
7.3	Architecture d'un réseau autonome chargé de l'auto-diagnostic	100
8.1	OpenIMS	106
8.2	Infrastructure IMS open source	112
8.3	Le graphe causal global	113
8.4	Le graphe causal restreint à HSS et AS	114
8.5	Le graphe causal utilisé	115

8.6	Interface administrative de <i>Mobicents</i>	116
8.7	Interface de gestion des <i>Sip Servlets</i> de <i>Mobicents</i>	117
8.8	Critères de filtrage initiaux (IFC)	118
8.9	Profil de service pour le service déployé sur <i>Mobicents</i>	119
8.10	Flux d'appel lors de l'invocation d'un serveur d'application	120
8.11	Enregistrement du client «a»	121
8.12	L'alarme déclenchée au niveau de <i>AS1</i>	122
8.13	Les 3 types de nœud	122
8.14	Hypothèse 2 au niveau de <i>AS1</i>	123
8.15	Hypothèse 3 au niveau de <i>AS1</i>	123
8.16	L'hypothèse 7 exige la communication avec <i>HSS</i>	124
8.17	Lancement des <i>Tests</i>	125
8.18	Le résultat concernant l'hypothèse 7	126
8.19	Les résultats concernant l'hypothèse 4	127
8.20	Les résultats concernant l'hypothèse 5	128
8.21	Les résultats concernant l'hypothèse 6	129
8.22	Le graphe causal pour le service IPTV	130

Table des matières

1	Introduction	1
1.1	Contexte Général	1
1.2	Objectif	3
1.3	Plan de la thèse	4
I	Etat de l’art	7
2	Réseaux autonomes	9
2.1	Introduction	9
2.2	Les principaux concepts	12
2.2.1	Historique de l’ <i>autonomic</i>	12
2.2.2	Les termes techniques	13
2.2.3	Les objectifs de la gestion autonome	16
2.3	Architecture de la gestion autonome	20
2.3.1	Les ressources gérées	21
2.3.2	Les interfaces de gestion	21
2.3.3	Les gestionnaires autonomes	22
2.4	Les projets de recherche	25
2.4.1	Projets européens	26
2.4.2	Projets industriels et académiques	28
2.4.3	Synthèse	30
2.5	Conclusion	30
3	Diagnostic dans les réseaux de télécommunications	31
3.1	Réseaux de télécommunications	31
3.1.1	L’Internet	32
3.1.2	La convergence des réseaux	33
3.2	La gestion des pannes	34
3.2.1	Définition d’une panne	35
3.2.2	Approches utilisées pour la gestion des pannes	36

3.3	Vision IA du Diagnostic	37
3.3.1	Introduction	37
3.3.2	Les deux principales approches	37
3.3.3	Intérêt du diagnostic à base de modèles pour les réseaux au- tonomes	45
3.4	Conclusion	45
4	Les services multimédia sur réseau IP de nouvelle génération	47
4.1	Introduction à IMS	47
4.1.1	Motivation pour l'utilisation de l'IMS	49
4.1.2	L'architecture IMS	50
4.1.3	Cœur du réseau IMS	52
4.1.4	Protocoles utilisés dans l'IMS	54
4.1.5	Conclusion	56
4.2	Service VoIP	56
4.3	Service IPTV	58
4.4	Conclusion	60
II	Proposition d'une méthode d'auto-diagnostic	63
5	Diagnostic avec Graphes Causaux	65
5.1	Introduction	65
5.2	Architecture du gestionnaire autonome	67
5.3	Modélisation de la connaissance par le graphe causal	68
5.3.1	Nœuds causaux	68
5.3.2	Liens causaux	70
5.3.3	Exemple	70
5.4	Le processus de diagnostic du diagnostiqueur	72
5.5	Définitions et notations	74
5.5.1	Relations de causalité, graphe causal	74
5.5.2	Pannes, Symptômes et Diagnostic	75
5.5.3	Diagnostics minimaux	76
5.6	Conclusion	77

6	L'Algorithme de Diagnostic de base	79
6.1	Introduction	79
6.2	Recherche des diagnostics	80
6.3	Algorithme de recherche des diagnostics minimaux	82
6.4	Alarmes absentes ou perdues	83
6.5	Introduction des Tests	84
6.5.1	Proposition initiale pour le lancement des <i>Tests</i>	84
6.5.2	Solution améliorée pour un diagnostic insuffisant	85
6.5.3	Politiques de lancement des <i>Tests</i>	88
6.6	Cause sûre et cause possible	88
6.7	Conclusion	89
7	Distribution de l'algorithme de diagnostic	91
7.1	Systèmes distribués	91
7.2	Architectures pour un système de diagnostic	94
7.3	Principe	95
7.4	Algorithme distribué	98
7.5	Architecture du réseau autonome	100
7.6	Conclusion	101
III	Implémentation et évaluation	103
8	Implémentation dans un environnement IP de nouvelle génération	105
8.1	Introduction	106
8.2	Description des composants	106
8.2.1	Le cœur de réseau Open IMS	106
8.2.2	Les terminaux	107
8.2.3	La plate-forme de développement	108
8.3	Implémentation	111
8.3.1	Scénario d'implémentation	111
8.3.2	Flux d'appel	120
8.3.3	Les résultats de tests	121
8.4	Extension proposée pour le service IPTV	130

8.5 Synthèse des résultats et analyse	131
8.6 Conclusion	131
9 Conclusion et perspectives	133
A	137
A.1 Représentation d'un graphe causal	137
Bibliographie	143

CHAPITRE 1 Introduction

1.1 Contexte Général

Au tout début, les opératrices du réseau téléphonique se chargeaient de mettre en relation manuellement les clients qui voulaient rentrer en communication. A partir des années 1920, le développement des réseaux est devenu de plus en plus rapide. Ce développement spectaculaire nécessita le déploiement, l'exploitation et l'administration d'infrastructures complexes et distribuées.

L'intervention manuelle étant devenue difficile voire impossible, la nécessité de décharger l'opérateur manuel de la fonction de contrôle a entraîné l'introduction des autocommutateurs.

Le monde des technologies de la communication, de l'information en général et des réseaux en particulier, a connu une deuxième grande révolution, au même titre que ce qu'elle a connu lors de l'automatisation du plan de contrôle, avec le réseau Internet. Depuis la fin du XX^e siècle, nous vivons un développement sans précédent des nouvelles technologies. Les réseaux sont devenus accessibles de partout (réseaux ubiquitaires). Les services disponibles deviennent de plus en plus variés : la VoIP (voix sur IP), l'IPTV (télévision sur IP), la VOD (vidéo à la demande), etc. Ces services dépendent fortement de la QoS (Qualité de Service). Les terminaux de communication sont de plus en plus nombreux et hétérogènes (PC, Mobiles et Smartphones, etc.).

La tendance actuelle est désormais à la convergence entre les différents types de réseaux : réseaux fixes et mobiles, réseaux téléphoniques et réseaux informatiques, réseaux cellulaires et réseaux sans fil (WiFi). Cette convergence ne signifie pas l'uniformisation par une seule technologie, mais l'interconnexion entre celles-ci via des tentatives d'uniformisation des protocoles (IP, SIP, ...). Afin d'assurer la qualité

de service, la nécessité de techniques innovantes est donc cruciale.

Ces évolutions rendent la gestion du cœur du réseau de plus en plus complexe, il faut en effet prendre en compte les exigences des différents types de services utilisés par les terminaux communicants. Alors que la gestion traditionnelle des réseaux se faisait de manière centralisée et que l'optimisation des paramètres de configuration était réalisée par des experts humains, l'ensemble des paramètres à considérer est maintenant devenu tellement important et complexe qu'il est impossible de le gérer manuellement. Le diagnostic et la correction des pannes pour la gestion de réseaux deviennent de véritables défis avec un cœur de réseau complexifié par la traversée de différents services et l'hétérogénéité des équipements.

Avec le développement rapide des technologies informatiques et de télécommunication, les capacités de traitement des données des ordinateurs ont sensiblement augmenté et l'architecture de connexion entre les équipements s'est complexifiée (réseaux fixes, mobiles, sans fil). Quand une panne ou une erreur se produit à un endroit d'un réseau, il est possible qu'elle ait des répercussions dans le réseau entier. Ceci rend difficile la tâche qui consiste à trouver la cause de la panne et donc de la dépanner en temps voulu. Une panne peut influencer les différents composants ou les autres opérations du réseau. Il est évident que l'accroissement des délais pour identifier les raisons d'une panne et des délais de réparation vont augmenter le coût de réparation. La seule véritable option pour protéger les équipements du réseau et pour améliorer le service du réseau contre ces pannes, est la prévention et la détection automatique et évolutive.

Les administrateurs ne doivent pas seulement configurer les équipements pour interconnecter les réseaux, mais ils doivent aussi assurer la stabilité des systèmes. Chaque mise à jour d'un équipement peut engendrer des perturbations et des pannes. Cette complexité conduit souvent à retarder le changement de systèmes, même s'ils sont désuets. D'autre part, la formation pour les techniciens et les experts capables de gérer les équipements réseaux et d'optimiser les configurations est certainement une source de charge importante pour les entreprises et les opérateurs réseau. Pourtant, l'un des objectifs de la gestion du réseau est de réduire les coûts (*Total Cost Ownership*). La gestion du réseau par des techniciens et experts semble également augmenter les temps d'indisponibilité durant les diagnostics et les optimisations. En contrepartie, la fiabilité du système permet de réduire les coûts.

Ainsi, 40% des pannes de services sont causées par des problèmes de configuration réalisées par les opérateurs [Patterson 2002] qui doivent prendre des décisions trop rapidement [Brown & Patterson. 2001].

La réalisation de la gestion autonome est un nouveau concept. Le but est de permettre aux équipements de se gérer automatiquement tout en respectant des politiques de haut niveau fournies par l'administrateur. Ce nouveau concept doit permettre de diminuer la complexité de la gestion des équipements et réduire l'activité humaine dans la boucle de gestion des réseaux. C'est l'équivalent de la révolution du téléphone avec l'automatisation du plan de contrôle dans les années 20, mais cette fois-ci sur la gestion des réseaux de communications. Ce concept se base sur quatre fonctions autonomes principales que sont l'auto-configuration, l'auto-optimisation, l'auto-protection et l'auto-rétablissement.

L'idée d'utiliser les systèmes autonomes, qui ont la capacité de s'adapter à leur environnement sans interventions humaines pour réaliser des fonctions automatiques, est entrain d'être développée et de s'affiner continuellement. Le plan de connaissance proposé est basé sur cette idée. En effet, il s'agit d'une structure qui permet de gérer le réseau en utilisant la connaissance du réseau et des processus autonomes.

Durant ces dernières années, le diagnostic a joué un rôle de plus en plus important dans la gestion du réseau. La réalisation de l'auto-diagnostic a un attrait considérable pour les opérateurs, puisqu'ils peuvent réduire les coûts d'exploitation comme mentionnés ci-dessus, ainsi qu'améliorer l'efficacité et la rapidité du dépannage. Alors que le diagnostic traditionnel basé souvent sur le système expert ne répond pas aux besoins d'évolution constante des réseaux d'aujourd'hui, au contraire, le diagnostic à base de modèle peut être modifié et mis à jour en exploitant le plan de connaissance et correspond à la demande de développement des réseaux futurs. Mes travaux de thèse se sont centrés sur cette problématique.

1.2 Objectif

Cette thèse s'intéresse à la problématique de l'auto-diagnostic, et plus précisément à étudier dans quelle mesure les approches du diagnostic peuvent être utilisées dans le cadre des réseaux autonomes. Nous proposons ainsi une étude et une série de

propositions pour utiliser le diagnostic à base de modèle dans les réseaux autonomes. Notre cas d'étude à porté sur l'auto-diagnostic pour les applications s'appuyant sur un réseau IMS pour lequel nous avons développé un algorithme d'auto-diagnostic distribué. Dans ce contexte, le travail inclut :

- La détection au plus tôt de situations anormales dégradant le niveau de service qui pourraient se manifester par une alarme ;
- Le diagnostic des causes primaires de ces défaillances (dysfonction d'équipement, panne de logiciel, problème de base de données, panne de synchronisation, panne de configuration, etc.) ;
- La définition du modèle (graphe causal) pour expliciter le plan de connaissance concernant les pannes, les causes et les relations causales entre elles ;
- La réalisation de l'algorithme d'auto-diagnostic distribué capable de tenir compte de l'architecture du réseau autonome.

Ce travail se situe principalement dans un cadre d'exécution simple de service VoIP, mais peut facilement être étendu pour d'autres services tels que l'IPTV. L'auto-diagnostic des pannes liées à des alarmes rentre dans le cadre de ce travail de thèse.

1.3 Plan de la thèse

Le présent document de thèse est organisé en 8 chapitres dont les contenus respectifs sont listés ci-dessous.

Chapitre 2 : Dans ce chapitre, nous présentons un état de l'art des réseaux autonomes en commençant par l'historique et les motivations conduisant à la mise en œuvre de réseaux autonomes. Cet historique permet de montrer l'importance et la nécessité de réaliser des réseaux autonomes dans les réseaux de futures générations. Ce chapitre précise également l'architecture de base de la gestion autonome ainsi que celle des éléments autonomes selon la vision d'IBM. Ce chapitre se termine par la présentation des projets européens concernant la thématique des réseaux autonomes. Cette présentation permet ainsi d'introduire les concepts des réseaux autonomes pour réaliser les fonctions de gestion autonome.

Chapitre 3 : Ce chapitre introduit d'abord l'architecture des réseaux de télécommunications. Ensuite, ce chapitre présente nos travaux de recherche qui se situent dans l'environnement des réseaux de télécommunications. Ce chapitre présente également l'aire fonctionnelle de gestion des pannes car nos travaux se focalisent sur cette aire pour réaliser l'auto-diagnostic.

Nous introduisons ensuite le diagnostic, présentons son objectif et proposons une définition d'une panne. Nous comparons ensuite les deux principales approches pour le diagnostic s'appuyant sur l'Intelligence Artificielle (IA) en montrant les avantages et les inconvénients de chacune de ces approches. Nous montrons également que le diagnostic à base de modèles est mieux adapté pour notre architecture de réseau.

Chapitre 4 : Ce chapitre présente les services multimédia dans un environnement IMS. D'abord, nous introduisons les motivations pour l'utilisation de l'IMS, puis présentons et illustrons l'architecture IMS et la structure de cœur IMS. Le cœur du réseau IMS est utilisé pour le contrôle des sessions média. Ensuite, les protocoles utilisés dans IMS (comme *SIP*, *AAA*, etc.), standardisés par l'IETF, sont présentés. Enfin, ce chapitre se consacre, dans la partie suivante, à l'étude des services *VoIP* et *IPTV* dans un environnement IMS afin d'y introduire la réalisation de l'auto-diagnostic.

Chapitre 5 : Dans ce chapitre, nous présentons la modélisation de la causalité des pannes par un graphe causal et présentons aussi le principe général du diagnostic qui sera formellement mis en œuvre par les algorithmes introduit aux chapitres 6 et 7. Ce chapitre décrit également les définitions et notations nécessaires à la formalisation du processus de diagnostic.

