

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

Matthieu Nicolas (matthieu.nicolas@loria.fr)

20 décembre 2022

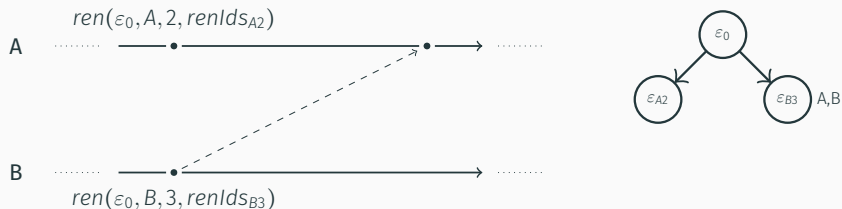
<i>Rapporteurs :</i>	Hanifa Boucheneb Davide Frey	Professeure, Polytechnique Montréal Chargé de recherche, HdR, Inria Rennes Bretagne-Atlantique
<i>Examineurs :</i>	Hala Skaf-Molli Stephan Merz	Professeure des Universités, Nantes Université, LS2N Directeur de Recherche, Inria Nancy - Grand Est
<i>Encadrants :</i>	Olivier Perrin Gérald Oster	Professeur des Universités, Université de Lorraine, LORIA Maître de conférences, Université de Lorraine, LORIA

RenamableLogootSplit

Mais ça sert à quoi de renommer ?

Puisqu'on doit conserver les *renIds* pour gérer les opérations concurrentes...

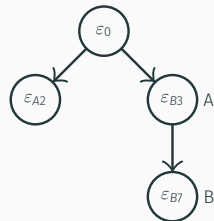
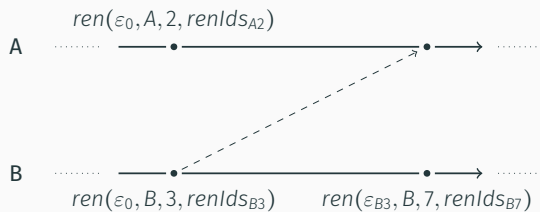
Suppression des *renlds* obsolètes (Garbage Collection)



- Besoin de garder *renlds* pour transformer les opérations

[1]. 10.1007/978-3-662-43352-2_11.

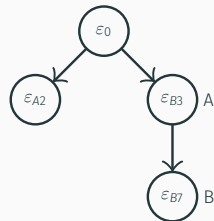
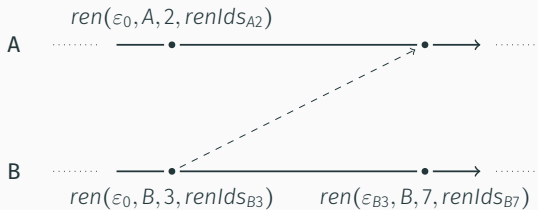
Suppression des *renIds* obsolètes (Garbage Collection)



- Besoin de garder *renIds* pour transformer les opérations

[1]. 10.1007/978-3-662-43352-2_11.

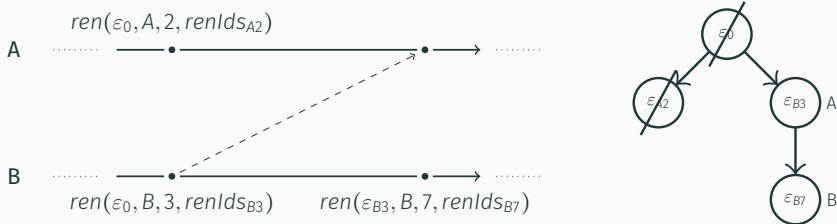
Suppression des *renIds* obsolètes (Garbage Collection)



- Besoin de garder *renIds* pour transformer les opérations
- Si plus d'opérations nécessitant des transformations vers une époque donnée...

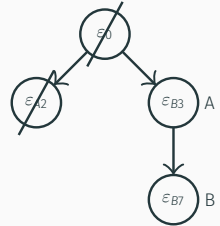
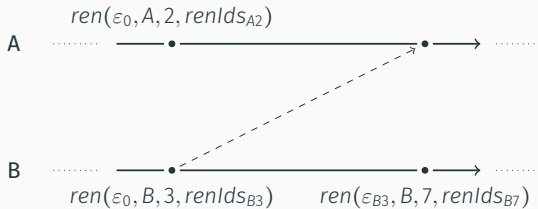
[1]. 10.1007/978-3-662-43352-2_11.

