

Simulation d'algorithmes d'équilibrage de charge dans un environnement distribué

Identifications des besoins

Kevin Barreau

Guillaume Marques

Corentin Salingue

26 février 2015

Résumé

Dans une première partie, nous présenterons le projet, le contexte et les hypothèses. Ensuite, nous développerons les besoins fonctionnels et les besoins non-fonctionnels. De plus, nous dégagerons une première version de la planification du projet (GANTT). Enfin, nous présenterons les livrables.

Sommaire

1	Présentation du projet	4
1.1	Utilisation d'une base de données par un client	4
1.2	Base de données distribuée	4
1.3	Gestion des requêtes dans la base de données distribuée	5
1.3.1	Requêtes de lecture	5
1.3.2	Requêtes d'écriture	6
1.4	Stockage des données	6
1.5	Protocoles de réaffectation des requêtes de lecture	7
1.6	Gestion de la popularité des objets	7
1.7	Gestion des copies d'un objet	8
1.8	Visualisation des statistiques de fonctionnement de la base de données distribuée	8
2	Définition du projet	8
2.1	Contexte	8
2.2	Finalité	9
2.3	Hypothèses	9
3	Ordonnancement des besoins	11
3.1	Priorité	11
3.2	Criticité	11
4	Besoins fonctionnels	12
4.1	Gestion d'un réseau	12
4.1.1	Gestion des noeuds	12
4.1.2	Réplication d'un objet	12
4.1.3	Popularité d'un objet	13
4.2	Protocoles d'affectation	13
4.3	Requêtes	14
4.3.1	Générations de requêtes	14
4.3.2	Importation d'un jeu de requête	14
4.4	Visualisation des données	14
4.4.1	Enregistrement des données	14
4.4.2	Affichage des données	14
5	Besoins non fonctionnels	16
5.1	Cassandra	16
5.2	Maintenabilité du projet	16
5.3	Gestion d'un réseau	16
5.3.1	Communication entre noeuds	16
5.3.2	Taille des données	16
5.4	Protocole de test	16
5.5	Visualisation des données	17
5.5.1	Etat du réseau	17
5.5.2	Actualisation de la vue	17
6	Répartitions des tâches	18
6.1	Diagramme de Gantt	18
6.2	Affectation des tâches	19

7 Livrables	20
7.1 Livrables intermédiaires	20
7.2 Livrable final	20

Table des figures

1	Intéractions client/base de données	4
2	Processus pour la visualisation des statistiques	8
3	Visualisation d'une base de données distribuée sous forme de cluster possédant trois data center	21
4	Exemple de partitionnement des données dans une base de données distribuée	22
5	Partitionnement des réplicas d'un objet avec une fonction de hachage pour chaque réplica	23
6	Cheminement d'une requête de lecture dans une base de données distribuée avec la prise en charge de l'affectation (un seul noeud traite la requête)	24
7	Cheminement d'une requête d'écriture dans une base de données distribuée	25
8	Passage d'une représentation des données pour le client à une représentation pour la base de données	25
9	Fonctionnement de l'algorithme de réaffectation des requêtes de lecture SLVO	26

1 Présentation du projet

Dans le présent document, nous considérons que le lecteur possède des notions en informatique et que chaque mot est défini par son sens commun. Cependant, si un terme utilisé présentant une définition différente que celle admise par tous, nous ne manquerons pas de le préciser et de le définir.

1.1 Utilisation d'une base de données par un client

Une *base de données* est un outil permettant de stocker et récupérer des *données*, : codage (une représentation sous forme binaire), propre au système de base de données, d'une information quelconque.

Dans un premier temps, le client se connecte à la base de données. Le client interagit avec celle-ci en lui envoyant des *requêtes*, message, dont la forme dépend de la base de données et permettant de stocker, récupérer ou modifier des données.

Selon les requêtes émises par le client, la base de données envoie des résultats.

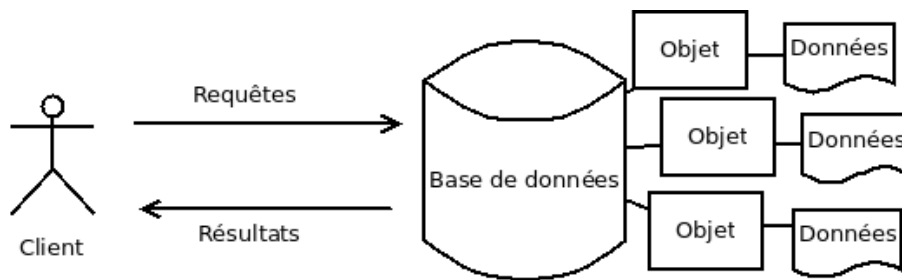


FIGURE 1 – Interactions client/base de données

On distingue deux types de requêtes :

- Les requêtes de **lecture** : requêtes ne modifiant pas les données contenues dans la base de données. Il s'agit de récupérer des objets contenus dans la base de données.
- Les requêtes d'**écriture** : requêtes modifiant les données contenues dans la base de données.

Le client peut être une personne physique ou un logiciel. Dans notre cas, il s'agit d'un logiciel permettant l'importation de fichiers contenant des requêtes ou de générer des requêtes pseudo-aléatoirement.

1.2 Base de données distribuée

La base de données utilisée par le client est plus précisément une base de données dites *distribuée*. Le client ne voit pas de différence, lorsqu'il l'utilise, entre une base de données classique et une base de données distribuée. On dit qu'une base de données est distribuée lorsque les données qu'elle stocke sont réparties sur plusieurs machines ou emplacements physiques, appelés *noeuds*. Les noeuds sont capables de communiquer entre eux afin de s'échanger des informations.

On peut rassembler des noeuds pour former un *data center*. Un rassemblement de data center correspond à un *cluster* (voir la figure 3).

La base de données va stocker les données sous forme d'*objets*. Un *objet* est composé d'une clé d'identification appelée *token* et d'un ensemble de *données*.

Pour savoir quel noeud doit stocker quelle donnée, on utilise une méthode de *partitionnement*. Cette méthode se base sur les *tokens*. Chaque noeud a un token qui lui est attribué. Un noeud prend en charge des objets dont le token est compris entre celui que le noeud possède et celui qui est le plus grand dans ses prédécesseurs (voir la figure 4). Ainsi dans cet exemple, le noeud 2 a le token 25 qui lui est attribué. Il s'occupe donc des objets dont le token est compris entre 25 et 0 (qui est le token le plus grand dans ses prédécesseurs). On parle alors de l'*intervalle* de tokens dont s'occupe le noeud.

En positionnant les noeuds suivant leur token, on obtient alors une forme d'anneau (ou de *ring*).

Afin de garantir une meilleure disponibilité, chaque objet possède des copies, appelées *réplicas*, disposées sur d'autres noeuds que le noeud initial (le noeud qui s'occupe du token de cet objet). La méthode pour choisir l'emplacement des copies d'un objet est variable. C'est ce que l'on appelle la *stratégie de réplication*.

1.3 Gestion des requêtes dans la base de données distribuée

1.3.1 Requêtes de lecture

Il est possible de réaliser des requêtes de lecture sur un objet, ce qui consiste à vouloir récupérer une donnée contenue dans un objet. Pour expliquer le cheminement d'une requête de lecture dans la base de données, nous allons prendre un exemple (voir la figure 6).