Chapitre 6 : Ce chapitre présente l'algorithme de base du diagnostic mis en œuvre par le diagnostiqueur ainsi que les différentes extensions proposées afin de prendre en compte les tests en cours de diagnostic. L'algorithme présenté dans ce chapitre permet de calculer l'ensemble de tous les diagnostics, il s'appuie sur un ensemble d'hypothèses et se termine lorsqu'il ne reste aucun nœud à expliquer dans les hypothèses (C'est-à-dire lorsque les hypothèses sont closes.). Nous avons aussi tenu compte du cas des alarmes absentes ou perdues et du cas où les observables

ne remontent pas spontanément mais correspondent à des tests à déclencher (ou à des requêtes à exécuter). Le développement de notre algorithme peut passer par le choix du test à exécuter, c'est-à-dire qu'une partie de la politique de diagnostic réside alors dans le choix du test à effectuer.

Chapitre 7 : Ce chapitre décrit le principe de la distribution de l'algorithme et explicite l'algorithme formel. La technique utilisée pour exécuter le diagnostic est dépendant de l'architecture du système de diagnostic, donc nous présentons d'abord les différentes architectures de diagnostic : centralisée, hiérarchique, distribuée et hiérarchique et distribuée. Dans ce chapitre, l'algorithme montré repose sur l'architecture distribuée. L'algorithme distribué que nous proposons permet de fournir des explications à des alarmes déclenchées en tenant compte des différents composants d'un système autonome. L'algorithme permet également de faire la communication entre les différents diagnostiqueurs locaux. La communication est réalisée par les nœuds de «connexion», il s'agit de trouver une explication distante pour ce nœud et de vérifier que l'état de ce nœud n'est pas compromis vis-à-vis d'un observable distant.

Chapitre 8 : Dans ce chapitre, nous présentons des descriptions des composants utilisés pour réaliser l'implémentation de notre algorithme d'auto-diagnostic sur une plate-forme IMS (OpenIMS). Nous présentons ensuite un descriptif des différentes étapes d'exécution de l'algorithme de diagnostic. Nous présentons, également, les différents éléments utilisés dans le cadre de cette démonstration de faisabilité : le graphe causal utilisé, le scénario de test considéré et la configuration du service de VoIP des tests. L'implémentation de notre approche est réalisée sur une plateforme IMS, et nous montrons que notre algorithme d'auto-diagnostic pourrait être utilisé pour un exemple de service VoIP. Les résultats obtenus correspondent bien à ce qui est attendu. Notre approche est ensuite étendue au service IPTV.

Ce mémoire de thèse se finit par une conclusion et décrit un ensemble de perspectives pour des recherches futures.

Première partie

Etat de l'art

CHAPITRE 2 Réseaux autonomes

Sommaire

2.1	Introduction	9
2.2	Les principaux concepts	12
2.2.1	Historique de l' <i>autonomic</i>	12
2.2.2	Les termes techniques	13
2.2.3	Les objectifs de la gestion autonome	16
2.3	Architecture de la gestion autonome	20
2.3.1	Les ressources gérées	21
2.3.2	Les interfaces de gestion	21
2.3.3	Les gestionnaires autonomes	22
2.4	Les projets de recherche	25
2.4.1	Projets européens	26
2.4.2	Projets industriels et académiques	28
2.4.3	Synthèse	30
2.5	Conclusion	30

2.1 Introduction

A partir des années 1920, le développement des réseaux est devenu de plus en plus rapide. Ce développement spectaculaire nécessita le déploiement, l'exploitation et l'administration d'infrastructures complexes et distribuées. L'intervention ma-

nuelle étant devenue difficile voire impossible, la nécessité de décharger l'opérateur manuel de la fonction de contrôle a entraîné l'introduction des autocommutateurs.

Le monde des technologies de la communication, de l'information en général et des réseaux en particulier, a connu une deuxième grande révolution, au même titre que ce qu'elle a connu lors de l'automatisation du plan de contrôle, avec le réseau Internet. Depuis la fin du XX^e siècle, nous vivons un développement sans précédent des nouvelles technologies. Les réseaux sont devenus accessibles de partout (réseaux ubiquitaires). Les services disponibles deviennent de plus en plus variés : la VoIP (voix sur IP), l'IPTV (télévision sur IP), la VOD (vidéo à la demande), etc. Ces services dépendent fortement de la QoS (Qualité de Service). Les terminaux de communication sont de plus en plus nombreux et hétérogènes (PC, Mobiles et Smartphones, etc.).

La tendance actuelle est désormais à la convergence entre les différents types de réseaux : réseaux fixes et mobiles, réseaux téléphoniques et réseaux informatiques, réseaux cellulaires et réseaux sans fils (WiFi). Cette convergence ne signifie pas l'uniformisation par une seule technologie, mais l'interconnexion entre celles-ci via des tentatives d'uniformisation des protocoles (IP, SIP, ...). Afin d'assurer la qualité de service, la nécessité de techniques innovantes est donc cruciale.

Ces évolutions rendent la gestion du cœur du réseau de plus en plus complexe, il faut en effet prendre en compte les exigences des différents types de services utilisés par les terminaux communicants. Alors que la gestion traditionnelle des réseaux se faisait de manière centralisée et que l'optimisation des paramètres de configuration était réalisée par des experts humains, le nombre de paramètres à considérer est maintenant devenu tellement important et complexe qu'il est impossible de le gérer manuellement. Le diagnostic et la correction des pannes pour la gestion de réseaux deviennent également de véritables défis avec un cœur de réseau complexifié par la traversée de différents services et l'hétérogénéité des équipements.

Avec le développement rapide des technologies informatiques et de télécommunication, les capacités de traitement des données des ordinateurs ont sensiblement augmenté et l'architecture de connexion entre les équipements s'est complexifiée (réseaux fixes, mobiles, sans fil). Quand une panne ou une erreur se produit à un endroit d'un réseau, il est possible qu'elle ait des répercussions dans le réseau entier. Ceci rend difficile la tâche qui consiste à trouver la cause de la panne et donc de la

dépanner en temps voulu. Une panne peut influencer les différents composants ou les autres opérations du réseau. Il est évident que l'accroissement des délais pour identifier les raisons d'une panne et les délais de réparation vont augmenter le coût de réparation. La seule véritable option pour protéger les équipements du réseau et pour améliorer le service du réseau contre ces pannes, est la prévention et la détection automatique et évolutive des pannes.

Les administrateurs ne doivent pas seulement configurer les équipements pour interconnecter les réseaux, mais ils doivent aussi assurer la stabilité des systèmes. Chaque mise à jour d'un équipement peut engendrer des perturbations et des pannes. Cette complexité conduit souvent à retarder le changement de systèmes, même s'ils sont désuets. D'autre part, la formation pour les techniciens et les experts capables de gérer les équipements réseaux et d'optimiser les configurations est certainement une source de charge importante pour les entreprises et les opérateurs réseau. Pourtant, l'un des objectifs de la gestion du réseau est de réduire les coûts (*Total Cost Ownership*). La gestion du réseau par des techniciens et experts semble également augmenter les temps d'indisponibilité durant les diagnostics et les optimisations. En contrepartie, la fiabilité du système permet de réduire les coûts. Ainsi, 40% des pannes de services sont causées par des problèmes de configuration réalisées par les opérateurs [Patterson 2002] qui doivent prendre des décisions trop rapidement [Brown & Patterson. 2001]. Tous les problèmes qui viennent d'être mentionnés montrent le besoin pour des fonctions de gestion comme :

- Le monitoring continu des ressources (matérielles et logicielles) ;
- Le diagnostic et la correction des pannes de manière automatique pour trouver les causes plus rapidement afin de maintenir les paramètres de configuration et de service à des valeurs désirées.

La réalisation de la gestion autonome est un nouveau concept. Le but est de permettre aux équipements de se gérer automatiquement et sans intervention humaine en respectant des politiques de haut niveau. Ce nouveau concept peut permettre de diminuer la complexité de la gestion des équipements et de réduire l'intervention humaine dans la boucle de contrôle des réseaux. C'est l'équivalent de la révolution du téléphone avec l'automatisation du plan de contrôle dans les années 20, mais cela concerne la gestion des réseaux de communications à présent.

Ce chapitre présente les concepts de la gestion autonome. La première section

introduit les principaux concepts pour la gestion autonome, la deuxième section présente les travaux existants dans ce domaine.

2.2 Les principaux concepts

2.2.1 Historique de l'*autonomic*

Le concept de l'*autonomic* est apparu dans la période où le domaine de la recherche a commencé à tenir compte du fonctionnement de la nature pour concevoir les systèmes futurs. Effectivement, à la fin des années 1990, particulièrement en 1997, le projet SUO SAS (*Small Unit Operations Situation Awareness System*) a été lancé par DARPA ¹. Bien que l'objectif de ce projet concerne les développements militaires, il est une des pierres de fondation dans le processus de la construction de l'*autonomic*. Ce projet est une preuve de concept de l'*autonomic* et aussi une orientation pour le domaine de la recherche.

Au début des années 2000, certains grands industriels, notamment IBM et HP, ont initié une réflexion sur la *vision du futur*. Cette nouvelle vision se basait sur l'observation du fonctionnement de la nature. En 2001, IBM a proposé le concept de l'*Autonomic Computing* pour la première fois inspiré de la biologie (Système nerveux autonome). Le modèle du système nerveux autonome utilise un système distribué et hiérarchique de contrôle qui s'appuie sur des capteurs et des réactions. IBM suggère que les systèmes informatiques complexes doivent posséder les propriétés de l'*autonomic*. C'est-à-dire que les systèmes doivent être capables de prendre en charge leur propre maintenance et optimisation sans intervention de l'administrateur. Ceci permet de réduire les tâches de l'administrateur qui pourra se concentrer sur les objectifs de l'exécution du système. Ce concept a été développé largement par le domaine de la recherche et les industries, puisque il est bien adapté aux systèmes de gestion et de contrôle.

1. <http://www.darpa.mil/default.aspx>

2.2.2 Les termes techniques

Depuis la proposition du concept de l'*autonomic*, le domaine de la recherche et les industries ont porté un grand intérêt pour la nécessité d'une gestion autonome. Ce paradigme est devenu rapidement un objectif de différenciation marketing pour les grandes entreprises. Il y a donc des nombreux termes techniques pour ce concept. Ces nouveaux termes, dans le domaine de la recherche, sont présentés comme des mots clés :

- Autonomic Computing
- Autonomic Communication
- Autonomic Networking
- Autonomic and Autonomous Systems
- Cognitive Networks
- Ambient Intelligent (AmI) ecosystems
- Global service Ecosystems
- ...

Bien qu'il y ait des variations sur le sens et la portée des termes cités précédemment, la vision et l'objectif sont partagés par tous : rendre les systèmes capables de s'auto-gérer. Pourtant, il faut noter que le terme *Autonomic Communication* couvre la plupart des domaines liés aux réseaux, comme l'indique l'annotation sur le site Autonomic-Communication² :

Autonomic Communication (AC) - a new communication paradigm to assist the evolution of communication networks towards functional adaptability, extensibility and resilience to a wide range of possible faults and attacks. Special emphasis is given on the grounding principles to achieve purposeful behavior on top of self-organization (including self-management, self-healing, self-awareness, etc.). Papers are solicited that study network element's autonomic behavior exposed by innovative (cross-layer optimized, context-aware, and securely programmable) protocol stack in its interaction with numerous often dynamic network groups and communities. The goals are to understand how autonomic behaviors are learned, influenced or changed, and how, in turn,

2. <http://www.autonomic-communication.org/Annales/>

these affect other elements, groups and the network.

Les objectifs principaux de ce concept sont définis par l'ACI (*Autonomic Computing Initiative*) créé par IBM en 2001 [Horn 2001]. L'ACI regroupe les caractéristiques d'un système autonome en huit points :

1. Un système autonome doit avoir des connaissances détaillées sur lui-même (ses composants, l'état actuel, sa capacité et toutes les connexions avec les autres systèmes) pour se gérer (*Know itself*).
2. Un système autonome doit se configurer et se reconfigurer sans intervention externe, afin de s'adapter aux conditions et aux changements variables de son environnement.
3. Un système autonome doit être capable d'optimiser son fonctionnement et de surveiller son travail de façon continue pour atteindre des objectifs de haut niveau.
4. Un système autonome doit pouvoir réparer ses pannes et ses dysfonctionnements, sans endommager les données ni le délai de traitement.
5. Un système autonome doit être capable de se protéger contre les différents types d'attaques, afin d'assurer la sécurité du fonctionnement.
6. Un système autonome doit prendre connaissance de son environnement et du contexte de ses activités et faciliter ainsi son adaptation.
7. Un système autonome doit être capable de fonctionner dans un environnement hétérogène et implémenter des standards ouverts.
8. Un système autonome doit cacher la complexité de son fonctionnement aux utilisateurs, les prises de décision doivent être transparentes.

Ces huit éléments sont les conditions nécessaires pour qu'un système soit autonome. Ils peuvent être traduits en propriétés [Sterritt & Bustard 2003] représentées par la figure 2.1 [Sterritt & Bustard 2003]. Ces propriétés décrivent la vision, les objectifs, les attributs et les moyens de ce nouveau concept.

