Entretien technique

 ${\bf Matthieu\ Nicolas\ ({\tt nicolasmatthieu570gmail.com})}$

19 octobre 2023

Déroulement de carrière

Ingénieur Recherche & Développement

2014-2017

Inria, Loria, équipe Coast, Nancy

- MUTE : conception et développement d'un éditeur de texte collaboratif temps réel web pair-à-pair (https://mute.loria.fr)
- PLM: webification d'un environnement d'apprentissage de la programmation (http://people.irisa.fr/Martin.Quinson/Teaching/PLM/)

Doctorat en informatique

2017-2022

Université de Lorraine, Loria, équipe Coast, Nancy

 (Ré)Identification efficace dans les types de données répliquées sans conflit (CRDTs)

Déroulement de carrière

Ingénieur Recherche & Développement

2014-2017

Inria, Loria, équipe Coast, Nancy

- MUTE : conception et développement d'un éditeur de texte collaboratif temps réel web pair-à-pair (https://mute.loria.fr)
- PLM: webification d'un environnement d'apprentissage de la programmation (http://people.irisa.fr/Martin.Quinson/Teaching/PLM/)

Doctorat en informatique

2017-2022

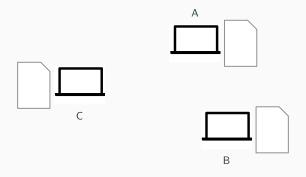
Université de Lorraine, Loria, équipe Coast, Nancy

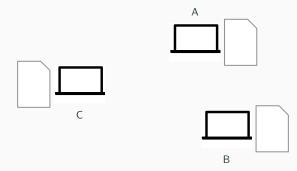
 (Ré)Identification efficace dans les types de données répliquées sans conflit (CRDTs)

MUTE, un exemple de Local-First Software (LFS) [Kle+19]

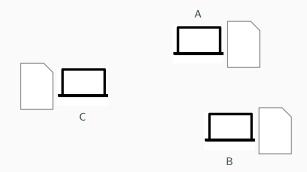


- Application pair-à-pair
- Permet de rédiger collaborativement des documents texte
- Garantit la confidentialité & souveraineté des données

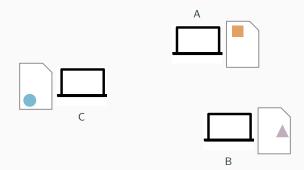




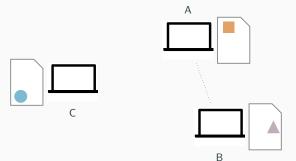
• Noeuds peuvent être déconnectés



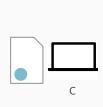
- Noeuds peuvent être déconnectés
- Doivent pouvoir travailler sans coordination synchrone préalable (par ex. consensus)

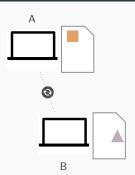


- Noeuds peuvent être déconnectés
- Doivent pouvoir travailler sans coordination synchrone préalable (par ex. consensus)

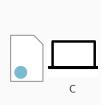


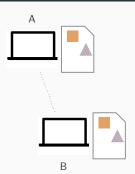
- Noeuds peuvent être déconnectés
- Doivent pouvoir travailler sans coordination synchrone préalable (par ex. consensus)



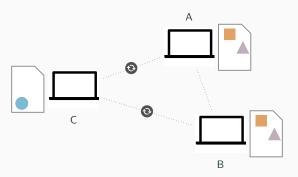


- Noeuds peuvent être déconnectés
- Doivent pouvoir travailler sans coordination synchrone préalable (par ex. consensus)

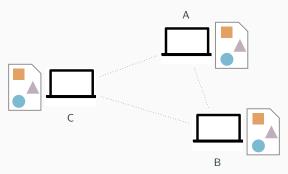




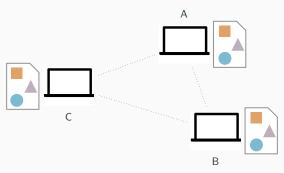
- Noeuds peuvent être déconnectés
- Doivent pouvoir travailler sans coordination synchrone préalable (par ex. consensus)