Suppression des *renlds* obsolètes (Garbage Collection)



- Besoin de garder *renlds* pour transformer les opérations
- Si plus d'opérations nécessitant des transformations vers une époque donnée...
- ...alors époque et *renlds* correspondant obsolètes

Suppression des *renIds* obsolètes (Garbage Collection)



- Besoin de garder *renIds* pour transformer les opérations
- Si plus d'opérations nécessitant des transformations vers une époque donnée...
- ...alors époque et *renIds* correspondant obsolètes

Besoins

- Détecter stabilité causale^[1] des opérations *rename*
- Connaître noeuds appartenant au groupe

[1]. 10.1007/978-3-662-43352-2_11.

RenamableLogootSplit

Validation

- Montrer convergence des noeuds
- Montrer que mécanisme de renommage améliore performances de la séquence répliquée (mémoire, calculs, bande-passante)

- Montrer convergence des noeuds
- Montrer que mécanisme de renommage améliore performances de la séquence répliquée (mémoire, calculs, bande-passante)

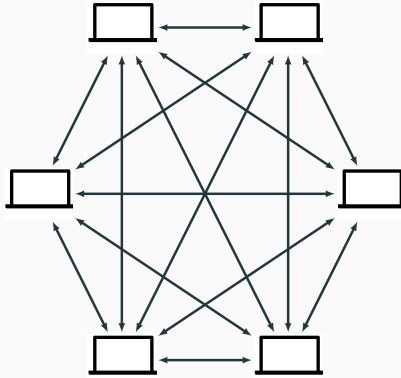
Conduite d'une évaluation expérimentale

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

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

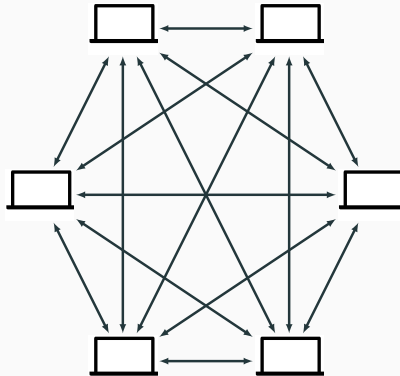
Mise en place de simulations pour générer un
jeu de données

Simulations - Architecture



- 10 noeuds éditent collaborativement un document
- Utilisent soit LogootSplit (LS), soit RenamableLogootSplit (RLS)

Simulations - Architecture



- 10 noeuds éditent collaborativement un document
- Utilisent soit LogootSplit (LS), soit RenamableLogootSplit (RLS)
- Topologie réseau entièrement maillée
- Ne considère pas pannes ou pertes de message

- Phase 1 (génération du contenu) : Beaucoup d'insertions, quelques suppressions (80/20%)
- Phase 2 (édition) : Équilibre insertions/suppressions (50/50%)
- Noeuds passent à la phase 2 quand document atteint taille donnée (15 pages - 60k caractères)

Simulations - Modifications

- Phase 1 (génération du contenu) : Beaucoup d'insertions, quelques suppressions (80/20%)
- Phase 2 (édition) : Équilibre insertions/suppressions (50/50%)
- Noeuds passent à la phase 2 quand document atteint taille donnée (15 pages - 60k caractères)
- Noeuds terminent quand ensemble des noeuds a effectué nombre donné de modifications (15k)...
- ...et intégré celles des autres (150k au total)

- Noeuds désignés comme *noeuds de renommage* (1 à 4)
- Noeuds de renommage effectue un renommage à toutes les 30k opérations qu'ils intègrent (5 opérations *rename* par noeud de renommage)
- Opérations *rename* générées à un point donné sont **concurrentes**

- 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>

- 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>

Convergence

Intuition

Comparer l'état final des différents noeuds d'une session pour confirmer l'absence de divergence

Intuition

Comparer l'état final des différents noeuds d'une session pour confirmer l'absence de divergence

- Ensemble des noeuds convergent

Intuition

Comparer l'état final des différents noeuds d'une session pour confirmer l'absence de divergence

- Ensemble des noeuds convergent
- Un résultat empirique, pas une preuve...
- ...mais un premier pas vers la validation de RLS