Un client réalise une requête de lecture R. Il envoie la requête à n'importe quel noeud du réseau. On appelle alors ce noeud le noeud *coordinateur* pour cette requête. Ce noeud ne contient pas forcément l'objet de la requête, mais il va faire la liaison entre le réseau et le client.

Le noeud coordinateur va avoir cette requête dans une file d'attente dédiée aux requêtes des clients. Il les traite les unes à la suite des autres. Lorsque le noeud commence à traiter cette requête, il va d'abord identifier les noeuds responsables de l'objet de la requête. Cela inclut le noeud possédant l'objet *original* (dont le token est géré par ce noeud) ainsi que les noeuds possédant un réplica. Cette étape exige une connaissance complète du réseau sur chaque noeud et une connaissance de la stratégie de réplication mise en place.

Dès que les noeuds sont identifiés, le noeud coordinateur leur envoie un message pour traiter la requête de lecture (les flèches rouges sur le schéma entre le noeud coordinateur et les autres noeuds). Ce message est mis dans la file d'attente des requêtes de lecture de ces noeuds.

A un moment, l'un des noeuds qui possède cette requête dans sa file d'attente va la défiler et la traiter. Ce noeud s'*affecte* la requête. Il avertit les autres noeuds possédant cette même requête dans leur file d'attente (c.à.d tous les autres noeuds possédant une copie de l'objet de la requête) qu'ils n'auront pas besoin de la traiter, et qu'ils peuvent

la supprimer de leur file d'attente (les flèches oranges sur le schéma). Le noeud qui s'est affecté la requête la traite et renvoie le résultat au noeud coordinateur, qui peut transmettre le résultat obtenu au client (les flèches vertes sur le schéma).

1.3.2 Requêtes d'écriture

Il est possible de réaliser des requêtes d'écriture d'un objet, ce qui consiste à stocker des données dans la base de données, sous forme d'objet. Pour expliquer le cheminement d'une requête d'écriture dans la base de données, nous allons prendre un exemple (voir la figure 7). Le cheminement est plus simple que pour une requête de lecture car il n'y a pas le mécanisme d'affectation.

Un client réalise une requête d'écriture R. Il envoie la requête à n'importe quel noeud du réseau. On appelle alors ce noeud le noeud *coordinateur* pour cette requête. Ce noeud n'est pas forcément celui qui va stocker les données, mais il va faire la liaison entre le réseau et le client.

Le noeud coordinateur va avoir cette requête dans une file d'attente dédiée aux requêtes des clients. Il les traite les unes à la suite des autres. Lorsque le noeud commence à traiter cette requête, il va d'abord identifier les noeuds responsables de l'objet de la requête. Cela inclut le noeud qui se charge de l'objet *original* (dont le token est géré par ce noeud) ainsi que les noeuds devant posséder un réplica. Cette étape exige une connaissance complète du réseau sur chaque noeud et une connaissance de la stratégie de réplication mise en place.

Dès que les noeuds sont identifiés, le noeud coordinateur leur envoie un message à tous pour traiter la requête d'écriture (les flèches rouges sur le schéma entre le noeud coordinateur et les autres noeuds). Ce message est mis dans la file d'attente des requêtes d'écriture de ces noeuds.

Tous les noeuds recevant le message vont alors stocker les données envoyées par la requête. Le noeud coordinateur peut demander un certain nombre de message de retour pour s'assurer que les requêtes d'écritures se sont bien déroulées. Dans l'exemple, le noeud coordinateur demande 1 retour. L'un des messages envoyés aux noeuds demandera donc un message de retour pour confirmer que l'écriture s'est bien passée (la flèche verte entre les noeuds sur le schéma). Dès que le noeud coordinateur reçoit le message, il indique au client que sa requête s'est terminée et bien passée.

1.4 Stockage des données

Chaque base de données possède sa propre manière de stocker les données dans un espace de stockage. Pour le projet, la méthode de stockage n'est pas un problème sur lequel nous allons travailler. La seule contrainte imposée pour la base de données est qu'elle stocke les données sous la forme d'objet. C'est à dire qu'un objet est identifiable par son token, une clé d'identification générée le plus souvent par une fonction de hachage.

Le token est généré par la base de donnée à partir de la *clé primaire* d'une table. Une clé primaire est, comme le token, une donnée permettant d'identifier un objet. Sauf que que la clé primaire est une donnée choisit par le client. Elle peut être un entier, une chaîne de caractères, toutes les représentations possibles d'une donnée au sein de la base de données (voir la figure 8).

1.5 Protocoles de réaffectation des requêtes de lecture

Lorsqu'une requête de lecture est envoyée par un client, on a vu précédemment que cette requête était transmise à tous les noeuds possédant une copie de l'objet à lire. Si un nombre important de requêtes de lecture arrivent en même temps, les files d'attente dans les noeuds pour les requêtes de lecture vont commencer à se remplir plus vite que les requêtes ne sont traitées. Les charges des files d'attente ne seront pas forcément uniformes entre les noeuds, certains pouvant avoir plus de requêtes à traiter que d'autre.

C'est pourquoi on met en place un système de *réaffectation* des requêtes de lecture, afin de rééquilibrer la charge des noeuds. La réaffectation consiste, de manière périodique, à stopper le traitement des requêtes de lecture et enclencher un processus permettant de décider de l'affectation des requêtes suivant l'état actuel du réseau.

Les algorithmes de réaffectation à implémenter, **SLVO** et **AverageDegree**, ont un comportement similaire qui se base sur la connaissance des charges de chaque noeud du réseau. L'algorithme consiste à comparer, pour tous les noeuds, sa propre charge par rapport à une certaine valeur.

Pour SLVO, la valeur est la charge minimale sur le réseau. Pour AverageDegree, la valeur est la charge moyenne sur le réseau.

Si la valeur est inférieure ou égale (strictement égale dans le cas de SLVO), alors le noeud s'affecte toutes les requêtes de sa file d'attente et avertit tous les autres noeuds. Les noeuds possédant les requêtes qui ont été affectées les suppriment de leur file d'attente, modifiant ainsi leur charge. L'opération est renouvelée jusqu'à ce que tous les noeuds se sont attribués leurs requêtes (voir la figure 9 pour un exemple avec SLVO).

1.6 Gestion de la popularité des objets

Pour mieux équilibrer la charge du réseau, nous nous intéressons à la *popularité* des objets. En effet, plus un objet va recevoir de requêtes, plus il sera populaire et occasionnera une grande charge pour les noeuds qui s'en occupent. Afin de répartir cette charge, il faudra alors augmenter ou diminuer le nombre de répliques. Si un objet est populaire, il suffira de créer de nouveaux répliques, ce qui permettra d'envoyer une partie de la charge sur d'autres noeuds. A l'inverse, si un objet n'est pas populaire, diminuer le nombre de copies fera gagner de l'espace mémoire et du temps (quand on a besoin de contacter tous les noeuds qui gèrent un objet, le nombre de noeuds joue sur le temps nécessaire à réaliser l'action...).