- Noeuds peuvent être déconnectés
- Doivent pouvoir travailler sans coordination synchrone préalable (par ex. consensus)



- Noeuds peuvent être déconnectés
- Doivent pouvoir travailler sans coordination synchrone préalable (par ex. consensus)
- Doit garantir cohérence à terme [Ter+95]...
- ... malgré ordres différents d'intégration des modifications



- Noeuds peuvent être déconnectés
- Doivent pouvoir travailler sans coordination synchrone préalable (par ex. consensus)
- Doit garantir cohérence à terme [Ter+95]...
- ... malgré ordres différents d'intégration des modifications

Nécessite des mécanismes de résolution de conflits

Conflict-free Replicated Data Types (CRDTs) [Sha+11]

- Nouvelles spécifications des types de données, e.g. Ensemble ou Séquence
- Incorpore nativement mécanisme de résolution de conflits

Conflict-free Replicated Data Types (CRDTs) [Sha+11]

- Nouvelles spécifications des types de données, e.g. Ensemble ou Séquence
- Incorpore nativement mécanisme de résolution de conflits

Propriétés des CRDTs

- Permettent modifications sans coordination
- Garantissent la cohérence forte à terme

Conflict-free Replicated Data Types (CRDTs) [Sha+11]

- Nouvelles spécifications des types de données, e.g. Ensemble ou Séquence
- Incorpore nativement mécanisme de résolution de conflits

Propriétés des CRDTs

- Permettent modifications sans coordination
- Garantissent la cohérence forte à terme

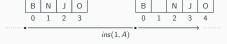
Cohérence forte à terme

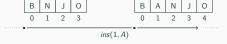
Ensemble des noeuds ayant intégrés le même ensemble de modifications obtient des états équivalents, sans nécessiter d'actions ou messages supplémentaires











Type Séquence usuel



• Changements d'indices sont source de conflits

Type Séquence usuel

CRDTs pour Séquence





- Changements d'indices sont source de conflits
- CRDTs assignent des identifiants de position [Pre+09] à chaque élément
- Identifiants permettent d'ordonner les élements

Type Séquence usuel

CRDTs pour Séquence





- Changements d'indices sont source de conflits
- CRDTs assignent des identifiants de position [Pre+09] à chaque élément
- Identifiants permettent d'ordonner les élements

$$id_0 <_{id} id_1 <_{id} id_2 <_{id} id_3$$

Type Séquence usuel

CRDTs pour Séquence



- Changements d'indices sont source de conflits
- CRDTs assignent des identifiants de position [Pre+09] à chaque élément
- Identifiants permettent d'ordonner les élements

$$id_0 <_{id} id_1 <_{id} id_2 <_{id} id_3$$

Type Séquence usuel

CRDTs pour Séquence



- Changements d'indices sont source de conflits
- CRDTs assignent des identifiants de position [Pre+09] à chaque élément
- Identifiants permettent d'ordonner les élements

$$id_0 <_{id} id_1 <_{id} id_2 <_{id} id_3$$

• Identifiants appartiennent à un espace dense

Type Séquence usuel

CRDTs pour Séquence



- Changements d'indices sont source de conflits
- CRDTs assignent des identifiants de position [Pre+09] à chaque élément
- Identifiants permettent d'ordonner les élements

$$id_0 <_{id} id_1 <_{id} id_2 <_{id} id_3$$

• Identifiants appartiennent à un espace dense

$$id_0 <_{id} id_{0.5} <_{id} id_1$$

Type Séquence usuel

CRDTs pour Séquence



- Changements d'indices sont source de conflits
- CRDTs assignent des identifiants de position [Pre+09] à chaque élément
- Identifiants permettent d'ordonner les élements

