Vers des politiques de contrôle d'usage pour régir l'usage et le partage des données dans les environnements cross-cloud

Karim Laouchedi, Romain Laborde, Mohamed Ali Kandi, Abir Laraba, Abdelmalek Benzekri, Afonso Ferreira
IRIT, Université de Toulouse, CNRS
Toulouse, France
Karim.Laouchedi@irit.fr, Romain.Laborde@irit.fr, Mohamed-Ali.Kandi@irit.fr, Abir.Laraba@irit.fr,
Abdelmalek.Benzekri@irit.fr, Afonso.Ferreira@irit.fr

***Résumé*—L'utilisation croissante de solutions Cloud a suscité de nombreux problèmes concernant la protection des données mais aussi la trop forte dépendance des clients vis-à-vis de leur fournisseur de service Cloud. L'approche cross-cloud, en permettant l'utilisation transparente de plusieurs solutions Cloud, a apporté une réponse à certains de ces problèmes comme l'enfermement technologique (en anglais, vendor lock-in). Par contre, il a également exacerbé le problème de protection des données. En effet, les clients dépendent alors de plusieurs fournisseurs pour protéger leurs données et doivent prendre en compte l'hétérogénéité des solutions de sécurité mises en oeuvre par chaque fournisseur. Dans cet article, nous analysons précisément la problématique du contrôle de l'usage des données dans les environnements cross-cloud et présentons les pistes de recherche que nous étudions actuellement pour le résoudre. Nous y étudions plusieurs catégories de mécanismes de sécurité en présentant leurs limites en terme de maturité, mais aussi le niveau de dépendance et donc de confiance vis-à-vis des fournisseurs de services Cloud. Nous montrons ainsi qu'une solution de contrôle d'usage des données nécessite la combinaison de plusieurs catégories de mécanismes de sécurité.**

***Index Terms*—Contrôle d'usage des données, cross-cloud, multi-cloud.**

I. Introduction

Une stratégie récente utilisée par les entreprises, appelée cross-cloud ou encore multi-cloud, consiste à combiner les offres des plusieurs fournisseurs de cloud. Cela leur permet de répondre à différentes limitations telles que le respect des contraintes légales nécessitant des certifications spéciales de la part du fournisseur de cloud (e.g. SecNumCloud), l'enfermement technologique dans un unique fournisseur, la capacité de combiner différents services offerts par des fournisseurs, etc. Malheureusement, cette stratégie n'a pas encore de définition précise et universellement acceptée, mais est plutôt définie en fonction des besoins et des perspectives [1]. Cette difficulté se retrouve dans la définition de Barolli et al. [2] où le multi-cloud est défini comme étant : "les systèmes multi-cloud utilisent simplement plus d'un fournisseur de services cloud (CSP), bien qu'ils se déclinent en de nombreuses sous-catégories."

La distinction entre le terme multi-cloud et l'autre terme couramment utilisé, qui est le cross-cloud, est floue car les deux font l'intégration de systèmes cloud [2]. Cependant selon [2], une distinction est faite concernant la nature plus fluidifiée et cohésive du cross-cloud, contrairement au multi-cloud. Pour une analyse plus poussée sur les appellations relatives à l'utilisation de plusieurs fournisseurs de cloud, nous redirigeons le lecteur vers l'article de Ben Haj Salah et al. [1]. Dans la suite de cet article, nous utiliserons le terme cross-cloud pour désigner un environnement qui combine plusieurs fournisseurs de service cloud et garantit une expérience homogène.

Si la stratégie de cross-cloud offre de nombreux avantages, elle exacerbe en même temps les problèmes de gestion de la confiance dans les fournisseurs de service cloud et leur capacité à protéger les données. En effet, les récents scandales impliquant des entreprises offrant des services cloud qui ont été révélés remettent en cause l'hypothèse de confiance que l'on peut avoir dans ces environnements : infiltration dans l'infrastructure de fournisseur de cloud avec obtention de privilèges permettant d'attaquer ses clients [3], un ancien employé d'AWS qui vole des données des clients[1], Meta qui veut utiliser les données personnelles des utilisateurs pour entraîner ses IA[2], ou encore Microsoft qui surveille ses utilisateurs pour identifier des hackers[3]. Combiner les offres de plusieurs services cloud ne fait que complexifier ce problème de confiance.