Il y a plusieurs méthodes pour calculer la popularité des objets durant un intervalle de temps T défini par l'utilisateur :

- La première consiste à ce que chaque noeud possède un vecteur de la taille du nombre d'objets dont il a la gestion. A chaque nouvelle requête, la case de l'objet est incrémentée. Au début de chaque période T , les noeuds envoient la popularité aux autres noeuds et décident du nombre de copies à faire.
- La seconde méthode est une variante visant à réduire la taille du vecteur d'objets et est défini par le Space-Saving Algorithm [ADA05].

1 schéma : (je vois pas quoi faire non plus là)

1.7 Gestion des copies d'un objet

Explication de la méthode de fonction de hash pour chaque index de copie. Mettre en comparaison avec une méthode connu de placement de réplica (mettre sur le noeud suivant). Distribution plus homogène des copies de cette manière.

1 schéma : partitionnement des copies d'un objet avec plusieurs fonctions de hachages

1.8 Visualisation des statistiques de fonctionnement de la base de données distribuée

Le but est de visualiser les statistiques de fonctionnement de la base de données pour permettre une comparaison de l'efficacité des algorithmes d'équilibrage de charge.

On souhaite récupérer :

- la charge effective de chaque noeud ou taille de la file d'attente des requêtes de lecture.
- une représentation de la file d'attente des requêtes de lecture
- la popularité de chaque objet
- la requête en cours de traitement

On enregistre les statistiques de fonctionnement de la base de données distribuée dans des fichiers. Un outil de visualisation traite ces fichiers et affiche ensuite les statistiques.

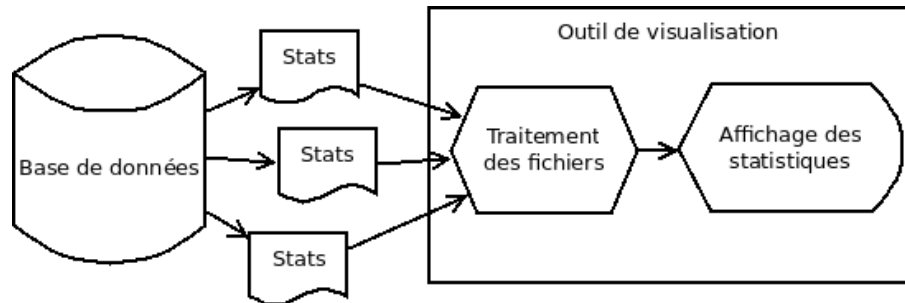


FIGURE 2 – Processus pour la visualisation des statistiques

2 Définition du projet

2.1 Contexte

Définition Un *environnement distribué* est constitué de plusieurs machines (ordinateurs généralement), appelées *noeuds*, sur lesquelles sont stockées des données et pouvant traiter des requêtes. Chaque noeud possède des informations locales propres à son fonctionnement (exemple : une liste des noeuds voisins).

Définition Une *base de données* est une entité permettant de stocker des données afin d'en faciliter l'exploitation (ajout, mise à jour, recherche de données).

Définition Une *charge* est associée à un noeud et désigne le nombre de requêtes restantes que le noeud doit traiter à l'instant T .

Définition La *réplication* d'une donnée est une action qui réalise des copies de cette donnée sur d'autres noeuds.

Définition Un *réseau* est un ensemble de noeuds qui sont reliés entre eux (par exemple par Internet) et qui communiquent ensemble.

L'expansion, au cours des deux dernières décennies, des réseaux et notamment d'Internet a engendré une importante création de données, massives par leur nombre et leur taille. Stocker ces informations sur un seul point de stockage (ordinateur par exemple) n'est bien sûr plus envisageable, que ce soit pour des raisons techniques ou pour des raisons de sûreté (pannes potentielles par exemple). Pour cela des systèmes de stockages dits distribués sont utilisés en pratique afin des les répartir sur différentes unités de stockages.

Pour répartir toutes ces données, notre client a développé de nouveaux algorithmes d'équilibrage de charge basés sur la réplication qu'il souhaite tester dans un environnement distribué.

2.2 Finalité

Nous devons développer une solution logicielle permettant de tester ces nouveaux algorithmes d'équilibrage de charge et de réplication proposés par le client dans un environnement distribué.

Définition Une *charge minimum* d'un environnement distribué est la plus petite charge trouvée sur un noeud parmi l'ensemble des noeuds. La *charge moyenne*, est la moyenne des charges de l'ensemble des noeuds.

Il s'agira d'implémenter les algorithmes développés par le client. On distingue les algorithmes d'affectation de requête :

- **SLVO** Si la charge du noeud est inférieure ou égale à la charge minimum, il s'affecte toutes les requêtes en attente et en avertit les autres noeuds.
- **AverageDegree** Si la charge du noeud est inférieure ou égale à la charge moyenne, il s'affecte toutes les requêtes en attente et en avertit les autres noeuds.

Définition La *popularité* d'un objet est le nombre de requêtes que va recevoir un objet durant un intervalle de temps T défini par l'utilisateur.

Ainsi que l'algorithme de gestion de copie, permettant d'établir le nombre de répliques d'un objet en fonction de sa popularité.

Pour comparer l'efficacité de ces algorithmes, on peut visualiser l'état du réseau *en temps réel*.

2.3 Hypothèses

Nous évoluerons dans un environnement distribué constitué de n noeuds de stockage dans lequel on souhaite stocker m objets. C'est un réseau statique : on ne peut pas ajouter ou supprimer de noeuds après création du réseau.

Définition Un *message* est un envoi d'information d'un noeud vers un autre noeud pour mettre à jour ses données locales ou effectuer des actions particulières (autre que des requêtes, comme par exemple, "donne à tel noeud ta charge actuelle").

Données locales d'un noeud Un noeud contient les données locales suivantes :

- la charge de tous les noeuds du réseau
- la popularité de chaque objet stocké sur ce noeud
- une file d'attente de message à traiter
- la requête en cours de traitement

Requêtes Nous supposons que les requêtes seront effectuées en un temps fixe.

3 Ordonnement des besoins

Nous avons dégagé une liste de besoins fonctionnels et non-fonctionnels. Pour mieux les comparer, nous les avons ordonnés en fonction de leur priorité et de leur criticité.

3.1 Priorité

La priorité est un indicateur de l'ordre dans lequel nous devons implémenter les fonctionnalités afin de satisfaire les besoins du client.

Valeur	Signification	Description
1	Priorité haute	A implémenter en premier.
2	Priorité moyenne	A implémenter.
3	Priorité faible	A implémenter (en fonction du temps restant).

3.2 Criticité

Le niveau de criticité d'un besoin est un indicateur de l'impact qu'aura la non-implémentation de ce besoin sur le bon fonctionnement de l'application.

Valeur	Signification	Description
1	Criticité extrême	L'application ne sera pas utilisable par le client.
2	Criticité haute	L'application est utilisable par le client. En revanche, certaines fonctionnalités de l'application ne seront pas utilisables.
3	Criticité moyenne	L'application est utilisable par le client. En revanche, certaines fonctionnalités de l'application n'amèneront pas au résultat attendu.
4	Criticité faible	L'application peut fonctionner sans l'ajout de ces fonctionnalités

4 Besoins fonctionnels

4.1 Gestion d'un réseau

Un réseau est un ensemble de noeuds qui sont reliés entre eux (par exemple par Internet) et qui communiquent ensemble afin de traiter toutes les requêtes reçues. On peut rassembler des noeuds pour former un *data center*, et rassembler des data center pour former un *cluster* (voir la figure 3).