$$id_0 <_{id} id_1 <_{id} id_2 <_{id} id_3$$

Identifiants appartiennent à un espace dense

$$id_0 <_{id} id_{0.5} <_{id} id_1$$

Utilise LogootSplit [And+13] comme base

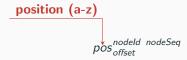
Identifiant

• Composé d'un ou plusieurs tuples de la forme

 $pos_{offset}^{nodeld\ nodeSeq}$

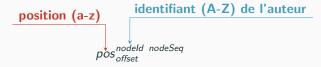
Identifiant

• Composé d'un ou plusieurs tuples de la forme



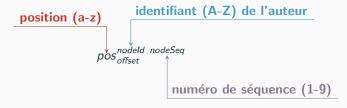
Identifiant

• Composé d'un ou plusieurs tuples de la forme



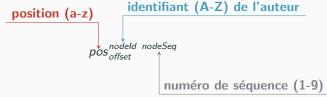
Identifiant

• Composé d'un ou plusieurs tuples de la forme



Identifiant

• Composé d'un ou plusieurs tuples de la forme

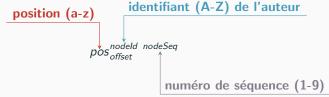


Relation d'ordre < id

Se base sur l'ordre lexicographique sur les éléments des tuples

Identifiant

• Composé d'un ou plusieurs tuples de la forme



Relation d'ordre < id

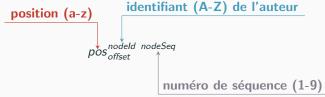
• Se base sur l'ordre lexicographique sur les éléments des tuples

Exemples

$$d_0^{F5}$$

Identifiant

Composé d'un ou plusieurs tuples de la forme



Relation d'ordre < id

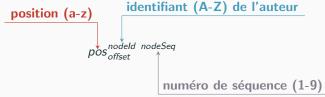
• Se base sur l'ordre lexicographique sur les éléments des tuples

Exemples

$$d_0^{F5} <_{id} m_0^{C1}$$

Identifiant

Composé d'un ou plusieurs tuples de la forme



Relation d'ordre < id

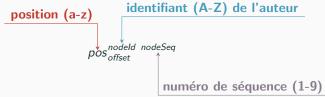
• Se base sur l'ordre lexicographique sur les éléments des tuples

Exemples

$$\mathbf{d}_{0}^{F5} <_{id} \mathbf{m}_{0}^{C1}$$

Identifiant

Composé d'un ou plusieurs tuples de la forme



Relation d'ordre < id

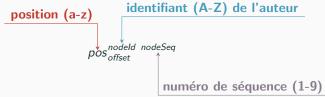
• Se base sur l'ordre lexicographique sur les éléments des tuples

Exemples

$$d_0^{F5} <_{id} m_0^{C1} <_{id} m_0^{C1} f_0^{E1}$$

Identifiant

Composé d'un ou plusieurs tuples de la forme



Relation d'ordre < id

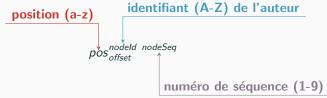
• Se base sur l'ordre lexicographique sur les éléments des tuples

Exemples

$$d_0^{F5} <_{id} m_0^{C1} <_{id} m_0^{C1} \mathbf{f_0^{E1}}$$

Identifiant

Composé d'un ou plusieurs tuples de la forme



Relation d'ordre < id

• Se base sur l'ordre lexicographique sur les éléments des tuples

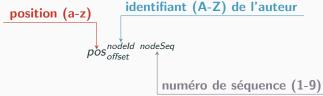
Exemples

$$d_0^{F5} <_{id} m_0^{C1} <_{id} m_0^{C1} f_0^{E1}$$

$$i_0^{B1} <_{id}$$
 ? $<_{id} i_1^{B1}$

Identifiant

Composé d'un ou plusieurs tuples de la forme



Relation d'ordre < id

• Se base sur l'ordre lexicographique sur les éléments des tuples

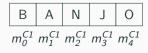
Exemples

$$d_0^{F5} <_{id} m_0^{C1} <_{id} m_0^{C1} f_0^{E1}$$

$$i_0^{B1} <_{id} i_0^{B1} f_0^{A1} <_{id} i_1^{B1}$$

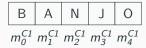
Bloc LogootSplit

• Coûteux de stocker les identifiants de chaque élément



Bloc LogootSplit

• Coûteux de stocker les identifiants de chaque élément



Aggrège en un bloc éléments ayant identifiants contigus

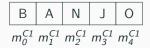
Identifiants contigus

Deux identifiants sont contigus si et seulement si :

