IN PARTNERSHIP WITH:
Université Versailles Saint-Quentin

Activity Report 2017

Project-Team PETRUS

PErsonal & TRUSted cloud

RESEARCH CENTER
Saclay - Île-de-France

THEME
Data and Knowledge Representation and Processing

Table of contents

Project-Team PETRUS

Creation of the Team: 2016 December 01, updated into Project-Team: 2017 July 01

Keywords:

Computer Science and Digital Science:

A1.1.8. - Security of architectures
A1.4. - Ubiquitous Systems
A3.1.2. - Data management, quering and storage
A3.1.3. - Distributed data
A3.1.5. - Control access, privacy
A3.1.6. - Query optimization
A3.1.8. - Big data (production, storage, transfer)
A3.1.9. - Database
A4.3. - Cryptography
A4.5. - Formal methods for security
A4.7. - Access control
A4.8. - Privacy-enhancing technologies

Other Research Topics and Application Domains:

B2.5.3. - Assistance for elderly
B6.4. - Internet of things
B6.6. - Embedded systems
B9.8. - Privacy

1. Personnel

Research Scientists

Nicolas Anciaux [Team leader, Inria, Researcher, HDR]
Luc Bouganim [Inria, Senior Researcher, HDR]

Faculty Members

Guillaume Scerri [Univ de Versailles Saint-Quentin-en-Yvelines, Associate Professor]
Philippe Pucheral [Univ de Versailles Saint-Quentin-en-Yvelines, Professor, HDR]
Iulian Sandu Popa [Univ de Versailles Saint-Quentin-en-Yvelines, Associate Professor]

PhD Students

Paul Tran Van [Cozy Cloud (CIFRE), from 2014]
Axel Michel [Insa CVL, from 2015]
Riad Ladjel [Inria, from 2016]
Julien Loudet [Cozy Cloud (CIFRE), from 2016]
Dimitrios Tsolovos [Inria, from 2017]

Technical staff

Aydogan Ersoz [Inria]
Oana Manea [Inria, until Oct 2017]
Laurent Schneider [Inria, from Aug 2017]

Interns

Martin Boyer [Inria, from Jun 2017 until Aug 2017]

Robin Carpentier [Inria, from Jul 2017 until Aug 2017]
Baptiste Crepin [Inria, from Jul 2017 until Aug 2017]
Kai Huang [Inria, from Jun 2017 until Aug 2017]
Poulmanogo Illy [Inria, from Mar 2017 until Dec 2017]
Yiqun Liu [Inria, from May 2017 until Aug 2017]
Lorin Mace [Inria, from Jun 2017 until Aug 2017]
Nawfel Mestoui [Inria, from Jun 2017 until Aug 2017]
Manuel Ory [Inria, from Jun 2017 until Aug 2017]
Nicolas Querhammer [Inria, from Jul 2017 until Aug 2017]

Administrative Assistant

Emmanuelle Perrot [Inria]

External Collaborator

Benjamin Nguyen [INSA CVL, Professor, HDR]

2. Overall Objectives

2.1. Overall Objectives

We are witnessing an exponential accumulation of personal data on central servers: data automatically gathered by administrations and companies but also data produced by individuals themselves (e.g., photos, agendas, data produced by smart appliances and quantified-self devices) and deliberately stored in the cloud for convenience. The net effect is, on the one hand, an unprecedented threat on data privacy due to abusive usage and attacks and, on the other hand, difficulties in providing powerful user-centric services (e.g. personal big data) which require crossing data stored today in isolated silos. The Personal Cloud paradigm holds the promise of a Privacy-by-Design storage and computing platform, where each individual can gather her complete digital environment in one place and share it with applications and users, while preserving her control. However, this paradigm leaves the privacy and security issues in user's hands, which leads to a paradox if we consider the weaknesses of individuals' autonomy in terms of computer security, ability and willingness to administer sharing policies. The challenge is however paramount in a society where emerging economic models are all based - directly or indirectly - on exploiting personal data.

