Construction auto-stabilisante d'arbre couvrant en dépit d'actions malicieuses

Swan Dubois[1], Toshimitsu Masuzawa[2] et Sébastien Tixeuil[1]

[1]*Université Pierre et Marie Curie & INRIA (France), {swan.dubois,sebastien.tixeuil}@lip6.fr*
[2]*Université d'Osaka (Japon), masuzawa@ist.osaka-u.ac.jp*

Un protocole *auto-stabilisant* est par nature tolérant aux fautes *transitoires* (*i.e.* de durée finie). Ces dernières années ont vu apparaître une nouvelle classe de protocoles qui, en plus d'être auto-stabilisants, tolèrent un nombre limité de fautes *permanentes*. Dans cet article, nous nous intéressons aux protocoles auto-stabilisants tolérant des fautes permanentes très sévères : les fautes *byzantines*. Nous montrons que, pour certains problèmes n'admettant pas de solution *strictement stabilisante* ([NA02]), il existe toutefois des solutions tolérant des fautes byzantines si nous définissons un critère de tolérance moins contraignant.

Keywords: Arbre couvrant, auto-stabilisation, stabilisation forte, tolérance byzantine

1 Introduction

Le développement des systèmes distribués à grande échelle a démontré que la tolérance aux différents types de fautes doit être incluse dans les premières étapes du développement d'un tel système. L'*auto-stabilisation* permet de tolérer des fautes *transitoires* tandis que la tolérance aux fautes traditionnelle permet de masquer l'effet de fautes *permanentes*. Il est alors naturel de s'intéresser à des systèmes qui regrouperaient ces deux formes de tolérance. Cet article s'inscrit dans cette voie de recherche.

Auto-stabilisation Dans cet article, nous considérons un système distribué asynchrone anonyme, *i.e.* un graphe non orienté connexe G où les sommets représentent les processus et les arêtes représentent les liens de communication. Deux processus u et v sont *voisins* si l'arête (u,v) existe dans G. Les variables d'un processus définissent son *état*. L'ensemble des états des processus du système à un instant donné forme la *configuration* du système. Nous souhaitons résoudre une classe particulière de problèmes sur ce système : les problèmes *statiques* (*i.e.* les problèmes où le système doit atteindre un état donné et y rester). Par exemple, la construction d'arbre couvrant est un problème statique. De plus, nous considérons des problèmes pouvant être spécifiés de manière locale (*i.e.* il existe, pour chaque processus v, un prédicat $spec(v)$ qui est vrai lorsque la configuration locale de v est conforme au problème). Les variables apparaissant dans $spec(v)$ sont appelées *variables de sortie* ou *S-variables*.

Un système auto-stabilisant ([Dij74]) est un système atteignant en un temps fini une configuration légitime (*i.e.* $spec(v)$ est vraie pour tout v) indépendamment de la configuration initiale. Une fois cette configuration légitime atteinte, tout processus v vérifie $spec(v)$ pour le restant de l'exécution (et donc, dans le cas d'un problème statique, le système ne modifie plus ses S-variables). Par définition, un tel système peut tolérer un nombre arbitraire de fautes *transitoires*, *i.e.* des fautes de durée finie (la configuration initiale arbitraire modélisant le résultat de ces fautes). Cependant, la stabilisation du système n'est en général garantie que si tous les processus exécutent correctement leur protocole.

Stabilisation stricte Si certains processus exhibent un comportement byzantin (*i.e.* ont un comportement arbitraire, et donc potentiellement malicieux), ils peuvent perturber le système au point que certains processus corrects ne vérifient jamais $spec(v)$. Pour gérer ce type de fautes, [NA02] définit un protocole *strictement stabilisant* comme un protocole auto-stabilisant tolérant des fautes byzantines permanentes. Pour en donner une définition formelle (dans le cas des problèmes statiques), nous devons introduire quelques notations et définitions.

Nous prenons comme modèle de calcul le *modèle à états* : Les variables des processus sont partagées : chaque processus a un accès direct en lecture aux variables de ses voisins. En une *étape* atomique, chaque processus peut lire son état et ceux de ses voisins et modifier son propre état. Un *protocole* est constitué d'un ensemble de règles de la forme $< garde > \longrightarrow < action >$. La *garde* est un prédicat sur l'état du processus et de ses voisins tandis que l'*action* est une séquence d'instructions modifiant l'état du processus. A chaque étape, chaque processus évalue ses gardes. Il est dit *activable* si l'une d'elles est vraie. Il est alors autorisé à exécuter son *action* correspondante. Les *exécutions* du système (séquences d'étapes) sont gérées par un *ordonnanceur* : à chaque étape, il sélectionne au moins un processus activable pour que celui-ci exécute sa règle. Cet ordonnanceur permet de modéliser l'asynchronisme du système. La seule hypothèse que nous faisons sur l'ordonnancement est qu'il est *faiblement équitable, i.e.* qu'aucun processus ne peut rester infiniment longtemps activable sans être choisi par l'ordonnanceur.