- 1. les deux identifiants sont identiques à l'exception de leur dernier offset
- 2. ces deux derniers offsets sont consécutifs

Bloc LogootSplit

• Coûteux de stocker les identifiants de chaque élément



Aggrège en un bloc éléments ayant identifiants contigus

Identifiants contigus

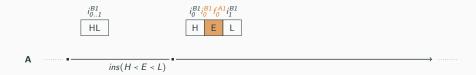
Deux identifiants sont contigus si et seulement si :

- 1. les deux identifiants sont identiques à l'exception de leur dernier offset
- 2. ces deux derniers offsets sont consécutifs
- Note l'intervalle d'identifiants d'un bloc : $pos_{begin..end}^{nodeSeq}$

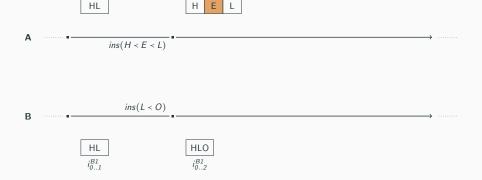
$$\begin{array}{c} \mathsf{BANJO} \\ m_{0..4}^{C1} \end{array}$$

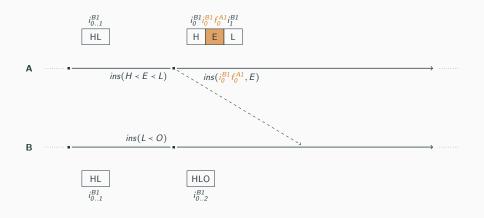


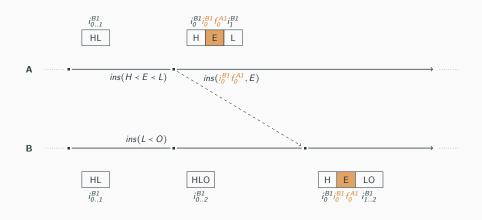
A

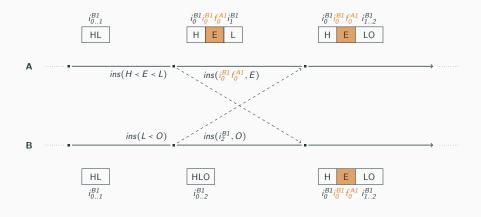


B $\stackrel{\blacksquare}{}$





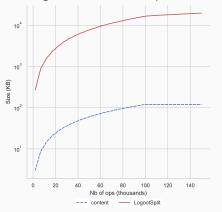




Limites de LogootSplit

Sources de la croissance des métadonnées

- Augmentation non-bornée de la taille des identifiants
- Fragmentation de la séquence en un nombre croissant de blocs



Diminution des performances du point de vue mémoire, calculs et bande-passante

Figure 1 – Taille du contenu comparée à la taille de la séquence LogootSplit

Comment réduire le surcoût?

Solution naïve



• Convertir l'état actuel...

Comment réduire le surcoût?

Solution naïve



- Convertir l'état actuel...
- ... en un état optimisé (identifiants de taille minimale, moins de blocs)...

Comment réduire le surcoût?