While many research works tackle the organization of the user's workspace, the semantic unification of personal information, the personal data analytics problems, the objective of the PETRUS project-team is to tackle the privacy and security challenges from an architectural point of view. More precisely, our objective is to help providing a technical solution to the personal cloud paradox. More precisely, our goals are (i) to propose new architectures (encompassing both software and hardware aspects) for secure personal cloud data management and formally prove important bricks of the architecture, (ii) propose new data administration models reaching the main requirements of a personal cloud (decentralized access and usage control models, data sharing, data collection and retention models, etc.) and study the enforcement of the resulting privacy policies based on secure hardware and formally proven architectural components, (iii) propose new secure distributed database indexing models, privacy preserving query processing strategies and data anonymization techniques for the personal cloud.

3. Research Program

3.1. Research program

To tackle the challenge introduced above, we identify four main lines of research:

- (Axis 1) Personal cloud server architectures. Based on the intuition that user control, security and privacy are key properties in the definition of trusted personal cloud solutions, our objective is to propose new architectures (encompassing both software and hardware aspects) for secure personal cloud data management and formally prove important bricks of the architecture.
- (Axis 2) Privacy preserving administration models and enforcement. This research axis is devoted to the definition of sharing rules that are easily manageable for the individual and enforced by default (i.e., secure implementation). Complementary to the definition of sharing policies, it is mandatory to help the average user regulate the complete lifecycle of her data, from its capture, to its dissemination and up to its deletion. Our objective is to propose new data administration models reaching the main requirements of a personal cloud (decentralized access and usage control models, data sharing, data collection and retention models, etc.) and study the enforcement of the resulting privacy policies based on secure hardware and formally proven architectural components.
- (Axis 3) Global query evaluation. The goal of this line of research is to provide capabilities for crossing data belonging to multiple individuals (e.g., performing statistical queries over personal data, computing queries on social graphs or organizing participatory data collection) in a fully decentralized setting while providing strong and personalized privacy guarantees. This means proposing new secure distributed database indexing models, privacy preserving query processing strategies and data anonymization techniques for the personal cloud.
- (Axis 4) Economic, legal and societal issues. This research axis is more transversal and entails multidisciplinary research, addressing the links between economic, legal, societal and technological aspects. We will follow here a multi-disciplinary approach based on a 3-step methodology: i) identifying important common issues related to privacy and to the exploitation of personal data; ii) characterizing their dimensions in all relevant disciplines and jointly study their entanglement; iii) validating the proposed analysis, models and trade-offs thanks to in vivo experiments.

These contributions will also rely on tools (algorithms, protocols, proofs, etc.) from other communities, namely security (cryptography, secure multiparty computations, formal methods, differential privacy, etc.) and distributed systems (distributed hash tables, gossip protocols, etc.). Beyond the research actions, we structure our software activity around a single common platform (rather than isolated demonstrators), integrating our main research contributions, called PlugDB. This platform is the cornerstone to help validating our research results through accurate performance measurements on a real platform, a common practice in the DB community, and target the best conferences. It is also a strong vector to federate the team, simplify the bootstrapping of new PhD or master students, conduct multi-disciplinary research and open the way to industrial collaborations and technological transfers.

4. Application Domains

4.1. Application Domains

As stated in the software section, the Petrus research strategy aims at materializing its scientific contributions in an advanced hardware/software platform with the expectation to produce a real societal impact. Hence, our software activity is structured around a common Secure Personal Cloud platform rather than several isolated demonstrators. This platform will serve as the foundation to develop a few emblematic applications. Several privacy-preserving applications can actually be targeted by a Personal Cloud platform, like: (i) smart disclosure applications allowing the individual to recover her personal data from external sources (e.g., bank, online shopping activity, insurance, etc.), integrate them and cross them to perform personal big data tasks (e.g., to improve her budget management) ; (ii) management of personal medical records for care coordination and well-being improvement; (iii) privacy-aware data management for the IoT (e.g., in sensors, quantified-self devices, smart meters); (iv) community-based sensing and community data sharing; (v) privacy-preserving studies (e.g., cohorts, public surveys, privacy-preserving data publishing). Such applications overlap with all the research axes described above but each of them also presents its own specificities. For instance, the smart disclosure applications will focus primarily on sharing models and enforcement, the IoT applications require to look with priority at the embedded data management and sustainability issues, while community-based sensing and privacy-preserving studies demand to study secure and efficient global query processing. Among these applications domains, one is already receiving a particular attention from our team. Indeed, we gained a strong expertise in the management and protection of healthcare data through our past DMSP (Dossier Medico-Social Partagé) experiment in the field. This expertise is being exploited to develop a dedicated healthcare and well-being personal cloud platform. We are currently deploying 10000 boxes equipped with PlugDB in the context of the DomYcile project. In this context, we are currently setting up an Inria Innovation Lab with the Hippocad company to industrialize this platform and deploy it at large scale (see Section the bilateral contract OwnCare II-Lab).