Définition 1 *Un processus correct est c-*correct *s'il est situé à au moins c sauts du byzantin le plus proche.*

Définition 2 *Une configuration* ρ *est* (c,f)-contenue *pour spec si, étant donné au plus* f *byzantins, tout processus c-correct v vérifie spec*(v) *et ne modifie pas ses S-variables dans toute exécution issue de* ρ.

Le paramètre *c* de la définition 2 fait référence au *rayon de confinement* défini dans [NA02]. Le paramètre f fait référence au nombre de byzantins alors que [NA02] traite d'un nombre non borné de byzantins.

Définition 3 *Un protocole est* (c,f)-strictement stabilisant *pour spec si, étant donné au plus* f *byzantins, toute exécution (issue d'une configuration arbitraire) contient une configuration* (c,f)*-contenue pour spec.*

Une limite importante du modèle de tolérance de [NA02] est la notion de spécification *r-restrictive*. Intuitivement, il s'agit d'une spécification interdisant des combinaisons d'états de deux processus distants de *r* sauts l'un de l'autre. Un résultat important de [NA02] est le suivant : s'il existe une solution (c,f)-strictement stabilisante à un problème admettant une spécification *r*-restrictive alors $c \geq r$. Pour certains problèmes, dont la construction d'arbre couvrant, on peut montrer que *r* ne peut pas être borné. En conséquence, il n'existe pas de solution (c,f)-strictement stabilisante pour tout rayon de confinement *c* pour le problème de la construction d'un arbre couvrant.

Stabilisation forte Pour contourner de tels résultats d'impossibilité, nous définissons ici un modèle de tolérance plus faible : la *stabilisation forte* ([MT06]). Intuitivement, nous affaiblissons les contraintes relatives au rayon de confinement. En effet, nous autorisons certains processus à l'extérieur de ce rayon à ne pas respecter la spécification en raison des byzantins. Cependant, ces perturbations sont limitées dans le temps : les processus ne peuvent être perturbés par les byzantins qu'un nombre fini de fois et toujours pendant un temps limité même si les byzantins agissent infiniment longtemps. En voici la définition formelle dans le cas des problèmes statiques.

Définition 4 *Une configuration est c-*légitime *pour spec si tout processus c-correct v vérifie spec*(v).

Définition 5 *Une configuration est c-*stable *si tout processus c-correct ne modifie pas ses S-variables tant que les byzantins n'effectuent aucune action.*

Définition 6 *Une portion d'exécution* $e = \rho_0, \rho_1, \ldots, \rho_t$ $(t > 1)$ *est une c-*perturbation *si : (1) e est finie, (2) e contient au moins une action d'un processus c-correct modifiant une S-variable, (3)* ρ_0 *est c-légitime pour spec et c-stable, et (4)* ρ_t *est la première configuration c-légitime pour spec et c-stable après* ρ_0.

Définition 7 *Une configuration* ρ_0 *est* (t,k,c,f)-temporellement contenue *pour spec si, étant donné au plus* f *byzantins : (1)* ρ_0 *est c-légitime pour spec et c-stable, (2) toute exécution issue de* ρ_0 *contient une configuration c-légitime pour spec après laquelle les S-variables de tout processus c-correct ne sont pas modifiées (même si les byzantins exécutent une infinité d'actions), (3) toute exécution issue de* ρ_0 *contient au plus t c-perturbations, et (4) toute exécution issue de* ρ_0 *contient au plus k modifications des S-variables de chaque processus c-correct.*

Remarquons qu'une configuration (t,k,c,f)-temporellement contenue est (c,f)-contenue si $t = k = 0$. Le premier concept est donc une généralisation du second.

Définition 8 *Un protocole* $\mathcal{P}$ *est* (t,c,f)-fortement stabilisant *pour spec si, étant donné au plus* f *byzantins, toute exécution (issue d'une configuration arbitraire) contient une configuration* (t,k,c,f)*-temporellement contenue pour spec atteinte en au plus l unités de temps. l et k désignent respectivement le* temps de stabilisation *et le* temps de perturbation *de* $\mathcal{P}$.

Par définition, un protocole fortement stabilisant est plus faible qu'un protocole strictement stabilisant (car un processus en dehors du rayon de confinement peut subir l'influence des byzantins). Cependant, il est plus puissant qu'un protocole auto-stabilisant (qui peut ne jamais stabiliser en présence de byzantins).