Le deuxième problème introduit par le cross-cloud est lié à l'hétérogénéité des solutions de sécurité des services cloud [1]. La solution actuelle pour traiter ce problème est de passer par un courtier (en anglais, cloud broker) dont l'objectif est d'offrir une couche d'abstraction afin de faciliter les opérations avec les systèmes appelants [1]. Cette approche permet de répondre au problème d'hétérogénéité des solutions, mais place à nouveau le client/utilisateur dans une situation d'enfermement technologique (problème que le cross-cloud/multi-cloud est censé résoudre) vis-à-vis cette fois-ci du courtier. De plus, le niveau de confiance placé dans le courtier est très important

[1]https://www.theverge.com/2022/6/18/23173727/former-amazon-employee-convicted-over-2019-capital-one-hack-paige-thompson

[2]https ://www.forbes.com/sites/emmawoollacott/2024/06/10/meta-faces-legal-complaints-over-new-ai-training-data-plans/

[3]https://spectrum.ieee.org/online-privacy

car le client lui délègue la gestion de ses actifs sur toutes les plate-formes cloud utilisées. Il est donc nécessaire d'envisager d'autres approches plus décentralisées.

Notre objectif est de fournir un système permettant à la fois d'exprimer des politiques de contrôle d'usage et de les mettre en oeuvre dans un environnement cross-cloud. Dans cet article, nous présentons un travail préliminaire où nous décrivons et analysons les différentes options pouvant être suivies pour réaliser un tel système.

La suite de cet article est structurée ainsi. Dans la section 2, nous analysons le problème en analysant les relations de confiance entre les parties prenantes. La section 3 présente le travail en cours, c'est-à-dire l'étude et l'analyse des différentes approches pour définir et mettre en oeuvre les politiques de contrôle d'usage. Enfin, nous concluons dans la section 4.

II. Définition du problème

Afin de mieux appréhender le problème du partage de données en environnement cross-cloud, nous proposons de le représenter comme plusieurs relations de confiance [4] entre les parties prenantes (figure 1) : le propriétaire des données, l'environnement cross-cloud qui est en fait une relation composite avec chaque fournisseur impliqué et les sujets demandeurs (utilisateurs, entreprises, etc). Au moment où le propriétaire partage une donnée avec un demandeur, il doit considérer deux questions pour construire sa relation de confiance avec le demandeur. Tout d'abord, le propriétaire des données doit évaluer la *volonté* du demandeur à utiliser les données partagées en respectant l'accord avec le propriétaire ainsi que les réglementations en vigueur. Autrement dit, le demandeur ne doit pas outrepasser les droits qu'on lui a accordé. Le deuxième élément à considérer pour construire la relation de confiance est la *capacité* du demandeur à effectivement appliquer les contraintes d'utilisation imposées. A-t-il la maturité ou encore la compétence pour respecter les contraintes d'usage ? La deuxième relation de confiance implique les différents fournisseurs intégrés dans l'environnement cross-cloud et porte sur leur *volonté* et leur *capacité* à gérer le stockage et le partage des données du propriétaire avec des demandeurs au sein de l'environnement cross-cloud. Comment assurer que les données stockées dans l'environnement cross-cloud ne soient pas accédées par un des fournisseurs de services cloud sans autorisation ? Quels mécanismes de contrôle d'usage sont offerts par chaque fournisseur cloud et comment les combiner efficacement ?