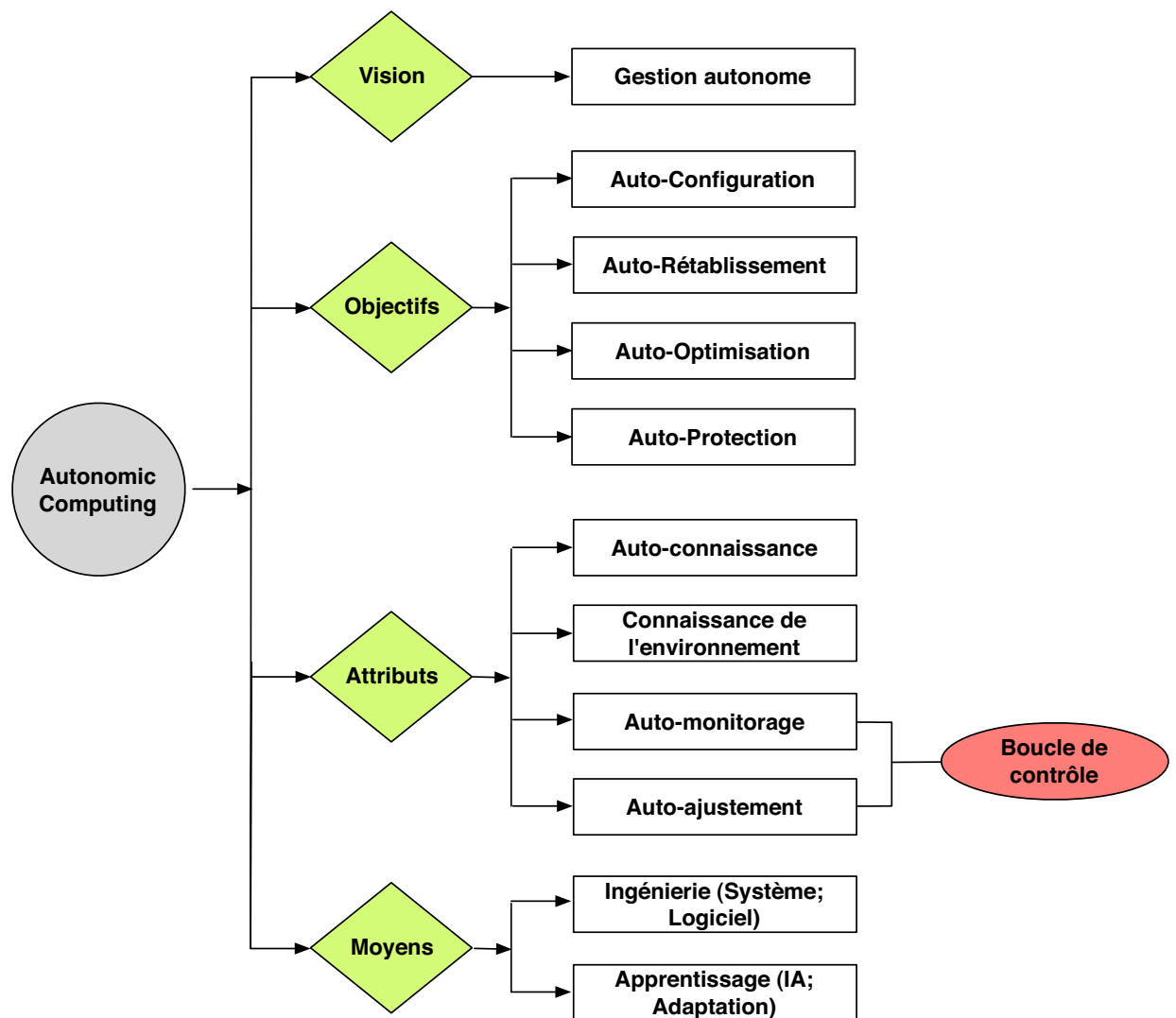


Figure 2.1 – Les propriétés de la gestion autonome

IBM a étendu sa vision [Kephart & Chess 2003, Jacob *et al.* 2002] du concept d'*Autonomic Computing* afin de couvrir les domaines des nouvelles technologies de l'informatique et des télécommunications. D'autre part, l'AS³ (Autonomic Communication), ACF⁴ (Autonomic Communication Forum) et TF-AAS⁵ (Task Force Autonomous & Autonomic Systems) tentent de rassembler les idées concernant l'*Autonomic networking* afin de proposer des standards.

2.2.3 Les objectifs de la gestion autonome

2.2.3.1 Les fonctions des réseaux autonomes

Les réseaux autonomes représentent une coupure technologique, qui peut offrir aux opérateurs et aux équipements des possibilités de réduction des coûts de gestion et d'intégration des réseaux et services, ainsi qu'une intégration rapide de nouveaux services. Cette coupure s'appuie essentiellement sur la prise de décision par des éléments autonomes situés dans le réseau et partageant leurs connaissances et leurs décisions avec d'autres éléments autonomes.

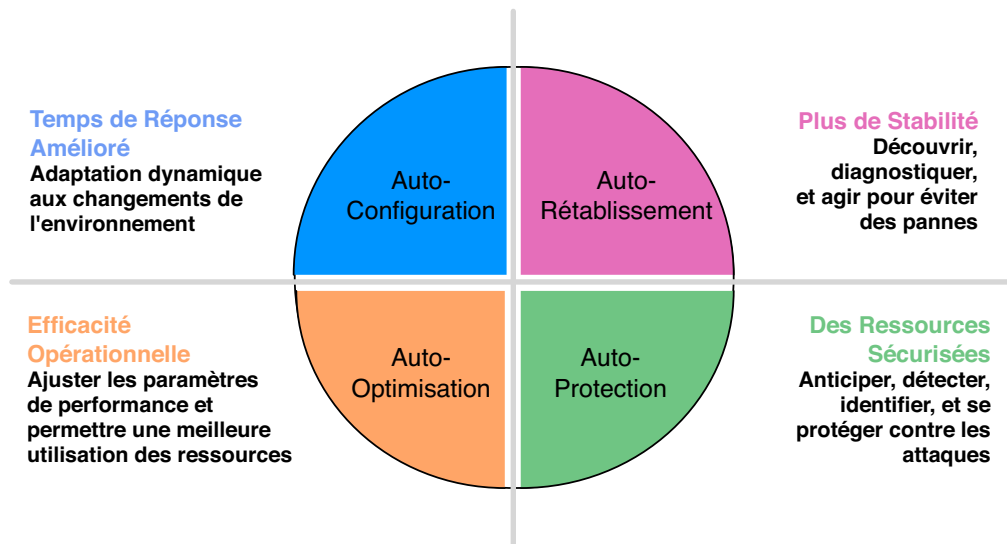


Figure 2.2 – Les 4 objectifs principaux d'un système autonome

3. <http://www.autonomic-communication.org/>

4. <http://www.autonomic-communication-forum.org/>

5. http://www.infj.ulst.ac.uk/~autonomic/IEEE_TF-AAS/index.html

Actuellement, la communauté scientifique et industrielle travaille sur l'application des systèmes autonomes aux réseaux avec l'objectif principal de construire des réseaux qui peuvent atteindre l'auto-configuration, l'auto-rétablissement (en cas de défaillance ou de panne), l'auto-optimisation, et l'auto-protection (en cas d'attaques extérieures et de malveillances) (voir la Figure 2.2).

2.2.3.2 Auto-Configuration : *Self-Configuring*

Il s'agit d'installer, de configurer et d'intégrer de grands systèmes complexes dans des délais très courts et sans erreur. C'est une tâche très complexe dans l'environnement hétérogène des réseaux actuels. Pour la gestion traditionnelle, la fonction de configuration des systèmes avec de nombreux composants hétérogènes est vraiment difficile et demande beaucoup de temps aux experts. Mais les systèmes autonomes agissent d'une manière transparente vis-à-vis de l'opérateur qui peut effectuer les politiques et les directives de haut niveau afin de réaliser l'auto-configuration. Les politiques spécifient la cible à atteindre, mais non la manière avec laquelle les composants doivent se configurer pour l'atteindre. Lorsqu'un composant autonome est introduit, il s'intègre de manière homogène dans son environnement et le reste du système s'adapte automatiquement à sa présence. Il a aussi la capacité de se réajuster automatiquement.

Les erreurs de configuration peuvent donc être évitées. Une fois les politiques de configuration spécifiées, les administrateurs peuvent gagner un temps considérable.

2.2.3.3 Auto-Rétablissement : *Self-Healing*

L'auto-rétablissement désigne la capacité du système à examiner, diagnostiquer et réagir à des dysfonctionnements du système. Les éléments et les applications d'auto-rétablissement doivent pouvoir observer les défaillances du système, évaluer les contraintes imposées par l'extérieur, et appliquer les corrections appropriées. Afin de découvrir automatiquement les dysfonctionnements du système ou les défaillances possibles, il est nécessaire de connaître le comportement prévu du système. Un système autonome doit avoir des connaissances sur son comportement propre, puis il doit avoir une connaissance pour déterminer si le comportement actuel est cohérent et prévisible dans le contexte de l'environnement. Dans de nou-

veaux contextes ou dans des scénarios différents, de nouveaux comportements du système peuvent être observés et l'entité autonome doit pouvoir réagir avec le changement d'environnement grâce à son plan de connaissance.

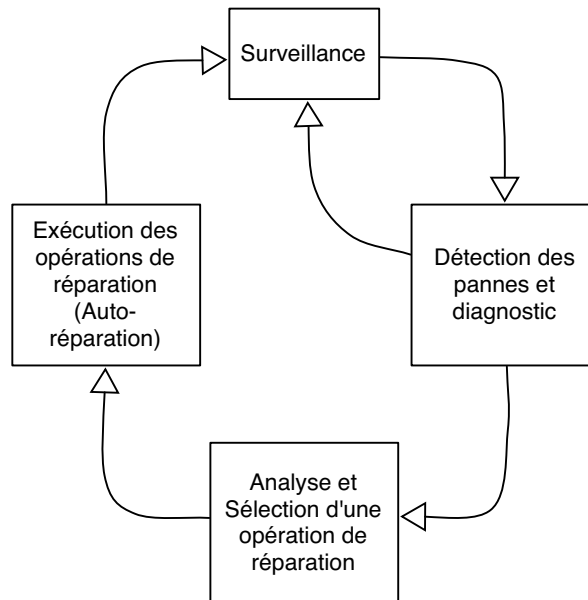


Figure 2.3 – Le processus de l'auto-rétablissement

Les systèmes d'auto-rétablissement assurent un processus afin de maintenir une qualité satisfaisante de service du système y compris en présence d'une panne. La première étape du processus est appelée "Surveillance". Pendant l'étape de "Surveillance", la supervision du système devra inspecter l'information d'environnement pour détecter toute conduite inappropriée. Les inspections de supervision terminées, il envoie les données des observations actuelles recueillies à l'étape suivante. La deuxième étape du processus est appelée "Détection des pannes et diagnostic", si le diagnostic indique qu'il n'y a pas de panne dans le système, il retournera à l'étape de "Surveillance". S'il y a une panne détectée par le superviseur, l'étape de détection d'erreur l'indiquera à l'étape suivante. La troisième étape du processus est d'analyser et de sélectionner une opération de réparation. Dans cette étape, les erreurs sont analysées et une méthode de réparation est déterminée. Une fois l'opération de réparation déterminée, la dernière étape consiste à exécuter la réparation (L'auto-réparation). Toute réparation nécessaire est réalisée dans cette étape. Une

fois les pannes réparées, le processus recommence. Puisque ce processus est une boucle fermée, il sera continu comme le montre la figure 2.3.

Le système d’auto-rétablissement doit avoir un modèle de diagnostic pour bien trouver les causes correspondant aux pannes. Sinon il n’y a aucun moyen d’évaluer si un système actuel peut se rétablir en situation d’intérêt. C’est pourquoi nos travaux se focalisent sur la deuxième étape “Détection des pannes et diagnostic”.

Nos travaux dissocient la fonction d’auto-diagnostic et d’auto-réparation qui sont contenues dans la fonction d’auto-rétablissement.

2.2.3.4 Auto-Optimisation : *Self-Optimizing*

Cet objectif permet au système de surveiller son environnement et apprendre à améliorer ses choix en terme d’optimisation. Pour bien réaliser ceci, les systèmes autonomes vont continuellement chercher à améliorer leurs fonctionnements et saisir les opportunités de devenir plus performants. En même temps, ils vont surveiller leur environnement pour trouver les ressources adaptées, afin d’assurer un fonctionnement optimal par rapport aux besoins spécifiés.

L’auto-optimisation permet une meilleure utilisation des ressources, et avec le temps, vise à obtenir une certaine expérience pour réaliser l’auto-régulation et atteindre un objectif qui permet de réaliser une QoS optimale pour les utilisateurs avec une utilisation optimale des ressources pour les systèmes [Mbaye 2009].

2.2.3.5 Auto-Protection : *Self-Protecting*

Le système autonome va détecter les situations d’attaques et se protéger pour ne pas perturber les utilisateurs. Ceci correspond à la mise en place de mécanismes et d’architectures pour la détection et la protection de l’ensemble des éléments de réseaux.

De plus, les systèmes autonomes s’auto-protègent de deux façons :

- Le système va se défendre contre les attaques malveillantes ou les défaillances en isolant ou réparant les sources de pannes qui ne sont pas traitées par l’auto-rétablissement.
- Le système doit anticiper les attaques en analysant les rapports d’événements et en lançant les actions adéquates.

2.3 Architecture de la gestion autonome

Un système autonome peut s'adapter aux changements de l'environnement. Il se gère lui-même sans intervention humaine [Horn 2001]. En plus de ces concepts, le système autonome doit être "Environment-Awareness", c'est-à-dire être capable d'appréhender et de comprendre l'environnement dans lequel il agit. Cette connaissance est indispensable au bon fonctionnement du système qui doit prendre les bonnes décisions et pouvoir interagir avec son environnement. Fournir des décisions rapides et adaptées à la situation est crucial dans des environnements fortement dynamiques comme c'est le cas aujourd'hui pour les demandes de communication. Cette prise de décision, qui peut concerner les changements de configuration ou le traitement du trafic, doit être coordonnée.

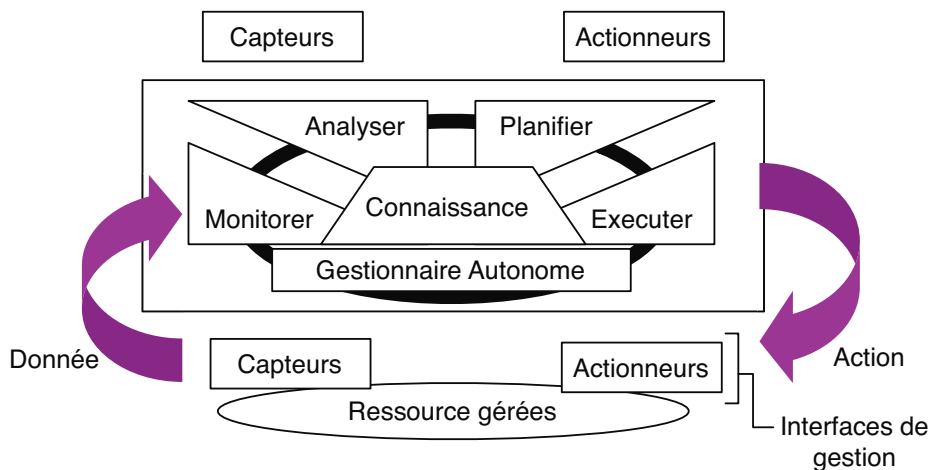


Figure 2.4 – L'architecture d'une gestion autonome (Boucle autonome)

Comme la figure 2.4 [Kephart & Chess 2003, Jacob *et al.* 2002] le montre, un élément autonome est essentiellement composé d'une ou de plusieurs ressources gérées, de capteurs, d'actionneurs, d'un gestionnaire autonome et, également d'une boucle de contrôle qui doit être implémentée par ce dernier afin de réaliser les quatre objectifs de gestion autonome.

2.3.1 Les ressources gérées

Une ressource gérée est un composant matériel ou bien un composant logiciel, qui peut être contrôlé. Par rapport à une ressource traditionnelle, un élément géré dans l'architecture autonome doit permettre au gestionnaire de le surveiller et de le contrôler grâce à l'utilisation des interfaces de gestion. Les points de contact pour implémenter ces interfaces correspondent à des capteurs et des actionneurs.

Un élément géré peut être une ressource unique telle qu'un serveur, une base de données, une application, un service, une mémoire ou encore un ensemble de ressources tel qu'un ensemble de serveurs. Un gestionnaire autonome peut lui-même être une ressource gérée ou faire partie d'un ensemble de ressources à gérer. C'est pourquoi il dispose également de capteurs et d'actionneurs.

2.3.2 Les interfaces de gestion

L'interface de gestion, pour contrôler un élément géré, est organisée en capteurs et actionneurs. L'implémentation du comportement d'un capteur et d'un actionneur spécifique à une ressource est assurée par des points de contacts appelés, également, point de communication qui permettent de réduire la complexité des différentes technologies pouvant être utilisées par les ressources, et ce en offrant aux gestionnaires autonomes une interface standardisée afin de contrôler les ressources dont il est responsable.