Surcoût en métadonnées - 1 noeud de renommage

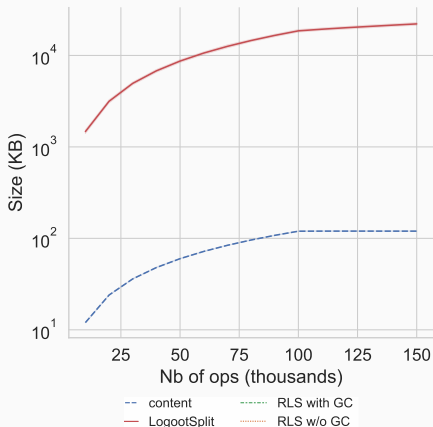
Intuition

Mesurer évolution de la taille de la structure de données à partir des instantanés des sessions avec 1 seul noeud de renommage

Surcoût en métadonnées - 1 noeud de renommage

Intuition

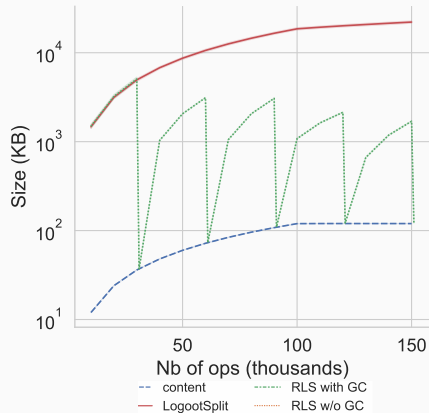
Mesurer évolution de la taille de la structure de données à partir des instantanés des sessions avec 1 seul noeud de renommage



Surcoût en métadonnées - 1 noeud de renommage

Intuition

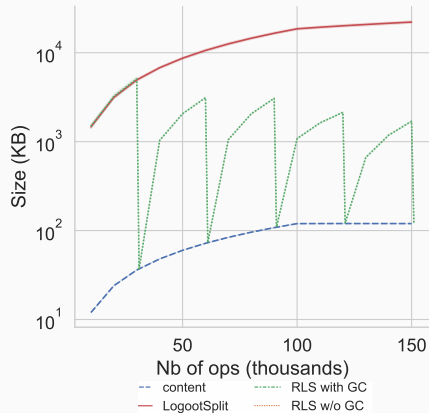
Mesurer évolution de la taille de la structure de données à partir des instantanés des sessions avec 1 seul noeud de renommage



Surcoût en métadonnées - 1 noeud de renommage

Intuition

Mesurer évolution de la taille de la structure de données à partir des instantanés des sessions avec 1 seul noeud de renommage



Observations

- Opération *rename* réinitialise surcoût du CRDT, si GC possible

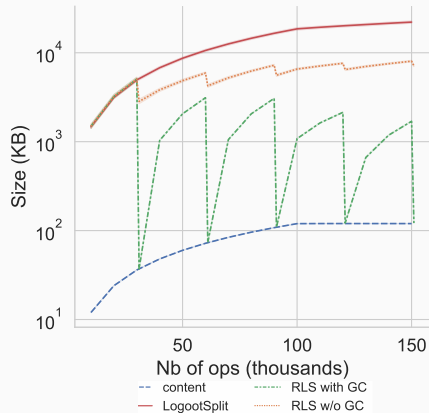
Surcoût en métadonnées - 1 noeud de renommage

Intuition

Mesurer évolution de la taille de la structure de données à partir des instantanés des sessions avec 1 seul noeud de renommage

Observations

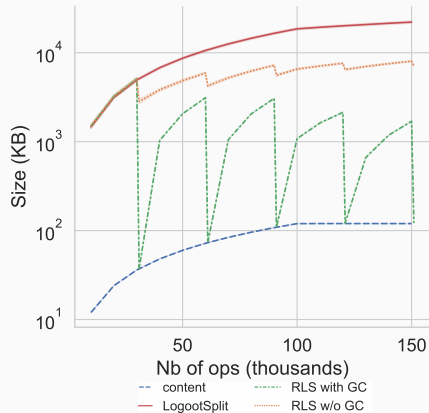
- Opération *rename* réinitialise surcoût du CRDT, si GC possible



Surcoût en métadonnées - 1 noeud de renommage

Intuition

Mesurer évolution de la taille de la structure de données à partir des instantanés des sessions avec 1 seul noeud de renommage