4.1.1 Gestion des noeuds

Création d'un noeud (*Priorité* : 1, *criticité* : 1) Il est possible de séparer ce besoin en plusieurs sous-besoins :

- créer un noeud dans l'environnement
- initialiser les données locales d'un noeud

Mise à jour des données locales (*Priorité* : 1, *criticité* : 1) Afin de connaître l'état du réseau de manière précise, les données locales doivent être mise à jour à chaque action, c'est à dire lors du traitement d'un message dans la file d'attente.

Communication des données locales (*Priorité* : 1, *criticité* : 1) Un noeud doit être capable de communiquer ses données locales à d'autres noeuds du réseau.

Récupération de l'état du réseau (*Priorité* : 1, *criticité* : 1) L'application doit permettre la description de l'état du réseau. On souhaite connaître :

- le nombre de requêtes en attente
- la popularité des objets

4.1.2 Réplication d'un objet

Définition Une *fonction de hachage* est une fonction mathématique déterministe (c'est à dire, si on lui donne la même entrée, elle renvoie la même sortie). Nous définissons ses entrées et sorties dans le paragraphe suivant.

Définition Un *token* permet comme une étiquette sur un produit, de désigner une donnée.

Il s'agit de copier un objet sur un autre noeud. Il est possible de définir le nombre de copies d'un objet au sein d'un ensemble de noeuds, appelé *data center*. Pour savoir quel noeud stocke l'objet, on utilise une fonction de hachage dans laquelle on fait passer la clé de l'objet (la clé de l'objet est une donnée permettant d'identifier un objet de manière unique). On obtient ainsi un *token*.

Tous les noeuds possèdent un intervalle (ou ensemble) de tokens dont ils sont responsables. On regarde le token de l'objet pour savoir quel noeud va le prendre en charge (voir la figure 4).

Une stratégie de réplication est la méthode qui permet de placer les copies d'un objet dans un data center. La stratégie consiste à utiliser une fonction de hachage différente pour chaque copie (voir la figure 5). Le numéro de la copie définit la fonction à utiliser. Ainsi, sur le schéma, la deuxième copie de tous les objets utilisera la fonction de hachage *Hash2* pour obtenir un token et placer la copie.

Définition des fonctions de hachage (*Priorité : 1, criticité : 3*)

Mise en place de la stratégie de réplication (*Priorité : 1, criticité : 3*)

4.1.3 Popularité d'un objet

Les algorithmes à implémenter nécessitent de connaître la popularité d'un objet dans le réseau. La popularité d'un objet est défini par le nombre de requêtes sur cet objet. Plus le nombre de requêtes est grand, plus l'objet est populaire.

Calcul de la popularité (*Priorité : 1, criticité : 1*)

Stockage de la popularité (*Priorité : 1, criticité : 1*) Chaque noeud stocke la popularité des objets qu'il contient.

Communication de la popularité (*Priorité : 1, criticité : 1*) Un noeud stockant des copies d'un objet doit communiquer la popularité de ces derniers au noeud possédant l'objet original.

4.2 Protocoles d'affectation

Une *affectation* consiste, pour un noeud, à effectuer le traitement d'une requête. Les requêtes peuvent arriver sur n'importe quel noeud. On dit alors que ce noeud devient le noeud *coordinateur* pour cette requête. Il transmet la requête aux noeuds possédant l'objet de la requête (voir la figure 6).

Les noeuds possédant l'objet mettent la requête dans leur file d'attente. Dès qu'un noeud aura à traiter cette requête, il communique aux autres noeuds possédant l'objet qu'il se charge de la traiter. Les noeuds suppriment la requête de leur file d'attente. Le noeud qui prend en charge la requête la traite et donne le résultat de la requête au noeud coordinateur, qui peut ainsi renvoyer le résultat.

Communication d'un message d'ajout d'une requête dans la file d'attente (*Priorité : 1, criticité : 1*)

Communication d'un message de suppression d'une requête de la file d'attente (*Priorité : 1, criticité : 1*)

Ajout d'une requête dans la file d'attente (*Priorité : 1, criticité : 1*)

Suppression d'une requête de la file d'attente (*Priorité : 1, criticité : 1*)

Des protocoles (équivalent algorithmes) d'affectation plus spécifiques peuvent être implémentés. Ils utilisent les données locales du noeud pour décider de l'affectation des requêtes.

Implémentation de l'algorithme SLVO (*Priorité : 1, criticité : 1*)

Implémentation de l'algorithme AverageDegree (*Priorité : 1, criticité : 1*)

4.3 Requêtes

4.3.1 Générations de requêtes

Pour tester la validité des algorithmes, l'application devra posséder une fonction de génération de requêtes. Si l'utilisateur ne détient pas de suites de requêtes prêtes, il pourra demander à l'application d'en créer pour lui. L'application, ne connaissant pas la nature des données, ne pourra qu'effectuer un nombre restreint de requêtes différentes. Elle pourra par exemple, compter le nombre de données sauvegardées, chercher si une donnée existe réellement, mais ne pourra pas en modifier une.

4.3.2 Importation d'un jeu de requête

Définition La notion d'efficacité est laissée à l'appréciation du client. Une brève approche serait de comparer les temps d'exécution.

Pour comparer l'efficacité des algorithmes, il doit être possible d'envoyer sur le réseau une même suite de requêtes : un jeu de requêtes.

Importation (*Priorité : 1, criticité : 2*) L'application doit pouvoir lire un fichier contenant une suite de requêtes et envoyer ces requêtes sur le réseau.

4.4 Visualisation des données

Afin de suivre l'évolution des charges de chaque noeud lors de l'exécution des algorithmes, on enregistre les données locales de chaque noeud à chaque modifications de celles-ci.

4.4.1 Enregistrement des données

Ecriture dans un fichier (*Priorité : 1, criticité : 2*) Lorsque les données locales d'un noeud sont modifiées, on les enregistre dans un fichier. L'écriture est de la forme `itération de l'algorithme; identifiant du noeud; charge du noeud;`

4.4.2 Affichage des données

Définition Un *graphe* est un ensemble de points appelés *sommets*, dont certaines paires sont directement reliées par un (ou plusieurs) lien(s) appelé(s) *arêtes* [com15].

Noeuds (*Priorité : 3, criticité : 2*) L'application doit permettre la représentation de chaque noeud par un sommet.

Analyse syntaxique (*Priorité : 2, criticité : 1*) Lors de l'exécution d'un algorithme, la charge de chaque noeud est enregistrée dans un fichier. Un analyseur syntaxique (un programme qui possède des règles et qui agit sur un fichier donné en entrée selon celles-ci) découpe chaque ligne du fichier pour récupérer le moment auquel a été enregistrée l'information (`itération de l'algorithme`), le noeud concerné (`identifiant du noeud`) et la charge de ce noeud à ce moment (`charge du noeud`).