- **Capteur** : C'est l'intégration des différents mécanismes qui sont capables de collecter les informations des éléments gérés. Un capteur peut traiter l'ensemble des opérations qui permettent de retrouver les informations sur l'état de l'élément géré. Ainsi, un capteur peut traiter la série des événements de gestion qui se traduit par l'envoi de messages asynchrones informant les changements intervenus sur l'état de l'équipement géré.
- **Actionneur** : C'est l'agrégation des différents mécanismes qui sont utilisés pour changer le comportement d'un élément géré. Un actionneur peut traiter l'ensemble des opérations qui permettent la modification de la configuration de l'élément géré.

Dans l'architecture de gestion autonome, les opérations des capteurs et des actionneurs sont associées à la *boucle de contrôle*. Effectivement, un changement de

configuration effectué par un actionneur doit être accompagné d'une notification de ce changement grâce à l'interface du capteur. Cette notification qui peut être considérée comme un type de *feed-back*, permet aux processus d'évaluer l'impact des actions.

2.3.3 Les gestionnaires autonomes

Le gestionnaire autonome est la partie la plus importante dans l'architecture qui exécute le concept de boucle de contrôle sans l'intervention humaine. En fait, si un système est capable de s'auto-gérer, il doit utiliser une méthode automatique qui peut permettre de collecter toutes les informations dont il a besoin et de les analyser pour identifier si un changement doit être opéré sur le système. Si c'est le cas, le gestionnaire autonome doit établir un plan contenant une séquence d'actions qui va être exécutée afin de réaliser les changements nécessaires.

2.3.3.1 Les fonctions d'un gestionnaire autonome

Une représentation de la *boucle de contrôle* est déjà montée dans la figure 2.4. Cette boucle permet de réaliser les actions de manière automatique. Elle est composée des quatre fonctions indépendantes reliées par le *plan de connaissance*.

- **Monitoring** : cette fonction offre des mécanismes qui réalisent la collecte, l'agrégation, la corrélation et le filtrage des données. Ces données peuvent porter sur la topologie, les paramètres de configuration, l'état fonctionnel, la capacité offerte ou encore les mesures provenant des ressources gérées dont le gestionnaire est responsable. La collecte des informations se fait par l'intermédiaire des points de contact, et plus précisément de l'interface capteur de la ressource. Ensuite, une quantité importante de données, qui peut être soit statique soit dynamique, est regroupée en symptômes qui vont être analysés dans une seconde étape.
- **Analyse** : cette fonction offre des mécanismes permettant d'observer et d'analyser des situations afin de déterminer si un changement doit être effectué pour satisfaire un objectif ou une directive de haut niveau. Ainsi, la fonctionnalité d'analyse est responsable du respect des politiques de haut niveau durant toute l'activité des composants réseau. Pour réaliser cette tâche, cette

fonction implique l'utilisation de modèles comportementaux complexes et de techniques de prédiction afin de permettre au gestionnaire autonome d'apprendre à connaître l'environnement dans lequel se trouvent les ressources gérées. Ces outils permettent, aussi, d'anticiper sur l'évolution de l'environnement et sur les besoins en terme de ressources afin de s'y adapter au mieux. Le gestionnaire autonome doit alors être capable de faire des analyses, des raisonnements et des traitements complexes sur les symptômes reçus de la fonction de monitoring. Si des modifications doivent être effectuées, la fonction d'analyse va passer une requête à la fonction de planification. Cette requête contient la description des modifications qui doivent être effectuées afin d'atteindre un fonctionnement respectueux des objectifs de gestion.

- **Planification** : cette fonction offre des mécanismes permettant de planifier les actions nécessaires pour atteindre les objectifs fixés. Ces mécanismes de planification peuvent prendre la forme de règles de politique. Cette fonction permet la création ou la sélection de procédures capables de mettre en œuvre les changements souhaités par la fonction d'analyse sur les ressources gérées. Le résultat de la planification peut avoir plusieurs formes allant d'une simple commande à un enchaînement ordonné de tâches. Les modifications ainsi planifiées peuvent être passées à la fonction d'exécution.
- **Exécution** : Cette fonction offre des mécanismes permettant de contrôler l'exécution d'un plan d'actions afin de réaliser les changements nécessaires sur le système. En effet, lorsque le gestionnaire autonome génère un plan (fonction planification), un ensemble d'actions doit être entrepris pour modifier l'état d'une ou de plusieurs ressources gérées dont il a le contrôle. La fonction d'exécution de la boucle de contrôle d'un gestionnaire autonome est responsable de la mise en œuvre de la procédure générée par la fonction de planification. Ces actions sont exécutées en utilisant le point de contact actionneur de la ressource concernée par les modifications. Le résultat de ces plans d'exécution peut permettre ensuite d'enrichir la connaissance partagée entre les différentes fonctions [Mbaye 2009].

2.3.3.2 La Connaissance

Les données utilisées par les quatre fonctions assurées par le gestionnaire autonome sont stockées et partagées telle qu'illustré dans la figure 2.4 (Connaissance).

La connaissance utilisée par le gestionnaire autonome peut être obtenue de diverses manières. En effet, les informations qui enrichissent cette connaissance peuvent être passées au gestionnaire par l'intermédiaire de son interface de type actionneur, et ce sous forme par exemple de règles de politique qui ne sont autre qu'un ensemble de règles de comportement ou des préférences qui influencent les décisions du gestionnaire. Un autre moyen d'enrichir cette base de connaissance est la récupération d'information d'un service externe susceptible de fournir des informations sur le système sous forme de rapports précisant les différents événements qui sont survenus sur les différents composants.

Enfin, le gestionnaire peut créer lui même sa connaissance par le biais des informations récupérées par la fonction monitoring et la procédure de planification qui en a résulté. Grâce aux informations remontées par les capteurs, il pourra récupérer les conséquences des actions planifiées et exécutées.

La connaissance peut être de plusieurs natures :

- les informations sur la topologie du réseau. Celles-ci permettent de spécifier les différents composants formant le système autonome. Ainsi, elles contiennent les constructions de ces composants avec leurs capacités, ce qui facilitera la planification.
- les politiques qui sont les règles qui peuvent être consultées pour déterminer si des changements sont nécessaires ou non pour atteindre les objectifs du système autonome. Ces règles doivent être définies de manière standardisée pour qu'elles puissent être partagées par plusieurs gestionnaires autonomes. Ce partage permet au système de fonctionner en se basant sur un ensemble de règles de politique cohérentes. Cependant, dans ce contexte il y a plusieurs niveaux de cohérence qui sont plus ou moins nécessaires suivant les objectifs du système : une cohérence interne, du voisinage locale et globale.
- La connaissance qui permet la détermination et la résolution de problème en s'appuyant sur une adaptation de la configuration des ressources gérées de manière automatique et suivant les changements de l'environnement. Ces

informations évolutives permettent au système autonome de reconnaître les symptômes de mauvais fonctionnement du système.

Dans nos travaux, nous modéliserons la connaissance en utilisant les graphes causaux pour réaliser l'auto-diagnostic. Le chapitre 5 décrira cette modélisation plus en détail.

2.3.3.3 Le plan de connaissance

Le plan de connaissance qui était né dans l'esprit de régler des limites spécifiques à Internet va être adopté (ou intégré), par la suite, dans les architectures de gestion autonomes pour, finalement, y prendre une place assez importante. En parallèle, le concept de connaissance a fait naître de nouveaux concepts telles que la gestion orientée connaissances qui est une gestion de réseaux basée sur les connaissances acquises par le plan de connaissance.

La vision de base du plan de connaissance est celle de la conception de nouveaux systèmes réseaux avec de nouvelles techniques avancées permettant d'avoir plus d'intelligence distribuée au niveau des équipements. Cependant, la contrainte qui est posée est celle de garder les avantages des systèmes classiques (extensibilité, facilités de déploiement de nouvelles applications). Les équipements au cœur du réseau auront ainsi une "*connaissance*" du comportement des applications aux extrémités du réseau et s'adapteront en fonction de ces connaissances mais aussi des objectifs de haut niveau qui leur sont fixés par les administrateurs. Cette vision permet d'avoir une adaptation automatique de la configuration des éléments réseaux sous la contrainte des configurations "acceptables" fournies par les administrateurs. Cette adaptation est basée sur une vue globale de la connaissance du réseau qui permet aux équipements de prendre des décisions de bas niveau sans violer les objectifs de haut niveau qu'ils doivent atteindre [Mbaye 2009].

2.4 Les projets de recherche

Plusieurs travaux ont mis en avant des systèmes et des architectures qui ont pour objectif de réaliser une ou plusieurs fonctions de gestion autonome avec des mécanismes et concepts offrant des degrés différents d'autonomie.

2.4.1 Projets européens

Au niveau européen, une action de coordination appelée ACCA (*Autonomic Communication Coordination Action*) qui fait partie des projets ouverts du programme FP6 (*Framework Programme*) concernant les technologies émergentes FET (*Future and Emergent Network Technologies*) dans l'IST (*Information Society Technologies*) a stimulé de nouvelles initiatives de recherche dans le domaine de la gestion autonome. Ainsi, dans le cadre du SAC (*Situated and Autonomic Communication*), quatre projets portant sur la gestion autonome ont été proposés : BIONETs (*Biologically-Inspired Autonomic Networks*), ANA (*Autonomic Network Architecture*), CASCADAS (*Componentware for Autonomic, Situated-aware Communication and Dynamically Adaptable Services*) et enfin HAGGLE (*A novel Communication Paradigm for Autonomic Opportunistic Communication*). Le programme FP7 a propulsé d'autres projets tels que AutoI (*Autonomic Internet*), UniverSelf.

Le projet **BIONETs** (2006-2009) [Pellegrini *et al.* 2006] suppose que les environnements de communications futurs seront aussi complexes que les organismes ou encore les écosystèmes biologiques. Ainsi, BIONETS cherche à s'inspirer des systèmes biologiques dans le but de concevoir un environnement qui permet de faciliter l'intégration d'un grand nombre de composants ou encore de ressources hétérogènes, et qui est capable de s'adapter et évoluer de façon autonome. Dans ce cadre, le projet BIONETS vise plutôt la recherche fondamentale en prenant comme modèle les systèmes biologiques.

Le projet **ANA** (2006-2009) [Tschudin *et al.* 2008] explore de nouvelles méthodes d'organisation et d'utilisation au delà des technologies traditionnelles de l'Internet. Ces travaux visent la conception et le développement d'une nouvelle architecture réseau qui permet la formation d'entités globales mais aussi de réseaux entiers de façon flexible, dynamique et totalement autonome. L'objectif scientifique de ce projet est l'identification des principes fondamentaux des réseaux autonomes. De plus, ce dernier a développé une plateforme de démonstration pour tester l'architecture de gestion autonome proposée.

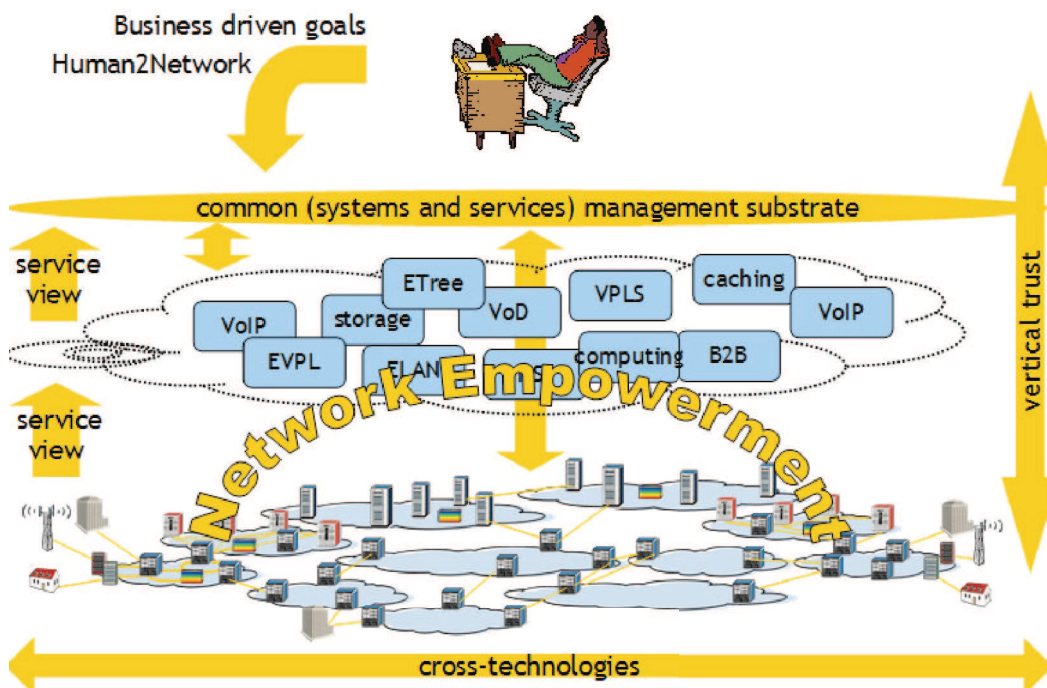
Le projet **CASCADAS** [CASCADAS 2006] 2006-2008 a comme objectif l'identification, le développement et l'évaluation d'une abstraction (ACE : *Autonomic Communication Element*) d'usage universel pour les services de communications

autonomes. Dans ce cadre, ces éléments réalisent d'une façon autonome des fonctions d'auto-organisation et d'auto-adaptation pour la fourniture de services adaptés et situés. Enfin, ce projet vise à introduire de nouvelles technologies dans le domaine de l'*autonomic networking*.

Le projet **HAGGLE** [Giordano & Puiatti 2006] (2006-2010) propose une nouvelle architecture de gestion autonome des réseaux pour permettre la communication en présence de connectivité intermittente dans les réseaux, et ce en exploitant le concept de communication opportuniste autonome (AOC : *Autonomic and Opportunistic Communication*), en l'absence d'infrastructures de communication de bout en bout. Les travaux de recherches réalisés dans ce contexte vont au delà des approches innovante de Cross Layer en définissant un système qui utilise le *best effort* et les informations de contexte (context-aware) pour l'acheminement des messages entre dispositifs mobiles omniprésents, afin de fournir des services quand la connectivité est locale et l'intermittente. Enfin, ce projet présente un environnement ouvert afin de faciliter la prolifération des applications et des services.

Le projet **AutoI** (*Autonomic Internet*) (2008-2010) [AutoI 2008] a pour ambition d'avoir une solution autonome qui pourra être déployée à court terme. Il développe des overlays virtuels de ressources qui peuvent couvrir des réseaux hétérogènes, supporter la mobilité, la fiabilité et la QoS. Les services d'adaptation se basent sur des informations et modèles de données ontologiques, ceci dans le but de faciliter le déploiement de nouveaux services et ainsi prendre en compte les NGN (Next Generation Networks). Le but attendu est l'unification des efforts de recherche dans le domaine de "l'autonomic" pour produire des standards, définir un framework de référence pour "l'autonomic" pour garantir l'interopérabilité et enfin, créer une structure organisationnelle pour maintenir ces deux premiers objectifs.

