



HAL
open science

Confidentialité et disponibilité des données entreposées dans les nuages

Kawthar Karkouda, Nouria Harbi, Jérôme Darmont, Gérald Gavin

► **To cite this version:**

Kawthar Karkouda, Nouria Harbi, Jérôme Darmont, Gérald Gavin. Confidentialité et disponibilité des données entreposées dans les nuages. 12ème Conférence Internationale Francophone sur l'Extraction et la Gestion de Connaissance (EGC 2012), 2012, Bordeaux, France. , 12ème Conférence Internationale Francophone sur l'Extraction et la Gestion de Connaissance (EGC 2012), 2012. hal-01448385

HAL Id: hal-01448385

<https://hal.science/hal-01448385>

Submitted on 27 Jan 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

CONFIDENTIALITE ET DISPONIBILITE DES DONNEES DANS LES NUAGES

Kawthar Karkouda, Nouria Harbi, Jérôme Darmont, Gérald Gavin

Laboratoire Eric, Université Lumière Lyon 2, 5 avenue Mendès France, Bât K, 69767 Bron Cedex

kawthar.karkouda@gmail.com, nouria.harbi@univ-lyon2.fr, jerome.darmont@univ-lyon2.fr, gerald.gavin@univ-lyon1.fr

1 Motivation :

Pour ce type de service l'utilisation de l'informatique dans les nuages est basée sur la confiance des fournisseurs

- **Existant** : Utilisation d'une architecture traditionnelle du Cloud reposant sur un seul fournisseur
- **Problème** : Menace de la confidentialité des données des clients (hébergées chez un seul prestataire externe qui risque de les exploiter)

2 Proposition : Partager chaque donnée à stocker chez plusieurs fournisseurs des nuages

Utilisation de l'algorithme de secret Sharing inspirée de (Danwei Chen et Yanjun He).

- **Idée** : Stocker le n-uplet chez plusieurs fournisseurs permettant ainsi de :
 - **Stocker au niveau de chaque fournisseur une partie de l'information non compréhensibles et non exploitables par un utilisateur malveillant en cas d'intrusion**
 - **Ne pas dépendre d'un seul fournisseur minimisant ainsi le risque de non disponibilité des données.**
- **Etapas** :
 - Chaque fournisseur des nuages possède une copie de l'architecture de l'entrepôt du client.
 - Chaque donnée de l'entreprise est partagée et stockée chez les différents fournisseurs de manière à la rendre inexploitable par chaque fournisseur car non significative.
 - Le nombre de fragments dépend du nombre de fournisseurs choisis par le client.
 - Pour la restauration d'une donnée, le client doit récupérer les fragments stockés chez les différents fournisseurs pour reconstituer la donnée initiale (figure 1).

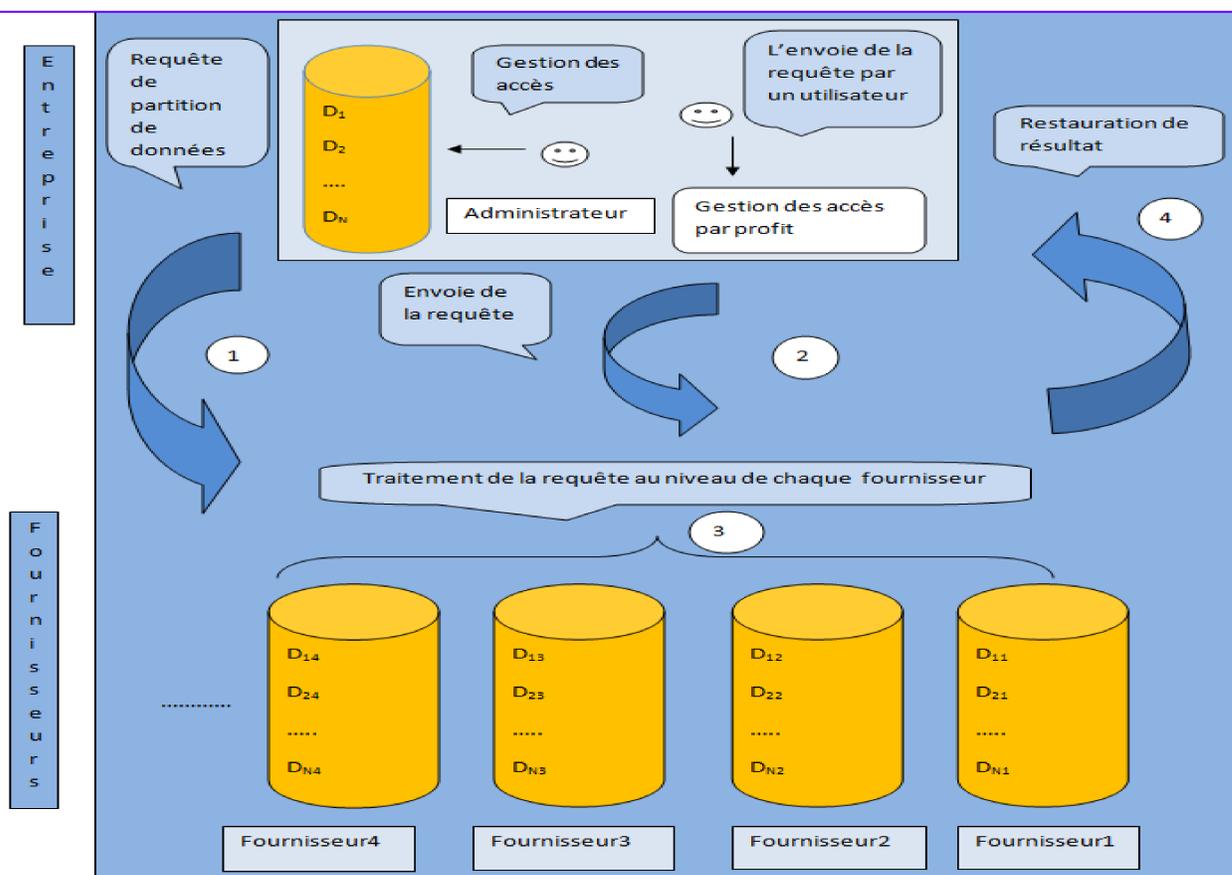


Figure 1 : Scénario d'un entrepôt de données partagé dans les nuages

3 Le niveau de sécurité attendu:

Notre solution assure trois niveaux de sécurité :

- **Capacité de restauration des données en cas de non disponibilité du service ou de la disparition d'un fournisseur** (Algorithme de secret sharing).
- **Sécurité des transactions entre le client et les fournisseurs (données partielles et inexploitables).**
- **Sécurité des données stockées chez les différents fournisseurs (chacun d'eux n'a qu'une partie d'une donnée non significative).**

4 Conclusion et perspectives:

Création d'un prototype assurant **les traitements nécessaires pour le partage et la restitution des données** et **intègre quelques opérateurs d'agrégation** (Max, Count, variance, moyenne) pour l'analyse OLAP.

Perspectives:

- Intégrer la cryptographie traditionnelle pour sécuriser le transfert des requêtes sur les réseaux.
- Implémenter d'autres opérateurs d'agrégation : Min....nécessaires pour les analyses OLAP.
- Appliquer une méthode de gestion des risques pour déterminer le coût de risque réel.