5. New Software and Platforms

5.1. PLUG-DB ENGINE

KEYWORDS: Databases - Personal information - Privacy - Hardware and Software Platform
FUNCTIONAL DESCRIPTION: PlugDB is a complete platform dedicated to a secure and ubiquitous management of personal data. It aims at providing an alternative to a systematic centralization of personal data. The PlugDB engine is a personal database server capable of storing data (tuples and documents) in tables and BLOBs, indexing them, querying them in SQL, sharing them through assertional access control policies and enforcing transactional properties (atomicity, integrity, durability). The PlugDB engine is embedded in a tamper-resistant hardware device combining the security of smartcard with the storage capacity of NAND Flash. The personal database is hosted encrypted in NAND Flash and the PlugDB engine code runs in the microcontroller. Complementary modules allow to pre-compile SQL queries for the applications, communicate with the DBMS from a remote Java program, synchronize local data with remote servers (typically used for recovering the database in the case of a broken or lost devices) and participate in distributed computation (e.g., global queries). PlugDB runs both on secure devices provided by Gemalto and on specific secure devices designed by PETRUS and assembled by electronic SMEs. Mastering the hardware platform opens up new research and experiment opportunities (e.g., we have recently integrated a Bluetooth module to communicate wirelessly with PlugDB and a fingerprint module to strongly authenticate users). PlugDB engine has been registered first at APP (Agence de Protection des Programmes) in 2009 - a new version being registered every two years and the hardware datasheets in 2015. PlugDB has been experimented in the field, notably in the healthcare domain. We also recently set up an educational platform on top of PlugDB, named SIPD (Système d'Information privacy-by-Design) and used at ENSIIE, INSA CVL and UVSQ through the Versailles Sciences Lab fablab, to raise students awareness of privacy protection problems and embedded programming. As a conclusion, PlugDB combines several research contributions from the team, at the crossroads of flash data management, embedded data processing and secure distributed computations. It then strongly federates all members of our team (permanent members, PhD students and engineers). It is also a vector of visibility, technological transfer and dissemination and gives us the opportunity to collaborate with researchers from other disciplines around a concrete privacy-enhancing platform. PlugDB is now being industrialized in the context of the OwnCare Inria Innovation Lab (II-Lab).

- Participants: Aydogan Ersoz, Laurent Schneider, Luc Bouganim, Nicolas Anciaux and Philippe Pucheral
- Contact: Nicolas Anciaux
- URL: https://project.inria.fr/plugdb/

6. New Results

6.1. Personal Cloud Architecture Based on Trusted Execution Environments (Axis 1)

Participants: Nicolas Anciaux [correspondent], Luc Bouganim, Riad Ladjel, Julien Loudet, Benjamin Nguyen, Philippe Pucheral, Iulian Sandu Popa, Guillaume Scerri, Paul Tran Van.

The Personal Cloud paradigm and its challenges: The time of individualized management and control over one's personal data is upon us. Thanks to smart disclosure initiatives, we can access our personal data from the companies or government agencies that collected them. Concurrently, Personal Cloud solutions are flourishing. Their goal is to empower us to leverage our personal data for our own good. However, managing our own personal data constitutes a considerable burden. We must now: (1) ensure the security of the data we gather; and (2) manage the disclosed data and control its usage. We inherit the combined responsibility of an information security expert and a database administrator. Since very few users are actually IT experts, personal cloud providers propose solutions to manage personal data on behalf of their customers. Thus, paradoxically, instead of empowering users, smart disclosure and personal clouds create new privacy risks. In this work, we formulate this paradox and the problems it creates. Our central contribution is a reference architecture for the Personal Cloud, instantiated on several hardware configuration using trusted execution environments (paper in preparation).