Charge des noeuds (*Priorité : 3, criticité : 3*) A chaque sommet est associée une valeur correspondant à la charge de ce noeud. Ces données sont récupérées grâce à l'analyseur syntaxique.

Film de l'exécution (*Priorité : 3, criticité : 3*) Cela consiste à afficher la charge des noeuds dans l'ordre chronologique, c'est à dire dans l'ordre des itérations croissant.

5 Besoins non fonctionnels

5.1 Cassandra

Cassandra est une base de données distribuée. Nous créons notre environnement distribué à partir de la dernière version stable de Cassandra.

Le choix de cette solution nous a été fortement recommandé par le client. En effet, celui-ci dispose de connaissances sur cette application et pourra donc plus facilement intervenir s’il souhaite faire évoluer le projet en implémentant par exemple de nouveaux algorithmes.

5.2 Maintenabilité du projet

L’envergure du projet fait qu’il est possible que d’autres personnes travaillent sur la finalité de ce projet, peu importe son état d’avancement. Afin de faciliter la compréhension, nous avons défini quelques normes pour que le projet puisse être repris :

- documentation dans le code source suivant la norme du langage utilisé ;
- document externe spécifiant les fichiers modifiés par rapport au code source original ;
- guide d’installation pour utiliser le projet et pour modifier le projet.

5.3 Gestion d’un réseau

5.3.1 Communication entre noeuds

Algorithme Le calcul de la popularité nécessite l’implémentation de l’algorithme d’approximation Space-Saving Algorithm [ADA05].

Pour connaître l’état du réseau, il faut regrouper les données locales des noeuds. Nous cherchons donc à récupérer ces données en un temps raisonnable ($O(\log(n))$ pour n noeuds).

Pour cela, nous nous appuyons sur le protocole **Gossip** [Fou14a]. Périodiquement, chaque noeud choisit n noeuds aléatoirement dont un noeud *seed* [Fou14b], noeud en mesure d’avoir une connaissance globale du système, et il communique à ces noeuds ses statistiques (valeur de sa charge, objets les plus populaires...).

Ainsi, la connaissance globale du système se fait, dans la théorie, en $O(\log(n))$.

5.3.2 Taille des données

La taille de chaque donnée est laissée à l’appréciation de l’équipe. Néanmoins, celle-ci doit être suffisamment importante, afin de permettre de créer des requêtes qui “stressent” le système pour avoir des résultats cohérents (sur la base de l’hypothèse : chaque requête prend un même temps à être traitée).

5.4 Protocole de test

La conformité des algorithmes implémentés est assurée par un protocole de test suivant la démarche :

- Définir un réseau R , un ensemble d’objets O et un ensemble de requêtes Q
- Faire tourner l’algorithme à la main avec R , O et Q
- Stocker l’état final du réseau

- Faire valider ce processus par le client
- Exécuter l'algorithme sur ordinateur avec R , O et Q
- Vérifier les résultats constatés avec les résultats attendus

S'il y a une différence entre les deux résultats, une vérification par le client peut être envisagée dans le cas de résultats *presque* similaires. La notion de similitude est laissée à l'appréciation de l'équipe en charge du projet, lors de la vérification.

5.5 Visualisation des données

5.5.1 Etat du réseau

La vue permet de montrer l'état du réseau.

Le réseau est représenté par un graphe, les machines par des noeuds. Pour chaque machine, les données affichées sont la charge ainsi que le contenu de la file d'attente.

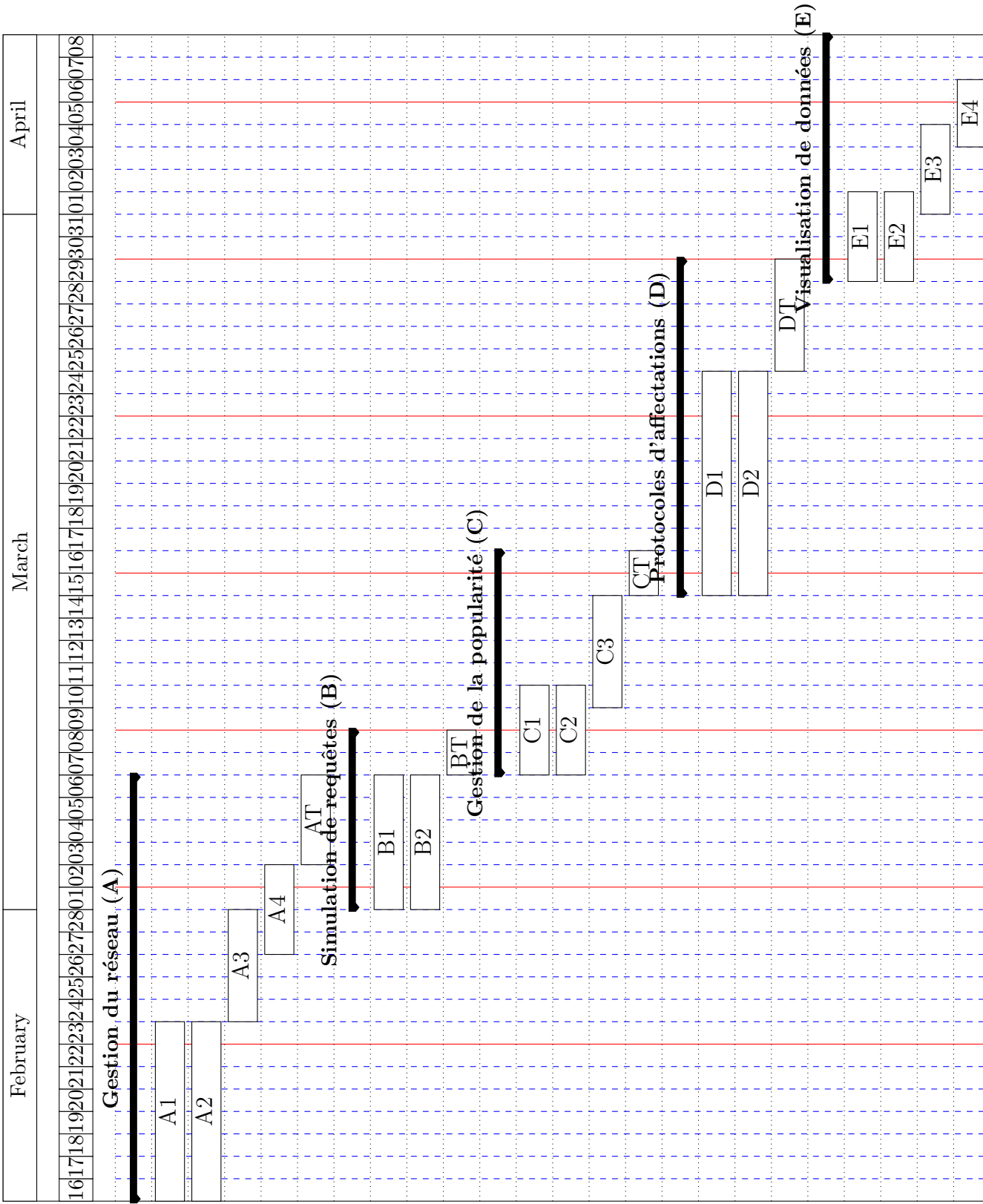
5.5.2 Actualisation de la vue

L'état du réseau doit être visible en temps réel.