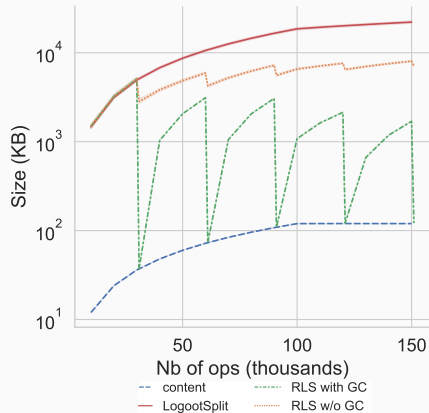
Observations

- Opération *rename* réinitialise surcoût du CRDT, si GC possible
- Opération *rename* réduit de 66% surcoût du CRDT, si GC impossible

Surcoût en métadonnées - 1 noeud de renommage

Intuition

Mesurer évolution de la taille de la structure de données à partir des instantanés des sessions avec 1 seul noeud de renommage



Observations

- Opération *rename* réinitialise surcoût du CRDT, si GC possible
- Opération *rename* réduit de 66% surcoût du CRDT, si GC impossible

Explications

- Un seul bloc après une opération *rename*
- Réduction du nombre de blocs réduit les métadonnées utilisée par l'implémentation de la Séquence (arbre AVL)

Surcoût en métadonnées - 1 à 4 noeuds de renommage

Intuition

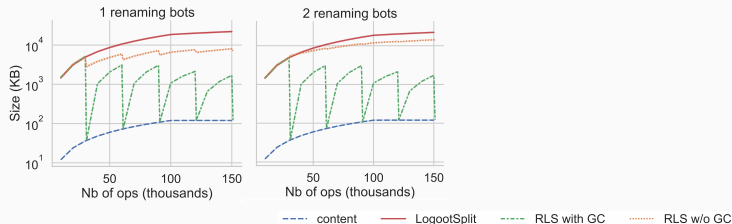
Mesurer évolution de la taille de la structure de données en fonction du nombre de noeuds de renommage



Surcoût en métadonnées - 1 à 4 noeuds de renommage

Intuition

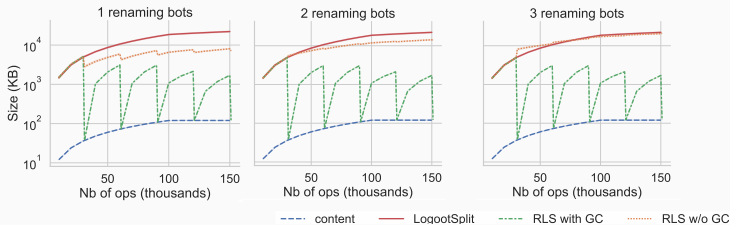
Mesurer évolution de la taille de la structure de données en fonction du nombre de noeuds de renommage



Surcoût en métadonnées - 1 à 4 noeuds de renommage

Intuition

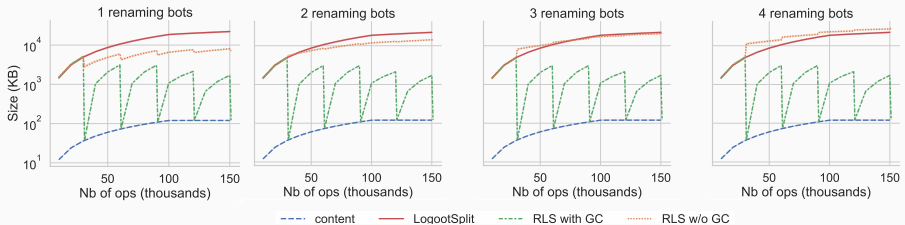
Mesurer évolution de la taille de la structure de données en fonction du nombre de noeuds de renommage



Surcoût en métadonnées - 1 à 4 noeuds de renommage

Intuition

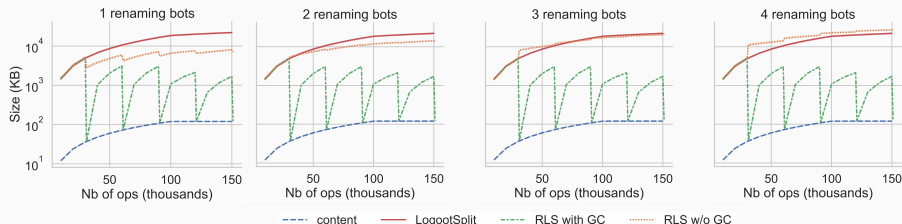
Mesurer évolution de la taille de la structure de données en fonction du nombre de noeuds de renommage



Surcoût en métadonnées - 1 à 4 noeuds de renommage

Intuition

Mesurer évolution de la taille de la structure de données en fonction du nombre de noeuds de renommage