Solution naïve



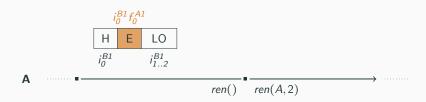
- Convertir l'état actuel...
- ... en un état optimisé (identifiants de taille minimale, moins de blocs)...
- ... à l'aide d'une nouvelle opération

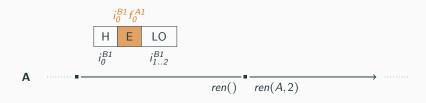
Contribution: RenamableLogootSplit

- CRDT pour le type Séquence qui incorpore un mécanisme de renommage
- Prend la forme d'une nouvelle opération : rename

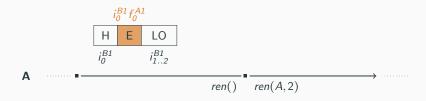
Propriétés de l'opération rename

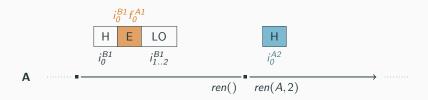
- Est déterministe
- Préserve l'intention des utilisateur-rices
- Préserve les propriétés de la séquence, c.-à-d. l'unicité et l'ordre de ses identifiants
- Commute avec les opérations insert, remove mais aussi rename concurrentes





• Génère nouvel identifiant pour le 1er élément :





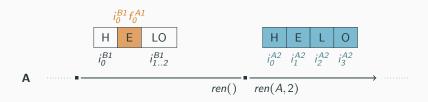
- Génère nouvel identifiant pour le 1er élément : $i_0^{B1} \rightarrow i_0^{A2}$
- Puis génère identifiants contigus pour éléments suivants :



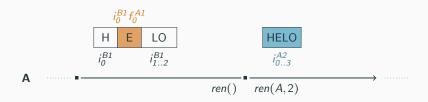
- Puis génère identifiants contigus pour éléments suivants : i_1^{A2}



- Génère nouvel identifiant pour le 1er élément : $i_0^{B1} \rightarrow i_0^{A2}$
- Puis génère identifiants contigus pour éléments suivants : i_1^{A2} , i_2^{A2}

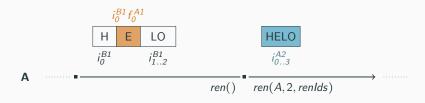


- Génère nouvel identifiant pour le 1er élément : $i_0^{B1} \rightarrow i_0^{A2}$
- Puis génère identifiants contigus pour éléments suivants : i_1^{A2} , i_2^{A2} , ...



- Génère nouvel identifiant pour le 1er élément : $i_0^{B1} \rightarrow i_0^{A2}$
- Puis génère identifiants contigus pour éléments suivants : i_1^{A2} , i_2^{A2} , ...

Regroupe tous les éléments en 1 unique bloc

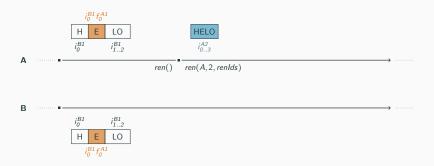


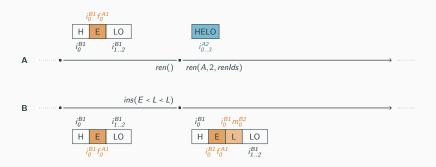
- Génère nouvel identifiant pour le 1er élément : $i_0^{B1} \to i_0^{A2}$
- Puis génère identifiants contigus pour éléments suivants : i_1^{A2} , i_2^{A2} , ...

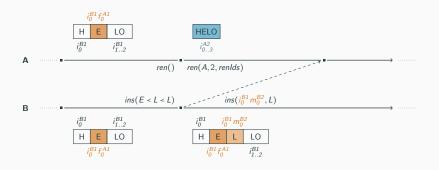
Regroupe tous les éléments en 1 unique bloc

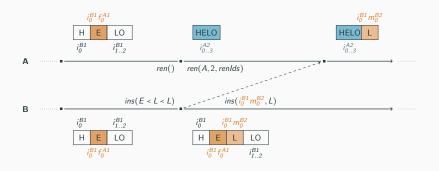
Pour plus tard:

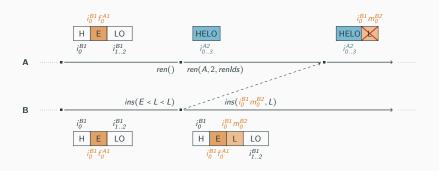
• Stocke identifiants ($[i_0^{B1}, i_0^{B1}f_0^{A1}, \dots]$) de l'état d'origine : renlds



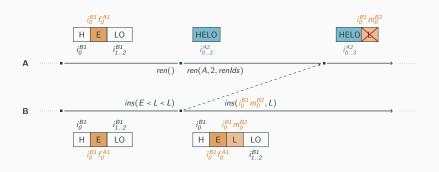








- Noeuds peuvent générer opérations concurrentes aux opérations rename
- Opérations produisent anomalies si intégrées naïvement



- Noeuds peuvent générer opérations concurrentes aux opérations rename
- Opérations produisent anomalies si intégrées naïvement

Nécessité d'un mécanisme dédié

Mécanisme de résolution de conflits entre une opération rename et une opération insert ou remove

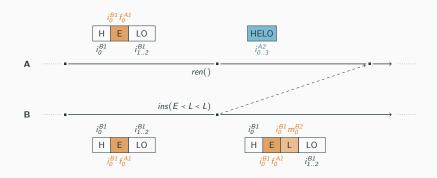
Besoins

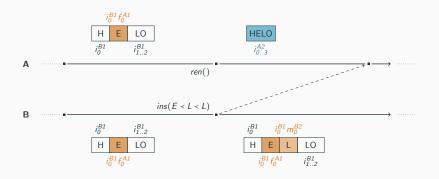
- 1. Détecter les opérations concurrentes aux opérations rename
- 2. Prendre en compte effet des opérations *rename* lors de l'intégration des opérations concurrentes

Mécanisme de résolution de conflits entre une opération rename et une opération insert ou remove

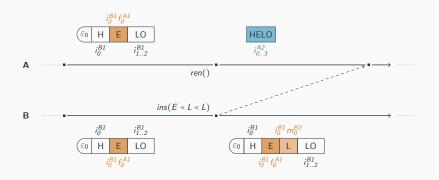
Besoins

- 1. Détecter les opérations concurrentes aux opérations rename
- 2. Prendre en compte effet des opérations *rename* lors de l'intégration des opérations concurrentes



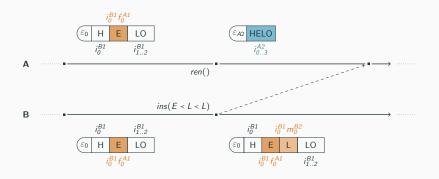


Ajout mécanisme d'époques



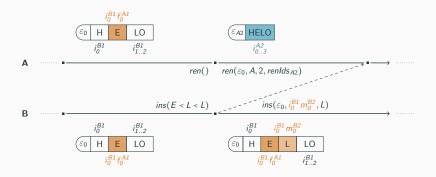
Ajout mécanisme d'époques

• Séquence commence à époque d'origine, notée ε_0



Ajout mécanisme d'époques

- Séquence commence à époque d'origine, notée ε_0
- rename font progresser à nouvelle époque, $\varepsilon_{nodeld\ nodeSeq}$



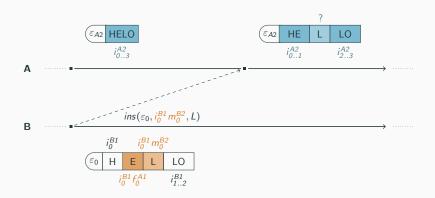
Ajout mécanisme d'époques

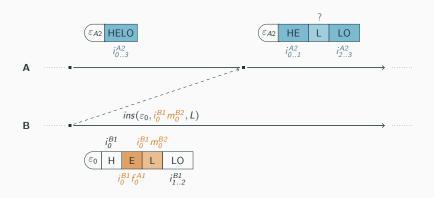
- Séquence commence à époque d'origine, notée ε_0
- ullet rename font progresser à nouvelle époque, $arepsilon_{nodeld\ nodeSeq}$
- Opérations labellisées avec époque de génération