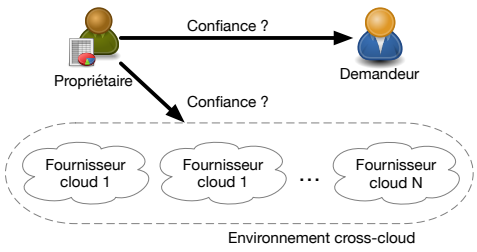

Figure 1. Représentation de la relation de confiance

III. Analyse des approches existantes

Un système de contrôle de l'usage des données cross-cloud a donc pour objectif de permettre au propriétaire des données d'exprimer l'usage autorisé de ses données ainsi que d'avoir un niveau d'assurance important sur sa mise en oeuvre en limitant sa dépendance à la volonté et la capacité des demandeurs et des fournisseurs de cloud à se conformer aux exigences. Dans cet article, nous nous focalisons principalement sur la relation de confiance entre le propriétaire et les fournisseurs de cloud, i.e. les mécanismes limitant la dépendance des propriétaires de données vis-à-vis de la volonté et la capacité des fournisseurs de cloud à stocker et gérer les accès à leurs données. Nous nous intéressons donc aux possibilités de définir une politique de contrôle des usages et à sa mise en oeuvre par les différents fournisseurs clouds (figure 2).

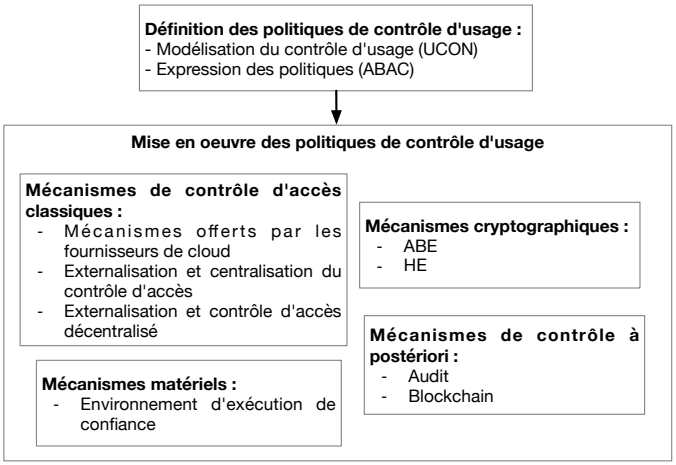

Figure 2. Fonctions attendues par un système de contrôle de l'usage des données cross-cloud et approches possibles

A. Définition des politiques de contrôle d'usage

La définition de politiques de contrôle d'usage nécessite de s'appuyer sur un modèle théorique définissant les concepts du contrôle d'usage ainsi que d'un langage permettant d'exprimer des contraintes selon le modèle. Nous présentons donc le modèle de politique d'usage UCON_{ABC} et expliquons la nécessité d'exprimer les politiques suivant un schéma à base d'attributs (ABAC pour Attribute Based Access Control), ainsi que différents langages ayant la capacité à exprimer les politiques d'usage par leur richesse ou leur extensibilité.

1) Modélisation du contrôle d'usage (UCON): Le modèle de contrôle d'usage UCON_{ABC} (Usage CONtrol Authorization, oBligation, Condition) [5] [6] est un cadre théorique qui couvre de nombreux aspects, tels que le contrôle d'accès traditionnel et le contrôle sur l'usage des ressources numériques avant, pendant et même après leur accès [5]. Ainsi, UCON_{ABC} étend fondamentalement le contrôle d'accès avec un contrôle continu des accès autorisés à une ressource donnée.