6.2. Data Management in Secure Hardware (Axis 1)

Participants: Nicolas Anciaux, Philippe Pucheral, Iulian Sandu Popa [correspondent].

Secure keyword search in the Personal Cloud: The Personal Cloud paradigm has emerged as a solution that allows individuals to manage under their control the collection, usage and sharing of their data. However, by regaining the full control over their data, the users also inherit the burden of protecting it against all forms of attacks and abusive usages. The Secure Personal Cloud architecture relieves the individual from this security task by employing a secure token (i.e., a tamper-resistant hardware device) to control all the sensitive information (e.g., encryption keys, metadata, indexes) and operations (e.g., authentication, data encryption/decryption, access control, and query processing). However, secure tokens are usually equipped with extremely low RAM but have significant Flash storage capacity (Gigabytes), which raises important barriers for embedded data management. This work [11] proposes a new embedded search engine specifically designed for secure tokens, which applies to the important use-case of managing and securing documents in the Personal Cloud context. Conventional search engines privilege either insertion or query scalability but cannot meet both requirements at the same time. Moreover, very few solutions support data deletions and updates in this context. In this work, we introduce three design principles, namely Write-Once Partitioning, Linear Pipelining and Background Linear Merging, and show how they can be combined to produce an embedded search engine matching the hardware constraints of secure tokens and reconciling high insert/delete/update rate and query scalability. Our experimental results, obtained with a prototype running on a representative hardware platform, demonstrate the scalability of the approach on large datasets and its superiority compared to state of the art methods. Finally, the integration of our solution in another important real use-case related to performing information retrieval in smart objects has been previously discussed in [5] and demonstrated at [25].

6.3. Data Management in Flash Memory (Axis 1)

Participant: Luc Bouganim [correspondent].

Understanding Flash I/O Patterns on Open-Channel Solid-State Drives: Solid-State Drives (SSDs) have gained acceptance by providing the same block device abstraction as magnetic hard drives, at the cost of suboptimal resource utilization and unpredictable performance. Recently, Open-Channel SSDs have emerged as a means to obtain predictably high performance, based on a clean break from the block device abstraction. Open-channel SSDs embed a minimal flash translation layer (FTL) and expose their internals to the host. The Linux open-channel SSD subsystem, LightNVM, lets kernel modules as well as user-space applications control data placement and I/O scheduling. This way, it is the host that is responsible for SSD management. But what kind of performance model should the host rely on to guide the way it manages data placement and I/O scheduling? For addressing this question we have defined uFLIP- OC, a benchmark designed to identify the I/O patterns that are best suited for a given open-channel SSD. Our experiments on a Dragon- Fire Card (DFC) SSD, equipped with the OX controller, illustrate the performance impact of media characteristics and parallelism. In [17], we present uFLIP-OC and how it can be used to guide the design of host-based data systems on open-channel SSDs.

6.4. Data Sharing architecture for the Personal Cloud (Axis 2)

Participants: Nicolas Anciaux [correspondent], Philippe Pucheral, Paul Tran Van.

SWYSWYK Architecture: Pushed by recent legislation and smart disclosure initiatives, Personal Cloud platforms emerge and hold the promise of giving the control back to the individual on her data. However, this shift leaves the privacy and security issues in user's hands, a role that few people can properly endorse. Indeed, existing sharing models are difficult to administrate and securing their implementation in user's computing environment is an unresolved challenge. This study advocates the definition of a Privacy-by-Design sharing architecture, called SWYSWYK (Share What You See with Who You Know), dedicated to the Personal Cloud context. This architecture allows each user to physically visualize the net effects of sharing rules on her Personal Cloud and automatically provides tangible guarantees about the enforcement of the defined sharing policies. The architecture relies on a secure reference monitor, a set of user defined functions only interacting with the secure monitor and isolated from the unsecure environment, and an unsecure personal cloud platform managing encrypted personal data. The SWYSWYK architecture is presented in [20]. A validation of this architecture combining PlugDB to host the secure reference monitor, a RaspberryPI to launch the isolated user defined functions and a personal computer to host the untrusted personal cloud software was demonstrated in [19]. It shows the practicality of the approach and a performance evaluation on a real Personal Cloud platform.