Mécanisme de résolution de conflits

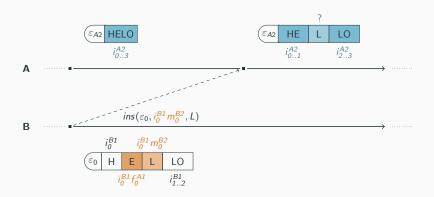
Besoins

- 1. Détecter les opérations concurrentes aux opérations rename
- 2. Prendre en compte effet des opérations *rename* lors de l'intégration des opérations concurrentes



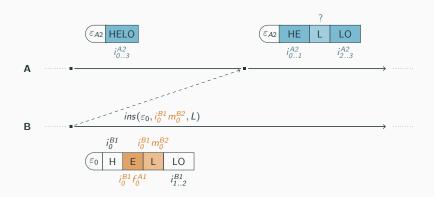


Ajout d'un mécanisme de transformation des opérations insert et remove concurrentes



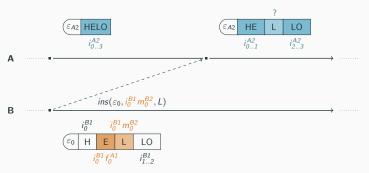
Ajout d'un mécanisme de transformation des opérations insert et remove concurrentes

Prend la forme de l'algorithme renameId



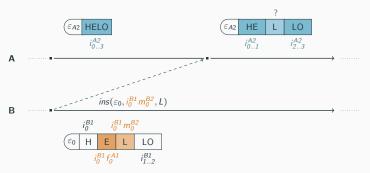
Ajout d'un mécanisme de transformation des opérations insert et remove concurrentes

- Prend la forme de l'algorithme renameId
- Inclure l'effet de l'opération rename dans l'opération transformée



Rappel:

$$renlds_{A2} = \left[i_0^{B1}, i_0^{B1}f_0^{A1}, i_1^{B1}, i_2^{B1}\right]$$

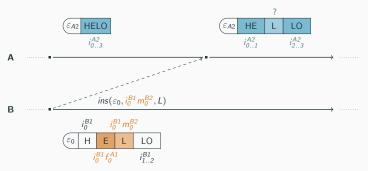


Rappel:

$$renlds_{A2} = [i_0^{B1}, \mathbf{i_0^{B1}f_0^{A1}}, i_1^{B1}, i_2^{B1}]$$

Exemple avec $i_0^{B1} m_0^{B2}$

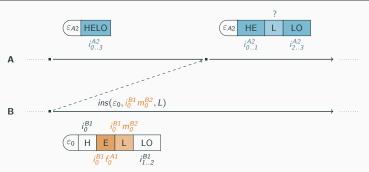
• Rechercher son prédecesseur dans $renlds_{A2}: i_0^{B1} f_0^{A1}$



Rappel:

$$renlds_{A2} = \left[i_0^{B1}, i_0^{B1} f_0^{A1}, i_1^{B1}, i_2^{B1}\right]$$

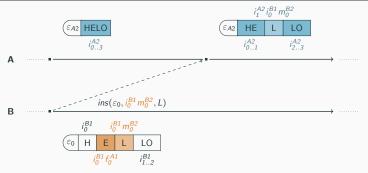
- Rechercher son prédecesseur dans $renlds_{A2}: i_0^{B1} f_0^{A1}$
- Utiliser son index (1) pour calculer équivalent à époque ε_{A2} : i_1^{A2}



Rappel:

$$renlds_{A2} = [i_0^{B1}, i_0^{B1} f_0^{A1}, i_1^{B1}, i_2^{B1}]$$

- Rechercher son prédecesseur dans $renlds_{A2}: i_0^{B1} f_0^{A1}$
- Utiliser son index (1) pour calculer équivalent à époque ε_{A2} : i_1^{A2}
- Préfixer $i_0^{B1} m_0^{B2}$ par ce dernier : $i_1^{A2} i_0^{B1} m_0^{B2}$