Discussion Il existe une analogie entre la puissance respective de la (c, f)-stabilisation stricte et de la (t, k, c, f)-stabilisation forte d'une part et l'auto-stabilisation et la pseudo-stabilisation d'autre part. Un protocole *pseudo-stabilisant* ([BGM93]) garantit que toute exécution (issue d'une configuration arbitraire) a un suffixe qui vérifie la spécification. Cependant, il est possible qu'une exécution n'atteigne jamais une configuration légitime à partir de laquelle toute exécution vérifie la spécification. En effet, un protocole pseudo-stabilisant peut se comporter suivant la spécification mais en ayant la possibilité de l'invalider dans le futur. Un ordonnancement particulier peut empêcher un tel protocole d'avoir un comportement correct pendant un temps arbitrairement long. Toutefois, un protocole pseudo-stabilisant peut être utile car un comportement correct est ultimement atteint. De manière similaire, toute exécution d'un protocole (t, k, c, f)-fortement stabilisant a un suffixe tel que tout processus c-correct a un comportement correct. Cependant, ce protocole peut ne jamais atteindre une configuration après laquelle les byzantins ne peuvent plus perturber les processus c-corrects. En effet, tous les processus c-corrects peuvent avoir un comportement correct pendant un temps arbitrairement long tout en ayant la possibilité d'effectuer au plus k actions incorrectes (en au plus t perturbations au niveau du système). Une différence importante (mais subtile) est que les perturbations d'un protocole fortement stabilisant sont dûes uniquement aux byzantins alors que les invalidations de spécification d'un protocole pseudo-stabilisant sont dûes à l'ordonnancement.

2 Construction d'arbre couvrant

Problème Dans cette section, nous nous intéressons au problème de la construction d'un arbre couvrant du système. Il s'agit d'un problème fondamental car il permet de mettre en œuvre de nombreux protocoles de communication (par exemple, diffusion, routage par les plus courts chemins, etc.).

Nous considérons que le système est *semi-uniforme* (il existe un processus r distingué comme la racine de l'arbre à construire). Chaque processus v dispose de deux S-variables : P_v et H_v qui désignent respectivement le parent et la hauteur de v dans l'arbre en construction. Le but de la construction est que l'ensemble des pointeurs P_v forme un arbre couvrant du système enraciné en r. Nous considérons un système dans lequel un certain nombre de processus peuvent être byzantins et donc exhiber un comportement arbitraire (ces processus peuvent donc se comporter comme des nœuds internes de l'arbre ou encore comme le processus racine). C'est pourquoi nous devons faire certaines hypothèses sur le système. La première est que le processus r est toujours correct (il n'aura jamais de comportement byzantin). La seconde est que l'ensemble des processus corrects reste toujours connecté. En d'autres termes, les byzantins ne partitionnent jamais le sous-système des processus corrects.

Dans ces conditions, il est impossible, pour un processus correct, de distinguer la racine réelle r d'un byzantin se comportant comme une racine. Nous devons donc autoriser le système à construire une forêt couvrante du système (donc un ensemble d'arbres couvrant le système) dans laquelle chaque racine est soit r soit un byzantin. Voici la spécification formelle du problème que nous considèrons dans cet article :

$$spec(v) : \begin{cases} (P_v = \perp) \wedge (H_v = 0) \text{ si } v \text{ est la racine } r \\ (P_v \in N_v) \wedge (P_v \text{ correct} \Rightarrow H_v = H_{P_v} + 1) \text{ dans le cas contraire} \end{cases}$$

Il est possible de remarquer que, dans le cas où aucun processus n'est byzantin et où tout processus v vérifie $spec(v)$, il existe un arbre couvrant du système au sens "classique". De plus, il faut noter que cette spécification n'impose aucune contrainte sur l'arbre construit (arbre en largeur, degré des nœuds...).

Solution fortement stabilisante Dans la majorité des constructions d'arbre couvrant auto-stabilisantes, chaque processus vérifie localement la cohérence de sa hauteur par rapport à celle de ses voisins. Quand il détecte une incohérence, il modifie sa variable parent pour choisir un "meilleur" voisin. Le critère de qualité d'un voisin dépend en réalité de la propriété globale souhaitée pour l'arbre. Lorsque que le système contient des processus byzantins, ceux-ci peuvent perturber un nombre non borné de processus en prenant successivement des états "meilleurs" et "pires" que leurs voisins. C'est pourquoi le protocole que nous

Algorithme 1 $\mathcal{CAFS}$: Construction d'arbre couvrant fortement stabilisante pour le processus v.