6.5. Data sharing model for the Personal Cloud (Axis 2)

Participants: Nicolas Anciaux [correspondent], Paul Tran Van, Philippe Pucheral.

SWYSWYK Semantics: The personal cloud content intrinsically describes the individual's acquaintances under different forms (e.g., contact files, agendas, identity pictures, address book entries, etc.). Conversely, acquaintances are associated with pieces of information in the user's space (e.g., photos on which a friend appears). New sharing models should be thus able to map personal data to acquaintances (or subjects) and exploit their links with the stored documents (or objects) to produce authorizations satisfying users' sharing desires such as those expressed above. Interesting and common sharing rules could also be published and adopted by the members of a community of interest. In [18], we propose SWYSWYK, a new data sharing model which builds upon the transversal nature of the content of a personal cloud and makes easy and intuitive the definition and administration of sharing policies. Beyond the definition of the sharing policy, SWYSWYK provides means to the personal cloud owner to easily understand the net effects of a sharing policy, identify suspicious permissions and sanitize the sharing policy accordingly, and finally, to trust the way the policy is practically enforced. In [21] we demonstrate the semantics of the model with the goal to assess its practical interest for the personal cloud owner. To this end, we have integrated SWYSWYK in a real personal cloud platform (namely Cozy) and apply it to a smart surrounding scenario.

6.6. Privacy-preserving Computation Protocols on Asymmetric Architectures (Axis 3)

Participant: Iulian Sandu Popa [correspondent].

Distributed Vehicular Traffic Re-routing System for Congestion Avoidance: Centralized solutions for vehicular traffic re-routing to alleviate congestion suffer from two intrinsic problems: scalability, as the central server has to perform intensive computation and communication with the vehicles in real-time; and privacy, as the drivers have to share their location as well as the origins and destinations of their trips with the server. In this work [12], we proposed DIVERT, a distributed vehicular re-routing system for congestion avoidance. DIVERT offloads a large part of the re-routing computation at the vehicles, and thus, the re-routing process becomes practical in real-time. To take collaborative re-routing decisions, the vehicles exchange messages over vehicular ad hoc networks. DIVERT is a hybrid system because it still uses a server and Internet communication to determine an accurate global view of the traffic. In addition, DIVERT balances the user privacy with the re-routing effectiveness. The simulation results demonstrate that, compared with a centralized system, the proposed hybrid system increases the user privacy by 92 percent on average. In terms of average travel time, DIVERT's performance is slightly less than that of the centralized system, but it still achieves substantial gains compared to the no re-routing case. In addition, DIVERT reduces the CPU and network load on the server by 99.99 and 95 percent, respectively.

6.7. Privacy-preserving Anonymization Protocols on Asymmetric Architectures (Axis 3)

Participants: Axel Michel, Benjamin Nguyen [correspondent], Philippe Pucheral.

Managing Distributed Queries under Personalized Anonymity Constraints The benefit of performing Big data computations over individual's microdata is manifold, in the medical, energy or transportation fields to cite only a few, and this interest is growing with the emergence of smart disclosure initiatives around the world. However, these computations often expose microdata to privacy leakages, explaining the reluctance of individuals to participate in studies despite the privacy guarantees promised by statistical institutes. In this work [22], we propose a novel approach to push personalized privacy guarantees in the processing of database queries so that individuals can disclose different amounts of information (i.e. data at different levels of accuracy) depending on their own perception of the risk. Moreover, we propose a decentralized computing infrastructure based on secure hardware enforcing these personalized privacy guarantees all along the query execution process. A performance analysis conducted on a real platform shows the effectiveness of the approach.

