

Ré-identification sans coordination dans les types de données répliquées sans conflits (CRDTs)

THÈSE

présentée et soutenue publiquement le TODO : Définir une date

pour l'obtention du

Doctorat de l'Université de Lorraine
(mention informatique)

par

Matthieu Nicolas

Composition du jury

<i>Président :</i>	À déterminer	
<i>Rapporteurs :</i>	Hanifa Boucheneb	Professeure, Polytechnique Montréal
	Davide Frey	Chargé de recherche, HdR, Inria Rennes Bretagne-Atlantique
<i>Examineurs :</i>	Hala Skaf-Molli	Maîtresse de conférences, HdR, Nantes Université
	Stephan Merz	Directeur de Recherche, Inria Nancy - Grand Est
<i>Encadrants :</i>	Olivier Perrin	Professeur des Universités, Université de Lorraine, LORIA
	Gérald Oster	Maître de conférences, Université de Lorraine, LORIA

Mis en page avec la classe thesul.

Remerciements

WIP

WIP

Sommaire

Chapitre 1

Introduction 1

1.1	Contexte	1
1.2	Questions de recherche et contributions	5
1.2.1	Ré-identification sans coordination synchrone pour les Conflict-free Replicated Data Types (CRDTs) pour le type Séquence	5
1.2.2	Éditeur de texte collaboratif Pair-à-Pair (P2P) temps réel chiffré de bout en bout	6
1.3	Plan du manuscrit	8
1.4	Publications	9

Chapitre 2

Conclusions et perspectives 11

2.1	Résumés des contributions	11
2.1.1	Ré-identification sans coordination pour les CRDTs pour Séquence	11
2.1.2	Éditeur de texte collaboratif P2P chiffré de bout en bout	13
2.2	Perspectives	15
2.2.1	Définition de relations de priorité pour minimiser les traitements . .	15
2.2.2	Détection et fusion manuelle de versions distantes	16
2.2.3	Étude comparative des différents modèles de synchronisation pour CRDTs	19
2.2.4	Approfondissement du patron de conception de Pure Operation-Based CRDTs	21

Annexe A

Entrelacement d'insertions concurrentes dans Treedoc

Annexe B

Algorithmes `RENAMEID`

Annexe C

Algorithmes `REVERTRENAMEID`

Index

31

Bibliographie

Table des figures

1.1	Caption for lfs-properties	3
A.1	Modifications concurrentes d'une séquence Treedoc résultant en un entre- lacement	25

Chapitre 1

Introduction

Sommaire

1.1	Contexte	1
1.2	Questions de recherche et contributions	5
1.2.1	Ré-identification sans coordination synchrone pour les CRDTs pour le type Séquence	5
1.2.2	Éditeur de texte collaboratif P2P temps réel chiffré de bout en bout	6
1.3	Plan du manuscrit	8
1.4	Publications	9

1.1 Contexte

- L'évolution des technologies du web a conduit à l'avènement de ce qui est communément appelé le Web 2.0. La principale caractéristique de ce média est la possibilité aux utilisateur-rices non plus seulement de le consulter, mais aussi d'y contribuer.
- Ces nouvelles fonctionnalités ont permis l'apparition d'applications incitant les utilisateur-rices à créer et partager leur propre contenu, ainsi que d'échanger avec d'autres utilisateur-rices à ce sujet. [1] définit ce type d'applications, c.-à-d. les *réseaux sociaux*, de la manière suivante :

Définition 1. Un réseau social est une application respectant les critères suivants :

- (i) Elle prend la forme d'une application interactive.
- (ii) Elle permet à ses utilisateur-rices de produire et de partager leur propre contenu. Ce contenu généré par les utilisateur-rices constitue le contenu principal de l'application.
- (iii) Elle permet à ses utilisateur-rices de posséder et de maintenir leur propre profil sur la plateforme.
- (iv) Elle encourage ses utilisateur-rices à étendre leur réseau social en se connectant à d'autres utilisateur-rices, voire en créant ou en s'intégrant à des communautés.

De nos jours, les réseaux sociaux représentent les applications les plus populaires du paysage internet, e.g. Facebook compte 2,9 milliards d'utilisateur-rices par mois [2], Youtube 2,5 milliards [2], Wikipedia 788 millions [3] ou encore Quora 300 millions [2].

- La démocratisation de ces applications présente plusieurs bienfaits. Notamment, nous notons que les réseaux sociaux permettent, en diminuant le coût individuel pour tout à chacun-e de contribuer, d'améliorer la diffusion de l'information [4, 5]. Aussi, ils contribuent à la création de communautés [6]. Finalement, en permettant à chacun-e de partager son savoir, ils permettent la création de bases de connaissances complètes [7, 8].
- Cependant, la conception de réseaux sociaux fait face à de nombreux défis. Notamment, ces applications doivent assurer leur haute disponibilité, tolérance aux pannes et capacité de passage à l'échelle en raison de leur popularité et importance dans notre quotidien.

Définition 2 (Disponibilité).

Définition 3 (Tolérance aux pannes).

Définition 4 (Capacité de passage à l'échelle).

Pour cela, ces systèmes adoptent une architecture décentralisée¹, c.-à-d. une architecture reposant sur un ensemble de serveurs qui se répartissent la charge de travail et les tâches. Malgré ce que le nom de cette architecture peut suggérer, il convient de noter que de manière globale les serveurs jouent toujours un rôle central dans les systèmes décentralisés. Par exemple, les serveurs servent à authentifier les utilisateur-rices, à stocker leurs données ou encore à assurer la communication entre utilisateur-rices.

- Additionnellement, il convient de préciser que ces serveurs ne sont pas une ressource libre. En effet, ils sont mis à disposition et maintenus par la ou les organisations qui proposent le réseau social. Ces organisations font alors office d'*autorités centrales* du système, e.g. en se portant garantes de l'identité des utilisateur-rices ou encore de l'authenticité d'un contenu.
- De part le rôle que jouent les serveurs dans les systèmes décentralisés, ces derniers échouent à assurer un second ensemble de propriétés, que nous jugeons néanmoins fondamentales :

Définition 5 (Confidentialité des données).

Définition 6 (Souveraineté des données).

Définition 7 (Pérennité).

Définition 8 (Résistance à la censure).

1. Nous utilisons la classification présentée dans [9] pour distinguer les architectures systèmes centralisées, décentralisées et distribuées.

Ainsi, les utilisateur-rices des réseaux sociaux prennent, de manière consciente ou non, le risque que ces propriétés soient transgressées par les autorités auxquelles appartiennent ces applications ou par des tiers avec lesquelles ces autorités interagissent, e.g. des gouvernements.

- Qui plus est, il convient de noter que les autorités auxquelles appartiennent ces applications, e.g. des entreprises, possèdent leurs propres intérêts, e.g. leur propre profit. De nombreux faits d’actualité ont malheureusement montré que ces autorités ont tendance à privilégier leurs intérêts, quitte à prendre des décisions s’avérant nocives pour une partie de leurs utilisateur-rices. Nous pouvons par exemple noter la mise en avant de contenu à caractère raciste [10], la non-modération de contenu misogyne [11], l’encouragement à la production de contenu toxique [12] ou l’inaction et même entrave à la correction des effets délétères de son réseau social sur la santé mentale de ses utilisatrices [13].
- Ainsi, la présence d’autorités centrales dans les réseaux sociaux représente un danger pour les utilisateur-rices. Il nous paraît alors fondamental de proposer des moyens technologiques pour concevoir et déployer des réseaux sociaux alternatifs qui minimisent le rôle des autorités centrales, voire l’éliminent.
- Dans cette optique, une piste de recherche que nous jugeons intéressante à étudier est celle présentée dans [14]. Dans ce travail, les auteurs proposent un nouveau paradigme de conception d’applications, nommées Local-First Softwares (LFS). Ce type d’applications se démarque par la place centrale donnée aux utilisateur-rices et leurs propres appareils, les éventuels serveurs étant relegués qu’à de simples rôles de support.
- [14] identifie 7 propriétés que doivent satisfaire les applications LFS, illustrées dans la Figure 1.1. Ces propriétés se recoupent en grande partie avec celles que nous

	1. Fast	2. Multi-device	3. Offline	4. Collaboration	5. Longevity	6. Privacy	7. User control
???	✓	✓	✓	✓	✓	✓	✓

FIGURE 1.1 – Liste des propriétés visées par les LFS²

avons identifiées précédemment :

- (i) Le fonctionnement en mode hors-ligne et le fonctionnement avec une latence minimale sont tous deux assurés en privilégiant la disponibilité [15] (Définition 2).
- (ii) Le respect de la vie privée des utilisateur-rices correspond à la propriété de confidentialité (Définition 5).
- (iii) Le contrôle des utilisateur-rices sur leurs données correspond à la propriété de souveraineté (Définition 6).

2. Source : <https://www.inkandswitch.com/local-first/#towards-a-better-future>

- (iv) La longévit   de l'application correspond aux propri  t  s de p  rennit   (D  finition 7) et de r  sistance    la censure (D  finition 8).

Notons qu'il d  coule de ces propri  t  s que les applications LFS sont fondamentalement P2P.

- De mani  re similaire, observons que ce paradigme met en lumi  re une derni  re propri  t   des applications LFS :

D  finition 9 (Collaborative). Une application collaborative est une application permettant    plusieurs utilisateur-rices ou robots de travailler ensemble pour r  aliser une t  che. Dans ce manuscrit, nous consid  rons qu'une collaboration peut prendre la forme d'une discussion, e.g. en participant    un fil de discussion d'une plateforme de questions et r  ponses, ou d'une   dition collaborative d'un contenu, e.g. en participant    la r  daction d'une page de wiki.

- Ainsi, [14]   tablit un paradigme de conception d'applications correspondant    notre vision. Cependant, de nombreuses probl  matiques de recherche identifi  es dans ce travail sont encore non r  solues et entravent la d  mocratisation des applications LFS.
- Notamment, les applications LFS se doivent de r  pliquer les donn  es entre appareils pour permettre :
 - (i) Le fonctionnement en mode hors-ligne et le fonctionnement avec une faible latence.
 - (ii) Le partage de contenu entre appareils d'un-e m  me utilisateur-ric  .
 - (iii) Le partage de contenu entre utilisateur-rices pour la collaboration.
- Cependant, compte tenu des propri  t  s vis  es par les applications LFS, plusieurs contraintes restreignent le choix des m  thodes de r  plication possibles. Ainsi, pour permettre le fonctionnement en mode hors-ligne de l'application, c.-  -d. la consultation et la modification de contenu, les applications LFS doivent relaxer la propri  t   de coh  rence des donn  es.

D  finition 10 (Coh  rence).

Les applications LFS ne peuvent donc pas reposer sur des m  thodes de r  plication dites pessimistes, c.-  -d. qui emp  chent toutes modifications concurrentes d'une m  me donn  e.

- Les applications LFS doivent donc adopter des m  thodes de r  plication dites optimistes [16]. Ces m  thodes autorisent chaque noeud poss  dant une copie de la donn  e de la consulter et de la modifier sans coordination au pr  alable avec les autres noeuds. L'  tat des copies des noeuds peut donc diverger temporairement. Un m  canisme de synchronisation permet ensuite aux noeuds de partager les modifications effectu  es et de les int  grer de fa  on    converger    terme [17], c.-  -d. obtenir    terme de nouveau des   tats   quivalents.