Le projet **UniverSelf** [UniverSelf 2009] est une des composantes des Réseaux du futur dans le cadre du 7ème programme-cadre de l'Union Européenne. Il est doté d'un budget de 10 millions d'euros, pour une période de trois ans (2010-2013). Ce projet de recherche a pour but de simplifier la gestion des réseaux de télécommunication, notamment à travers l'auto-gestion. Ce projet permet de consolider les méthodes autonomes de l'Internet du Futur pour le business, le service et l'administration du réseau dans un nouvel environnement appelé *Unified Management Framework* (UMF) qui évolue par cognition. Cet UMF vise à résoudre les problèmes

Figure 2.5 – Les défis clés du projet *UniverSelf*

de conception de l'Internet et de sa mosaïque de protocoles arrivée plus tard, en unifiant parfaitement le contrôle et la gestion et en permettant l'auto-organisation et l'autorisation avec cognition. Cela mettra les tâches de gestion par l'humain au niveau de la gouvernance du réseau entier et des services écosystémiques. Ce projet de recherche européen s'articule autour de trois objectifs : concevoir une plate-forme de gestion unifiée pour les architectures existantes, élaborer des fonctions permettant aux réseaux de s'auto-gérer et permettre le déploiement de solutions autonomes dans les réseaux des opérateurs. Les utilisateurs finaux bénéficieront également directement des résultats du projet *UniverSelf*, puisque les solutions ainsi créées devraient permettre d'améliorer la qualité de service et les performances des réseaux de télécommunications.

2.4.2 Projets industriels et académiques

La plupart des grands industriels ont également lancé des initiatives dans la foulée d'IBM [Horn 2001] : par exemple "Autonomic System" de *Fujitsu Siemens Com-*

puter [Fujitsu-Siemens 2003], “Dynamic Systems” de Microsoft [Microsoft 2007], “Harmonious Computing” de Hitachi [Hitachi 2007] et “N1 Architecture de” de Sun Microsystems. Le monde académique s’est également intéressé de prêt au domaine et plusieurs projets ont vu le jour : Bio-Networking [BIO-NETWORKING 2007] de l’Université de Californie, AUtonomia [Autonomia 2007] de l’université d’Arizona, JSPOON [Konstantinou & Yemini 2003] de l’université de Columbia, et CPN : Cognitive Packets Networks de l’impérial College.

Le projet AUTONETS (*AUTO*nomie*C* *N*etworking for *E*nd To End *S*upervision) est un projet interne de France Télécom, repris maintenant dans le projet UniverSelf. Il considère l’amélioration de la supervision de VoIP (plus généralement des services IP) par l’introduction de comportements autonomes. Il rentre dans le cadre de l’administration de services. Notre travaux pour l’auto-diagnostic se situent dans le cadre de ce projet.

Par conséquent, les capacités d’auto-gestion, et en particulier l’auto-rétablissement, représentent une solution très attrayante pour résoudre les problèmes de gestion du réseau. Comme mentionné ci-dessus, il est de plus en plus crucial d’assister l’homme dans les opérations de surveillance. Cette aide se déploie dans la fonction d’auto-rétablissement, incluant donc les problématiques d’auto-diagnostic et d’auto-réparation.

D’autres travaux de recherche portant sur le domaine des réseaux autonomes ont été réalisés et mettent en avant des fonctions particulières de gestion autonome. Ainsi l’architecture OpenWings [Bieber & Carpenter 2003] fait de l’auto-restauration de services en réaction aux problèmes réseaux et SELFCON [Boutaba *et al.* 2001] préconise l’auto-configuration de composants grâce aux informations de configuration et de politique associées aux données par un serveur d’annuaire.

Dans self-Star [SELF-STAR 2007], l’auto-optimisation se fait par l’évaluation de performances de systèmes de composants et d’équilibrage de charge alors que les acteurs dans [Menasce *et al.* 2001] proposent une solution d’auto-optimisation par le monitoring dynamique et la reconfiguration des paramètres basée sur une métrique de QoS agrégée.

2.4.3 Synthèse

De nombreux travaux de recherche en réseau autonome ont déjà été effectués depuis de nombreuses années. Nous avons présenté plusieurs projets concernant les réseaux autonomes dans les sections précédentes. Nous avons pu constater que la plupart des projets se consacrent à l'architecture autonome et à la fonction d'auto-configuration (ANA, AutoI). Il y a très peu de recherches sur la fonction d'auto-rétablissement, notamment l'auto-diagnostic. C'est pourquoi nous avons choisi l'auto-diagnostic comme sujet d'étude.

2.5 Conclusion

Nous avons, dans ce chapitre 2, présenté un état de l'art se rapportant aux réseaux autonomes. Nous avons d'abord présenté l'historique et les motivations qui ont conduit à ce nouveau paradigme. Cet historique permet de montrer l'importance sinon la nécessité de faire évoluer les architectures réseaux vers plus d'autonomie pour répondre aux besoins de communication de demain et proposer une architecture adaptée aux réseaux de future génération. Ce chapitre décrit aussi, l'architecture de base de la gestion autonome ainsi que l'architecture des éléments autonomes selon la vision d'IBM qui peut être vue comme modèle de référence dans ce domaine. Ce chapitre se termine par la présentation des projets de recherche concernant la thématique des réseaux autonomes. Peu de travaux sur "l'autonomic" traitent spécifiquement de l'auto-rétablissement. Notre travaux portent plus précisément sur l'auto-diagnostic et son application pour la supervision de services multimédia dans un réseau IP de nouvelle génération. C'est pourquoi dans le chapitre suivant, nous introduisons le diagnostic dans les réseaux de télécommunications.

CHAPITRE 3 Diagnostic dans les réseaux de télécommunications

Sommaire

3.1 Réseaux de télécommunications	31
3.1.1 L'Internet	32
3.1.2 La convergence des réseaux	33
3.2 La gestion des pannes	34
3.2.1 Définition d'une panne	35
3.2.2 Approches utilisées pour la gestion des pannes	36
3.3 Vision IA du Diagnostic	37
3.3.1 Introduction	37
3.3.2 Les deux principales approches	37
3.3.3 Intérêt du diagnostic à base de modèles pour les réseaux autonomes	45
3.4 Conclusion	45

3.1 Réseaux de télécommunications

Jusqu'au milieu des années 80, les services de télécommunications traditionnels possédaient chacun leur réseau dédié optimisé pour le transport d'un type d'infor-

32 Chapitre 3. Diagnostic dans les réseaux de télécommunications

mation pour lequel il a été conçu, et l'interconnexion entre ces réseaux était très limitée ou inexistante. Ainsi, le monde des réseaux était constitué de trois sous-réseaux distincts et indépendants, à savoir les réseaux de diffusion, les réseaux de télécommunications et les réseaux de données. En effet, la voix est transportée sur les réseaux téléphoniques commutés (RTC) qui répondent alors parfaitement aux impératifs de QoS, de fiabilité et d'interactivité en se basant sur un plan de contrôle rigide. La télévision est diffusée par satellite ou par voie hertziennne qui apparaît comme le mode de transmission le plus adapté vue leur nature diffusive intrinsèque caractérisée par une communication unidirectionnelle qui permet la scalabilité en supportant un grand nombre d'utilisateurs. Enfin, les réseaux de données, basés sur X.25, représentent un intermédiaire entre la fiabilité du réseau de télécommunications et la mise à l'échelle du réseau de diffusion tout en permettant une certaine interactivité.

Cette vision des réseaux de communication se révèle progressivement étroite puisqu'elle entraîne des situations de «monopole naturel» caractérisées par la mise en œuvre systématique de solutions propriétaires et d'infrastructures qu'il est difficile de faire évoluer ou d'interconnecter. Par conséquent, l'infrastructure dédiée au transfert de données va proposer une nouvelle vision des architectures de communication ; il s'agit de l'architecture Internet qui se fonde sur le modèle TCP/IP.

3.1.1 L'Internet

Au cours des années 90, l'explosion du volume du trafic véhiculé sur Internet et l'accroissement de son hétérogénéité en termes de services (données, voix, vidéo), ont favorisé l'adoption de l'architecture d'Internet.

Basé sur le modèle TCP/IP et la commutation de paquets, le développement de protocoles et de services sur Internet est largement simplifié puisque ce modèle est ouvert, indépendant d'une architecture particulière et propose une hiérarchie protocolaire qui permet à chaque couche de s'abstraire des difficultés soulevées et résolues par la couche inférieure. En effet, le protocole IP, offre une connectivité en mode paquet indépendante du réseau sous-jacent permettant d'interconnecter tout type de réseau. Avec les protocoles de niveau «Transport» (UDP et TCP), les applications disposent alors d'une interface standard pour transmettre sur un réseau

IP. Graduellement, des protocoles de niveau applicatif (FTP, SMTP, HTTP) ont été standardisés ; ce qui a permis le développement des services « classiques » d'Internet (mail, Web, FTP). Finalement, l'augmentation en puissance des terminaux utilisateurs, conjointement avec l'accroissement des débits et portées des réseaux d'accès, a permis d'envisager non seulement la réplication, sur Internet, des services RTC ou télévisuels mais aussi le développement de nouveaux services large bande multimédia.

3.1.2 La convergence des réseaux

Durant ces dernières années, les trois réseaux ont connu des évolutions conséquentes. Le réseau de télécommunications est devenu numérique, sans fil et mobile grâce à l'émergence des réseaux mobiles de 2^{ème} et 3^{ème} générations qui ont augmenté le débit de transmission, et par la même, le nombre de services. Le même constat peut être fait pour le réseau de diffusion qui migre actuellement vers la télévision et la radio numérique. Ainsi, les frontières entre les trois réseaux tendent à disparaître et les services se généralisent sur tous les réseaux. Par exemple, l'Internet et la TV sont disponibles sur les réseaux de télécommunications, les communications téléphoniques peuvent être effectuées sur Internet, etc.

Cette nouvelle mouvance a fait émerger la notion de convergence des réseaux qui a pour objectif de définir un cadre global pour le regroupement des trois types de réseaux sous une seule architecture.

Les réseaux NGN (*Next Generation Network*) représentent la prochaine génération de réseaux censée réaliser la convergence totale des services en une seule architecture.

IP Multimedia Subsystem (IMS) est une architecture standardisée NGN pour les opérateurs de téléphonie, qui permet de fournir des services multimédia fixes et mobiles. C'est sur ce type d'architecture réseau que nous appliquerons nos travaux de recherche. C'est pourquoi nous détaillerons IMS au chapitre 4.

3.2 La gestion des pannes

La gestion des pannes (ou fautes) est une des cinq aires fonctionnelles [ISO10040 1991] de gestion (voir la figure 3.1).

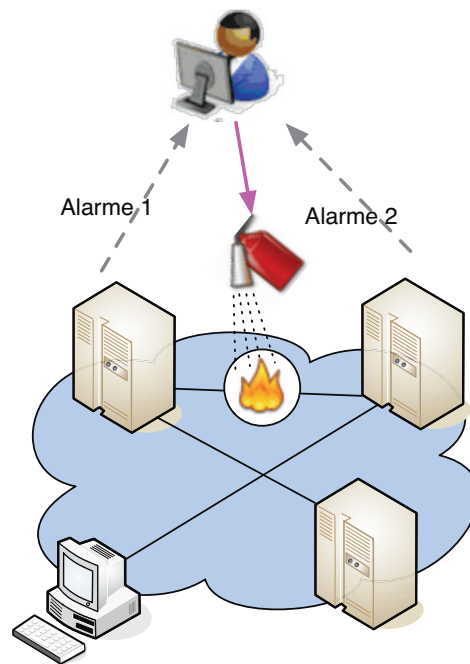


Figure 3.1 – La gestion des pannes

La gestion des pannes couvre l'ensemble des fonctionnalités suivantes :

- La détection des pannes : elle comprend la préparation de rapports d'incidents de fonctionnement, la gestion de compteurs ou des seuils d'alarme, le filtrage d'événements par filtrage en amont des informations, l'affichage des dysfonctionnements.
- La localisation : on y procède au moyen de rapports d'alarme, de mesures et de tests.
- La réparation : elle consiste à prendre les mesures correctives (réaffectation de ressources, «reroutage», limitation du trafic par filtrage, maintenance), ou encore à rétablir du service (tests de fonctionnement, gestion de systèmes de secours, etc.).

- L’enregistrement des historiques d’incidents et statistiques : la gestion des pannes ne peut se limiter à ces actions ponctuelles, nécessaires mais insuffisantes pour donner le service attendu. C’est la raison pour laquelle elle comporte aussi, d’une part, l’enregistrement d’historiques d’incidents et la compilation de statistiques qui peuvent porter sur la probabilité des pannes, leur durée, les délais de réparation et, d’autre part, un rôle d’interface avec les usagers qui consiste à les informer des problèmes réseau et à leur donner la possibilité de signaler eux-mêmes des incidents :
 - la déconnexion d’un câble ;
 - une mauvaise configuration d’un équipement ;
 - une interface défectueuse d’un routeur ;
 - la réinitialisation accidentelle.

3.2.1 Définition d’une panne

Nous trouvons qu’il y a une variété de termes pour définir «les pannes en général» dans les normes et dans la littérature scientifique : Faute, dysfonctionnement, défaillance, anomalie, incident etc. Dans [UIT-T 1992] par exemple, les pannes sont définies comme :

- *Erreur* : une déviation du système par rapport à l’opération normale ;
- *Faute* : une condition qui provoque un dysfonctionnement (et se manifeste par des erreurs).

Afin de simplifier et de bien comprendre les problèmes de la gestion des pannes dans les réseaux de télécommunications, une définition de la panne a été proposée par [Boubour 1997] pour ce type de réseau :

Définition 3.2.1 (Pannes) *Une panne est un état de non fonctionnement ou de dysfonctionnement, matériel ou logiciel pertinent pour l’opérateur, au sens où il souhaite en avoir une trace pour le suivi.*

Nous pouvons remarquer que cette définition est basée sur la perception subjective de l’opérateur, et dépend du niveau de détail auquel il s’intéresse.

Les pannes peuvent être classées en deux types : *permanentes* ou *intermittentes*. Les pannes permanentes exigent une action de réparation. Par exemple, un câble

fractionné entre deux équipements est une panne permanente. Une panne intermittente se manifeste de façon discontinue mais peut se répéter au cours du temps. Par exemple, la réinitialisation d'un équipement réseau est une panne intermittente. Elle peut se produire plusieurs fois au cours du fonctionnement du réseau. Le diagnostic des pannes intermittentes est une tâche plus complexe car les conséquences d'une panne de ce type peuvent disparaître.

3.2.2 Approches utilisées pour la gestion des pannes

Le développement rapide des réseaux à la fois en termes d'extension physique et de multiplicité des services offerts, signifie qu'ils deviennent de plus en plus complexes techniquement. Cela signifie également qu'il est de plus en plus difficile d'identifier ou de réparer les défauts de fonctionnement lorsqu'ils se produisent.