6.8. Economic, legal and societal issues (Axis 4)

Participants: Nicolas Anciaux [correspondent], Philippe Pucheral.

Data Portability and Users' Empowerment as a Privacy Incentive. The principle of 'data portability' recently introduced in regulations (smart disclosure in the US, data portability in France and EU) is tightly coupled with the notion of Personal Cloud. We conduct a study of this principle in common with the DANTE Lab at UVSQ, in particular with Prof. Celia Zolynski (jurist, member of the CNN), within the DATAIA convergence institute at Inria and in the SIHS CNRS federation at UVSQ. Our recent contributions analyze the technical conditions under which individuals can get their data back from service providers according to this data portability principle, and examine its technical feasibility and legal opportunity. We also explain how data portability favors a form of users' empowerment, which can be viewed as a potential privacy incentive. Our recent results are presented in multi-disciplinary papers appeared in prestigious French journals like DALLOZ [14] and 'Revue Contrats, Concurrence, Consommation' [13] [15].

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

7.1.1. OwnCare II-Lab (Jul 2017 - Dec 2020)

Partners: PETRUS (Inria-UVSQ), Hippocad (SME)
Funding: to be determined
End 2016, the Yvelines district lauched a public call for tender to deploy an industrial solution aiming at covering the whole distinct (10.000 patients). The Hippocad company, in partnership with Inria, won this call for tender with a solution called DomYcile in May 2017 and the project was launched in July 2017. DomYcile is based on a home box combining the PlugDB hardware/software technology developed by the Petus team and a communication layer based on SigFox. Hippocad and Petrus then decided to launch a joint II-Lab (Inria Innovation Lab) named OwnCare. The objective is threefold: (1) build an industrial solution based on PlugDB and deploy it in the Yvelines district in the short-term, (2) use this Yvelines testbed to improve the solution and try to deploy it at the national/international level in the medium-term and (3) design flexible/secure/mobile personal medical folder solutions targeting individual uses rather than professional uses in the long-term. The DomYcile project with the Yvelynes district has started in July 2017 and the II-Lab should be officially created in January 2018.

7.2. Bilateral Grants with Industry

7.2.1. Cozy Cloud CIFRE - Tran Van contract (Oct 2014 -Feb 2017)

Partners: Cozy Cloud, PETRUS (Inria-UVSQ)
SMIS funding: 30k€
In relation with the bilateral contract mentioned above, a CIFRE PhD thesis has been started by Paul Tran Van. The objective is to capitalize on the Cozy-PlugDB platform to devise new access and usage control models to exchange data among devices of the same user (devices may have different levels of trustworthiness) and among different users thanks to a user-friendly sharing model (see the work on the SWYSWYK - Share What You See with Who You Know - model presented above).

7.2.2. Cozy Cloud CIFRE - Loudet contract (Apr 2016 - Apr 2019

Partners: Cozy Cloud, Inria-SMIS
SMIS funding: 45k€
In relation with the bilateral contract mentioned above, a second CIFRE PhD thesis has been started by Julien Loudet. The objective is to allow for a secure execution of distributed queries on a set of personal clouds associated to users, depending on social links, user's localization or user's profile. The general idea is to build secure indexes, distributed on the users' personal clouds and to devise a secure execution protocol revealing solely the query result to the querier. Such highly distributed secure queries potentially enable new (social) applications fed by user's personal data which could be developed on the Cozy-PlugDB platform.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR PerSoCloud (Jan 2017 - Jan 2020)

Partners: Orange Labs (coordinator), PETRUS (Inria-UVSQ), Cozy Cloud, U. of Versailles.
PETRUS funding: 170k€. The objective of PerSoCloud is to design, implement and validate a fullâ€'fledged Privacy-by-Design Personal Cloud Sharing Platform. One of the major difficulties linked to the concept of personal cloud lies in organizing and enforcing the security of the data sharing while the data is no longer under the control of a central server. We identify three dimensions to this problem. Devices-sharing: assuming that the primary copy of user U1's personal data is hosted in a secure place, how to share and synchronize it with U1's multiple (mobile) devices without compromising security? Peers-sharing: how user U1 could exchange a subset of his-her data with an identified user U2 while providing to U1 tangible guarantees about the usage made by U2 of this data? Community-sharing: how user U1 could exchange a subset of his-her data with a large community of users and contribute to personal big data analytics while providing to U1 tangible guarantees about the preservation of his-her anonymity? In addition to tackling these three scientific and technical issues, a legal analysis will guarantee compliance of this platform with the security and privacy French and UE regulation, which firmly promotes the Privacy by Design principle, including the current reforms of personal data regulation.