La vue peut donc être actualisée toutes les 0.5 secondes. Un délai plus faible risquerait de ralentir le système, étant donné que l'obtention des données nécessaires à la visualisation se fait sur la même base de données que celle qui est testée.

6 Répartitions des tâches

6.1 Diagramme de Gantt



6.2 Affectation des tâches

Fct	Description	Développeur(s)	Commentaire
A1	Création des noeuds		
A2	Données locales des noeuds		Initialisation et implémentation
A3	Communication des données locales entre noeuds		
A4	Gestion des replicas		
AT	Tests groupe A		Vérification, tests, mémoire
B1	Générateur de requêtes		A détailler
B2	Importateur de jeu de requêtes		A détailler
BT	Tests groupe B		Vérification, tests, mémoire
C1	Popularité objet sur noeud		
C2	Space-Saving Algorithm		
C3	Popularité d'un objet dans le réseau		
CT	Tests groupe C		Vérification, tests, mémoire
D1	Implémentation SLVO		
D2	Implémentation AverageDegree		
DT	Tests groupe D		Avec client
E1	Prise en main Tulip		
E2	Ecriture des données dans un fichier		(+Analyseur syntaxique)
E2	Représentation réseau		
E3	Représentation données		
T	Tests finaux		Vérification, tests, mémoire

Remarque Il s'agit d'une première version de notre GANTT. Nous n'avons pas encore défini l'affectation des tâches aux développeurs.

7 Livrables

7.1 Livrables intermédiaires

Un livrable intermédiaire est une ébauche de l'application. C'est à dire que seulement quelques fonctionnalités sont implémentées.

Il n'a pas encore été décidé de remettre un ou plusieurs livrables intermédiaires au client.

7.2 Livrable final

Il devra être remis le 8 Avril 2015. Il comportera les besoins de priorité 1 et 2.

Références

- [ADA05] Metwally A, Agrawal D, and El Abbadi A. Efficient computation of frequent and top-k elements in data streams. 2005.
- [com15] Wikipedia community. Théorie des graphes - wikipédia. <http://fr.wikipedia.org/wiki/Th%C3%A9orie_des_graphes#D%C3%A9finition_de_graphe_et_vocabulaire>, 2015. [Accessed 5 February 2015].
- [Fou14a] The Apache Software Foundation. Architecturegossip - cassandra wiki. <<http://wiki.apache.org/cassandra/ArchitectureGossip>>, 2014. [Accessed 21 January 2015].
- [Fou14b] The Apache Software Foundation. Faq cassandra wiki. <<http://wiki.apache.org/cassandra/FAQ#seed>>, 2014. [Accessed 21 January 2015].

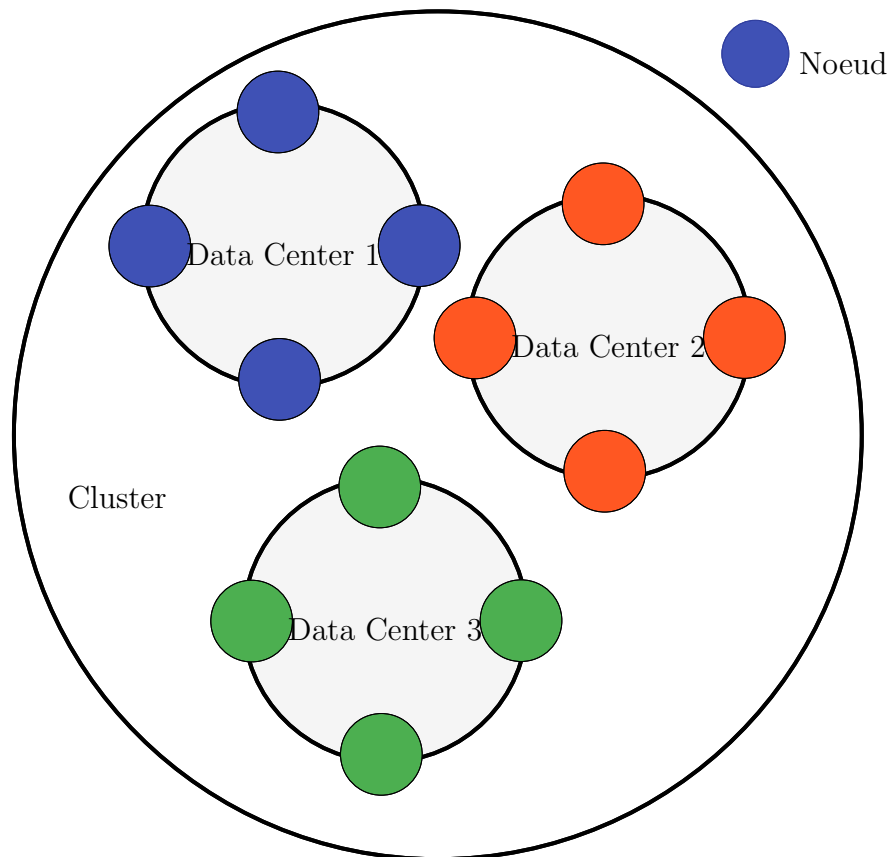


FIGURE 3 – Visualisation d’une base de données distribuée sous forme de cluster possédant trois data center

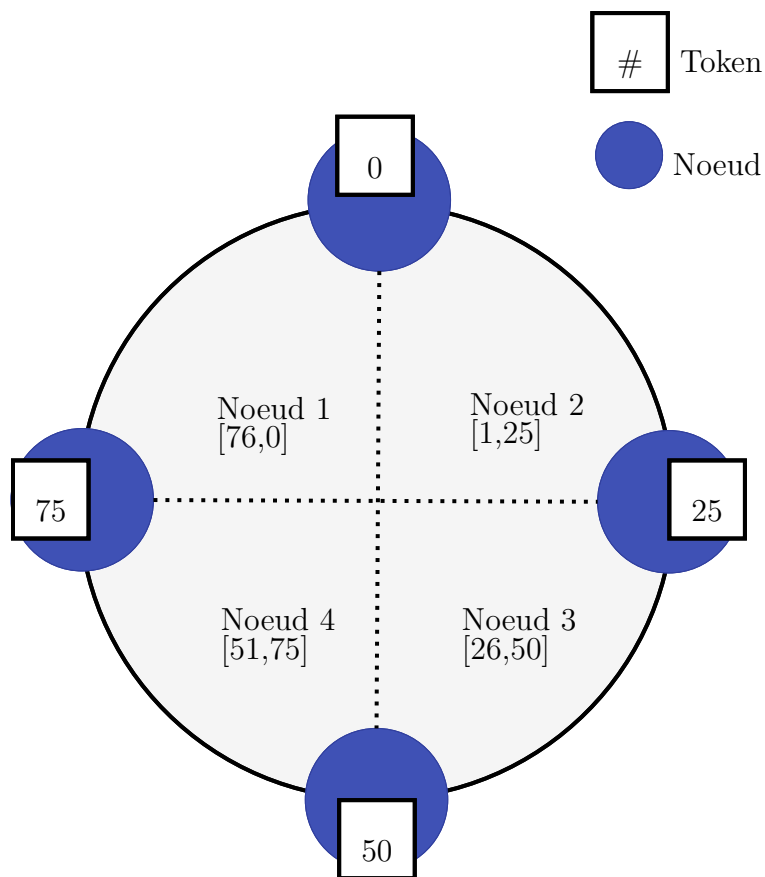


FIGURE 4 – Exemple de partitionnement des données dans une base de données distribuée