Constantes :

Δ_v le degré de v

N_v l'ensemble des voisins de v (pour $k \in N_v$, le numéro de canal de k est noté $\|k\|$)

S-variables :

$P_v \in N_v \cup \{\perp\}$: parent de v

$H_v \in \mathbb{N}$: hauteur de v

Règles :

$(v = r) \wedge ((P_v \neq \perp) \vee (H_v \neq 0)) \longrightarrow H_v := 0;\ P_v := \perp$

$(v \neq r) \wedge ((P_v \notin N_v) \vee (H_v \neq H_{P_v} + 1)) \longrightarrow P_v := \mathit{suivant}_v(P_v);\ H_v := H_{P_v} + 1$ où $\|\mathit{suivant}_v(k)\| = (\|k\| + 1) \bmod \Delta_v$

proposons ici suit une autre approche. L'idée principale est de contourner ce genre de perturbations par la stratégie suivante : lorsqu'un processus détecte une incohérence locale, il ne choisit pas un "meilleur" voisin mais choisit le voisin suivant son parent actuel selon un ordre cyclique sur ses voisins.

L'algorithme 1 présente notre protocole de construction d'arbre couvrant fortement stabilisant $\mathcal{CAFS}$ qui peut tolérer un nombre arbitraire de byzantins (autres que la racine) tant que le sous ensemble des processus corrects reste connexe. Son rayon de confinement est égal à 0 ce qui est évidement optimal.

Intuitivement, la correction de ce protocole repose sur les éléments suivants. Premièrement, tout processus correct v est activable si et seulement si $\mathit{spec}(v)$ est faux. Cela signifie que toute configuration dans laquelle aucun processus correct n'est activable est 0-légitime et 0-stable. Deuxièmement, il faut remarquer que toute exécution issue d'une configuration dans laquelle au moins un processus correct v ne vérifie pas $\mathit{spec}(v)$ atteint en un temps fini une configuration dans laquelle tout processus correct v vérifie $\mathit{spec}(v)$. Pour cela, il faut constater que la racine r fait au plus une action dans toute exécution. Ensuite, chaque voisin correct de r fait au plus 2Δ actions dans toute exécution où Δ est le degré du système (*i.e.* le degré maximal de ses processus). Enfin, nous pouvons généraliser ce raisonement et dire que tout processus correct v effectue au plus $O(\Delta^{\delta})$ actions dans toute exécution où δ est la distance entre v et r dans le sous-système des processus corrects. Ceci permet de déduire que le système atteindra toujours en un temps fini une configuration dans laquelle tout processus correct v vérifie $\mathit{spec}(v)$ (que ce soit en partant d'une configuration initiale ou après une action byzantine). Cela permet de conclure que le système ne peut connaître qu'un nombre fini de 0-pertubations, chacune ayant une durée finie. Le théorème 1 résume les propriétés de $\mathcal{CAFS}$.

Théorème 1 *Notons d le diamètre du sous-système constitué des processus corrects et f le nombre de byzantins. Le protocole $\mathcal{CAFS}$ est un protocole $(n\Delta^d, 0, n-1)$-fortement stabilisant pour la construction d'arbre couvrant. Le temps de stabilisation de $\mathcal{CAFS}$ est en $O((n-f)\Delta^d)$ étapes (de processus corrects) et le temps de perturbation de $\mathcal{CAFS}$ est Δ^d.*

3 Conclusion

Nous avons étudié la classe des protocoles auto-stabilisants tolérant de plus des fautes byzantines permanentes. Cette classe contient les protocoles strictement stabilisants et fortement stabilisants. Par l'étude du problème de la construction d'arbre couvrant, nous avons illustré le fait que les seconds permettent de résoudre un plus grand nombre de problèmes que les premiers. En contrepartie, les propriétés de tolérance atteintes sont plus faibles. Une voie de recherche future possible est la caractérisation de la classe des problèmes admettant une solution fortement stabilisante mais pas de solution strictement stabilisante.

Références

[BGM93] James E. Burns, Mohamed G. Gouda, and Raymond E. Miller. Stabilization and pseudo-stabilization. *Distributed Computing*, 7(1) :35–42, 1993.

[Dij74] Edsger Dijkstra. Self-stabilizing systems in spite of distributed control. *Com. ACM*, 17(11) :643–644, 1974.

[MT06] Toshimitsu Masuzawa and Sébastien Tixeuil. Bounding the impact of unbounded attacks in stabilization. In *SSS*, pages 440–453, 2006.

[NA02] Mikhail Nesterenko and Anish Arora. Tolerance to unbounded byzantine faults. In *21st Symposium on Reliable Distributed Systems (SRDS 2002)*, page 22. IEEE Computer Society, 2002.