UCON_{ABC} se compose de plusieurs concepts qui participent à la construction des règles formant une politique d'usage : les sujets, les objets, les autorisations, les obligations, les

conditions, les attributs des objets et des sujets, ainsi que les droits (Figure 3) :

- Les attributs qui caractérisent les sujets (utilisateur, détenteur de données ou demandeur de données) et les objets (ressources à protéger) sont utilisés pour formuler une politique d'usage, car le modèle UCON_{ABC} repose sur un schéma basé sur les attributs [6] pour l'écriture des politiques d'usage. Ces attributs sont dynamiques et peuvent changer au fil du temps, influençant ainsi les décisions d'accès.
- Les autorisations (A) sont des prédicats fonctionnels sur les attributs qui doivent être satisfaits pour accorder des droits sur un objet.
- Les obligations (B) sont des tâches obligatoires devant être accomplies par l'utilisateur demandeur afin d'obtenir un droit spécifique décrit par la politique ou le contrat. Un exemple serait l'agrément donné par un utilisateur pour fournir certaines données personnelles en échange d'un service.
- Les conditions (C) sont des contraintes sur l'environnement ou le système qui sont prises en compte dans le processus de décision.

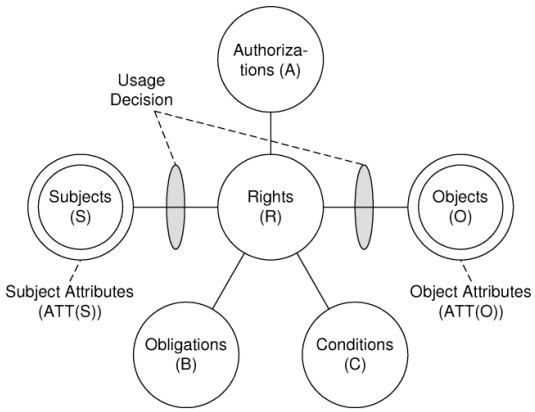

FIGURE 3. Composants du modèle UCON_{ABC} [5]

Les contraintes (autorisations, obligations et conditions) peuvent être évaluées avant l'accès ou pendant l'accès selon UCON_{ABC}.

2) Expression des politiques: Pour pouvoir exprimer les politiques de contrôle d'usage, il est nécessaire de disposer d'un langage offrant des capacités natives adaptées à la rédaction de ces politiques ou permettant l'ajout d'extensions pour prendre en charge les politiques de contrôle d'usage. Plusieurs langages permettent la définition de politiques basées sur les attributs, tels que Rego, Alfa et XACML. Ces langages sont principalement orientés vers l'écriture de politiques de contrôle d'accès ne prenant pas en compte la dynamicité des attributs. Toutefois, ils peuvent être étendus pour exprimer les contraintes du modèle UCON_{ABC}. Par exemple, Carniani et al. [7] ont proposé d'ajouter l'attribut DecisionTime dans les conditions XACML, avec les valeurs admises *pre* et *on* afin d'exprimer respectivement la pré-évaluation des conditions et l'évaluation des conditions lors de l'accès.

B. Mise en oeuvre des politiques de contrôle d'usage

Une fois la politique d'usage exprimée, il est nécessaire de la mettre en oeuvre au travers de mécanismes de sécurité. Il existe de nombreuses approches possibles comme les mécanismes de contrôle d'accès, les mécanismes cryptographiques, les mécanismes de d'isolation matérielles ou encore les mécanismes de contrôles à postériori. Nous analysons ces différentes approches dans le contexte du contrôle d'usage cross-cloud.

1) Mécanismes de contrôle d'accès: Les fournisseurs de services cloud offrent nativement des solutions de gestion des identités et des accès comme Cognito pour Amazon, Entra ID pour Azure ou Google IAM [1]. Cependant, cette approche pose plusieurs problèmes dans un contexte cross-cloud. Tout d'abord, il est nécessaire de prendre en charge l'hétérogénéité des solutions IAM proposées par les fournisseurs ainsi que leur incapacité à mettre en oeuvre des politiques UCON. De plus, cela crée une nouvelle dépendance vis-à-vis de chaque fournisseurs de cloud relative à la mise en oeuvre des politiques sans pour autant renforcer la relation de confiance avec le propriétaire des données.