- Cependant, il convient de noter que les méthodes de répllication optimistes autorisent la génération en concurrence de modifications provoquant un conflit, e.g. la modification et la suppression d’une même page dans un wiki. Un mécanisme de résolution de conflits est alors nécessaire pour assurer la convergence à terme des noeuds.
- De nouveau, le modèle du système des applications LFS limitent les choix possibles concernant les mécanismes de résolution de conflits. Notamment, les applications LFS ne disposent d’aucun contrôle sur le nombre de noeuds qui compose le système, c.-à-d. le nombre d’appareils utilisés par l’ensemble de leurs utilisateur-rices. Le nombre de noeuds peut donc croître de manière non-bornée. La complexité algorithmique des mécanismes de résolution de conflits doit donc être indépendante de ce paramètre, ou alors en être fonction uniquement de manière logarithmique.
- De plus, ces noeuds n’offrent aucune garantie sur leur stabilité. Des noeuds peuvent donc rejoindre et participer au système, mais uniquement de manière éphémère. Ce phénomène est connu sous le nom de *churn* [18]. Ainsi, de part l’absence de garantie sur le nombre de noeuds connectés de manière stable, les applications LFS ne peuvent pas utiliser des mécanismes de résolution de conflits reposant sur une coordination synchrone d’une proportion des noeuds du système, c.-à-d. sur des algorithmes de consensus [19, 20].
- Ainsi, pour permettre la conception d’applications LFS, il convient de disposer de mécanismes de résolution de conflits pour l’ensemble des types de données avec une complexité algorithmique efficace par rapport au nombre de noeuds et ne nécessitant pas de coordination synchrone entre une proportion des noeuds du système.

1.2 Questions de recherche et contributions

1.2.1 Ré-identification sans coordination synchrone pour les CRDTs pour le type Séquence

- Les Conflict-free Replicated Data Types (CRDTs) [21, 22] sont des types de données répliqués. Ils sont conçus pour permettre à un ensemble de noeuds d’un système de répliquer une donnée et pour leur permettre de la consulter et de la modifier sans aucune coordination préalable. Dans ce but, les CRDTs incorporent des mécanismes de résolution de conflits automatiques directement au sein leur spécification.
- Cependant, ces mécanismes induisent un surcoût, aussi bien en termes de métadonnées et de calculs que de bande-passante. Ces surcoûts sont néanmoins jugés acceptables par la communauté pour une variété de types de données, e.g. le Registre ou l’Ensemble. Cependant, le surcoût des CRDTs pour le type Séquence constitue toujours une problématique de recherche.
- En effet, la particularité des CRDTs pour le type Séquence est que leur surcoût croît de manière monotone au cours de la durée de vie de la donnée, c.-à-d. au fur et à mesure des modifications effectuées. Le surcoût introduit par les CRDTs pour

ce type de données se révèle donc handicapant dans le contexte de collaborations sur de longues durées ou à large échelle.

- De manière plus précise, le surcoût des CRDTs pour le type Séquence provient de la croissance des métadonnées utilisées par leur mécanisme de résolution de conflits automatique. Ces métadonnées correspondent à des identifiants qui sont associés aux éléments de la Séquence. Ces identifiants permettent de résoudre les conflits, e.g. en précisant quel est l'élément à supprimer ou en spécifiant la position d'un nouvel élément à insérer par rapport aux autres.
- Plusieurs approches ont été proposées pour réduire le coût induit par ces identifiants. Notamment, [23, 24] proposent un mécanisme de ré-assignation des identifiants pour réduire leur coût a posteriori. Ce mécanisme génère toutefois des conflits en cas de modifications concurrentes de la séquence, c.-à-d. l'insertion ou la suppression d'un élément. Les auteurs résolvent ce problème en proposant un mécanisme de transformation des modifications concurrentes par rapport à l'effet du mécanisme de ré-assignation des identifiants.
- Cependant, l'exécution en concurrence du mécanisme de ré-assignation des identifiants par plusieurs noeuds provoque elle-même un conflit. Pour éviter ce dernier type de conflit, les auteurs choisissent de subordonner à un algorithme de consensus l'exécution du mécanisme de ré-assignation des identifiants. Ainsi, le mécanisme de ré-assignation des identifiants ne peut être déclenché en concurrence par plusieurs noeuds du systèmes.
- Comme nous l'avons évoqué précédemment, reposer sur un algorithme de consensus qui requiert une coordination synchrone entre une proportion de noeuds du système est une contrainte incompatible avec les systèmes P2P à large échelle sujets au churn. Notre problématique de recherche est donc la suivante : *pouvons-nous proposer un mécanisme sans coordination synchrone de réduction du surcoût des CRDTs pour Séquence, c.-à-d. adapté aux applications LFS ?*
- Pour répondre à cette problématique, nous proposons RenamableLogootSplit, un nouveau CRDT pour le type Séquence. Ce CRDT intègre un mécanisme de ré-assignation des identifiants, dit de renommage, directement au sein de sa spécification. Nous associons au mécanisme de renommage un mécanisme de résolution de conflits automatique additionnel pour gérer ses exécutions concurrentes. Ainsi, nous proposons un CRDT pour le type Séquence dont le surcoût est périodiquement réduit par le biais d'un mécanisme n'introduisant aucune contrainte de coordination synchrone entre les noeuds du système.

1.2.2 Éditeur de texte collaboratif P2P temps réel chiffré de bout en bout

- Les systèmes collaboratifs permettent à plusieurs utilisateur-ices de collaborer pour la réalisation d'une tâche. Les systèmes collaboratifs actuels adoptent principalement une architecture décentralisée, c.-à-d. un ensemble de serveurs avec lesquels les utilisateur-ices interagissent pour réaliser leur tâche, e.g. Google Docs [25]. Par

rapport à une architecture centralisée, cette architecture leur permet d'améliorer leur disponibilité et tolérance aux pannes, notamment grâce aux méthodes de réplication de données. Cette architecture à base de serveurs facilite aussi la collaboration, les serveurs permettant d'intégrer les modifications effectuées par les utilisateur-rices, de stocker les données, d'assurer la communication entre les utilisateur-rices ou encore de les authentifier.

- De part le rôle qui leur incombe, ces serveurs occupent une place primordiale dans ces systèmes. Il en découle plusieurs problématiques :
 - (i) Ces serveurs manipulent et hébergent les données faisant l'objet de collaborations. Ces systèmes ont donc connaissance des données manipulées et de l'identité des auteur-rices des modifications. Les systèmes collaboratifs décentralisés demandent donc à leurs utilisateur-rices d'abandonner la souveraineté et la confidentialité de leur travail.
 - (ii) Ces serveurs sont gérés par des autorités centrales, e.g. Google. Les systèmes collaboratifs devenant non-fonctionnels en cas d'arrêt de leurs serveurs, les utilisateur-rices de ces systèmes dépendent de ces autorités centrales. Ainsi, de part leur pouvoir de vie et de mort sur les services qu'elles proposent, les autorités centrales représentent une menace pour la pérennité de ces systèmes, e.g. [26].
- Pour répondre à ces problématiques, c.-à-d. confidentialité et souveraineté des données, dépendance envers des tiers, pérennité des systèmes, un nouveau paradigme de conception d'applications propose de concevoir des applications LFS, c.-à-d. des applications mettant les utilisateur-rices et leurs appareils au coeur du système. Dans ce cadre d'applications P2P, les serveurs sont relégués seulement à un rôle de support à la collaboration.
- Dans le cadre de ses travaux, notre équipe de recherche étudie notamment la conception d'applications respectant ce paradigme. Ce changement de modèle, d'une architecture décentralisée appartenant à des autorités centrales à une architecture P2P sans autorités centrales, introduit un ensemble de problématiques de domaines variés, e.g.
 - (i) Comment permettre aux utilisateur-rices de collaborer en l'absence d'autorités centrales pour résoudre les conflits de modifications ?
 - (ii) Comment authentifier les utilisateur-rices en l'absence d'autorités centrales ?
 - (iii) Comment structurer le réseau de manière efficace, c.-à-d. en limitant le nombre de connexions par pair ?
- Cet ensemble de questions peut être résumé en la problématique suivante : *pouvons-nous concevoir une application collaborative P2P à large échelle, sûre et sans autorités centrales ?*
- Pour étudier cette problématique, l'équipe Coast développe l'application Multi User Text Editor (MUTE)³ [27]. Il s'agit d'un éditeur de texte web collaboratif P2P temps réel chiffré de bout en bout. Ce projet nous permet de présenter les travaux

3. Disponible à l'adresse : <https://mutehost.loria.fr>

de recherche de l'équipe portant sur les mécanismes de résolutions de conflits automatiques pour le type Séquence [28, 29, 30] et les mécanismes d'authentification des pairs dans les systèmes sans autorités centrales [31, 32]. Puis, ce projet nous donne l'opportunité d'étudier la littérature des nombreux domaines de recherche nécessaires à la conception d'un tel système, c.-à-d. le domaine des protocoles d'appartenance aux groupes [33, 34], des topologies réseaux P2P [35] ou encore des protocoles d'établissement de clés de chiffrement de groupe [36]. Ce projet nous permet ainsi de valoriser nos travaux et d'identifier de nouvelles perspectives de recherche. Finalement, il résulte de ce projet le Proof of Concept (PoC) le plus complet d'applications LFS, à notre connaissance. *Matthieu: TODO : Vérifier du côté des applis de IPFS*

1.3 Plan du manuscrit

- Ce manuscrit de thèse est organisé de la manière suivante :
- Dans le ??, nous introduisons le modèle du système que nous considérons, c.-à-d. les systèmes P2P à large échelle sujets au churn et sans autorités centrales. Puis nous présentons dans ce chapitre l'état de l'art des mécanismes de résolution de conflits automatiques utilisés dans les systèmes adoptant le paradigme de la réplique optimiste. À partir de cet état de l'art, nous identifions et motivons notre problématique de recherche, c.-à-d. l'absence de mécanisme adapté aux systèmes P2P à large échelle sujets au churn permettant de réduire le surcoût induit par les mécanismes de résolution de conflits automatiques pour le type Séquence.
- Dans le ??, nous présentons notre approche pour présenter un tel mécanisme, c.-à-d. un mécanisme de résolution de conflits automatiques pour le type Séquence auquel nous associons un mécanisme de Garbage Collection (GC) de son surcoût ne nécessitant pas de coordination synchrone entre les noeuds du système. Nous détaillons le fonctionnement de notre approche, sa validation par le biais d'une évaluation empirique puis comparons notre approche par rapport aux approches existantes. Finalement, nous concluons la présentation de notre approche en identifiant et en détaillant plusieurs de ses limites.
- Dans le ??, nous présentons MUTE, l'éditeur de texte collaboratif temps réel P2P chiffré de bout en bout que notre équipe de recherche développe dans le cadre de ses travaux de recherche. Nous présentons les différentes couches logicielles formant un pair et les services tiers avec lesquels les pairs interagissent, et détaillons nos travaux dans le cadre de ce projet, c.-à-d. l'intégration de notre mécanisme de résolution de conflits automatiques pour le type Séquence et le développement de la couche de livraison des messages associée. Pour chaque couche logicielle, nous identifions ses limites et présentons de potentielles pistes d'améliorations.
- Finalement, nous récapitulons dans le chapitre 2 les contributions réalisées dans le cadre de cette thèse. Puis nous clôturons ce manuscrit en introduisant plusieurs des pistes de recherches que nous souhaiterons explorer dans le cadre de nos travaux futurs.

1.4 Publications

Notre travail sur la problématique identifiée dans la sous-section 1.2.1, c.-à-d. la proposition d'un mécanisme ne nécessitant aucune coordination synchrone pour réduire le surcoût des CRDTs pour le type Séquence, a donné lieu à des publications à différents stades de son avancement :

- (i) Dans [37], nous motivons le problème identifié et présentons l'idée de notre approche pour y répondre.
- (ii) Dans [38], nous détaillons une première partie de notre approche et présentons notre protocole d'évaluation expérimentale ainsi que ses premiers résultats.
- (iii) Dans [30], nous détaillons notre proposition dans son entièreté. Nous accompagnons cette proposition d'une évaluation expérimentale poussée. Finalement, nous complétons notre travail d'une discussion identifiant plusieurs de ses limites et présentant des pistes de travail possibles pour y répondre.

Nous précisons ci-dessous les informations relatives à chacun de ces articles.