Pour pallier les défaillances d'un réseau de télécommunications, les grands opérateurs ont souvent développé ou fait développer à grands frais des solutions de gestion propriétaires parfois efficaces, mais peu évolutives et peu adaptées à l'explosion actuelle des réseaux de télécommunications et aux difficultés d'acquisition de l'expertise qui en résultent.

L'utilisation des techniques de l'intelligence artificielle pour la gestion n'est pas nouvelle [Gaiti & Pujolle 1993]. Si les systèmes experts ont fait l'objet de nombreuses utilisations en gestion, notamment pour la gestion des fautes, d'autres techniques de l'intelligence artificielle semblent aujourd'hui très prometteuses. Nous en avons retenues essentiellement une pour nos travaux de recherche que nous présentons dans la section 3.3.2.2 : le diagnostic à base de modèles. La section suivante est consacrée à l'introduction du diagnostic en *Intelligence Artificielle* (IA) et à la comparaison de deux grandes approches : le diagnostic à base de systèmes experts et le diagnostic à base de modèles.

3.3 Vision IA du Diagnostic

3.3.1 Introduction

Le diagnostic est la partie de l'intelligence artificielle visant à détecter, localiser et identifier les dysfonctionnements d'un système. Le diagnostic s'appuie sur des observations (symptômes) fournies par le système. Ces observations sont parfois imprécises, incertaines, voire partiellement fausses. L'objectif du diagnostic consiste alors à inférer les pannes possibles ou probables sur le système étant donné ces observations.

3.3.2 Les deux principales approches

3.3.2.1 Diagnostic à base de systèmes experts

Les systèmes experts traditionnels constituent le premier système d'aide au diagnostic. Ils sont à base de raisonnement ou règles de l'expert pour montrer les associations entre effets et causes. Ces associations sont généralement basées sur l'expérience de l'expert plutôt que sur une connaissance de la structure et du comportement du système. Communément, cette connaissance se présente sous la forme suivante : si *symptôme*₁ et *symptôme*₂ et ... *symptôme*_N alors *panne*_x

La fonctionnalité d'un système expert pour le diagnostic est de trouver la cause de ce qui a été observé en parcourant les règles par des techniques classiques en IA telles que le chaînage avant, le chaînage arrière ou encore le chaînage mixte.

Nous présentons suivante un exemple des systèmes experts pour la supervision d'un réseau de télécommunication à commutation de paquets bien connu : le réseau *Transpac*.

Un exemple de système expert : réseau *Transpac*

Le système d'aide à la supervision utilisé sur le réseau *Transpac* était à la base un système expert comportant environ 200 règles. Il effectue une synthèse des éléments provenant du réseau, propose des actions à l'opérateur, attire son attention sur des pannes trop fréquentes ou trop longues et met à jour une base de données des états des éléments du réseau. Cette base de données est mise à jour sans effectuer

38 Chapitre 3. Diagnostic dans les réseaux de télécommunications

de photographies d'état, c'est-à-dire sans aller demander à chaque composant dans quel état il se trouve.

Ce système expert utilise des règles de production mises en œuvre à l'aide d'un générateur de systèmes experts appelé Chronos. Ce générateur est un outil de développement de systèmes experts. Il a été choisi pour résoudre le problème d'efficacité et pour faciliter l'intégration de l'expertise. Une règle type est constituée d'une partie prémisse et d'une partie conclusion. La partie prémisse décrit une suite d'alarmes, précisant pour chacune d'entre elles leur nature, leur provenance, et éventuellement leur date de réception et des délais entre ces dates, ou encore l'indication d'un nombre minimum d'alarmes. La partie conclusion indique en général les éléments indésirables survenus dans le réseau et supposés responsables de l'émission des alarmes. Elle peut aussi mentionner des actions à entreprendre par le superviseur (envois de commandes). Dans l'exemple ci-dessous, nous présentons en langage naturel, une règle typique décrite dans le système expert de *Transpac*. Dans cette règle entrent en jeu un composant du réseau de type *CT* (Centre Technique), et des commutateurs (un centre technique gère un ensemble de commutateurs).

Exemple [Une règle du système expert]

Dès que :

On a reçu une alarme CVHS concernant un objet $\langle x \rangle$ du réseau du type *CT* au temps $T1$

et

On a reçu une alarme CVES concernant le même objet $\langle x \rangle$ au temps $T2$ avec $T2 > T1$

et

Durant la période $[T1, T2+30\text{secondes}]$, on a reçu plus de 3 alarmes de type N004 concernant des commutateurs dépendant de cet objet $\langle x \rangle$

Faire :

Afficher à l'intention de l'opérateur «Il y a eu arrêt du *CT* $\langle x \rangle$ du temps $T1$ au temps $T2$ »

La prémisse de cette règle exprime la réception d'un certain nombre d'alarmes (CVHS, CVES et N004) dans un intervalle de temps précis, provenant d'un centre

technique $\langle x \rangle$ et d'un sous-ensemble de ses commutateurs. La conclusion de cette règle est un diagnostic de panne envoyé à l'opérateur. Si une commande à activer existait dans ce cas de panne, la conclusion de cette règle serait augmentée d'un message informant l'opérateur que cette commande pourrait être activée pour résoudre le problème [Pencolé 2002].

Les principaux avantages des systèmes experts

Dans un système expert, les règles spécifient la partie du raisonnement que doit avoir l'opérateur de supervision. La qualité première d'un système fonctionnant avec de telles règles est son efficacité au niveau temps de calcul. Il est suffisant pour un tel système d'attendre que survienne un enchainement d'événements extérieurs facilement observables puis d'arriver directement aux conclusions [Ungauer 1993].

Il n'y a pas de raisonnement compliqué et coûteux en temps de calcul à réaliser. Ceci est possible car un expert n'a enregistré dans le système que les conditions initiales et les conclusions finales de son raisonnement. Donc, nous pouvons voir ces règles comme des raccourcis efficaces de raisonnements généralement beaucoup plus longs.

Puisque ces règles sont produites par l'expert humain, le résultat est compréhensible pour l'opérateur. Ainsi, une règle d'un système expert est directement interprétable par l'opérateur. Il peut aussi lui servir pour l'explication et la justification, face à la situation à laquelle il se trouve confrontée.

L'implantation d'un système expert est aussi très simple. En effet, le travail est facilité car il n'a pas besoin de développer des algorithmes complexes.

Les principaux inconvénients des systèmes experts

1. La difficulté d'acquisition de l'expertise. Ce point est particulièrement important lors de l'installation d'un nouveau réseau. En effet, puisque le réseau est nouveau, il n'y a pas ou peu d'expériences au sujet des pannes qui peuvent se produire et particulièrement des événements (observables en particulier) qui peuvent être les conséquences.
2. Le manque de généralité. C'est-à-dire que les règles pour un réseau ne peuvent pas être utilisées pour un autre réseau, parce qu'elles sont trop souvent dé-

pendantes de l'architecture du réseau.

3. Le problème de l'évolution du système. Si un système de supervision évolue (ce qui se produit souvent dans les réseaux de télécommunications) soit par remplacement de composants, soit par ajout de composants, les systèmes de règles doivent être modifiés. Une nouvelle expertise doit être mise en place afin que le système expert soit toujours pertinent face aux observations.

3.3.2.2 Diagnostic à base de modèles

Il y a un intérêt considérable pour le raisonnement à base de modèles, en particulier pour les applications du diagnostic et dépannage depuis plusieurs années.

La méthode appelée «Diagnostic à base de modèles» [Hamscher *et al.* 1992] a été proposée pour pallier les faiblesses des premiers systèmes de diagnostic. Le diagnostic à basé de modèles a une variété de noms, dont «le raisonnement à partir des principes premiers», parce qu'il est fondé principalement sur la base de la causalité, et du «raisonnement profond», il repose sur une modélisation explicite du comportement ou du fonctionnement du système et non plus sur la modélisation du raisonnement de l'expert pour effectuer le diagnostic.

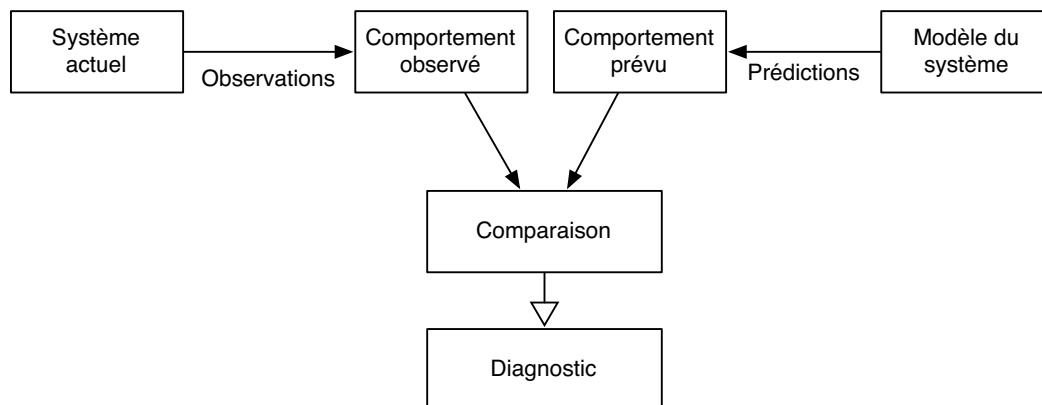


Figure 3.2 – Principe du diagnostic à base de modèles

Le paradigme principal du raisonnement à base de modèles pour le diagnostic est l'interaction entre l'observation et la prévision (figure 3.2). D'une part, pour les systèmes réels, particulièrement pour quelques artefacts physiques, nous pouvons

observer leurs comportements. D'autre part, nous avons un modèle du système qui permet de faire des prédictions sur son comportement prévu. L'observation indique ce que le système est entrain de faire, c'est-à-dire le comportement de l'état actuel du système ; la prédiction indique ce que le système est censé faire. Les événements intéressants, ceux sur lesquels nous devons nous concentrer pour retrouver les composants défectueux, sont ceux qui vont modifier l'état attendu.

Une présomption fondamentale concernant le diagnostic à base de modèles est que si le modèle est correct, toutes les différences entre l'observation et la prédiction découlent (et on peut faire remonter à) de défauts du système. Autrement dit, si le modèle est correct, le système doit être en panne, et les différences sont des pistes sur le caractère et la localisation des défauts.

Mais nous pouvons voir que c'est aussi un concept simplifié : en fait, l'hypothèse selon laquelle le modèle est correct est erronée dans tous les cas. Il est faux d'une manière qui est parfois assez évidente, et parfois assez subtile. Autrement dit, un modèle est un modèle, précisément parce qu'il n'est pas le système lui-même et doit donc être une approximation. Il y aura toujours des choses sur le système que le modèle ne peut pas capturer [Randall.Davis & Hamscher 1988].

Motivation du diagnostic à base de modèles

Le diagnostic à base de modèles peut utiliser des modèles différents pour décrire le système. Les plus habituels décrivent les fonctions, les comportements ou les topologies. Les caractères communs de ces modèles sont qu'ils montrent toujours un ensemble de relations entre les variables du système, chacun d'eux étant relatif aux composants du système. Lorsque le comportement du système est normal, toutes ces relations sont vérifiées par les observations provenant du système physique. Quand une erreur se produit, certaines de ces relations deviennent incompatibles avec le comportement observé du système réel.

Le succès de ces méthodes est fortement relatif à la qualité du modèle. En fait, pour que le modèle fournisse une description précise du système réel, il doit comprendre les effets de la connaissance imprécise et les perturbations inconnues ainsi que la prédiction de l'évolution naturelle du système en temps réel [R.Pons *et al.* 1999].

Pour le diagnostic à base de modèles, il existe deux grandes orientations, qui

42 Chapitre 3. Diagnostic dans les réseaux de télécommunications

utilisent des connaissances différentes sur le système. La première est le diagnostic basé sur la cohérence, qui utilise un modèle de bon fonctionnement du système. Le but de cette approche est de trouver l'état du système consistant avec les observations. La seconde est le diagnostic abductif [Console & Torasso 1991a, Console & Torasso 1991b], qui s'appuie sur le modèle de dysfonctionnement du système. Le but de cette approche est d'expliquer les observations.

Nous présentons ensuite ces deux approches afin de bien les comprendre et de pouvoir comparer leurs avantages et inconvénients. De cette façon, nous pouvons trouver les situations et les environnements correspondant à ces deux approches.

Diagnostic de cohérence

Une autre caractéristique du diagnostic à base de modèles est qu'il n'est nul besoin de savoir quoi que ce soit a priori sur les défauts ou dysfonctionnements pouvant affecter un système pour pouvoir le diagnostiquer : modéliser le comportement correct est suffisant. L'idée fondamentale est de comparer le comportement réel du système tel qu'il peut être observé par l'intermédiaire de capteurs et son comportement attendu tel qu'il peut être prédit grâce aux modèles de bon comportement. Le résultat de cette comparaison permet d'établir un diagnostic de cohérence. Si ces modèles sont corrects, en ce sens qu'ils sont effectivement vérifiés par un système en bon fonctionnement, toute contradiction entre les observations et les prédictions déduites des modèles est nécessairement la manifestation d'un dysfonctionnement, c'est-à-dire de la présence d'un ou plusieurs défauts. Dans le cadre de la théorie logique, DS mentionne un prédicat unaire $AN(x)$ où $x \in COMPS$ et qui est interprété comme signifiant anormal. Si, pour $\Delta \subseteq COMPS$, on note $D(\Delta) = (\wedge AN(c) \mid c \in \Delta) \wedge (\wedge \neg AN(c) \mid c \in COMPS - \Delta)$, le diagnostic est défini par :

Définition 3.3.1 (Diagnostic de cohérence) *Soit $(DS, COMPS, OBS)$ un système observé, son diagnostic est $D(\Delta)$ avec un $\Delta \subseteq COMPS$ tel que :*

$$DS \cup OBS \cup \{D(\Delta)\} \text{ est satisfiable.}$$

$COMPS$ décrit l'ensemble des composants du système à diagnostiquer, DS décrit le comportement des composants ainsi que la structure du système et OBS est un ensemble d'observations.

Ce type de raisonnement par l'absurde fait qu'un défaut est par définition n'importe quoi d'autre que le comportement attendu et il n'est pas recensé parmi une liste finie prédéterminée. Cette approche est une méthode de raisonnement très puissante, qui pallie la plupart des limites des approches traditionnelles et qui peut être logiquement fondée : la détection de dysfonctionnements par réfutation du bon comportement prédit est un raisonnement logiquement correct, ce que n'est pas le cas de la détection de dysfonctionnements par corroboration avec un mauvais comportement prédit [Pencolé 2002].

Le diagnostic à base de modèles, dans ses principes de base, traite principalement de la tâche de localisation des défauts et également, après extension, de celle d'identification de ces défauts.

