UN MOTEUR DE RECHERCHE SCALABLE POUR LE PERSONAL CLOUD PAR SALIHA LALLALI

Présentée par : Saliha Lallila Discipline : informatique Laboratoire : DAVID

Résumé :
Le nouveau paradigme du « Cloud Personnel » vise à offrir un espace de stockage et de traitement respectueux de la vie privée, où les données personnelles sont rendues accessibles aux applications autorisées, tout en restant sous le contrôle de l'individu propriétaire des données. Toutefois, rendre le contrôle de la gestion de données à l'utilisateur déporte les problèmes de sécurité au niveau de la plate-forme utilisateur, donc sur le serveur personnel. Ce serveur personnel est en charge d'organiser l'espace personnel de l'utilisateur sous forme d'une base de données orientée documents pour en faciliter la gestion, permettre de croiser des données issus de silos habituellement gérés séparément, et de les protéger contre la perte, le vol et l'utilisation abusive. Ainsi, les opérations de chiffrement/déchiffrement, la gestion des métadonnées (l'indexation et la recherche de documents) et la gestion des contrôles d'accès, sont sous la responsabilité du serveur personnel.
Dans cette thèse, nous proposons une plate-forme de «Cloud Personnel Sécurisé» basé sur un moteur de recherche et de contrôle d'accès embarqué dans un diapositif matériel personnel et sécurisé, relié à la plate-forme de l'utilisateur. Ce type de dispositifs est

généralement doté d’une très faible quantité de RAM et d'une grande capacité de stockage Flash, ce qui conduit à des contraintes matérielles contradictoires. Pour faire face à ces contraintes, les moteurs de recherche classique privilégient soit les insertions, soit les requêtes, mais ne peuvent répondre aux deux exigences simultanément. Ainsi, pour constituer une solution réaliste, le « Cloud Personnel Sécurisé » doit s'affranchir de deux difficultés principales. La première difficulté réside dans la conception d'un index inversé sur les documents adapté aux fortes contraintes matérielles du dispositif sécurisé, et permettant à la fois de supporter les mises à jours et l'interrogation de façon efficace sur de grands volumes de données. Deuxièmement, le contrôle d'accès doit être soigneusement intégré dans le moteur de recherche embarqué pour en assurer la sécurité, sans entraver les performances d'interrogation et de mise à jour.
Nous avons implanté notre moteur de recherche et de contrôle d’accès sur une carte de développement représentative de différents dispositifs personnels sécurisés, ainsi que sur un dispositif matériel réel que nous avons fait fabriqué. Nous avons mené des expérimentations approfondies sur de grandes collections de documents réelles et synthétiques. Les résultats expérimentaux ont démontré l'évolutivité de l'approche et sa supériorité par rapport aux méthodes d'état de l'art.

Abstract :

The emerging Personal Cloud paradigm holds the promise of a Privacy-by-Design storage and computing platform where personal data remains under the individual's control while being shared by valuable applications. However, leaving the data management control to user's hands pushes the security issues to the user's platform, i. e., the Personal server. This Personal server is in charge of organizing the personal dataspace in a document database style to ease its management, to allow crossing data from multiple "local silos" and to protect it against loss, theft and abusive use. Hence, encryption/decryption, metadata management (e.g., indexing and searching the documents) and access control management is under the responsibility of the Personal server.
In this thesis, we propose a Secure Personal Cloud platform relying on a query and access control engine embedded in a tamper resistant hardware device connected to the user's platform. Such devices are generally equipped with extremely low RAM and large Flash storage capacity, which lead to conflicting hardware constraints. To tackle these constraints, conventional search engines privilege either insertion or query scalability but cannot meet both requirements at the same time. Thus, to become reality, the Secure Personal Cloud has to overpass two main difficulties. The first difficulty lays in the design of an inverted document index capable of tackling the strong hardware constraints of secure devices, and reach update and query scalability at the same time. Second, the

access control has to be carefully integrated with the embedded search engine to ensure the security, but without hampering the query and update performance.
We have implemented our engine on a secure token having a hardware configuration representative of tamper resistant devices and have conducted extensive experiments using real, large document collections. The experimental results demonstrate the scalability of the approach and its superiority compared to state of the art methods.

INFORMATIONS COMPLÉMENTAIRES

Cédric DU MOUZA, Maître de Conférences, Habilité à Diriger des Recherches, au CNAM - Paris - Rapporteur
Esther PACITTI, Professeur des Universités, à l'Université de Montpellier 2/Laboratoire d'Informatique de Robotique et de Microélectronique de Montpellier (LIRMM) - UMR 5506 - Montpellier - Rapporteur
Philippe PUCHERAL, Professeur des Universités, à l'Université Versailles Saint-Quentin-en-Yvelines/Laboratoire Données et Algorithmes pour une Ville Intelligente et Durable de l'infrastructure à l'individu (DAVID) - Versailles - Directeur de thèse
Sophie CHABRIDON, Maître de Conférences, Habilitée à Diriger des Recherches, à Télécom SudParis/Département Informatique - Evry - Examinateur
Sergio ILARRI, Professeur, à l'Université de Zaragoza/Département d'Informatique et Systèmes d'Ingénierie - Zaragoza (Espagne) - Examinateur
Nicolas ANCIAUX, Chercheur, Habilité à Diriger des Recherches, à l'Université de Versailles Saint-Quentin-en-Yvelines/Laboratoire Données et Algorithmes pour une Ville Intelligente et Durable de l'infrastructure à l'individu (DAVID) - Versailles - Invité
Iulian SANDU POPA, Maître de Conférences, à l'Université de Versailles Saint-Quentin-en-Yvelines/Laboratoire Données et Algorithmes pour une Ville Intelligente et Durable de l'infrastructure à l'individu (DAVID) - Versailles - Invité

Contact : dredval service FED : theses@uvsq.fr