Efficient renaming in CRDTs [37]

Auteur Matthieu Nicolas

Article de position à Middleware 2018 - 19th ACM/IFIP International Middleware Conference (Doctoral Symposium), Dec 2018, Rennes, France.

Abstract *Sequence Conflict-free Replicated Data Types (CRDTs)* allow to replicate and edit, without any kind of coordination, sequences in distributed systems. To ensure convergence, existing works from the literature add metadata to each element but they do not bound its footprint, which impedes their adoption. Several approaches were proposed to address this issue but they do not fit a fully distributed setting. In this paper, we present our ongoing work on the design and validation of a fully distributed renaming mechanism, setting a bound to the metadata's footprint. Addressing this issue opens new perspectives of adoption of these CRDTs in distributed applications.

Efficient Renaming in Sequence CRDTs [38]

Auteurs Matthieu Nicolas, Gérald Oster, Olivier Perrin

Article de workshop à PaPoC 2020 - 7th Workshop on Principles and Practice of Consistency for Distributed Data, Apr 2020, Heraklion / Virtual, Greece.

Abstract To achieve high availability, large-scale distributed systems have to replicate data and to minimise coordination between nodes. Literature and industry increasingly adopt Conflict-free Replicated Data Types (CRDTs) to design such systems. CRDTs are data types which behave as traditional ones, e.g. the Set or the Sequence. However, unlike

traditional data types, they are designed to natively support concurrent modifications. To this end, they embed in their specification a conflict-resolution mechanism.

To resolve conflicts in a deterministic manner, CRDTs usually attach identifiers to elements stored in the data structure. Identifiers have to comply with several constraints, such as uniqueness or belonging to a dense order. These constraints may hinder the identifiers' size from being bounded. As the system progresses, identifiers tend to grow. This inflation deepens the overhead of the CRDT over time, leading to performance issues.

To address this issue, we propose a new CRDT for Sequence which embeds a renaming mechanism. It enables nodes to reassign shorter identifiers to elements in an uncoordinated manner. Experimental results demonstrate that this mechanism decreases the overhead of the replicated data structure and eventually limits it.

Efficient Renaming in Sequence CRDTs [30]

Auteurs Matthieu Nicolas, Gérald Oster, Olivier Perrin

Article de journal dans IEEE Transactions on Parallel and Distributed Systems, Institute of Electrical and Electronics Engineers, 2022, 33 (12), pp.3870-3885.

Abstract To achieve high availability, large-scale distributed systems have to replicate data and to minimise coordination between nodes. For these purposes, literature and industry increasingly adopt Conflict-free Replicated Data Types (CRDTs) to design such systems. CRDTs are new specifications of existing data types, e.g. Set or Sequence. While CRDTs have the same behaviour as previous specifications in sequential executions, they actually shine in distributed settings as they natively support concurrent updates. To this end, CRDTs embed in their specification conflict resolution mechanisms. These mechanisms usually rely on identifiers attached to elements of the data structure to resolve conflicts in a deterministic and coordination-free manner. Identifiers have to comply with several constraints, such as being unique or belonging to a dense total order. These constraints may hinder the identifier size from being bounded. Identifiers hence tend to grow as the system progresses, which increases the overhead of CRDTs over time and leads to performance issues. To address this issue, we propose a novel Sequence CRDT which embeds a renaming mechanism. It enables nodes to reassign shorter identifiers to elements in an uncoordinated manner. Experimental results demonstrate that this mechanism decreases the overhead of the replicated data structure and eventually minimises it.

Chapitre 2

Conclusions et perspectives

Sommaire

2.1	Résumés des contributions	11
2.1.1	Ré-identification sans coordination pour les CRDTs pour Séquence	11
2.1.2	Éditeur de texte collaboratif P2P chiffré de bout en bout	13
2.2	Perspectives	15
2.2.1	Définition de relations de priorité pour minimiser les traitements	15
2.2.2	Détection et fusion manuelle de versions distantes	16
2.2.3	Étude comparative des différents modèles de synchronisation pour CRDTs	19
2.2.4	Approfondissement du patron de conception de Pure Operation- Based CRDTs	21

Dans ce chapitre, nous revenons sur les contributions présentées dans cette thèse. Nous rappelons le contexte dans lequel elles s’inscrivent, récapitulons leurs spécificités et apports, et finalement présentons leurs limites que nous identifions. Puis, nous concluons ce manuscrit en présentant plusieurs pistes de recherche qui nous restent à explorer à l’issue de cette thèse. Les premières s’inscrivent dans la continuité directe de nos travaux sur un mécanisme de ré-identification pour CRDTs pour Séquence dans un système P2P à large échelle sujet au churn. Les dernières traduisent quant à elles notre volonté de recentrer nos travaux sur le domaine plus général des CRDTs.

2.1 Résumés des contributions

2.1.1 Ré-identification sans coordination pour les CRDTs pour Séquence

Pour privilégier leur disponibilité, latence et tolérance aux pannes, les systèmes distribués peuvent adopter le paradigme de la réplication optimiste [16]. Ce paradigme consiste à relaxer la cohérence de données entre les noeuds du système pour leur permettre de consulter et modifier leur copie locale sans se coordonner. Leur copies peuvent alors temporairement diverger avant de converger de nouveau une fois les modifications de chacun

propagées. Cependant, cette approche nécessite l'emploi d'un mécanisme de résolution pour assurer la convergence même en cas de modifications concurrentes. Pour cela, l'approche des CRDTs [21, 22] propose d'utiliser des types de données dont les modifications sont nativement commutatives.

Depuis la spécification des CRDTs, la littérature a proposé plusieurs de ces mécanismes résolution de conflits automatiques pour le type de données Séquence [39, 40, 41, 42]. Cependant, ces approches souffrent toutes d'un surcoût croissant de manière monotone. Ce problème a été identifié par la communauté, et celle-ci a proposé pour y répondre des mécanismes permettant soit de réduire la croissance du surcoût [43, 44], soit d'effectuer une GC du surcoût [40, 23, 24]. Nous avons cependant déterminé que ces mécanismes ne sont pas adaptés aux systèmes P2P à large échelle souffrant de churn et utilisant des CRDTs pour Séquence à granularité variable.

Dans le cadre de cette thèse, nous avons donc souhaité proposer un nouveau mécanisme adapté à ce type de systèmes. Pour cela, nous avons suivi l'approche proposée par [23, 24] : l'utilisation d'un mécanisme pour ré-assigner de nouveaux identifiants aux éléments stockés dans la séquence. Nous avons donc proposé un nouveau mécanisme appartenant à cette approche pour le CRDT LogootSplit [28].

Notre proposition prend la forme d'un nouvel CRDT pour Séquence à granularité variable : RenamableLogootSplit. Ce nouveau CRDT associe à LogootSplit un nouveau type de modification, *ren*, permettant de produire une nouvelle séquence équivalente à son état précédent. Cette nouvelle modification tire profit de la granularité variable de la séquence pour produire un état de taille minimale : elle assigne à tous les éléments des identifiants de position issus d'un même intervalle. Ceci nous permet de minimiser les métadonnées que la séquence doit stocker de manière effective. De plus, le passage à une représentation interne minimale de la séquence nous permet d'améliorer le coût des modifications suivantes en termes de calculs.

Afin de gérer les opérations concurrentes aux opérations *ren*, nous définissons pour ces dernières un algorithme de transformation. Pour cela, nous définissons un mécanisme d'époques nous permettant d'identifier la concurrence entre opérations. De plus, nous introduisons une relation d'ordre strict total, *priority*, pour résoudre de manière déterministe le conflit provoqué par deux opérations *ren*, c.-à-d. pour déterminer quelle opération *ren* privilégier. Finalement, nous définissons deux algorithmes, **renameId** et **revertRenameId**, qui permettent de transformer les opérations concurrentes à une opération *ren* pour prendre en compte l'effet de cette dernière. Ainsi, notre algorithme permet de détecter et de transformer les opérations concurrentes aux opérations *ren*, sans nécessiter une coordination synchrone entre les noeuds. Le surcoût induit par ce mécanisme, notamment en termes de calculs, est toutefois contrebalancé par l'amélioration précisée précédemment, c.-à-d. la réduction de la taille de la séquence.

Finalement, le mécanisme que nous proposons est partiellement générique : il peut être adapté à d'autres CRDTs pour Séquence à granularité variable, e.g. un CRDT pour Séquence appartenant à l'approche à pierres tombales. Dans le cadre d'une telle démarche, nous pourrions réutiliser le système d'époques, la relation *priority* et l'algorithme de contrôle qui identifie les transformations à effectuer. Pour compléter une telle adaptation, nous devrions cependant concevoir de nouveaux algorithmes **renameId** et **revertRenameId** spécifiques et adaptés au CRDT choisi.

Le mécanisme de renommage que nous présentons souffre néanmoins de plusieurs limites. La première d'entre elles concerne ses performances. En effet, notre évaluation expérimentale a mis en lumière le coût important en l'état de la modification *ren* par rapport aux autres modifications en termes de calculs (cf. ??, page ??). De plus, chaque opération *ren* comporte une représentation de l'ancien état qui doit être maintenue par les noeuds jusqu'à leur stabilité causale. Le surcoût en métadonnées introduit par un ensemble d'opérations *ren* concurrentes peut donc s'avérer important, voire pénalisant (cf. ??, page ??). Pour répondre à ces problèmes, nous identifions trois axes d'amélioration :

- (i) La définition de stratégies de déclenchement du renommage efficaces. Le but de ces stratégies serait de déclencher le mécanisme de renommage de manière fréquente, de façon à garder son temps d'exécution acceptable, mais tout visant à minimiser la probabilité que les noeuds produisent des opérations *ren* concurrentes, de façon à minimiser le surcoût en métadonnées.
- (ii) La définition de relations *priority* efficaces. Nous développons ce point dans la sous-section 2.2.1.
- (iii) La proposition d'algorithmes de renommage efficaces. Cette amélioration peut prendre la forme de nouveaux algorithmes pour `renameId` et `revertRenameId` offrant une meilleure complexité en temps. Il peut aussi s'agir de la conception d'une nouvelle approche pour renommer l'état et gérer les modifications concurrentes, e.g. un mécanisme de renommage basé sur le journal des opérations (cf. ??, page ??).

Une seconde limite de `RenamableLogootSplit` que nous identifions concerne son mécanisme de GC des métadonnées introduites par le mécanisme de renommage. En effet, pour fonctionner, ce dernier repose sur la stabilité causale des opérations *ren*. Pour rappel, la stabilité causale représente le contexte causal commun à l'ensemble des noeuds du système. Pour le déterminer, chaque noeud doit récupérer le contexte causal de l'ensemble des noeuds du système. Ainsi, l'utilisation de la stabilité causale comme pré-requis pour la GC de métadonnées constitue une contrainte forte, voire prohibitive, dans les systèmes P2P à large échelle sujet au churn. En effet, un noeud du système déconnecté de manière définitive suffit pour empêcher la stabilité causale de progresser, son contexte causal étant alors indéterminé du point de vue des autres noeuds. Il s'agit toutefois d'une limite récurrente des mécanismes de GC distribués et asynchrones [40, 45, 46]. Nous présentons une piste de travail possible pour pallier ce problème dans la sous-section 2.2.2.

2.1.2 Éditeur de texte collaboratif P2P chiffré de bout en bout

Les applications collaboratives permettent à des utilisateur-rices de réaliser collaborativement une tâche. Elles permettent à plusieurs utilisateur-rices de consulter la version actuelle du document, de la modifier et de partager leurs modifications avec les autres. Ceci permet de mettre en place une réflexion de groupe, ce qui améliore la qualité du résultat produit [47, 7].

Cependant, les applications collaboratives sont historiquement des applications centralisées, e.g. Google Docs [25]. Ce type d'architecture induit des défauts d'un point de vue technique, e.g. faible capacité de passage à l'échelle et faible tolérance aux pannes, mais

aussi d'un point de vue utilisateur, e.g. perte de la souveraineté des données et absence de garantie de pérennité.