Diagnostic abductif

Le premier objectif du diagnostic est de détecter/localiser voire identifier un dysfonctionnement à partir des observations du système. Mais parfois, il peut être intéressant que le diagnostic explique les observations. Dans ce cas, il faut passer à un raisonnement de type *abductif* où l'on cherche les causes qui expliquent les symptômes. En l'absence de modèle de comportements défectueux, le pouvoir explicatif est nul. Le diagnostic est uniquement fondé sur la restauration de la cohérence avec les observations. Pour établir un diagnostic abductif, il faut pouvoir disposer d'un modèle de dysfonctionnement du système. Formellement,

Définition 3.3.2 (Diagnostic abductif) *Soit $(DS, COMPS, OBS)$ un système observé et $OBS = E \cup S$ une partition de OBS , S correspondant aux observations que l'on veut expliquer. Un diagnostic abductif pour $(DS, COMPS, E \cup S)$ est un $D(\Delta)$ avec $\Delta \subseteq COMPS$ tel que :*

$$DS \cup E \cup \{D(\Delta)\} \text{ est satisfiable et } DS \cup E \cup \{D(\Delta)\} \models S.$$

Ici, on partage les observations OBS en deux sous-ensembles distincts E et S , et l'on cherche dans ce cadre les modes comportementaux qui, d'une part sont cohérents avec les observations E , et d'autre part *impliquent* S conjointement avec E . En faisant varier à volonté E et S , on a ainsi tout un spectre qui s'étend du diagnostic purement fondé sur la cohérence (cas où $S = \emptyset$) au cas du diagnostic

44 Chapitre 3. Diagnostic dans les réseaux de télécommunications

purement abductif (cas où $E = \emptyset$) [Console & Torasso 1991a].

Le diagnostic à base de modèles a également été utilisé pour modéliser le comportement de différents types de réseaux de télécommunications. Ainsi des outils ont été proposés pour diagnostiquer les fautes dans un réseau de télécommunications ATM [Osmani & Krief 1999] ainsi que SDH en utilisant une approche à base de modèles [Dague *et al.* 2001].

Les principaux avantages du diagnostic à base de modèles

Les techniques de diagnostic qui utilisent la méthode à base de modèles ont de nombreux avantages par rapport aux systèmes experts.

1. Elles n'ont pas besoin d'avoir nécessairement des connaissances sur les pannes. C'est-à-dire qu'un raisonnement peut être exécuté uniquement à partir d'un modèle de fonctionnement du système.
2. Elles sont plus flexibles et génériques. Quand il y a une modification du fonctionnement du système, nous n'avons pas besoin de mettre en œuvre une nouvelle expertise, mais la modification nécessite d'adapter le modèle. Elles peuvent être utilisés pour différents systèmes par apport à un modèle respectif de fonctionnement.
3. Elles possèdent un pouvoir d'explication clair des diagnostics assez important.

Les principaux inconvénients

Par contre, le diagnostic utilisant des raisonnements à base de modèles possède deux inconvénients principaux :

1. Ils nécessitent la construction d'un modèle du système. Cette construction est un processus complexe.
2. Ils sont souvent basés sur des techniques de raisonnement abductifs dont la complexité est exponentielle par rapport au nombre de composants.

Concernant le second point ci-dessus, la distribution du diagnostic permet de maîtriser cette complexité, c'est ce que nous montrerons dans le chapitre 7.

3.3.3 Intérêt du diagnostic à base de modèles pour les réseaux autonomes

Comme indiqué lors de la description des avantages du diagnostic à base de modèles, ce type de diagnostic permet de séparer la connaissance et le raisonnement. Le raisonnement est exécuté uniquement à partir d'un modèle de fonctionnement du système (la connaissance que l'on a du système). Cette séparation est bien adaptée à l'architecture des réseaux autonomes (boucle de contrôle fermée) dont les 4 fonctions sont indépendantes et reliées par la connaissance. Dans la suite de nos travaux, nous choisissons donc le diagnostic à base de modèles et nous nous intéresserons à la distribution de l'algorithme d'auto-diagnostic pour résoudre le problème de passage à l'échelle. La complexité de construction du modèle fera l'objet de travaux futurs.

3.4 Conclusion

Dans ce chapitre, nous avons présenté l'architecture des réseaux de télécommunications, puisque nos travaux de recherche sont situés dans l'environnement des télécommunications. Ce chapitre présente également l'aire fonctionnelle de gestion des pannes car nos travaux se focalisent sur cette aire pour réaliser l'auto-diagnostic. Ensuite, nous avons introduit le diagnostic, présenté son objectif et proposé une définition d'une panne. Nous avons comparé les deux principales approches pour le diagnostic s'appuyant sur l'Intelligence Artificielle (IA) en montrant les avantages et les inconvénients de chacune de ces approches. Nous avons montré également que le diagnostic à base de modèles est mieux adapté pour notre architecture de réseau. Dans le chapitre 4, nous présenterons les services multimédia sur réseaux IP de nouvelle génération qui nous a servi de cadre d'application à notre approche.

4 Les services multi-média sur réseau IP de nouvelle génération

Sommaire

4.1	Introduction à IMS	47
4.1.1	Motivation pour l'utilisation de l'IMS	49
4.1.2	L'architecture IMS	50
4.1.3	Cœur du réseau IMS	52
4.1.4	Protocoles utilisés dans l'IMS	54
4.1.5	Conclusion	56
4.2	Service VoIP	56
4.3	Service IPTV	58
4.4	Conclusion	60

4.1 Introduction à IMS

Alors que la première génération de l'Internet concernait principalement le transport de données non temps réel, les services ayant des exigences de qualité de service (QoS) sont aujourd'hui largement adoptées (Par exemple la voix sur IP

(VoIP), la vidéo à la demande (VoD)). Le mouvement vers une architecture tout-IP pour la délivrance des services est une tendance forte. Dans ce contexte, les clients veulent un accès aux services personnalisés interactifs, aux services multimédia, sur n'importe quel équipement et n'importe où. Cette tendance introduit de nouvelles exigences pour les infrastructures réseau. L'IMS (IP Multimedia Subsystem) est considéré comme une solution pour répondre à ces besoins.

L'IMS se réfère à une architecture fonctionnelle pour la délivrance de services multimédia, basée sur les protocoles de l'Internet. Son objectif est de fusionner l'Internet et le monde cellulaire, afin de permettre des échanges multimédia plus riches [Camarillo 2005, Tadault *et al.* 2003]. Il est spécifié dans le 3GPP (3rd Generation Project). La première version de l'IMS, se focalise sur la facilité de développement et de déploiement pour les nouveaux services dans le réseau mobile [3GPP 2003]. L'IMS a ensuite été développé par *European Telecommunication Standards Institute* (ETSI), dans le cadre de ses travaux sur les *Next Generation Networks* (NGN)¹. Un NGN, simplifié dans la Figure 4.1, est constitué principalement de deux couches : une couche transport qui fournit une connectivité IP aux différents composants d'un NGN tout en garantissant une QoS de bout-en-bout, et une couche de service qui fournit les fonctionnalités de base pour le fonctionnement des services avec ou sans session. Par exemple, des fonctionnalités relatives à l'enregistrement, la notification et la présence.

Un comité technique de normalisation de l'ETSI, appelé *Telecommunications and Internet converged Services and Protocols for Advanced Networking* (TISPAN) standardise IMS comme un sous-système de NGN. On peut dire que 3GPP décrit le point de vue des opérateurs de téléphonie mobile (le soutien de nouvelles applications), mais TISPAN ajoute les spécifications pour les opérateurs filaires (convergence). La plupart des protocoles IMS sont standardisés par l'*Internet Engineering Task Force* (IETF) (par exemple le *Session Initiation Protocol* (SIP)).

Nous devons distinguer entre le core IMS (vocabulaire de TISPAN) et l'IMS (vocabulaire de 3GPP). Nous nous concentrons sur le standard de l'ETSI TISPAN. Nous utilisons le terme d'architecture IMS pour nous référer à l'architecture NGN caractérisée par l'IMS.

1. 3GPP Release 7 (Mars 2007) fournit un IMS unifié qui supporte les technologies d'accès du réseau hétérogène (par exemple DSL, WLAN). [3GPP 2007]

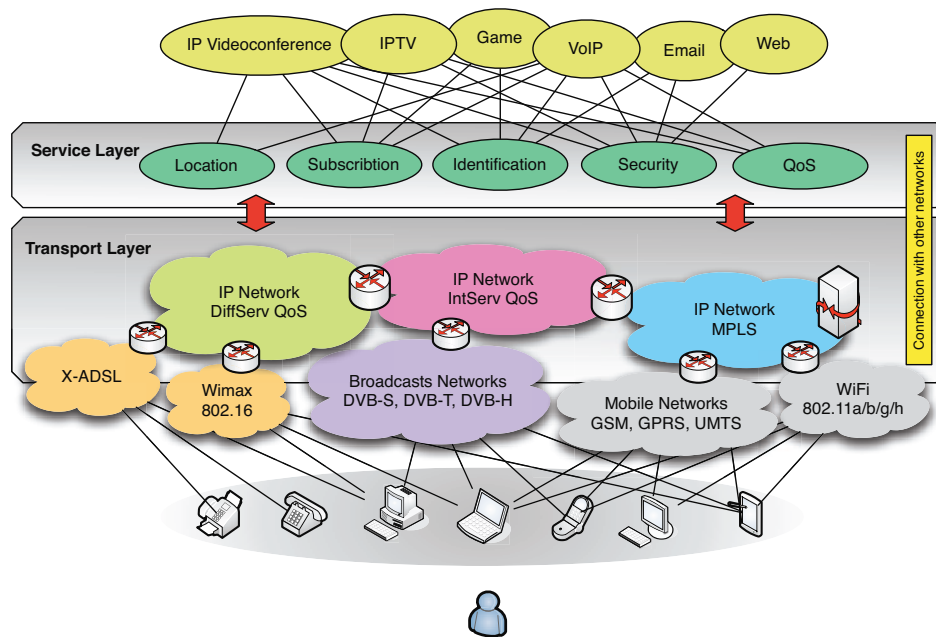


Figure 4.1 – L'architecture NGN

Cette section est divisée en trois sous sections. Après une courte introduction sur l'IMS, nous introduisons les principes, l'architecture IMS et les protocoles utilisés dans l'IMS.

4.1.1 Motivation pour l'utilisation de l'IMS

Anticipant la conversion des infrastructures de réseau de télécommunications vers un transport tout-IP, l'IMS est une architecture cible, dont la généralisation ne pourra être que progressive car complexe et qui ambitionne de remplacer les réseaux téléphoniques traditionnels. L'IMS est une architecture standardisée NGN pour les opérateurs de téléphonie, qui permet de fournir des services multimédia fixes et mobiles. Ce système utilise la technologie VoIP basée sur une implémentation 3GPP.

Une cible de l'IMS est de faire de la gestion de réseau plus facilement. IMS devrait simplifier l'administration du réseau, puisqu'un réseau intégré tout-IP est plus facile à gérer.

IMS est une architecture horizontale : il fournit un ensemble de fonctions communes qui peuvent être utilisées par plusieurs services (par exemple, la gestion de

groupe/liste, le provisioning, l'opération et la gestion ...). Cela rend l'implémentation de service plus facile et plus rapide.

En effet, l'IMS permet à l'opérateur du réseau de jouer un rôle central dans la délivrance des services, et le regroupement de services avec leurs offres d'accès. Par ailleurs, l'IMS devrait supporter la création et le déploiement de services innovants pour les opérateurs ou les tiers, afin de créer de nouvelles sources de revenus. Le développement plus rapide des services IMS devrait réduire le temps de marketing et le temps de stimuler l'innovation. La combinaison de plusieurs services dans une seule session peut susciter l'intérêt des clients et augmenter les opportunités de revenus.

En somme, l'IMS peut permettre le déploiement de nouveaux services plus riches. Il doit permettre la délivrance de communications en temps réel basées sur IP[Tadault *et al.* 2003]. Il doit faire l'intégration pour les applications temps réel et non temps réel plus facilement. Il doit permettre la délivrance de services conversationnels simultanés dans une seule session. Il doit être agnostique pour l'accès, c'est-à-dire qu'il peut permettre à l'utilisateur d'accéder à ses services par tous les média supportés [Bertrand 2007].

4.1.2 L'architecture IMS

L'architecture IMS est une infrastructure de contrôle de service. L'idée a été suggérée afin de répondre au besoin de convergence des services multimédia entre réseaux mobiles et filaires et d'en faciliter les interfaçages. Elle fournit une couche intermédiaire au cœur des réseaux pour passer du mode d'appel classique (circuit) au mode session. Il s'agit d'une architecture en couches, permettant d'agencer toutes les fonctions fondamentales communes à tout type de service de télécommunications. Pour cela, l'architecture IMS, illustrée par la figure 4.2, est découpée en trois plans distincts qui séparent les fonctionnalités liées aux contrôles, aux services et aux réseaux. Ci-dessous, nous donnons une brève description de chaque plan :

La couche d'accès : Elle peut représenter tout accès haut débit tel que : UTRAN (UMTS Terrestrial Radio Access Network), CDMA2000 (technologie d'accès large bande utilisée dans les réseaux mobiles aux Etats-Unis), xDSL, réseau câble, Wireless IP, WiFi, etc.

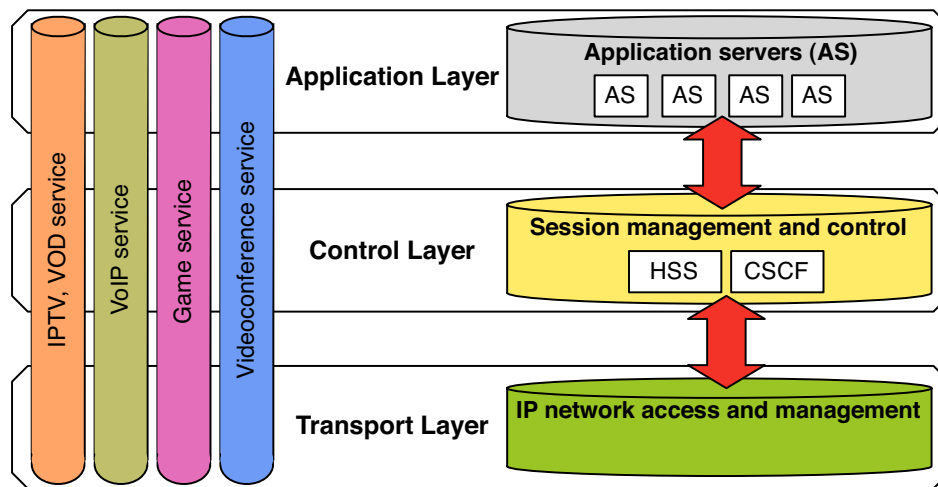


Figure 4.2 – L'architecture IMS

La couche de transport : Elle correspond au réseau IP. Le réseau IP peut intégrer des mécanismes de QoS avec MPLS, Diffserv, RSVP, etc. La couche transport consiste donc en des routeurs (edge router à l'accès et core router en transit) reliés par un réseau de transmission. Différentes piles de transmission peuvent être considérées pour le réseau IP : IP/ATM/SDH, IP/Ethernet, IP/SDH, etc.

La couche de contrôle : Elle authentifie, établit les sessions, interconnecte les opérateurs, permet l'interfonctionnement avec les réseaux existants et l'aiguillage vers les serveurs d'application. C'est au cœur de cette couche qu'est implémenté le protocole SIP.