8.1.2. PIA - PDP SECSi (May 2016 - Dec 2017)

Partners: Cozy Cloud (coordinator), Qwant, Inria (Inria-UVSQ), FING.
SMIS funding: 149k€. The objective of this PIA-PDP (Programme Investissement d'Avenir - Protection des DonnÃ©es Personnelles) SECSi project is to build a concrete Personal Cloud platform which can support a large scale deployment of Self Data services. Three major difficulties are identified and will be tackled in this project: (1) how to implement and enforce a fine control of the data flow when personal data are exploited by third party applications, (2) how to protect these same applications when processing is delegated to the personal cloud platform itself and (3) how to implement personalized search on the web without hurting user's privacy.

8.1.3. CityLab@Inria, Inria Project Lab (May 2014 -).

Inria Partners: ARLES-MIMOVE, CLIME, DICE, FUN, MYRIADS, OAK, PETRUS, URBANET, WILLOW.
External partners: UC Berkeley.
Funding: not associated to individual project teams. CityLab@Inria studies ICT solutions toward smart cities that promote both social and environmental sustainability. A strong emphasis of the Lab is on the undertaking of a multi-disciplinary research program through the integration of relevant scientific and technology studies, from sensing up to analytics and advanced applications, so as to actually enact the foreseen smart city Systems of Systems. SMIS contributes to Privacy-by-Design architectures for trusted smart objects so as to ensure privacy to citizens, which is critical for ensuring that urbanscale sensing contributes to social sustainability and does not become a threat. The PhD Thesis of Dimitris Tsoulovos, co-directed by MIMOVE and PETRUS, is funded by CityLab. http://citylab.inria.fr/

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Selection

9.1.1.1. Member of the Conference Program Committees

- Nicolas Anciaux: VLDB'18, SIGMOD'17, PAP'17, DATA'17
- Luc Bouganim: Associate Editor for VLDB'18, VLDB'17, BDA'17
- Philippe Pucheral: EDBT'17, DATA'17, MOBILITY'17
- Iulian Sandu Popa: APWeb-WAIM'17, DATA'17, IEEE MobileCloud'17

9.1.1.2. Reviewer

- Guillaume Scerri: CCS'17

9.1.2. Journal

9.1.2.1. Member of the Editorial Boards

- Nicolas Anciaux: Associate Editor of the VLDB Journal

9.1.2.2. Reviewer - Reviewing Activities

- Iulian Sandu Popa: ACM Transactions on Spatial Algorithms and Systems, International Journal of Geo-Information, Journal of Intelligent Transportation Systems

9.1.3. Invited Talks

- Nicolas Anciaux: "A new Approach for the Secure Personal Cloud", Security Seminar, Loria, 23 Mar. 2017. http://seminaire-securite.loria.fr/seances-passees.fr.html

9.1.4. Research Administration

Philippe Pucheral: Member of the HDR committee of the STV doctoral school (UVSQ) since 2014

Philippe Pucheral: Member of the steering committee of the ED STIC doctoral school of University Paris-Saclay, 'Data, Knowledge and Interactions' committee (about 250 PhD students) since 2014

Nicolas Anciaux: Member of the council of the doctoral college of the University Paris-Saclay since 2017

Nicolas Anciaux: Correspondent for Inria Saclay at ED STIC doctoral school of University Paris-Saclay since 2017

Nicolas Anciaux: Responsible for the 'Mission Jeunes Chercheurs' at Inria Saclay since 2017

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Licence : Iulian Sandu Popa, Bases de données (niveau L3), 96, UVSQ, France. Guillaume Scerri , Initiation aux bases de données (niveau L2), 63, UVSQ, France. Guillaume Scerri, Fondements de l'informatique (niveau L1), 72, UVSQ, France. Guillaume Scerri, Théorie des Langages (niveau L2), 36, UVSQ, France.