Les travaux de l'équipe Coast s'inscrivent dans une mouvance souhaitant résoudre ces problèmes et qui a conduit à la définition d'un nouveau paradigme d'applications : les LFS [14]. Le but de ce paradigme est la conception d'applications collaboratives, P2P, pérennes et rendant la souveraineté de leurs données aux utilisateur-rices.

Dans le cadre de cette démarche, l'équipe Coast développe depuis plusieurs années l'application Multi User Text Editor (MUTE), un éditeur de texte web collaboratif P2P temps réel chiffré de bout en bout. Cette application sert à la fois de plateforme de démonstration et de recherche pour les travaux de l'équipe, mais aussi de PoC pour les LFS.

Dans le cadre de cette thèse, nous avons implémenté dans MUTE nos travaux de recherche portant sur le nouvel CRDT pour le type Séquence : RenamableLogootSplit. MUTE a aussi servi à l'équipe pour présenter ses travaux concernant l'authentification des utilisateur-rices dans un système P2P [32]. Finalement, MUTE nous a permis de nous d'étudier et/ou de présenter les travaux de recherche existants concernant :

- (i) Les protocoles distribués d'appartenance au groupe [33].
- (ii) Les mécanismes d'anti-entropie [48].
- (iii) Les protocoles d'établissement de clés de chiffrement de groupe [36].
- (iv) Les protocoles d'établissement de topologies réseaux efficaces [35].
- (v) Les mécanismes de conscience de groupe.

MUTE offre donc, à notre connaissance, le tour d'horizon le plus complet des travaux de recherche permettant la conception d'applications LFS. Cependant, cela ne dispense pas MUTE de souffrir de plusieurs limites.

Tout d'abord, l'environnement web implique un certain nombre de contraintes, notamment au niveau des technologies et protocoles disponibles. Notamment, le protocole WebRTC repose sur l'utilisation de serveurs de signalisation, c.-à-d. de points de rendez-vous des pairs, et de serveurs de relai, c.-à-d. d'intermédiaires pour communiquer entre pairs lorsque les configurations de leur réseaux respectifs interdisent l'établissement d'une connexion directe. Ainsi, les applications P2P web doivent soit déployer et maintenir leur propre infrastructure de serveurs, soit reposer sur une infrastructure existante, e.g. celle proposée par OpenRelay [49]. Afin de minimiser l'effort requis aux applications P2P et la confiance exigée à leurs utilisateur-rices, nous devons supporter la mise en place d'une telle infrastructure transparente et pérenne.

Une autre limite de ce système que nous identifions concerne l'utilisabilité des systèmes P2P de manière générale. L'expérience vécue suivante constitue à notre avis un exemple éloquent des limites actuelles de l'application MUTE dans ce domaine. Après avoir rédigé une version initiale d'un document, nous avons envoyé le lien du document à notre collaborateur pour relecture et validation. Lorsque notre collaborateur a souhaité accéder au document, celui-ci s'est retrouvé devant une page blanche : comme nous nous étions déconnecté du système entretemps, c.-à-d. plus aucun pair n'était disponible pour effectuer une synchronisation. Notre collaborateur était donc dans l'incapacité de récupérer l'état et d'effectuer sa tâche. Afin de pallier ce problème, une solution possible est

de faire reposer MUTE sur un réseau P2P global, e.g. le réseau de InterPlanetary File System (IPFS) [50], et d'utiliser les pairs de ce dernier, potentiellement des pairs étrangers à l'application, comme pairs de stockage pour permettre une synchronisation future. Cette solution limite ainsi le risque qu'un pair ne puisse récupérer l'état du document faute de pairs disponibles. Cependant, elle nécessite de mettre en place un mécanisme de réplication de données additionnel. Ce mécanisme de réplication sur des pairs additionnels doit cependant garantir qu'il n'introduit pas de vulnérabilités, e.g. la possibilité pour les pairs de stockage sélectionnés de reconstruire et consulter le document.

2.2 Perspectives

2.2.1 Définition de relations de priorité pour minimiser les traitements

Dans la ??, nous avons spécifié la relation *priority* (cf. ??, page ??). Pour rappel, cette relation doit établir un ordre strict total sur les époques de notre mécanisme de renommage.

Cette relation nous permet ainsi de résoudre le conflit provoqué par la génération de modifications *ren* concurrentes en les ordonnant. Grâce à cette relation d'ordre, les noeuds peuvent déterminer vers quelle époque de l'ensemble des époques connues progresser. Cette relation permet ainsi aux noeuds de converger à une époque commune à terme.

La convergence à terme à une époque commune présente plusieurs avantages :

- (i) Réduire la distance entre les époques courantes des noeuds, et ainsi minimiser le surcoût en calculs par opération du mécanisme de renommage. En effet, il n'est pas nécessaire de transformer une opérations livrée avant de l'intégrer si celle-ci provient de la même époque que le noeud courant.
- (ii) Définir un nouveau Plus Petit Ancêtre Commun (PPAC) entre les époques courantes des noeuds. Cela permet aux noeuds d'appliquer le mécanisme de GC pour supprimer les époques devenues obsolètes et leur anciens états associés, pour ainsi minimiser le surcoût en métadonnées du mécanisme de renommage.

Il existe plusieurs manières pour définir la relation *priority* tout en satisfaisant les propriétés indiquées. Dans le cadre de ce manuscrit, nous avons utilisé l'ordre lexicographique sur les chemins des époques dans l'*arbre des époques* pour définir *priority*. Cette approche se démarque par :

- (i) Sa simplicité.
- (ii) Son surcoût limité, c.-à-d. cette approche n'introduit pas de métadonnées supplémentaires à stocker et diffuser, et l'algorithme de comparaison utilisé est simple.
- (iii) Sa propriété arrangeante sur les déplacements des noeuds dans l'arbre des époques. De manière plus précise, cette définition de *priority* impose aux noeuds de se déplacer que vers l'enfant le plus à droite de l'arbre des époques. Ceci empêche les

noeuds de faire un aller-retour entre deux époques données. Cette propriété permet de passer outre une contrainte concernant le couple de fonctions `renameId` et `revertRenameId` : leur réciprocity.

Cette définition présente cependant plusieurs limites. La limite que nous identifions est sa décorrélation avec le coût et le bénéfice de progresser vers l'époque cible désignée. En effet, l'époque cible est désignée de manière arbitraire par rapport à sa position dans l'arbre des époques. Il est ainsi possible que progresser vers cette époque détériore l'état de la séquence, c.-à-d. augmente la taille des identifiants et augmente le nombre de blocs. De plus, la transition de l'ensemble des noeuds depuis leur époque courante respective vers cette nouvelle époque cible induit un coût en calculs, potentiellement important (cf. ??, page ??).

Pour pallier ce problème, il est nécessaire de proposer une définition de *priority* prenant l'aspect efficacité en compte. L'approche considérée consisterait à inclure dans les opérations *ren* une ou plusieurs métriques qui représente le travail accumulé sur la branche courante de l'arbre des époques, e.g. le nombre d'opérations intégrées, les noeuds actuellement sur cette branche... L'ordre strict total entre les époques serait ainsi construit à partir de la comparaison entre les valeurs de ces métriques de leur opération *ren* respective.

Il conviendra d'adjoindre à cette nouvelle définition de *priority* un nouveau couple de fonctions `renameId` et `revertRenameId` respectant la contrainte de réciprocity de ces fonctions, ou de mettre en place une autre implémentation du mécanisme de renommage ne nécessitant pas cette contrainte, telle qu'une implémentation basée sur le journal des opérations (cf. ??, page ??).

Il conviendra aussi d'étudier la possibilité de combiner l'utilisation de plusieurs relations *priority* pour minimiser le surcoût global du mécanisme de renommage, e.g. en fonction de la distance entre deux époques.

Finalement, il sera nécessaire de valider l'approche proposée par une évaluation comparative par rapport à l'approche actuelle. Cette évaluation pourrait consister à monitorer le coût du système pour observer si l'approche proposée permet de réduire les calculs de manière globale. Plusieurs configurations de paramètres pourraient aussi être utilisées pour déterminer l'impact respectif de chaque paramètre sur les résultats.

2.2.2 Détection et fusion manuelle de versions distantes

À l'issue de cette thèse, nous constatons plusieurs limites des mécanismes de résolution de conflits automatiques dans les systèmes P2P à large échelle sujet au churn. La première d'entre elles est l'utilisation d'un contexte causal. Le contexte causal est utilisé par les mécanismes de résolution de conflits pour :

- (i) Satisfaire le modèle de cohérence causale, c.-à-d. assurer que si nous avons deux modifications m_1 et m_2 telles que $m_1 \rightarrow m_2$, alors l'effet de m_2 supplantera celui de m_1 . Ceci permet d'éviter des anomalies de comportement de la part de la structure de données du point de vue des utilisateur-rices, par exemple la résurgence d'un élément supprimé au préalable.
- (ii) Permettre de préserver l'intention d'une modification malgré l'intégration préalable de modifications concurrentes.

Le contexte causal est utilisé de manière différente en fonction du mécanisme de résolution de conflit. Dans l'approche Operational Transformation (OT), le contexte causal est utilisé par l'algorithme de contrôle pour déterminer les modifications concurrentes à une modification lors de son intégration, afin de prendre en compte leurs effets. Dans l'approche CRDT, le contexte causal est utilisé par la structure de données répliquée à la génération de la modification pour en faire une modification nativement commutative avec les modifications concurrentes, c.-à-d. pour en faire un élément du sup-demi-treillis représentant la structure de données répliquée.

Le contexte causal peut être représenté de différentes manières. Par exemple, le contexte causal peut prendre la forme d'un vecteur de version [51, 52] ou d'un Directed Acyclic Graph (DAG) des modifications [53]. Cependant, de manière intrinsèque, le contexte causal ne fait que de croître au fur et à mesure que des modifications sont effectuées ou que des noeuds rejoignent le système, incrémentant son surcoût en métadonnées, calculs et bande-passante.

La stabilité causale permet cependant de réduire le surcoût lié au contexte causal. En effet, la stabilité causale permet d'établir le contexte commun à l'ensemble des noeuds, c.-à-d. l'ensemble des modifications que l'ensemble des noeuds ont intégré. Ces modifications font alors partie de l'histoire commune et n'ont plus besoin d'être considérées par les mécanismes de résolution de conflits automatiques. La stabilité causale permet donc de déterminer et de tronquer la partie commune du contexte causal pour éviter que ce dernier ne pénalise les performances du système à terme.

La stabilité causale est cependant une contrainte forte dans les systèmes P2P dynamiques à large échelle sujet au churn. Il ne suffit en effet que d'un noeud déconnecté pour empêcher la stabilité causale de progresser. Pour répondre à ce problème, nous avons dès lors tout un spectre d'approches possibles, proposant chacune un compromis entre le surcoût du contexte causal et la probabilité de rejeter des modifications. Les extrémités de ce spectre d'approches sont les suivantes :

- (i) Considérer tout noeud déconnecté comme déconnecté de manière définitive, et donc les ignorer dans le calcul de la stabilité causale. Cette première approche permet à la stabilité causale de progresser, et ainsi aux noeuds connectés de travailler dans des conditions optimales. Mais elle implique cependant que les modifications potentielles des noeuds déconnectés soient perdues, c.-à-d. de ne plus pouvoir les intégrer en l'absence d'un lien entre leur contexte causal de génération et le contexte causal actuel de chaque autre noeud. Il s'agit là de la stratégie la plus agressive en terme de GC du contexte causal.
- (ii) Assurer en toutes circonstances la capacité d'intégration des modifications des noeuds, même ceux déconnectés. Cette seconde approche permet de garantir que les modifications potentielles des noeuds déconnectés pourront être intégrées automatiquement, dans l'éventualité où ces derniers se reconnectent à terme. Mais elle implique de bloquer potentiellement de manière définitive la stabilité causale et donc le mécanisme de GC du contexte causal. Il s'agit là de la stratégie la plus timide en terme de GC du contexte causal.