La couche d'application : Elle comprend les serveurs d'applications (AS - *Application Servers*) et les serveurs de contenu qui exécutent les applications et les services IMS. Le serveur d'application peut inclure des capacités web (HTTP) lui permettant de prendre d'un serveur de contenu des ressources tels que des fichiers multimédia et des scripts d'application.

Puisque l'IMS est encore en cours de définition pour notamment le support de l'IPTV, et puisqu'il décrit plusieurs interfaces et des entités fonctionnelles, un système IMS complet est assez difficile à représenter. Il faut néanmoins noter que l'IMS est une partie d'une architecture fonctionnelle, et que certains de ses composants

peuvent être mis en œuvre dans un seul matériel.

4.1.3 Cœur du réseau IMS

Le cœur IMS est utilisé pour la session et le contrôle de média. Sa structure est illustrée par la figure 4.3 [TISPAN 2006] et ses composants fonctionnels sont les suivants :

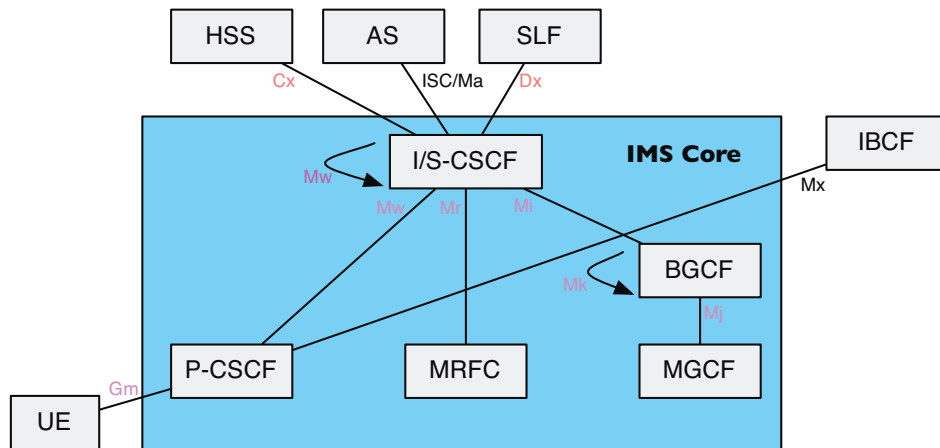


Figure 4.3 – La structure interne du cœur IMS et les interfaces

- *Call Session Control Function* (CSCF) peut établir, surveiller, supporter et communiquer les sessions multimédia, il peut également gérer les interactions de service de l'utilisateur [3GPP et al. 2002]. La fonction de CSCF se décompose en P-CSCF, S-CSCF et I-CSCF. Le «I-CSCF» (*Interrogating*) est le point d'aiguillage intermédiaire pour l'initialisation des connexions, et qui, *via* le DNS, fournit la destination recherchée pour les requêtes orientées vers les multiples S-CSCF des réseaux. Le «S-CSCF» (*Serving*) est utilisé pour la commutation vers l'application, l'enregistrement, le contrôle des session SIP, le service ou le réseau (*serving in charge*) demandés. Le P-CSCF (*Proxy*) sert d'extension logique vers le réseau de l'abonné ou vers le réseau visité et sert au contrôle du réseau d'accès. Il assure les fonctions de liaison aux réseaux de paquets et au PDF (*Policy Decision Function*) (recherche des profils de l'utilisateur) [Camarillo & Garcia-Martin 2004]. Dans la cinquième version de la norme TISPAN, le PDF est séparé de l'I-CSCF afin de permettre l'ouverture

de nouvelles applications liées à la qualité de service hors IMS. Cette interface P-CSCF existe dans tous les réseaux, fixes ou mobiles. En fixe, il sert à la voix sur IP et en réseau mobile, il est utilisé pour toutes les connexions [Beaufils *et al.* 2008]. Deux des CSCF (le I et le S) sont connectés à la base de données du réseau (HSS) afin de recevoir les informations nécessaires aux autorisations de connexion. Les I-CSCF de réseaux voisins sont reliés entre eux afin d'assurer les communications sortantes ou entrantes.

- *Multimedia Resource Function Controller* (MRFC) est utilisé pour contrôler un *Multimedia Resource Function Processor* (MRFP) qui fournit essentiellement la fonctionnalité de transcodage et d'adaptation de contenu [Cuevas *et al.* 2006].
- *Breakout Gateway Control Function* (BGCF) permet de sélectionner l'entité qui constituera la passerelle vers le domaine CS (*Circuit Switched*). C'est un serveur SIP qui possède des fonctionnalités de routage pour une session initiée par un terminal IMS à destination d'un client dans un réseau à commutation de circuit (cas de PSTN).
- *Media Gateway Controller Function* (MGCF) est, comme son nom l'indique, utilisé pour contrôler une passerelle média et leurs canaux dans le plan utilisateur afin d'établir, maintenir et libérer des connexions sous forme de canaux média. Il assure la conversion des messages de signalisation et sélectionne le CSCF approprié afin de remettre la signalisation SIP qu'il génère, au sous-système IMS.
- *Home Subscriber Server* (HSS) est le système de gestion des données des clients et des services auxquels ils ont souscrit. Parmi les données stockées figurent l'identité du client, les paramètres d'accès (authentification, autorisation de *roaming* et S-CSCF alloués) et les informations permettant l'invocation de ses services. HSS utilise le protocole *Diameter* afin d'interagir avec les autres entités du réseau [Beaufils *et al.* 2008].
- *Application Server* (AS) exécute des services (e.g., Push To Talk, Présence, Conférence, Instant messaging, etc.) et peut influencer le déroulement de la session à la demande du service.
- *Subscription Locator Function* (SLF) permet de localiser le HSS lorsque plusieurs HSS ont été déployés au sein du réseau de l'opérateur. Cette résolution

- d'adresse de HSS est requise par I-CSCF, S-CSCF ainsi que d'autres serveurs d'applications.
- *Interconnection Border Control Function* (IBCF) est utilisé comme passerelle vers les réseaux externes, et fournit également des fonctions NAT et Firewall.

4.1.4 Protocoles utilisés dans l'IMS

Comme mentionné précédemment, la plupart des protocoles utilisés dans l'IMS sont standardisés par l'IETF. Ils sont brièvement décrits dans les sections ci-dessous.

4.1.4.1 Signalisation et description de flux médias

Le protocole principal de signalisation utilisé dans l'IMS est appelé *Session Initiation Protocol* (SIP). Il a été principalement défini dans le *RFC 2543* et plus tard dans le *RFC 3261* [Handley et al. 1999, Rosenberg et al. 2002]. SIP profite de certains principes de HTTP (*Hypertext Transfer Protocol*) et SMTP (*Simple Mail Transfer Protocol*). SIP a été sélectionné pour être utilisé essentiellement dans l'IMS, parce qu'il peut fonctionner selon l'exigence d'IMS et il est flexible (plusieurs extensions sont standardisées) et sécurisé. En fait, le SIP d'IMS est une version améliorée du protocole SIP, il comprend plusieurs extensions qui sont décrites dans le standard 3GPP TS.24.299 [3GPP et al. 2005, Radvision 2006]. L'objectif principal du protocole SIP est l'établissement, la modification et la fermeture des sessions multimédia entre deux terminaux.

SIP est le protocole clé dans l'architecture IMS. Il est capable de gérer la gestion des clients, le contrôle de service, l'autorisation de QoS, la facturation et la gestion des ressources [Rosenberg et al. 2002].

4.1.4.2 Protocole AAA

Le protocole *Diameter* est un protocole AAA (*Authentication, Authorization and Accounting*) récent. Il a déjà remplacé le protocole RADIUS (*Remote Authentication Dial In User Service*) [Handley & Jacobson 1998] et il est défini dans le RFC 3588 [Calhoun et al. 2003]. La sécurité de *Diameter* est assurée par IPsec (*Internet Protocol Security*) ou TLS (*Transport Layer Security*). *Diameter* est utilisé dans le cadre d'IMS par I-CSCF, S-CSCF et les serveurs d'application (AS)

dans leurs échanges avec le HSS contenant les profils des utilisateurs.

4.1.4.3 Les protocoles additionnels

Le protocole *MeGaCo*, aussi appelé H248, est un successeur du protocole MGCP (*Media Gateway Control Protocol*) qui est utilisé pour contrôler les média au service des fonctions dans un environnement IMS. Il est spécifié dans le RFC 3015 [Cuervo *et al.* 2000].

Le protocole RTP (*Real Time Protocol*) fournit les fonctions de transport pour la transmission des données en temps réel. Il est spécifié dans le RFC 3550 [Schulzrinne *et al.* 2003]. Et il est utilisé en conjonction avec un protocole de contrôle appelé *Real Time Control Protocol* (RTCP) afin de permettre la supervision de la délivrance des données et fournir un contrôle minimum avec la fonctionnalité d'identification.

Le protocole de transport de signalisation SIGTRAN (*Signalling Transport*) est chargé de définir une infrastructure de signalisation au-dessus de IP. Le but principal est le transport des protocoles de la suite SS7 (*Signalling System 7*) d'un réseau IP [Beaufils *et al.* 2008].

Le protocole SDP (*Session Description Protocol*) est destiné à la description des sessions multimédia pour les annonces de session, l'invitation à une session, et d'autres formes d'initiation de session multimédia. Il permet de contrôler la conformité des types de médias et des codecs utilisés avec ceux autorisés à l'utilisateur.

Le protocole DNS (*Domain Name System*) est un service permettant d'établir une correspondance entre une adresse IP et un nom de domaine et, plus généralement, de trouver une information à partir d'un nom de domaine.

Le protocole COPS (*Common Open Policy Service*) est défini par l'IETF dans le RFC 2748. Il spécifie un simple modèle client/serveur pour soutenir la politique de contrôle sur le protocole de signalisation de QoS (e.g. RSVF (*Resource Reservation Protocol*)). Les politiques sont stockées sur des serveurs, et sollicitées par le PDP (*Policy Decision Points*), pour être appliquées sur les clients, connu sous le nom de PEP (*Policy Enforcement Points*).

4.1.5 Conclusion

Comme les introductions et les discussions des sections précédentes, l'IMS ouvre de nouvelles perspectives pour les opérateurs réseau. Mais plusieurs défis techniques et défis industriels devront être envisagés afin de permettre la sélection à grande échelle de cette technologie prometteuse. Par ailleurs, l'IMS peut résoudre certaines contradictions inhérentes : il s'appuie sur les technologies IP qui permettent la communication libre, mais vise à contrôler les services IP.

De plus, l'IMS conduit les opérateurs réseau à jouer un rôle central dans la distribution des services. Cela implique que les transporteurs devront obtenir des contenus [Tadault *et al.* 2003]. Le rôle des opérateurs, dans la facturation des services fournis par les tiers, doit aussi être clarifié. Avec l'IMS, un seul client peut souscrire aux services de plusieurs fournisseurs. L'IMS conduit donc les opérateurs réseau à une concurrence avec les acteurs de l'Internet mondial.

4.2 Service VoIP

La voix téléphonique sur IP, appelée VoIP (Voice over IP) est devenue une application classique grâce aux progrès de la numérisation et à la puissance des PC, qui permettent d'annuler les échos. L'élément le plus contraignant de cette application sur Internet reste le délai de bout en bout qui ne doit pas excéder 300ms pour permettre l'interactivité.

Les opérateurs de télécommunications doivent également contrôler le réseau de sorte que le temps total de transport de la parole, y compris la paquetsation et la dépaquetsation, soit limité.

De manière générale, la mise en place d'une communication téléphonique sur IP suit différentes étapes :

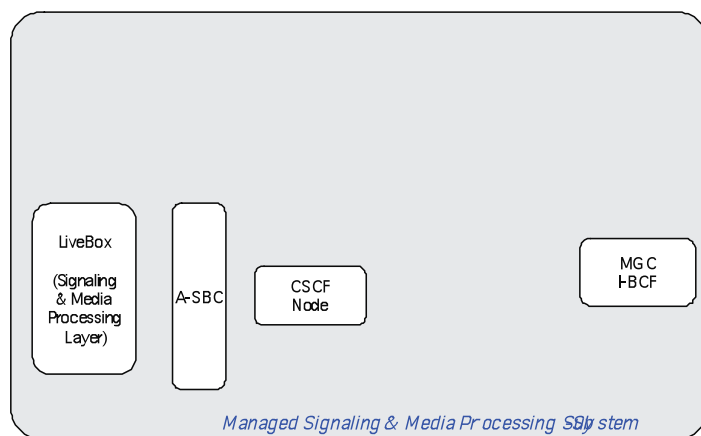
1. Pour mettre en place la communication, il faut d'abord utiliser une signalisation qui démarre la session. Le premier élément à considérer est la localisation du récepteur (*User Location*). Elle s'effectue par une conversion de l'adresse du destinataire en une adresse IP d'une machine qui puisse joindre le destinataire.
2. L'établissement de la communication passe par une acceptation du terminal

destinataire, que ce dernier soit un téléphone, une boîte vocale ou un serveur Web. Plusieurs protocoles de signalisation sont utilisés pour cela, en particulier le protocole SIP (*Session Initiation Protocol*) de l'IETF. Comme son nom l'indique, SIP est utilisé pour initialiser la session. Une requête SIP contient un ensemble d'en-têtes, qui décrivent l'appel, suivis du corps du message, qui contient la description de la demande de session. SIP est un protocole client-serveur, qui utilise la syntaxe et la sémantique de HTTP. Le serveur gère la demande et fournit une réponse au client. Trois types de serveurs gèrent différents éléments : un serveur d'enregistrement (*Registration Server*), un serveur relais (*Proxy Server*) et un serveur de redirection (*Redirect Server*). Ces serveurs travaillent à trouver la route : le serveur proxy détermine le prochain serveur (*Next-Hop Server*), qui à son tour trouve le suivant, et ainsi de suite. Des champs supplémentaires de l'en-tête gèrent des options, comme le transfert d'appel ou la gestion des conférences téléphoniques.

3. Le protocole RTP (*Real-time Transport Protocol*) prend le relais pour transporter l'information téléphonique proprement dite. Le rôle de ce protocole est organiser les paquets à l'entrée du réseau et de les contrôler à la sortie de façon à reformer le flot avec ses caractéristiques de départ (vérification du synchronisme, des pertes, etc.). C'est un protocole de niveau transport qui essaye de corriger les défauts apportés par le réseau.
4. Un autre lieu de transit important de la voix sur IP est constitué par les passerelles permettant de passer d'un réseau à transfert de paquets à un réseau à commutation de circuits, en prenant en charge les problèmes d'adressage de signalisation et de *transcodage* que cela pose. De nouveau, le protocole SIP propose des solutions pour permettre ces correspondances.

Dans le cadre de notre étude, nous nous intéressons principalement aux applications de service VoIP sur l'architecture IMS. Il s'agit plus précisément d'une simplification de l'architecture VoIP SIP utilisée dans le projet T3G SIP de France Télécom [Tuffin 2009].

La Figure 4.4 montre la signalisation du processus de VoIP dans ce cas. La LiveBox est représentée dans cette architecture comme un UE. Les lignes rouges sont les interfaces ou les liens de la signalisation.



ERROR: undefined
OFFENDING COMMAND: limitcheck

STACK:

32
-dictionary-