- Aucun impact si GC possible
- Métadonnées (époque et *renIds*) de chaque opération *rename* s'additionne si GC impossible

Surcoût en calculs - Opérations *insert* et *remove*

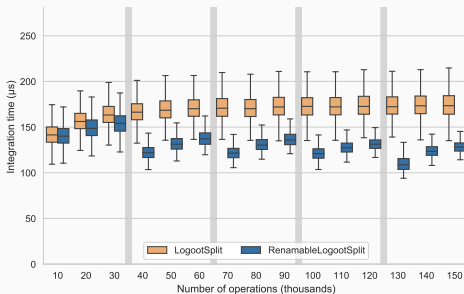
Intuition

Mesurer temps d'intégration **local** et **distant** d'opérations *insert* à différents stades de la collaboration

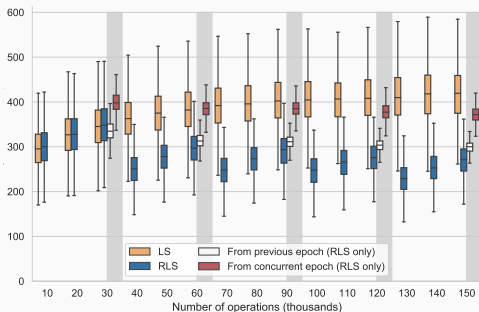
Surcoût en calculs - Opérations *insert* et *remove*

Intuition

Mesurer temps d'intégration **local** et **distant** d'opérations *insert* à différents stades de la collaboration



(a) Temps intégration modifs locales

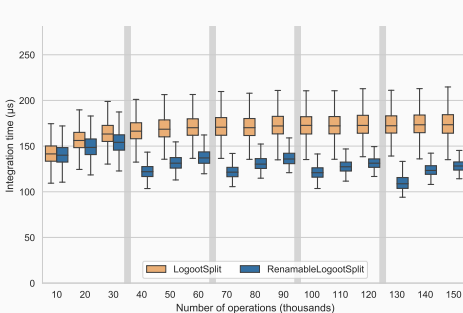


(b) Temps intégration modifs distantes

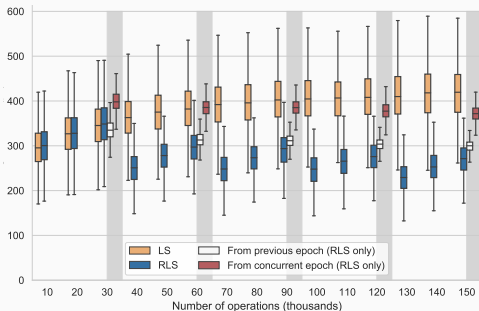
Surcoût en calculs - Opérations *insert* et *remove*

Intuition

Mesurer temps d'intégration **local** et **distant** d'opérations *insert* à différents stades de la collaboration



(a) Temps intégration modifs locales



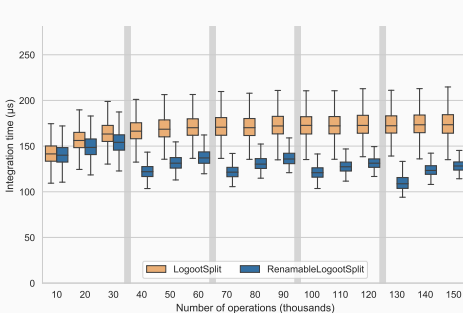
(b) Temps intégration modifs distantes

- Opération *rename* réduit les temps d'intégration

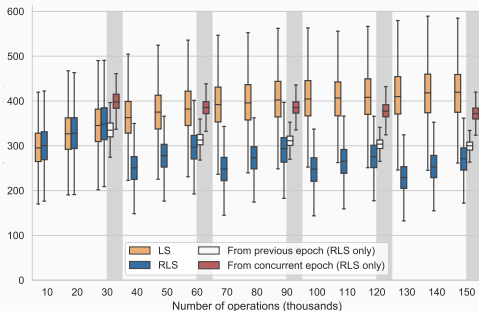
Surcoût en calculs - Opérations *insert* et *remove*

Intuition

Mesurer temps d'intégration **local** et **distant** d'opérations *insert* à différents stades de la collaboration



(a) Temps intégration modifs locales



(b) Temps intégration modifs distantes

- Opération *rename* réduit les temps d'intégration
- Réduction du nombre de blocs contrebalance le surcoût des transformations