La seconde limite que nous constatons est la limite des mécanismes actuels de résolution de conflits automatiques pour préserver l'intention des utilisateur-rices. Par exemple,

les mécanismes de résolution de conflits automatiques pour le type Séquence présentés dans ce manuscrit (cf. ??, page ??) définissent l'intention de la manière suivante : *l'intégration de la modification par les noeuds distants doit reproduire l'effet de la modification sur la copie d'origine*. Cette définition assure que chaque modification est porteuse d'une intention, mais limite voire ignore toute la dimension sémantique de la dite intention. Nous conjecturons que l'absence de dimension sémantique réduit les cas d'utilisation de ces mécanismes.

Considérons par exemple une édition collaborative d'un même texte par un ensemble de noeuds. Lors de la présence d'une faute de frappe dans le texte, e.g. le mot "HLLO", plusieurs utilisateur-rices peuvent la corriger en concurrence, c.-à-d. insérer l'élément "E" entre "H" et "L". Les mécanismes de résolution de conflits automatiques permettent aux noeuds d'obtenir des résultats qui convergent mais à notre sens insatisfaisant, e.g. "HEEEEEELLO". Nous considérons ce type de résultats comme des anomalies, au même titre que l'entrelacement [54]. Dans le cadre de collaborations temps réel à échelle limitée, nous conjecturons cependant qu'une granularité fine des modifications permet de pallier ce problème. En effet, les utilisateur-rices peuvent observer une anomalie produite par le mécanisme de résolution de conflits automatique, déduire l'intention initiale des modifications concernées et la restaurer par le biais d'actions supplémentaires de compensation.

Cependant, dans le cadre de collaborations asynchrones ou à large échelle, nous conjecturons que ces anomalies de résolution de conflits s'accumulent. Cette accumulation peut atteindre un seuil rendant laborieuse la déduction et le rétablissement de l'intention initiale des modifications. Le travail imposé aux utilisateur-rices pour résoudre ces anomalies par le biais d'actions de compensation peut alors entraver la collaboration. Pour reprendre l'exemple de l'édition collaborative de texte, nous pouvons constater de tels cas suite à de la duplication de contenu et/ou l'entrelacement de mots, phrases voire paragraphes nuisant à la clarté et correction du texte. Il convient alors de s'interroger sur le bien-fondé de l'utilisation de mécanismes de résolutions de conflits automatiques pour intégrer un ensemble de modifications dans l'ensemble des situations.

Ainsi, pour répondre aux limites des mécanismes de résolution conflits automatiques dans les systèmes P2P à large échelle, c.-à-d. l'augmentation de leur surcoût et la pertinence de leur résultat, nous souhaitons proposer une approche combinant un ou des mécanismes de résolution de conflits automatiques avec un ou des mécanismes de résolution de conflits manuels. L'idée derrière cette approche est de faire varier le mécanisme de résolution de conflits utilisé pour intégrer des modifications. Le choix du mécanisme de résolution de conflits utilisé peut se faire à partir de la valeur d'une distance calculée entre la version courante de la donnée répliquée et celle de la génération de la modification à intégrer, ou d'une évaluation de la qualité du résultat de l'intégration de la modification. Par exemple :

- (i) Si la distance calculée se trouve dans un intervalle de valeurs pour lequel nous disposons d'un mécanisme de résolution de conflits automatique satisfaisant, utiliser ce dernier. Ainsi, nous pouvons envisager de reposer sur plusieurs mécanismes de résolution de conflits automatiques, de plus en plus complexes et pertinents mais coûteux, sans dégrader les performances du système dans le cas de base.
- (ii) Si la distance calculée dépasse la distance seuil, c.-à-d. que nous ne disposons plus

à ce stade de mécanismes de résolution de conflits automatiques satisfaisants, faire intervenir les utilisateur-rices par le biais d'un mécanisme de résolution de conflits manuel. L'utilisation d'un mécanisme manuel n'exclut cependant pas tout pré-travail de notre part pour réduire la charge de travail des utilisateur-rices dans le processus de fusion.

Dans un premier temps, cette approche pourrait se focaliser sur un type d'application spécifique, e.g. l'édition collaborative de texte.

Cette approche nous permettrait de répondre aux limites soulevées précédemment. En effet, elle permettrait de limiter la génération d'anomalies par le mécanisme de résolution de conflits automatique en faisant intervenir les utilisateur-rices. Puis, puisque nous déléguons aux utilisateur-rices l'intégration des modifications à partir d'une distance seuil, nous pouvons dès lors reconsidérer les métadonnées conservées par les noeuds pour les mécanismes de résolution de conflits automatiques. Notamment, nous pouvons identifier les noeuds se trouvant au-delà de cette distance seuil d'après leur dernier contexte causal connu et ne plus les prendre en compte pour le calcul de la stabilité causale. Cette approche permettrait donc de réduire le surcoût lié au contexte causal et limiter la perte de modifications, tout en prenant en considération l'ajout de travail aux utilisateur-rices.

Pour mener à bien ce travail, il conviendra tout d'abord de définir la notion de distance entre versions de la donnée répliquée. Nous envisageons de baser cette dernière sur les deux aspects temporel et spatial, c.-à-d. en utilisant la distance entre contextes causaux et la distance entre contenus. Dans le cadre de l'édition collaborative, nous pourrions pour cela nous baser sur les travaux existants pour évaluer la distance entre deux textes. *Matthieu: TODO : Insérer refs distance de Hamming, Levenshtein, String-to-string correction problem (Tichy et al)*

Il conviendra ensuite de déterminer comment établir la valeur seuil à partir de laquelle la distance entre versions est jugée trop importante. Les approches d'évaluation de la qualité du résultat [55] pourront être utilisées pour déterminer un couple \langle méthode de calcul de la distance, valeur de distance \rangle spécifiant les cas pour lesquels les méthodes de résolution de conflits automatiques ne produisent plus un résultat satisfaisant. *Matthieu: TODO : Insérer refs travaux Claudia et Vinh* Le couple obtenu pourra ensuite être confirmé par le biais d'expériences utilisateurs inspirées de [56, 57].

Finalement, il conviendra de proposer un mécanisme de résolution de conflits adapté pour gérer les éventuelles fusions d'une même modification de façon concurrente par un mécanisme automatique et par un mécanisme manuel, ou à défaut un mécanisme de conscience de groupe invitant les utilisateur-rices à effectuer des actions de compensation.

2.2.3 Étude comparative des différents modèles de synchronisation pour CRDTs

Comme évoqué dans l'état de l'art (cf. ??, page ??), un nouveau modèle de synchronisation pour CRDT fut proposé récemment [58]. Ce dernier propose une synchronisation des noeuds par le biais de différences d'états.

Pour rappel, ce nouveau modèle de synchronisation se base sur le modèle de synchronisation par états. Il partage les mêmes pré-requis, à savoir la nécessité d'une fonction

`merge` associative, commutative et idempotente. Cette dernière doit permettre de la fusion toute paire d'états possible en calculant leur borne supérieure, c.-à-d. leur Least Upper Bound (LUB) [59].

La spécificité de ce nouveau modèle de synchronisation est de calculer pour chaque modification la différence d'état correspondante. Cette différence correspond à un élément irréductible du sup-demi-treillis du CRDT [60], c.-à-d. un état particulier de ce dernier. Cet élément irréductible peut donc être diffusé et intégré par les autres noeuds, toujours à l'aide de la fonction `merge`.

Ce modèle de synchronisation permet alors d'adopter une variété de stratégies de synchronisation, e.g. diffusion des différences de manière atomique, fusion de plusieurs différences puis diffusion du résultat..., et donc de répondre à une grande variété de cas d'utilisation.

Dans notre comparaison des modèles de synchronisation (cf. ??, page ??), nous avons justifié que les CRDTs synchronisés par différences d'états peuvent être utilisés dans les mêmes contextes que les CRDTs synchronisés par états et que les CRDTs synchronisés par opérations. Cette conclusion nous mène à reconsidérer l'intérêt des autres modèles de synchronisation de nos jours.

Par exemple, un CRDT synchronisé par différences d'états correspond à un CRDT synchronisé par états dont nous avons identifié les éléments irréductibles. La différence entre ces deux modèles de synchronisation semble reposer seulement sur la possibilité d'utiliser ces éléments irréductibles pour propager les modifications, en place et lieu des états complets. Nous conjecturons donc que le modèle de synchronisation par états est rendu obsolète par celui par différences d'états. Il serait intéressant de confirmer cette supposition.

En revanche, l'utilisation du modèle de synchronisation par opérations conduit généralement à une spécification différente du CRDT, les opérations permettant d'encoder plus librement les modifications. Notamment, l'utilisation d'opérations peut mener à des algorithmes d'intégration des modifications différents que ceux de la fonction `merge`. Il convient de comparer ces algorithmes pour déterminer si le modèle de synchronisation par opérations peut présenter un intérêt en termes de surcoût.

Au-delà de ce premier aspect, il convient d'explorer d'autres pistes pouvant induire des avantages et inconvénients pour chacun de ces modèles de synchronisation. À l'issue de cette thèse, nous identifions les pistes suivantes :

- (i) La composition de CRDTs, c.-à-d. la capacité de combiner et de mettre en relation plusieurs CRDTs au sein d'un même système, afin d'offrir des fonctionnalités plus complexes. Par exemple, une composition de CRDTs peut se traduire par l'ajout de dépendances entre les modifications des différents CRDTs composés. Le modèle de synchronisation par opérations nous apparaît plus adapté pour cette utilisation, de par le découplage qu'il induit entre les CRDTs et la couche de livraison de messages.
- (ii) L'utilisation de CRDTs au sein de systèmes non-sûrs, c.-à-d. pouvant compter un ou plusieurs adversaires byzantins [61]. Dans de tels systèmes, les adversaires byzantins peuvent générer des modifications différentes mais qui sont perçues comme identiques par les mécanismes de résolution de conflits. Cette attaque, nommée *équivoque*, peut provoquer la divergence définitive des copies. [46] propose une solution

adaptée aux systèmes P2P à large échelle. Celle-ci se base notamment sur l'utilisation de journaux infalsifiables. *Matthieu: TODO : Ajouter refs* Il convient alors d'étudier si l'utilisation de ces structures ne limite pas le potentiel du modèle de synchronisation par différences d'états, e.g. en interdisant la diffusion des modifications par états complets.

Un premier objectif de notre travail serait de proposer des directives sur le modèle de synchronisation à privilégier en fonction du contexte d'utilisation du CRDT.

Ce travail permettrait aussi d'étudier la combinaison des modèles de synchronisation par opérations et par différences d'états au sein d'un même CRDT. Le but serait notamment d'identifier les paramètres conduisant à privilégier un modèle de synchronisation par rapport à l'autre, de façon à permettre aux noeuds de basculer dynamiquement entre les deux.

2.2.4 Approfondissement du patron de conception de Pure Operation-Based CRDTs

BAQUERO et al. [45] proposent un framework pour concevoir des CRDTs synchronisés par opérations : Pure Operation-Based CRDTs. Ce framework a plusieurs objectifs :

- (i) Proposer une approche partiellement générique pour définir un CRDT synchronisé par opérations.
- (ii) Factoriser les métadonnées utilisées par le CRDT pour le mécanisme de résolution de conflits, notamment pour identifier les éléments, et celles utilisées par la couche livraison, notamment pour identifier les opérations.
- (iii) Inclure des mécanismes de GC de ces métadonnées pour réduire la taille de l'état.