Rappel:

$$renlds_{A2} = [i_0^{B1}, i_0^{B1} f_0^{A1}, i_1^{B1}, i_2^{B1}]$$

- Rechercher son prédecesseur dans $renlds_{A2}$: $i_0^{B1} f_0^{A1}$
- Utiliser son index (1) pour calculer équivalent à époque ε_{A2} : i_1^{A2}
- Préfixer $i_0^{B1} m_0^{B2}$ par ce dernier : $i_1^{A2} i_0^{B1} m_0^{B2}$

Objectifs

- Montrer que RenamableLogootSplit satisfait la convergence forte
- Montrer que le mécanisme de renommage améliore les performances de la séquence répliquée (mémoire, calculs, bande-passante)

Objectifs

- Montrer que RenamableLogootSplit satisfait la convergence forte
- Montrer que le mécanisme de renommage améliore les performances de la séquence répliquée (mémoire, calculs, bande-passante)

Conduite d'une évaluation expérimentale

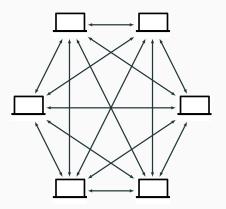
d'édition collaborative

Absence d'un jeu de données de sessions

Absence d'un jeu de données de sessions d'édition collaborative

Mise en place de simulations pour générer un jeu de données

Simulations - Architecture



- 10 noeuds éditent collaborativement un document
- Topologie réseau entièrement maillée
- Ne considère pas de pannes ou de pertes de message

Simulations - Modifications

Noeuds utilisent LogootSplit (LS) ou RenamableLogootSplit (RLS)

Simulations - Modifications

Noeuds utilisent LogootSplit (LS) ou RenamableLogootSplit (RLS)

Se décompose en 2 phases

- 1. Génération du contenu (80% d'insert, 20% de remove)
- 2. Édition (50/50%)

Noeuds passent à la phase 2 quand leur copie locale atteint une taille donnée (15 pages - 60k caractères)

Simulations - Modifications

Noeuds utilisent LogootSplit (LS) ou RenamableLogootSplit (RLS)

Se décompose en 2 phases

- 1. Génération du contenu (80% d'insert, 20% de remove)
- 2. Edition (50/50%)

Noeuds passent à la phase 2 quand leur copie locale atteint une taille donnée (15 pages - 60k caractères)

Nombre d'opérations : 15k par noeud, 150k au total

Simulations - Mécanisme de renommage

Noeuds de renommage

- 1 à 4 noeuds effectuent une opération rename toutes les 30k opérations
- Opérations rename générées à un point donné sont concurrentes

Simulations - Sorties

- Instantané de l'état de chaque noeud à différents points de la simulation (10k opérations et état final)
- Journal des opérations de chaque noeud

^{*.} Code des simulations et benchmarks :
https://github.com/coast-team/mute-bot-random

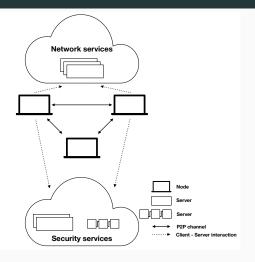
Simulations - Sorties

- Instantané de l'état de chaque noeud à différents points de la simulation (10k opérations et état final)
- Journal des opérations de chaque noeud

Permet de conduire évaluations sur ces données *

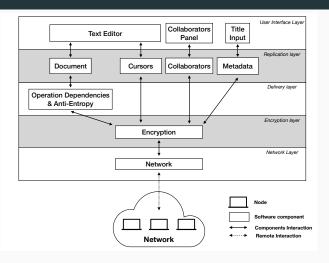
^{*.} Code des simulations et benchmarks :
https://github.com/coast-team/mute-bot-random

Architecture système de MUTE



• Toto

Architecture logicielle de MUTE



Toto