Pour résoudre ces problèmes, il est possible d'externaliser le mécanisme de contrôle d'accès en dehors des fournisseurs. Les approches de gestion à base de politique comme XACML permettent de dissocier les fonctions de prise de décision (appelée Policy Decision Point ou PDP) et de mise en oeuvre (appelée Policy Enforcement Point ou PEP). Centraliser les prises de décision dans un PDP résoud les problèmes d'hétérogénéité des solutions IAM des fournisseurs de cloud ainsi que leurs limites à exprimer le modèle UCON [8]. Cependant, cela ne répond pas au problème de confiance car les propriétaires doivent faire confiance à une entité tierce pour instancier le PDP et dépend encore des fournisseurs de cloud pour la mise en oeuvre des décisions de contrôle d'usages (instanciation des PEPs). Maesa et al. [9] ont proposé un système de contrôle d'accès externalisé où la prise de décision est décentralisée. Les auteurs ont exprimé les politiques sous forme de contrats intelligents exécutés dans une blockchain. Néanmoins, cette approche soulève de nombreuse questions en terme de performance mais aussi de confidentialité du contenu de la politique [10].

2) Mécanismes d'isolation matérielle: Les environnements d'exécution de confiance (en anglais TEE, pour Trusted Execution Environment) sont des zones isolées de la mémoire et du processeur, protégées grâce au chiffrement. Ils offrent une protection de l'intégrité et de la confidentialité des données et de l'exécution de code au niveau matériel. Toute donnée présente dans cette zone ne peut être ni lue ni modifiée par un code non autorisé[4]. L'industrie propose des environnements d'exécution de confiance physiques tels qu'Intel SGX, ARM TrustZone et AMD SEV [11] [12]. Plusieurs fournisseurs de services cloud proposent des solutions basées sur ces technologies, telles qu'Azure Confidential Computing et les machines

[4]https://learn.microsoft.com/fr-fr/azure/confidential-computing/trusted-execution-environment

virtuelles confidentielles de Google Cloud (Confidential VMs). Nous pouvons citer aussi [13] qui permet d'exécuter la logique de l'usage des données dans des TEE.

Si l'utilisation de TEE semble très prometteuse, il existe actuellement nombre de technologies TEE différentes avec des conceptions hétérogènes présentant chacune des forces et des faiblesses [14] ce qui rend son utilisation dans un contexte cross-cloud non triviale.

3) Mécanismes cryptographiques: Des mécanismes de contrôle d'accès peuvent être mis en oeuvre grâce à des techniques de chiffrement. Le chiffrement à base d'attributs (ABE pour Attribute-Based Encryption) est un schéma cryptographique où l'accès aux données chiffrées est contrôlé en fonction d'un ensemble d'attributs. Deux variantes ont été principalement proposées. Dans le Ciphertext-Policy Attribute-Based Encryption (CP-ABE) [15], la donnée est chiffrée à l'aide d'une clé construite par rapport à la politique d'accès à cette ressource. Les utilisateurs ont des clés construites sur des attributs les décrivant. Ils ne peuvent déchiffrer la données que si ces attributs correspondent à la politique de la donnée. L'approche Key-Policy Attribute-Based Encryption (KP-ABE) [16] propose de chiffrer une donnée par rapport à des attributs la décrivant ; les utilisateurs possédant eux des clés correspondant à leurs permissions. Cette solution présente toutefois certaines limitations, notamment la mise à jour de la politique, la dynamicité des attributs, l'impact sur les performances et l'absence de contrôle sur les données après l'octroi de l'accès.

Une autre approche repose sur l'utilisation du chiffrement homomorphe (HE pour Homomorphic Encryption) [17], un type de chiffrement permettant d'effectuer des opérations sur les données tout en les maintenant chiffrées. Cela signifie qu'il n'est pas nécessaire de les déchiffrer pour vérifier un certain résultat ou effectuer des calculs. Cette propriété garantit la confidentialité des données tout en permettant leur manipulation sécurisée. Cependant, cette méthode reste limitée en termes de performances, car elle est particulièrement coûteuse en ressources de calcul.