Pour cela, les auteurs se limitent aux CRDTs purs synchronisés par opérations, c.-à-d. les CRDTs dont les modifications enrichies de leurs arguments et d'une estampille fournie par la couche de livraison des messages sont commutatives. Pour ces CRDTs, les auteurs proposent un framework générique permettant leur spécification sous la forme d'un journal partiellement ordonné des opérations, nommé PO-Log. Les auteurs associent le PO-Log à une couche de livraison Reliable Causal Broadcast (RCB) des opérations.

Les auteurs définissent ensuite le concept de stabilité causale. Ce concept leur permet de retirer les métadonnées de causalité des opérations du PO-Log lorsque celles-ci sont déterminées comme étant causalement stables.

Finalement, les auteurs définissent un ensemble de relations, spécifiques à chaque CRDT, qui permettent d'exprimer la *redondance causale*. La redondance causale permet de spécifier quand retirer une opération du PO-Log, car rendue obsolète par une autre opération.

Cette approche souffre toutefois de plusieurs limites. Tout d'abord, elle repose sur l'utilisation d'une couche de livraison RCB. Cette couche satisfait le modèle de livraison causale. Mais pour rappel, ce modèle induit l'ajout de données de causalité précises à chaque opération, sous la forme d'un vecteur de version ou d'une barrière causale. Nous jugeons ce modèle trop coûteux pour un système P2P dynamique à large échelle sujet au churn.

En plus du coût induit en termes de métadonnées et de bande-passante, le modèle de livraison causale peut aussi introduire un délai superflu dans la livraison des opérations. En effet, ce modèle impose que tous les messages précédant un nouveau message d'après la relation *happens-before* soient eux-mêmes livrés avant de livrer ce dernier. Il en résulte que des opérations peuvent être mises en attente par la couche livraison, e.g. suite à la perte d'une de leurs dépendances d'après la relation *happens-before*, alors que leurs dépendances réelles ont déjà été livrées et que les opérations sont de fait intégrables en l'état. Plusieurs travaux [62, 63] ont noté ce problème. Pour y répondre et ainsi améliorer la réactivité du framework Pure Operation-Based, ils proposent d'exposer les opérations mises en attente par la couche livraison au CRDT. Bien que fonctionnelle, cette approche induit toujours le coût d'une couche de livraison respectant le modèle de livraison causale et nous fait considérer la raison de ce coût, le modèle de livraison n'étant dès lors plus respecté.

Ensuite, ce framework impose que la modification **prepare** ne puisse pas inspecter l'état courant du noeud. Cette contrainte est compatible avec les CRDTs pour les types de données simples qui sont considérés dans [45], e.g. le Compteur ou l'Ensemble. Elle empêche cependant l'expression de CRDTs pour des types de données plus complexes, e.g. la Séquence ou le Graphe. *Matthieu: TODO : À confirmer pour le graphe* Nous jugeons dommageable qu'un framework pour la conception de CRDTs limite de la sorte son champ d'application.

Finalement, les auteurs ne considèrent que des types de données avec des modifications à granularité fixe. Ainsi, ils définissent la notion de redondance causale en se limitant à ce type de modifications. Par exemple, ils définissent que la suppression d'un élément d'un ensemble rend obsolète les ajouts précédents de cet élément. Cependant, dans le cadre d'autres types de données, e.g. la Séquence, une modification peut concerner un ensemble d'éléments de taille variable. Une opération peut donc être rendue obsolète non pas par une opération, mais par un ensemble d'opérations. Par exemple, les suppressions d'éléments formant une sous-chaîne rendent obsolète l'insertion de cette sous-chaîne. Ainsi, la notion de redondance causale est incomplète et souffre de l'absence d'une notion d'obsolescence partielle d'une opération.

Pour répondre aux différents problèmes soulevés, nous souhaitons proposer une extension du framework Pure Operation-Based CRDTs. Nos objectifs sont les suivants :

- (i) Proposer un framework mettant en lumière la présence et le rôle de deux modèles de livraison :
 - (i) Le modèle de livraison minimal requis par le CRDT pour assurer la convergence forte à terme [22].
 - (ii) Le modèle de livraison utilisé par le système, qui doit être égal ou plus contraint que modèle de livraison minimal du CRDT. Ce second modèle de livraison est une stratégie permettant au système de respecter un modèle de cohérence donné et régissant les règles de compaction de l'état. Il peut être amené à évoluer en fonction de l'état du système et de ses besoins. Par exemple, un système peut par défaut utiliser le modèle de livraison causale pour assurer le modèle de cohérence causal. Puis, lorsque le nombre de noeuds atteint un seuil donné, le système peut passer au modèle de livraison FIFO pour assurer le modèle de cohérence PRAM afin de réduire les coûts en bande-passante.

- (ii) Étendre la notion de redondance causale pour prendre en compte la redondance partielle des opérations. De plus, nous souhaitons rendre cette notion accessible à la couche de livraison, pour détecter au plus tôt les opérations désormais obsolètes et prévenir leur diffusion.
- (iii) Identifier et classifier les mécanismes de résolution de conflits, pour déterminer lesquels sont indépendants de l'état courant pour la génération des opérations et lesquels nécessitent d'inspecter l'état courant dans **prepare**.

Annexe A

Entrelacement d'insertions concurrentes dans Treedoc

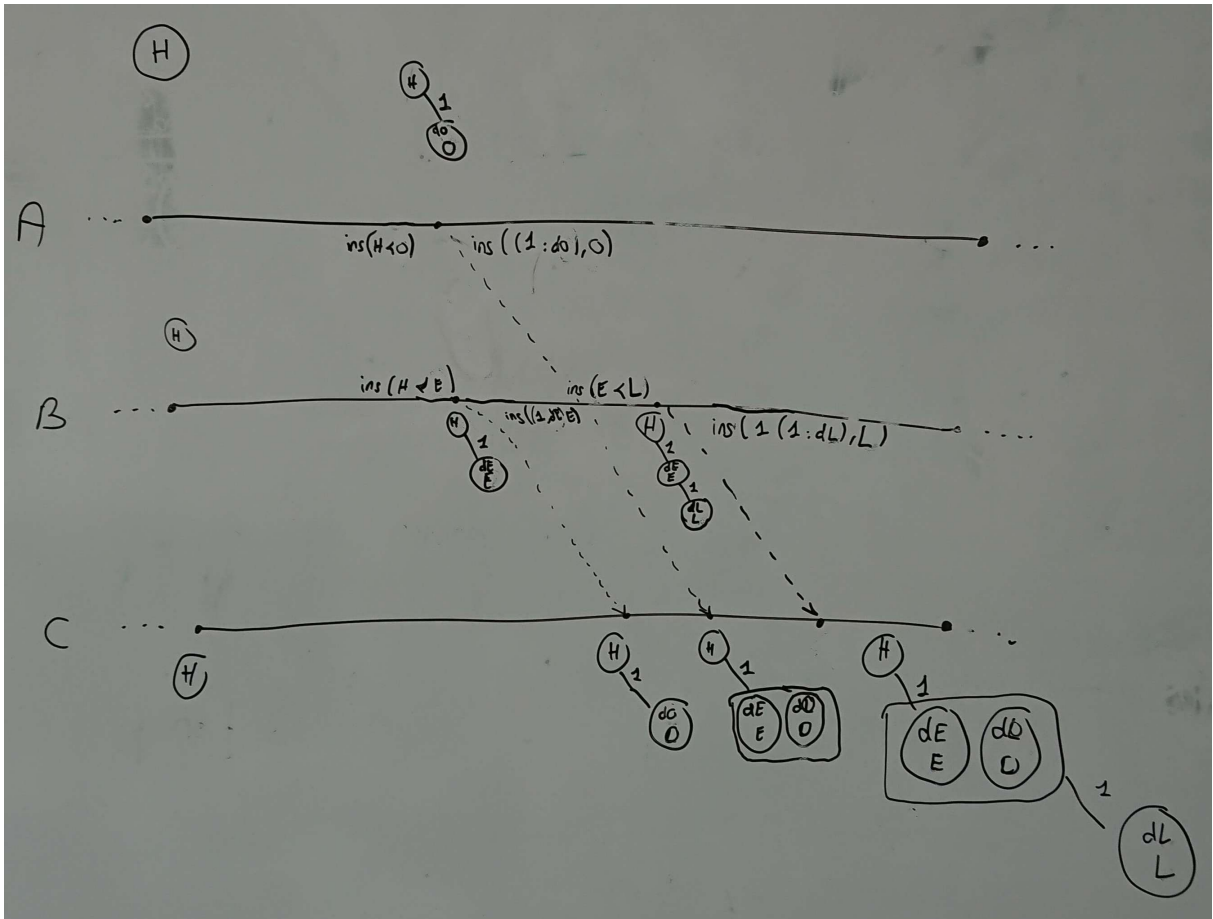


FIGURE A.1 – Modifications concurrentes d'une séquence Treedoc résultant en un entrelacement

Matthieu: TODO : Réaliser au propre contre-exemple. Nécessite que $d_E < d_O$, inverser A et B histoire d'éviter toute confusion. En soi, C pas nécessaire, à voir si le conserve.

Annexe B

Algorithmes RENAMEID

Algorithme 1 Remaining functions to rename an identifier

```
function RENIDLESTHANFIRSTID(id, newFirstId)
  if id < newFirstId then
    return id
  else
    pos ← position(newFirstId)
    nId ← nodeId(newFirstId)
    nSeq ← nodeSeq(newFirstId)
    predNewFirstId ← new Id(pos, nId, nSeq, -1)

    return concat(predNewFirstId, id)
  end if
end function

function RENIDGREATERTHANLASTID(id, newLastId)
  if id < newLastId then
    return concat(newLastId, id)
  else
    return id
  end if
end function
```

Annexe C

Algorithmes REVERTRENAMEID

Algorithme 2 Remaining functions to revert an identifier renaming

```
function REVRENIDLESTHANNEWFIRSTID(id, firstId, newFirstId)
  predNewFirstId  $\leftarrow$  createIdFromBase(newFirstId, -1)
  if isPrefix(predNewFirstId, id) then
    tail  $\leftarrow$  getTail(id, 1)
    if tail < firstId then
      return tail
    else
      ▷ id has been inserted causally after the rename op
      offset  $\leftarrow$  getLastOffset(firstId)
      predFirstId  $\leftarrow$  createIdFromBase(firstId, offset)
      return concat(predFirstId, MAX_TUPLE, tail)
    end if
  else
    return id
  end if
end function

function REVRENIDGREATERTHANNEWLASTID(id, lastId)
  if id < lastId then
    ▷ id has been inserted causally after the rename op
    return concat(lastId, MIN_TUPLE, id)
  else if isPrefix(newLastId, id) then
    tail  $\leftarrow$  getTail(id, 1)
    if tail < lastId then
      ▷ id has been inserted causally after the rename op
      return concat(lastId, MIN_TUPLE, tail)
    else if tail < newLastId then
      return tail
    else
      ▷ id has been inserted causally after the rename op
      return id
    end if
  else
    return id
  end if
end function
```

Index

Voici un index

FiXme :

Notes :

- 1 : Matthieu : TODO : Vérifier du côté des applis de IPFS, 8
- 2 : Matthieu : TODO : Insérer refs distance de Hamming, Levenstein, String-to-string correction problem (Tichy et al), 19
- 3 : Matthieu : TODO : Insérer refs travaux Claudia et Vinh, 19
- 4 : Matthieu : TODO : Ajouter refs, 21
- 5 : Matthieu : TODO : À confirmer pour le graphe, 22
- 6 : Matthieu : TODO : Réaliser au propre contre-exemple. Nécessite que $d_E < d_O$, inverser A et B histoire d'éviter toute confusion. En soi, C pas nécessaire, à voir si le conserve. , 25

FiXme (Matthieu) :

Notes :