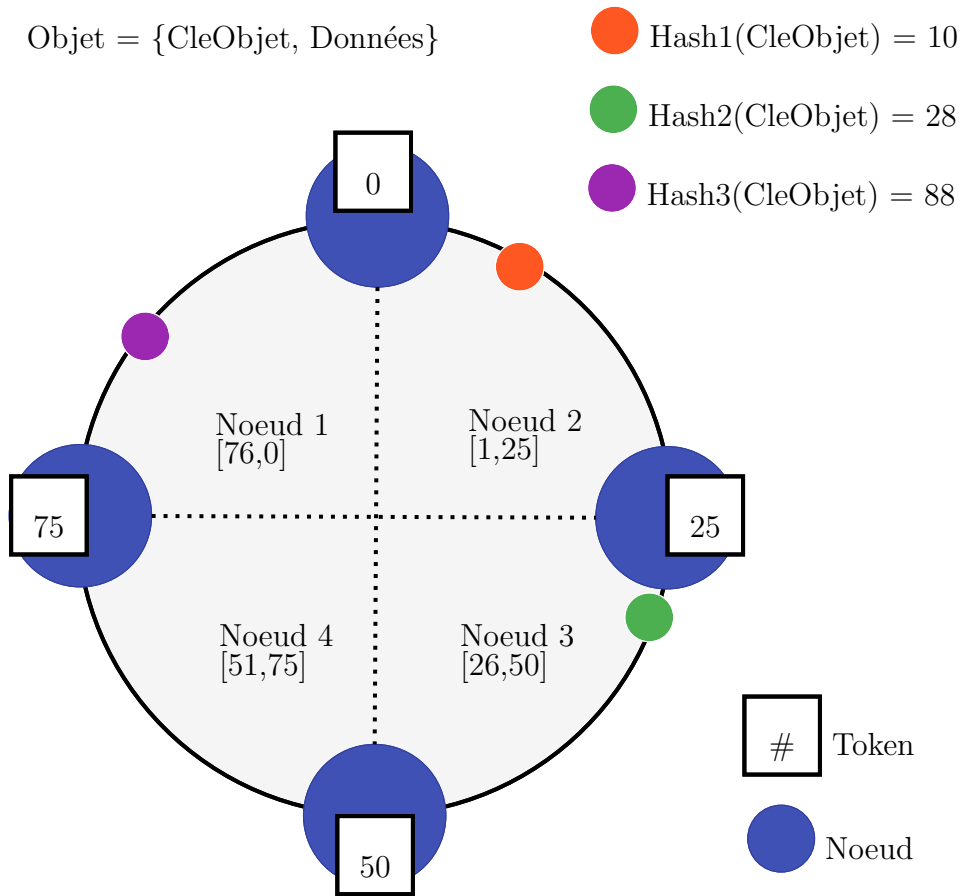


FIGURE 5 – Partitionnement des réplicas d'un objet avec une fonction de hachage pour chaque réplica

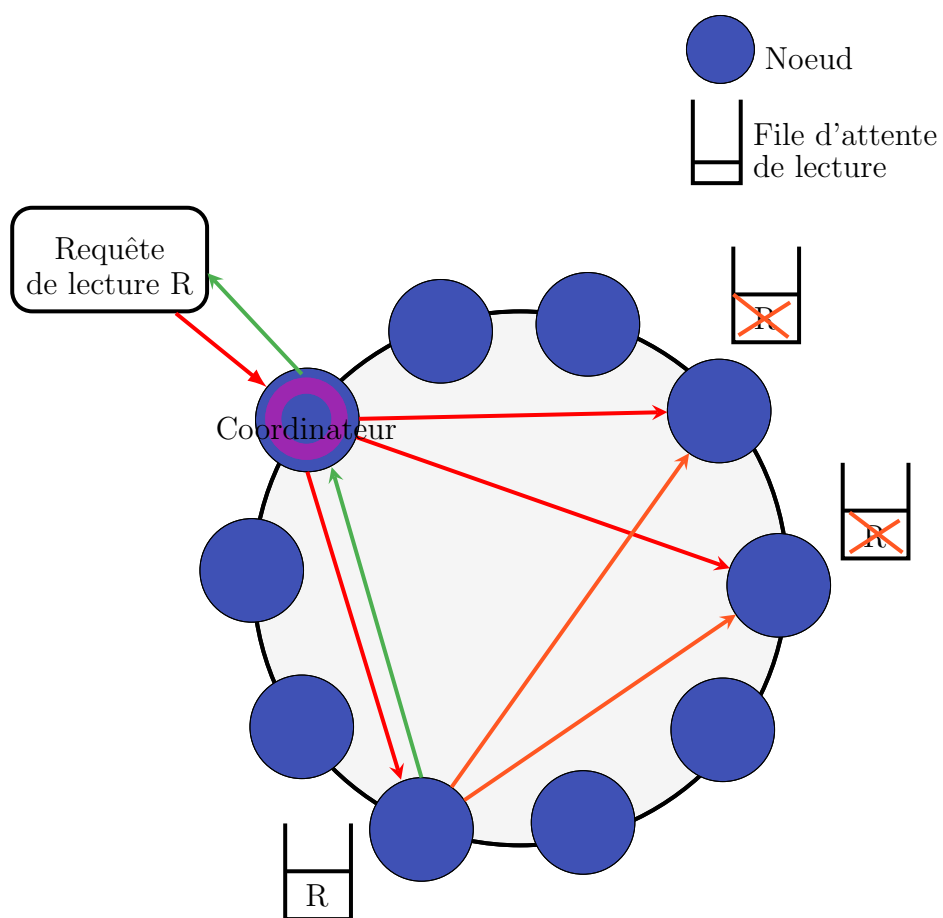


FIGURE 6 – Cheminement d’une requête de lecture dans une base de données distribuée avec la prise en charge de l’affectation (un seul noeud traite la requête)

