Acheminement Contraint de Données Critiques

Sajida Zouarhi
Grenoble INP, LIG, équipe HADAS - ORANGE LABS
28, chemin du vieux chêne - BP 98
38243 Meylan Cedex, FRANCE
sajida.zouarhi@orange.com

Directrice de Thèse :
Prof. Dr. Christine Collet - christine.collet@grenoble-inp.fr
Encadrants industriels :
Jean-Marc Temerson - jeanmarc.temerson@orange.com,
Philippe Genestier - philippe.genestier@orange.com.

Résumé

L'acheminement contraint de données critiques est une thématique essentielle dans les systèmes industriels de transmission.
Dans ces derniers, on retrouve de plus en plus de données sensibles en circulation.
Au niveau d'une chaîne de transmission, les questions de délai, de sécurité, de confidentialité, mais également de maîtrise de bout-en-bout de l'intégrité et de la traçabilité de la donnée se posent de manière immédiate. Dans ce papier nous présentons deux approches axées sur l'étude des données.
Nous proposons une vision selon laquelle une meilleure connaissance de l'écosystème via des modèles de dimensionnement orientés donnée couplée à une méthodologie de gestion du Risque permettrait à l'opérateur télécom de garantir une haute qualité de service.

Mots-clés

Donnée, transmission, réseau, sûreté de fonctionnement, qualité de service, événementiel, criticité, bout-en-bout.

1. CONTEXTE ET MOTIVATION

Cet article s'inscrit dans le cadre des travaux de la thèse intitulée "sûreté de fonctionnement des systèmes industriels de transmission de données critiques".
L'objectif de ces travaux est de permettre à l'opérateur télécom Orange d'être en mesure de maîtriser ses engagements en termes de **qualité de service** sur des offres de transmission ou de collecte de **données critiques** (privées, sensibles) au regard : du délai de transmission, de l'intégrité des données transmises, de la confidentialité de bout-en-bout etc.
En d'autres termes, il s'agit de chiffrer la capacité de l'opérateur à respecter une **QoS négociée** au préalable.
Bien que le besoin soit issu du domaine télécom, les travaux portent essentiellement sur la partie applicative. Il s'agit d'émettre des recommandations permettant de maximiser la sûreté de fonctionnement et de mettre au point des modèles qui permettront in fine à l'opérateur de garantir un seuil qualitatif à ses clients.
Pour répondre à ce besoin, différentes approches sont possibles. Dans ce papier seront présentées plus particulièrement l'approche physique de la donnée (partie 2) ainsi que le dimensionnement du risque via une approche événementielle (partie 3). Ces deux approches sont complémentaires.

BDA 14 octobre 2014, Grenoble-Autrans, France.

2. APPROCHE PHYSIQUE DE LA DONNÉE

L'idée de cette approche est de trouver un formalisme efficace qui permette à la fois de dimensionner et de visualiser la chaîne que l'on cherche à étudier et les données critiques qui y transitent. Cela afin d'avoir dans un second temps un support pertinent pour faire de la sûreté de fonctionnement (tolérance aux pannes, fiabilisation etc.)
Dans cette approche, nous considèrerons qu'une **Donnée** possède des *attributs* et que plusieurs données transitent entre différentes **Unités** (réceptrices et/ou émettrices).

- *Donnée* : il s'agit d'une mesure liée à une définition. (ex : 27 est une donnée, sa définition est *âge*).
- *Unité* : une unité est le composant élémentaire d'une chaîne de communication. Elles sont des abstractions permettant de représenter des capteurs, des smartphones, des ordinateurs, des serveurs, des bases de données, des équipements (médicaux ou autres) etc.

Dans nos travaux, nous chercherons à calculer l'attribut *masse* de la donnée. Cette masse est une grandeur proportionnelle à l'impact qualitatif et quantitatif que pourraient avoir : le non-acheminement de la donnée, le retard de la donnée, la corruption de la donnée etc.
Ainsi nous nous intéressons exclusivement aux conséquences **fâcheuses ou graves**.
Via un système de métriques nous serons en mesure de quantifier cette masse ou *Gravité* de la donnée.

Une Unité possède également des attributs :

- une Masse : calculée à partir de la masse des données qui transitent dans l'unité,

- un Poids : cela représente son influence vis-à-vis des autres unités qui constituent la chaîne, le poids dépendra de la masse de l'unité et des vecteurs d'interaction entre cette unité et ses voisines.

Ces listes d'attributs sont amenées à s'enrichir au fur et à mesure que des éléments des domaines physique et mathématique auront été adaptés à notre étude.

L'intérêt de ce formalisme est qu'il nous permet d'observer les *influences* inter-unité.
Par exemple si une Unité subit une défaillance les Unités tierces qui lui sont attenantes seront également affectées d'une manière que l'on pourra quantifier.
Il nous permet également de visualiser des *configurations d'équilibre* qui minimisent le poids de certaines unités sur d'autres et ainsi nous permettront dans une prochaine phase axée sur le dimensionnement du risque et la sûreté de fonctionnement de faire de la recommandation dès la conception de la chaîne pour augmenter la QoS.