- 1 : TODO : Vérifier du côté des applis de IPFS, 8
- 2 : TODO : Insérer refs distance de Hamming, Levenstein, String-to-string correction problem (Tichy et al), 19
- 3 : TODO : Insérer refs travaux Claudia et Vinh, 19
- 4 : TODO : Ajouter refs, 21
- 5 : TODO : À confirmer pour le graphe, 22
- 6 : TODO : Réaliser au propre contre-exemple. Nécessite que $d_E < d_O$, inverser A et B histoire d'éviter toute

confusion. En soi, C pas nécessaire, à voir si le conserve. , 25

Bibliographie

- [1] Jonathan A OBAR et Steven S WILDMAN. « Social Media Definition and the Governance Challenge - An Introduction to the Special Issue ». In : *Obar, JA and Wildman, S.(2015). Social media definition and the governance challenge : An introduction to the special issue. Telecommunications policy* 39.9 (2015), p. 745–750.
- [2] STATISTA. *Biggest social media platforms 2022*. Last Accessed : 2022-10-06. URL : <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>.
- [3] WIKIMEDIA. *Wikimedia Statistics - English Wikipedia*. Last Accessed : 2022-10-06. URL : <https://stats.wikimedia.org/#/en.wikipedia.org>.
- [4] David MEEK. « YouTube and Social Movements : A Phenomenological Analysis of Participation, Events and Cyberplace ». In : *Antipode* 44.4 (2012), p. 1429–1448. DOI : <https://doi.org/10.1111/j.1467-8330.2011.00942.x>. eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-8330.2011.00942.x>. URL : <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8330.2011.00942.x>.
- [5] Yannis THEOCHARIS. « The wealth of (occupation) networks ? Communication patterns and information distribution in a Twitter protest network ». In : *Journal of Information Technology & Politics* 10.1 (2013), p. 35–56.
- [6] José VAN DIJCK et Thomas POELL. « Understanding social media logic ». In : *Media and communication* 1.1 (2013), p. 2–14.
- [7] Jim GILES. « Special Report Internet encyclopaedias go head to head ». In : *nature* 438.15 (2005), p. 900–901.
- [8] Lada A ADAMIC, Jun ZHANG, Eytan BAKSHY et Mark S ACKERMAN. « Knowledge sharing and yahoo answers : everyone knows something ». In : *Proceedings of the 17th international conference on World Wide Web*. 2008, p. 665–674.
- [9] Paul BARAN. « On distributed communications networks ». In : *IEEE transactions on Communications Systems* 12.1 (1964), p. 1–9.
- [10] Safiya Umoja NOBLE. *Algorithms of Oppression : How Search Engines Reinforce Racism*. NYU Press, 2018. ISBN : 9781479849949.
- [11] Amnesty INTERNATIONAL. *#Toxictwitter : Violence and abuse against women online*. Last Accessed : 2022-10-07. URL : <https://www.amnesty.org/en/documents/act30/8070/2018/en/>.

- [12] Wall Street JOURNAL. *Facebook Tried to Make Its Platform a Healthier Place. It Got Angrier Instead*. Last Accessed : 2022-10-07. URL : <https://t.co/P6JohMdhQE>.
- [13] Wall Street JOURNAL. *Facebook Knows Instagram Is Toxic for Teen Girls, Company Documents Show*. Last Accessed : 2022-10-07. URL : <https://t.co/JAvzKFc61q>.
- [14] Martin KLEPPMANN, Adam WIGGINS, Peter van HARDENBERG et Mark MCGRANAGHAN. « Local-First Software : You Own Your Data, in Spite of the Cloud ». In : *Proceedings of the 2019 ACM SIGPLAN International Symposium on New Ideas, New Paradigms, and Reflections on Programming and Software*. Onward! 2019. Athens, Greece : Association for Computing Machinery, 2019, p. 154–178. ISBN : 9781450369954. DOI : 10.1145/3359591.3359737. URL : <https://doi.org/10.1145/3359591.3359737>.
- [15] Daniel ABADI. « Consistency Tradeoffs in Modern Distributed Database System Design : CAP is Only Part of the Story ». In : *Computer* 45.2 (2012), p. 37–42. DOI : 10.1109/MC.2012.33.
- [16] Yasushi SAITO et Marc SHAPIRO. « Optimistic Replication ». In : *ACM Comput. Surv.* 37.1 (mar. 2005), p. 42–81. ISSN : 0360-0300. DOI : 10.1145/1057977.1057980. URL : <https://doi.org/10.1145/1057977.1057980>.
- [17] Douglas B TERRY, Marvin M THEIMER, Karin PETERSEN, Alan J DEMERS, Mike J SPREITZER et Carl H HAUSER. « Managing Update Conflicts in Bayou, a Weakly Connected Replicated Storage System ». In : *SIGOPS Oper. Syst. Rev.* 29.5 (déc. 1995), p. 172–182. ISSN : 0163-5980. DOI : 10.1145/224057.224070. URL : <https://doi.org/10.1145/224057.224070>.
- [18] Daniel STUTZBACH et Reza REJAIE. « Understanding Churn in Peer-to-Peer Networks ». In : *Proceedings of the 6th ACM SIGCOMM Conference on Internet Measurement*. IMC '06. Rio de Janeiro, Brazil : Association for Computing Machinery, 2006, p. 189–202. ISBN : 1595935614. DOI : 10.1145/1177080.1177105. URL : <https://doi.org/10.1145/1177080.1177105>.
- [19] Leslie LAMPORT. « The part-time parliament ». In : *Concurrency : the Works of Leslie Lamport*. 2019, p. 277–317.
- [20] Diego ONGARO et John OUSTERHOUT. « In search of an understandable consensus algorithm ». In : *2014 USENIX Annual Technical Conference (Usenix ATC 14)*. 2014, p. 305–319.
- [21] Marc SHAPIRO et Nuno PREGUIÇA. *Designing a commutative replicated data type*. Research Report RR-6320. INRIA, 2007. URL : <https://hal.inria.fr/inria-00177693>.
- [22] Marc SHAPIRO, Nuno M. PREGUIÇA, Carlos BAQUERO et Marek ZAWIRSKI. « Conflict-Free Replicated Data Types ». In : *Proceedings of the 13th International Symposium on Stabilization, Safety, and Security of Distributed Systems*. SSS 2011. 2011, p. 386–400. DOI : 10.1007/978-3-642-24550-3_29.

-
- [23] Mihai LETIA, Nuno PREGUIÇA et Marc SHAPIRO. « Consistency without concurrency control in large, dynamic systems ». In : *LADIS 2009 - 3rd ACM SIGOPS International Workshop on Large Scale Distributed Systems and Middleware*. T. 44. Operating Systems Review 2. Big Sky, MT, United States : Assoc. for Computing Machinery, oct. 2009, p. 29–34. DOI : 10.1145/1773912.1773921. URL : <https://hal.inria.fr/hal-01248270>.
 - [24] Marek ZAWIRSKI, Marc SHAPIRO et Nuno PREGUIÇA. « Asynchronous rebalancing of a replicated tree ». In : *Conférence Française en Systèmes d'Exploitation (CFSE)*. Saint-Malo, France, mai 2011, p. 12. URL : <https://hal.inria.fr/hal-01248197>.
 - [25] GOOGLE. *Google Docs*. Last Accessed : 2022-10-07. URL : <https://docs.google.com/>.
 - [26] Cody ODGEN. *Google Graveyard*. Last Accessed : 2022-10-11. URL : <https://killedbygoogle.com/>.
 - [27] Matthieu NICOLAS, Victorien ELVINGER, G  rald OSTER, Claudia-Lavinia IGNAT et Fran  ois CHAROY. « MUTE : A Peer-to-Peer Web-based Real-time Collaborative Editor ». In : *ECSCW 2017 - 15th European Conference on Computer-Supported Cooperative Work*. T. 1. Proceedings of 15th European Conference on Computer-Supported Cooperative Work - Panels, Posters and Demos 3. Sheffield, United Kingdom : EUSSET, ao  t 2017, p. 1–4. DOI : 10.18420/ecscw2017_p5. URL : <https://hal.inria.fr/hal-01655438>.
 - [28] Luc ANDR  , St  phane MARTIN, G  rald OSTER et Claudia-Lavinia IGNAT. « Supporting Adaptable Granularity of Changes for Massive-Scale Collaborative Editing ». In : *International Conference on Collaborative Computing : Networking, Applications and Worksharing - CollaborateCom 2013*. Austin, TX, USA : IEEE Computer Society, oct. 2013, p. 50–59. DOI : 10.4108/icst.collaboratecom.2013.254123.
 - [29] Victorien ELVINGER. « R  plication s  curis  e dans les infrastructures pair-  -pair de collaboration ». Theses. Universit   de Lorraine, juin 2021. URL : <https://hal.univ-lorraine.fr/tel-03284806>.
 - [30] Matthieu NICOLAS, Gerald OSTER et Olivier PERRIN. « Efficient Renaming in Sequence CRDTs ». In : *IEEE Transactions on Parallel and Distributed Systems* 33.12 (d  c. 2022), p. 3870–3885. DOI : 10.1109/TPDS.2022.3172570. URL : <https://hal.inria.fr/hal-03772633>.
 - [31] Hoang-Long NGUYEN, Claudia-Lavinia IGNAT et Olivier PERRIN. « Trusternity : Auditing Transparent Log Server with Blockchain ». In : *Companion of the The Web Conference 2018*. Lyon, France, avr. 2018. DOI : 10.1145/3184558.3186938. URL : <https://hal.inria.fr/hal-01883589>.
 - [32] Hoang-Long NGUYEN, Jean-Philippe EISENBARTH, Claudia-Lavinia IGNAT et Olivier PERRIN. « Blockchain-Based Auditing of Transparent Log Servers ». In : *32th IFIP Annual Conference on Data and Applications Security and Privacy (DBSec)*. Sous la dir. de Florian KERSCHBAUM et Stefano PARABOSCHI. T. LNCS-10980.

- Data and Applications Security and Privacy XXXII. Part 1 : Administration. Bergamo, Italy : Springer International Publishing, juil. 2018, p. 21–37. DOI : 10.1007/978-3-319-95729-6_2. URL : <https://hal.archives-ouvertes.fr/hal-01917636>.
- [33] Abhinandan DAS, Indranil GUPTA et Ashish MOTIVALA. « SWIM : scalable weakly-consistent infection-style process group membership protocol ». In : *Proceedings International Conference on Dependable Systems and Networks*. 2002, p. 303–312. DOI : 10.1109/DSN.2002.1028914.
- [34] Armon DADGAR, James PHILLIPS et Jon CURREY. « Lifeguard : Local health awareness for more accurate failure detection ». In : *2018 48th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W)*. IEEE. 2018, p. 22–25.
- [35] Brice NÉDELEC, Julian TANKE, Davide FREY, Pascal MOLLI et Achour MOSTÉFAOUI. « An adaptive peer-sampling protocol for building networks of browsers ». In : *World Wide Web* 21.3 (2018), p. 629–661.
- [36] Mike BURMESTER et Yvo DESMEDT. « A secure and efficient conference key distribution system ». In : *Advances in Cryptology — EUROCRYPT’94*. Sous la dir. d’Alfredo DE SANTIS. Berlin, Heidelberg : Springer Berlin Heidelberg, 1995, p. 275–286. ISBN : 978-3-540-44717-7.
- [37] Matthieu NICOLAS. « Efficient renaming in CRDTs ». In : *Middleware 2018 - 19th ACM/IFIP International Middleware Conference (Doctoral Symposium)*. Rennes, France, déc. 2018. URL : <https://hal.inria.fr/hal-01932552>.
- [38] Matthieu NICOLAS, Gérald OSTER et Olivier PERRIN. « Efficient Renaming in Sequence CRDTs ». In : *7th Workshop on Principles and Practice of Consistency for Distributed Data (PaPoC’20)*. Heraklion, Greece, avr. 2020. URL : <https://hal.inria.fr/hal-02526724>.
- [39] Gérald OSTER, Pascal URSO, Pascal MOLLI et Abdessamad IMINE. « Data Consistency for P2P Collaborative Editing ». In : *ACM Conference on Computer-Supported Cooperative Work - CSCW 2006*. Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work. Banff, Alberta, Canada : ACM Press, nov. 2006, p. 259–268. URL : <https://hal.inria.fr/inria-00108523>.
- [40] Hyun-Gul ROH, Myeongjae JEON, Jin-Soo KIM et Joonwon LEE. « Replicated abstract data types : Building blocks for collaborative applications ». In : *Journal of Parallel and Distributed Computing* 71.3 (2011), p. 354–368. ISSN : 0743-7315. DOI : <https://doi.org/10.1016/j.jpdc.2010.12.006>. URL : <http://www.sciencedirect.com/science/article/pii/S0743731510002716>.
- [41] Nuno PREGUICA, Joan Manuel MARQUES, Marc SHAPIRO et Mihai LETIA. « A Commutative Replicated Data Type for Cooperative Editing ». In : *2009 29th IEEE International Conference on Distributed Computing Systems*. Juin 2009, p. 395–403. DOI : 10.1109/ICDCS.2009.20.