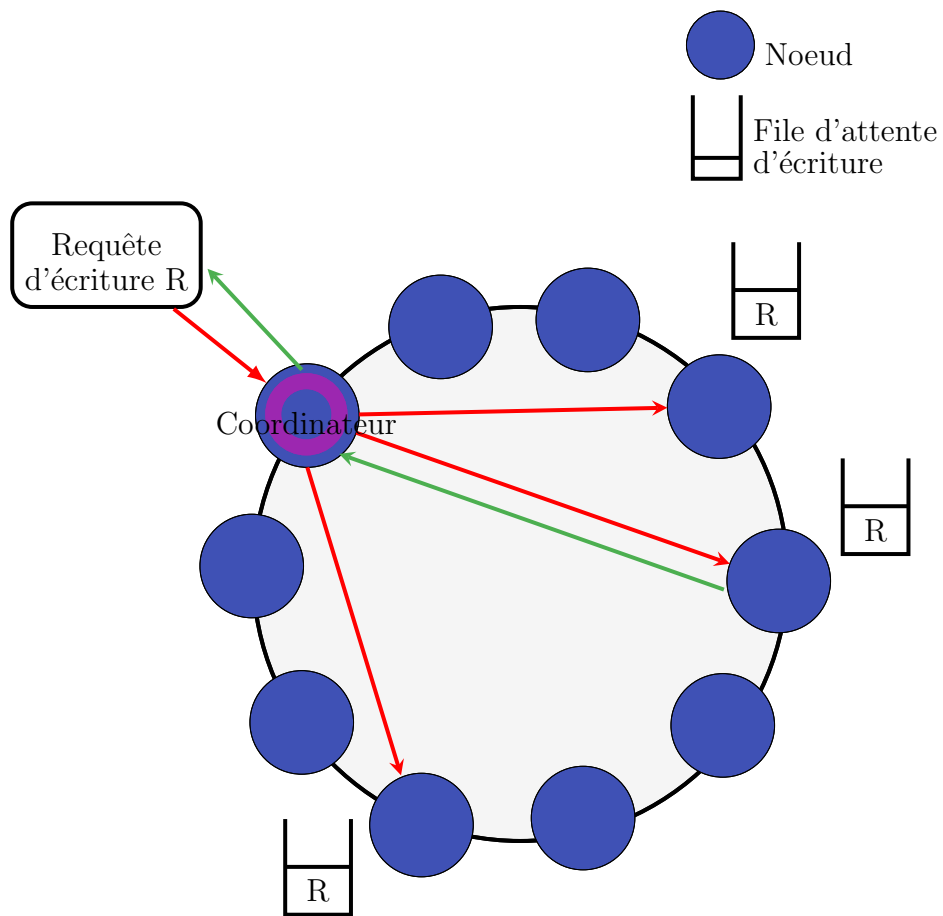


FIGURE 7 – Cheminement d’une requête d’écriture dans une base de données distribuée

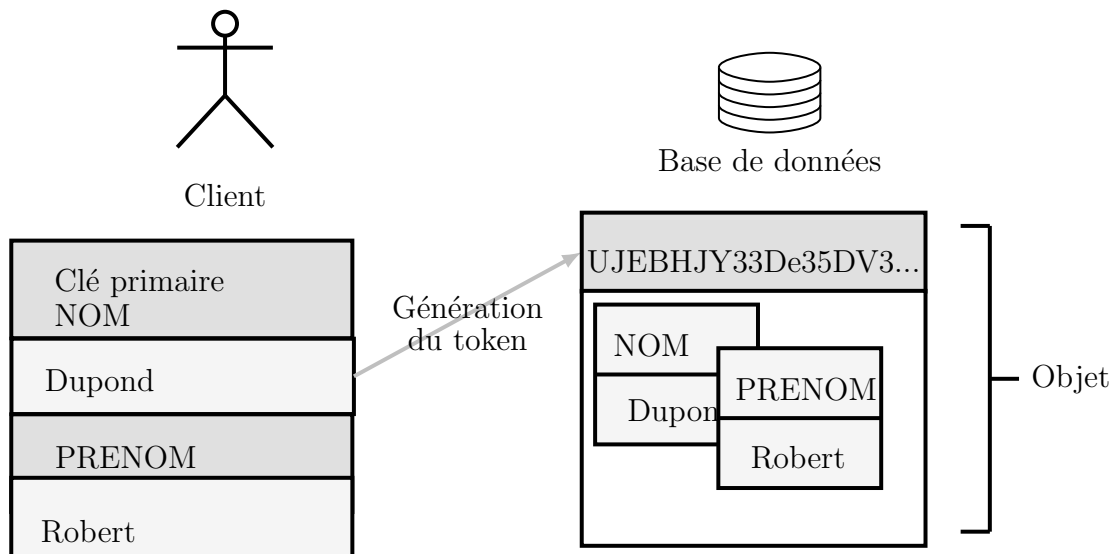


FIGURE 8 – Passage d’une représentation des données pour le client à une représentation pour la base de données

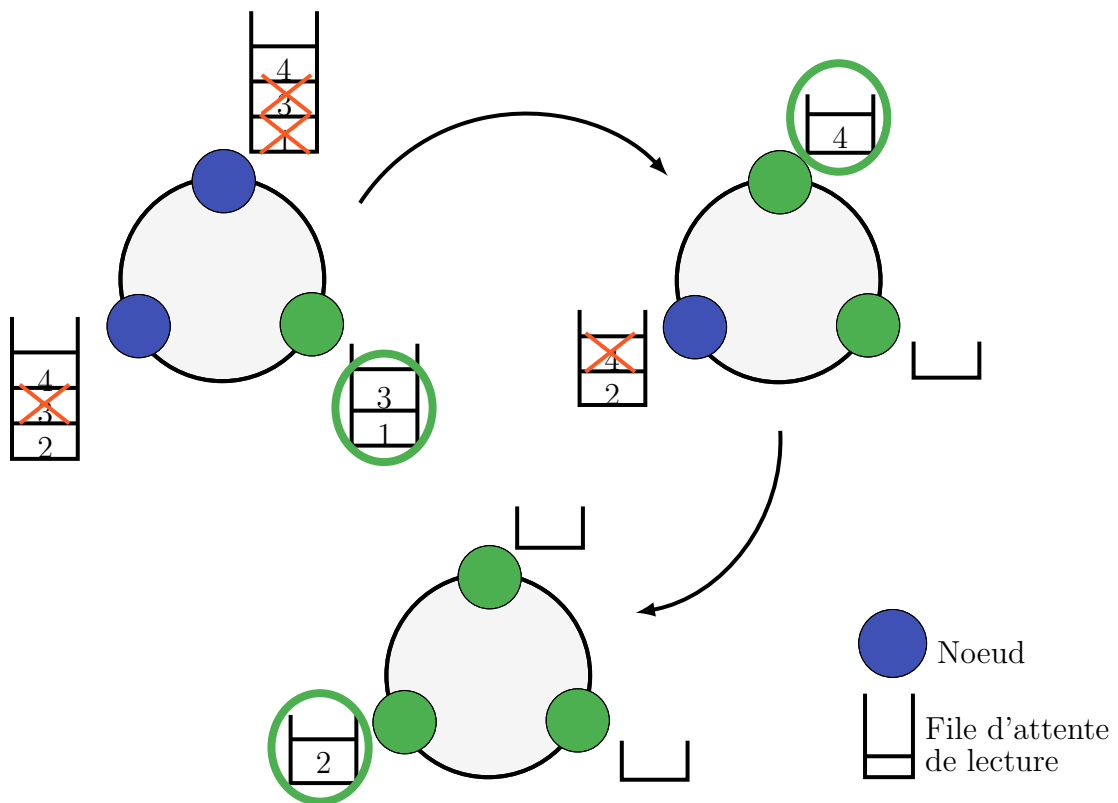


FIGURE 9 – Fonctionnement de l'algorithme de réaffectation des requêtes de lecture SLVO