C. Mécanismes de contrôle à posteriori

L'application de politiques de contrôle d'usage nécessite en complément des mécanismes de contrôle à posteriori afin de vérifier s'il n'y a pas eu de violation ou pour tout autre cas de litige entre le propriétaire et le fournisseur de cloud [18].

Les mécanismes de contrôle à posteriori sont aussi nécessaires pour s'assurer que les différentes réglementations (RGPD[5], HIPAA[6], ou EUCS[7]) sont correctement appliquées. Un tel mécanisme d'audit pourrait servir de base à de nouvelles approches de partage de données personnelles comme les "personal data licenses" introduites par Wayne Chang[8]. Dans un contexte impliquant de multiples fournisseurs de cloud à auditer, la technologie blockchain avec son approche décentralisée est très adaptée. De plus, son immuabilité confère un haut niveau de fiabilité sur la préservation des enregistrements. Wang et al. [19] ont d'ailleurs présenté un moteur de conformité RGPD basé sur cette technologie. Liang et al. [20] aussi ont proposé de combiner utilisation de blockchain et du schéma de chiffrement CP-ABE (section III-B3) pour obtenir à la fois un système de contrôle d'accès et de traçabilité.

[5]https://gdpr-info.eu/

[6]https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html

[7]https://ec.europa.eu/newsroom/cipr/items/713799/en

[8]https://blog.spruceid.com/how-personal-data-licenses-can-keep-digital-identity-private-part-1/

IV. Conclusion et travaux futurs

Les environnements cross-cloud offrent de nouvelles opportunités en s'affranchissant des contraintes liées à l'utilisation d'un seul fournisseur de cloud. Cependant, ces environnements apportent leurs propres problématiques en particulier dans le contrôle d'usage des données. Dans cet article, nous avons exprimé ce problème d'un point de vue gestion de la confiance. Nous avons aussi analysé différentes approches possibles pour exprimer et mettre en oeuvre des politiques de contrôle d'usage en montrant leur bénéfices mais aussi leur limites. Cette première étude montre le besoin de proposer de nouvelles solutions de contrôle d'usage ne reposant pas sur la confiance dans les fournisseurs de cloud. Nos prochains efforts porteront sur la possibilité d'exploiter et d'agencer de manière optimale différentes catégories de mécanismes de sécurité existants selon l'usage et le type de données considérés dans la politique de contrôle d'usage, activité que nous comptons développer dans le cadre de ce travail de doctorat.

V. Remerciement

Ce travail est financé par le projet France 2030 TRUSTIN-CloudS (ANR project ANR-23-PECL-0009)

Références

[1] M.-A. B. H. Salah, R. Laborde, A. Benzekri, M. A. Kandi, and A. Ferreira, "Identity management in cross-cloud environments : Towards self-sovereign identities using current solutions," 2024.

[2] J. Hong, T. Dreibholz, J. A. Schenkel, and J. A. Hu, "An Overview of Multi-cloud Computing," in *Web, Artificial Intelligence and Network Applications* (L. Barolli, M. Takizawa, F. Xhafa, and T. Enokido, eds.), vol. 927, pp. 1055–1068, Cham : Springer International Publishing, 2019. Series Title : Advances in Intelligent Systems and Computing.

[3] M. Chatterjee, P. Datta, F. Abri, A. S. Namin, and K. S. Jones, "Abuse of the cloud as an attack platform," in *2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*, pp. 1091–1092, IEEE, 2020.

[4] A. S. Wazan, R. Laborde, D. W. Chadwick, F. Barrere, A. Benzekri, M. Kaiiali, and A. Habbal, "Trust management for public key infrastructures : Implementing the x. 509 trust broker," *Security and Communication Networks*, vol. 2017, no. 1, p. 6907146, 2017.