-
- [42] Stéphane WEISS, Pascal URSO et Pascal MOLLI. « Logoot : A Scalable Optimistic Replication Algorithm for Collaborative Editing on P2P Networks ». In : *Proceedings of the 29th International Conference on Distributed Computing Systems - ICDCS 2009*. Montreal, QC, Canada : IEEE Computer Society, juin 2009, p. 404–412. DOI : 10.1109/ICDCS.2009.75. URL : <http://doi.ieeeecomputersociety.org/10.1109/ICDCS.2009.75>.
 - [43] Brice NÉDELEC, Pascal MOLLI, Achour MOSTÉFAOUI et Emmanuel DESMONTILS. « LSEQ : an adaptive structure for sequences in distributed collaborative editing ». In : *Proceedings of the 2013 ACM Symposium on Document Engineering*. DocEng 2013. Sept. 2013, p. 37–46. DOI : 10.1145/2494266.2494278.
 - [44] Brice NÉDELEC, Pascal MOLLI et Achour MOSTÉFAOUI. « A scalable sequence encoding for collaborative editing ». In : *Concurrency and Computation : Practice and Experience* (), e4108. DOI : 10.1002/cpe.4108. eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1002/cpe.4108>. URL : <https://onlinelibrary.wiley.com/doi/abs/10.1002/cpe.4108>.
 - [45] Carlos BAQUERO, Paulo Sergio ALMEIDA et Ali SHOKER. *Pure Operation-Based Replicated Data Types*. 2017. arXiv : 1710.04469 [cs.DC].
 - [46] Victorien ELVINGER, Gérald OSTER et Francois CHAROY. « Prunable Authenticated Log and Authenticable Snapshot in Distributed Collaborative Systems ». In : *2018 IEEE 4th International Conference on Collaboration and Internet Computing (CIC)*. 2018, p. 156–165. DOI : 10.1109/CIC.2018.00031.
 - [47] Sylvie NOËL et Jean-Marc ROBERT. « Empirical study on collaborative writing : What do co-authors do, use, and like ? ». In : *Computer Supported Cooperative Work (CSCW)* 13.1 (2004), p. 63–89.
 - [48] D. S. PARKER, G. J. POPEK, G. RUDISIN, A. STOUGHTON, B. J. WALKER, E. WALTON, J. M. CHOW, D. EDWARDS, S. KISER et C. KLINE. « Detection of Mutual Inconsistency in Distributed Systems ». In : *IEEE Trans. Softw. Eng.* 9.3 (mai 1983), p. 240–247. ISSN : 0098-5589. DOI : 10.1109/TSE.1983.236733. URL : <https://doi.org/10.1109/TSE.1983.236733>.
 - [49] OPENRELAY. *OpenRelay*. Last Accessed : 2022-10-07. URL : <https://openrelay.xyz/>.
 - [50] Protocol LABS. *IPFS*. Last Accessed : 2022-10-07. URL : <https://ipfs.io/>.
 - [51] Friedemann MATTERN et al. *Virtual time and global states of distributed systems*. Univ., Department of Computer Science, 1988.
 - [52] Colin FIDGE. « Logical Time in Distributed Computing Systems ». In : *Computer* 24.8 (août 1991), p. 28–33. ISSN : 0018-9162. DOI : 10.1109/2.84874. URL : <https://doi.org/10.1109/2.84874>.

- [53] Ravi PRAKASH, Michel RAYNAL et Mukesh SINGHAL. « An Adaptive Causal Ordering Algorithm Suited to Mobile Computing Environments ». In : *Journal of Parallel and Distributed Computing* 41.2 (1997), p. 190–204. ISSN : 0743-7315. DOI : <https://doi.org/10.1006/jpdc.1996.1300>. URL : <https://www.sciencedirect.com/science/article/pii/S0743731596913003>.
- [54] Martin KLEPPMANN, Victor B. F. GOMES, Dominic P. MULLIGAN et Alastair R. BERESFORD. « Interleaving Anomalies in Collaborative Text Editors ». In : *Proceedings of the 6th Workshop on Principles and Practice of Consistency for Distributed Data*. PaPoC '19. Dresden, Germany : Association for Computing Machinery, 2019. ISBN : 9781450362764. DOI : 10.1145/3301419.3323972. URL : <https://doi.org/10.1145/3301419.3323972>.
- [55] Quang Vinh DANG et Claudia-Lavinia IGNAT. « Quality Assessment of Wikipedia Articles : A Deep Learning Approach by Quang Vinh Dang and Claudia-Lavinia Ignat with Martin Vesely as Coordinator ». In : *SIGWEB Newsl.* Autumn (nov. 2016). ISSN : 1931-1745. DOI : 10.1145/2996442.2996447. URL : <https://doi.org/10.1145/2996442.2996447>.
- [56] Claudia-Lavinia IGNAT, G  rald OSTER, Meagan NEWMAN, Valerie SHALIN et Fran  ois CHAROY. « Studying the Effect of Delay on Group Performance in Collaborative Editing ». In : *Proceedings of 11th International Conference on Cooperative Design, Visualization, and Engineering, CDVE 2014, Springer 2014 Lecture Notes in Computer Science*. Proceedings of 11th International Conference on Cooperative Design, Visualization, and Engineering, CDVE 2014. Seattle, WA, United States, sept. 2014, p. 191–198. DOI : 10.1007/978-3-319-10831-5_29. URL : <https://hal.archives-ouvertes.fr/hal-01088815>.
- [57] Claudia-Lavinia IGNAT, G  rald OSTER, Olivia FOX, Fran  ois CHAROY et Valerie SHALIN. « How Do User Groups Cope with Delay in Real-Time Collaborative Note Taking ». In : *European Conference on Computer Supported Cooperative Work 2015*. Sous la dir. de Nina BOULUS-RODJE, Gunnar ELLINGSEN, Tone BRATTETEIG, Margunn AANESTAD et Pernille BJORN. Proceedings of the 14th European Conference on Computer Supported Cooperative Work. Oslo, Norway : Springer International Publishing, sept. 2015, p. 223–242. DOI : 10.1007/978-3-319-20499-4_12. URL : <https://hal.inria.fr/hal-01238831>.
- [58] Paulo S  rgio ALMEIDA, Ali SHOKER et Carlos BAQUERO. « Efficient State-Based CRDTs by Delta-Mutation ». In : *Networked Systems*. Sous la dir. d'Ahmed BOUAJJANI et Hugues FAUCONNIER. Cham : Springer International Publishing, 2015, p. 62–76. ISBN : 978-3-319-26850-7.
- [59] Nuno M. PREGUI  A, Carlos BAQUERO et Marc SHAPIRO. « Conflict-free Replicated Data Types (CRDTs) ». In : *CoRR* abs/1805.06358 (2018). arXiv : 1805.06358. URL : <http://arxiv.org/abs/1805.06358>.
- [60] Vitor ENES, Paulo S  rgio ALMEIDA, Carlos BAQUERO et Jo  o LEIT  O. « Efficient Synchronization of State-Based CRDTs ». In : *2019 IEEE 35th International Conference on Data Engineering (ICDE)*. 2019, p. 148–159. DOI : 10.1109/ICDE.2019.00022.

-
- [61] Leslie LAMPORT, Robert SHOSTAK et Marshall PEASE. « The Byzantine Generals Problem ». In : *Concurrency : The Works of Leslie Lamport*. New York, NY, USA : Association for Computing Machinery, 2019, p. 203–226. ISBN : 9781450372701. URL : <https://doi.org/10.1145/3335772.3335936>.
- [62] Jim BAUWENS et Elisa Gonzalez BOIX. « Flec : A Versatile Programming Framework for Eventually Consistent Systems ». In : *Proceedings of the 7th Workshop on Principles and Practice of Consistency for Distributed Data*. PaPoC '20. Heraklion, Greece : Association for Computing Machinery, 2020. ISBN : 9781450375245. DOI : 10.1145/3380787.3393685. URL : <https://doi.org/10.1145/3380787.3393685>.
- [63] Jim BAUWENS et Elisa Gonzalez BOIX. « Improving the Reactivity of Pure Operation-Based CRDTs ». In : *Proceedings of the 8th Workshop on Principles and Practice of Consistency for Distributed Data*. PaPoC '21. Online, United Kingdom : Association for Computing Machinery, 2021. ISBN : 9781450383387. DOI : 10.1145/3447865.3457968. URL : <https://doi.org/10.1145/3447865.3457968>.

Résumé

Afin d'assurer leur haute disponibilité, les systèmes distribués à large échelle se doivent de répliquer leurs données tout en minimisant les coordinations nécessaires entre noeuds. Pour concevoir de tels systèmes, la littérature et l'industrie adoptent de plus en plus l'utilisation de types de données répliquées sans conflits (CRDTs). Les CRDTs sont des types de données qui offrent des comportements similaires aux types existants, tel l'Ensemble ou la Séquence. Ils se distinguent cependant des types traditionnels par leur spécification, qui supporte nativement les modifications concurrentes. À cette fin, les CRDTs incorporent un mécanisme de résolution de conflits au sein de leur spécification.

Afin de résoudre les conflits de manière déterministe, les CRDTs associent généralement des identifiants aux éléments stockés au sein de la structure de données. Les identifiants doivent respecter un ensemble de contraintes en fonction du CRDT, telles que l'unicité ou l'appartenance à un ordre dense. Ces contraintes empêchent de borner la taille des identifiants. La taille des identifiants utilisés croît alors continuellement avec le nombre de modifications effectuées, aggravant le surcoût lié à l'utilisation des CRDTs par rapport aux structures de données traditionnelles. Le but de cette thèse est de proposer des solutions pour pallier ce problème.

Nous présentons dans cette thèse deux contributions visant à répondre à ce problème : (i) Un nouveau CRDT pour Séquence, *RenamableLogootSplit*, qui intègre un mécanisme de renommage à sa spécification. Ce mécanisme de renommage permet aux noeuds du système de réattribuer des identifiants de taille minimale aux éléments de la séquence. Cependant, cette première version requiert une coordination entre les noeuds pour effectuer un renommage. L'évaluation expérimentale montre que le mécanisme de renommage permet de réinitialiser à chaque renommage le surcoût lié à l'utilisation du CRDT. (ii) Une seconde version de *RenamableLogootSplit* conçue pour une utilisation dans un système distribué. Cette nouvelle version permet aux noeuds de déclencher un renommage sans coordination préalable. L'évaluation expérimentale montre que cette nouvelle version présente un surcoût temporaire en cas de renommages concurrents, mais que ce surcoût est à terme.

Mots-clés: CRDTs, édition collaborative en temps réel, cohérence à terme, optimisation mémoire, performance

Abstract

Keywords: CRDTs, real-time collaborative editing, eventual consistency, memory-wise optimisation, performance