3. DIMENSIONNEMENT DU RISQUE

3.1 De la *Gravité* au Risque

D'après le rapport de l'INERIS [2] "le risque prend la forme d'une combinaison de la probabilité que survienne un événement et de la gravité de ses conséquences". Notre approche physique nous permet de dimensionner la gravité. Il reste donc à prendre en compte l'Occurrence. On adopte ainsi le raisonnement suivant pour passer de la Gravité au Risque :

$$\text{Risque} = \text{Occurrence x Gravité}$$

L'Occurrence pourra être établie en se basant sur :
- des études statistiques issues de rapports de défaillance,
- des études probabilistiques,
- des avis d'expert.

3.2 Gestion du risque : Méthodologie

Dimensionner le Risque nous permet de mieux l'appréhender et de mettre au point des **barrières** de protection ou de prévention. [4, p. 60]
Pour cela nous proposons de mettre en place des Arbres de défaillances [3] faisant le lien de bout-en-bout : de l'incident sur la donnée jusqu'aux conséquences finales.
Différentes étapes doivent être observées : identifier les incidents qui peuvent apparaître dans la chaîne, identifier les données, les scénarii de défaillance etc.

3.3 Notion de donnée-événement

Dans notre approche, donnée et événement sont très liés. Un événement est l'apparition, suite à un incident dans la chaîne, d'un état dégradé de la donnée. Cet état provoquera l'entrée dans un mode de défaillance qui sera décrit par un scenario.
Pour définir le comportement de la chaîne suite à un incident, nous utiliserons la logique des "ECA rules" (règle événement-condition-action) [1, p. 3] de type :

on event *if* condition *do* actions.

Un événement peut se traduire par la génération des états suivants : donnée corrompue, donnée non acheminée, donnée retardée etc.
Dès lors, nous préférerons considérer que l'événement n'est en fait qu'une instance de la donnée en question. L'événement représente l'état de la donnée à un instant t.
Prenons l'exemple d'une prise de mesure dans le cadre d'un télésuivi de patient diabétique. Des instances pour ce même objet peuvent apparaître à tout instant t sur la chaîne :

- Objet de la **donnée D** : mesure de glycémie

1. Instance de D apparue à t : non effectuée.
2. Instance' de D apparue à t' : effectuée.

On voit que les instances de D sont des événements. Si ces instances ne remplissent pas la condition (ici effectuée ou non), elles peuvent déclencher une action via un système de règles actives [1, p. 3] prédéfinies :
Si les instances de D reste à l'état "non effectuée" pendant un délai maximum fixé par le médecin, alors l'action sera de faire une nouvelle demande de prise de mesure au patient en envisageant une notification ou un rappel téléphonique jusqu'à ce que l'instance indique "effectuée".
Il s'agit là d'une barrière -au sens gestion du risque- qui permettrait d'éviter les scenarii liés à cette donnée (connus via l'arbre) de se déclencher et d'entraîner l'ensemble des conséquences fâcheuses qui les composent (ex : malaise du patient).

4. CONCLUSION ET PERSPECTIVES

L'approche physique ouvre un champ intéressant dans le domaine de l'étude de la donnée.
De plus le formalisme du domaine physique permet une compréhension plus intuitive de certains concepts.
Les dimensionnements de la *Gravité* et du Risque sont des étapes essentielles dans la compréhension de l'écosystème que l'on cherche à étudier – ici une chaîne de transmission de données.
La notion de Gravité nous permet de porter notre attention sur l'aspect critique de ces données en transit.
À terme, nous nous baserons sur des **modèles orientés données** afin d'émettre des préconisations sur la mise en œuvre sûre d'un acheminement contraint de données critiques.
L'approche retenue est de concevoir dans un premier temps un outil de dimensionnement et de visualisation d'une chaîne de communication. Puis, via cet outil, d'agir sur cette chaîne et d'observer son comportement via des métriques de QoS.
La transversalité du sujet nous permet de nous intéresser au secteur de la Santé, où les données vitales sont souvent critiques, mais aussi aux secteurs bancaire et environnemental, où la transmission maîtrisée de l'information est également un enjeu de taille.

5. REFERENCES

[1] A. Hinze and A. Voisard. Eva: An event algrebra supporting complex event specification. 2014.
[2] M. Merad. Ineris - analyse de l'état de l'art sur les grilles de criticité. Technical report, 2004.
[3] Y. Mortureux. La sûreté de fonctionnement - méthodes pour maîtriser les risques. Technical report, 2002.
[4] NF-EN-31010. Gestion des risques - techniques d'évaluation des risques. Technical report, France, 2010.