[5] J. Park and R. Sandhu, "The UCON_{abc} usage control model," *ACM Transactions on Information and System Security*, vol. 7, pp. 128–174, Feb. 2004.

[6] A. Lazouski, F. Martinelli, and P. Mori, "Usage control in computer security : A survey," *Computer Science Review*, vol. 4, no. 2, pp. 81–99, 2010. Publisher : Elsevier.

[7] E. Carniani, D. D'Arenzo, A. Lazouski, F. Martinelli, and P. Mori, "Usage control on cloud systems," *Future Generation Computer Systems*, vol. 63, pp. 37–55, 2016. Publisher : Elsevier.

[8] R. Laborde, A. Oglaza, A. S. Wazan, F. Barrère, and A. Benzekri, "A situation-driven framework for dynamic security management," *Annals of Telecommunications*, vol. 74, pp. 185–196, 2019.

[9] D. D. F. Maesa, P. Mori, and L. Ricci, "A blockchain based approach for the definition of auditable access control systems," *Computers & Security*, vol. 84, pp. 93–119, 2019. Publisher : Elsevier.

[10] W. Liang, Y. Liu, C. Yang, S. Xie, K. Li, and W. Susilo, "On identity, transaction, and smart contract privacy on permissioned and permissionless blockchain : A comprehensive survey," *ACM Computing Surveys*, vol. 56, no. 12, pp. 1–35, 2024.

[11] P. Jauernig, A.-R. Sadeghi, and E. Stapf, "Trusted Execution Environments : Properties, Applications, and Challenges," *IEEE Security & Privacy*, vol. 18, pp. 56–60, Mar. 2020. Conference Name : IEEE Security & Privacy.

[12] D. Lee, D. Kohlbrenner, S. Shinde, K. Asanović, and D. Song, "Keystone : an open framework for architecting trusted execution environments," in *Proceedings of the Fifteenth European Conference on Computer Systems*, (Heraklion Greece), pp. 1–16, ACM, Apr. 2020.

[13] H. Wang, J. Wang, C. Ge, Y. Li, L. Zhou, Z. Liu, W. Wu, and M. Cao, "Adss : An available-but-invisible data service scheme for fine-grained usage control," *IEEE Transactions on Services Computing*, 2024.

[14] M. Li, Y. Yang, G. Chen, M. Yan, and Y. Zhang, "Sok : Understanding design choices and pitfalls of trusted execution environments," in *Proceedings of the 19th ACM Asia Conference on Computer and Communications Security*, pp. 1600–1616, 2024.

[15] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy attribute-based encryption," in *2007 IEEE symposium on security and privacy (SP'07)*, pp. 321–334, IEEE, 2007.

[16] A. Sahai and B. Waters, "Fuzzy identity-based encryption," in *Advances in Cryptology–EUROCRYPT 2005 : 24th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Aarhus, Denmark, May 22-26, 2005. Proceedings 24*, pp. 457–473, Springer, 2005.

[17] M. M. P. Mr, C. A. Dhote, and D. H. S. Mr, "Homomorphic encryption for security of cloud data," *Procedia Computer Science*, vol. 79, pp. 175–181, 2016. Publisher : Elsevier.

[18] R. Y. Seetharamarao, "A unified approach towards security audit and compliance in cloud computing environment," in *2023 16th International Conference on Developments in eSystems Engineering (DeSE)*, pp. 623–629, IEEE, 2023.

[19] L. Wang, Z. Guan, Z. Chen, and M. Hu, "Enabling integrity and compliance auditing in blockchain-based gdpr-compliant data management," *IEEE Internet of Things Journal*, vol. 10, no. 23, pp. 20955–20968, 2023.

[20] W. Liang, Y. Yang, C. Yang, Y. Hu, S. Xie, K.-C. Li, and J. Cao, "Pdpchain : A consortium blockchain-based privacy protection scheme for personal data," *IEEE Transactions on Reliability*, vol. 72, no. 2, pp. 586–598, 2022.