Master : Iulian Sandu Popa, Bases de données relationnelles (niveau M1), Gestion des données spatiotemporelles (niveau M2), Sécurité des bases de données (niveau M2), 96, UVSQ, France. Philippe Pucheral, responsible of the DataScale master, courses in M1 and M2 in databases and in security, introductory courses for jurists,UVSQ, France. Luc Bouganim, Bases de données relationnelles et XML (niveau M1 et M2), 20, AFTI, France. Guillaume Scerri, Bases de données relationnelles (niveau M1), 36, UVSQ, France. Guillaume Scerri, Sécurité et bases de données pour juristes, 4.5, UVSQ, France.

Engineers school : Nicolas Anciaux, courses on Databases (module IN206, niveau M1), 21, and Advanced databases (module ASI13, niveau M2), 24, at ENSTA ParisTech. Nicolas Anciaux, Systèmes d'Information "privacy by design" (niveau M1), 15, at ENSIIE Evry, France. Luc Bouganim, Systèmes d'Information "privacy by design" (niveau M1), 38, ENSIIE Evry et INSA CVL, France. Luc Bouganim, Bases de données et sécurité (niveau M2), 20, Telecom ParisTech.

9.2.2. Supervision

PhD in progress : Paul Tran Van, Partage de documents sécurisé dans le Cloud Personnel, October 2014, Nicolas Anciaux and Philippe Pucheral

PhD in progress : Axel Michel, Secure Distributed Computations, October 2015, Benjamin Nguyen and Philippe Pucheral

PhD in progress : Julien Loudet, Personal Queries on Personal Clouds, July 2016, Luc Bouganim and Iulian Sandu Popa

PhD in progress : Riad Ladjel, Secure Distributed Computation for the Personal Cloud, October 2016, Nicolas Anciaux and Philippe Pucheral

PhD in progress : Dimitri Tsoulovos, Privacy-by-design Middleware for Urban-scale Mobile Crowdsensing, April 2017, Nicolas Anciaux and Valérie Issarny (Inria Mimove)

9.2.3. Juries

Philippe Pucheral : reviewer of the PhD thesis of Yifan Li (CNAM Paris, 15/12/2017)

Nicolas Anciaux : Jury member of the PhD thesis of David Montoya (ENS Paris-Saclay, 6/3/2017)

Luc Bouganim : Reviewer of the HDR of Jalil Boukhobza (University of Bretagne Sud, 4/7/2017)

Luc Bouganim : Reviewer of the PhD of Levent Demir (University of Grenobles Alpes, 7/12/2017)

9.3. Popularization

MOOC "Défis technologiques des villes intelligentes participatives", du 6 Nov. 2017 au 31 Dec. 2017. Auteur(s): Nicolas Anciaux, Stéphane Grumbach, Valérie Issarny, Nathalie Mitton, Christine Morin, Animesh Pathak, Hervé Rivano. http://www.sup-numerique.gouv.fr/cid94187/mooc-villes-intelligentes-defis-technologiques-et-societaux.html

EIT Digital Professional School course "Technological challenges of participatory Smart Cities". Course for companies and cities. By Nicolas Anciaux, Stéphane Grumbach, Valérie Issarny, Nathalie Mitton, Christine Morin, Animesh Pathak, Hervé Rivano. https://www.eitdigital.eu/newsroom/news/article/eit-digital-professional-school-launches-a-smart-cities-course-for-companies-and-cities/

Nicolas Anciaux: interview pour le magazine La Recherche. Numéro de Juin 2017, dossier "Données personnelles: Notre vie privée est-elle en danger?", pp.81-86, par Vincent Glavieux et Denis Delbecq.

Rencontre Inria - Industrie, Les données et leurs applications, Conférence "Sécurité des données", Philippe Pucheral et Eric Léandri (fondateur et PDG de Qwant), 6 Juin 2017. https://www.inria.fr/centre/paris/innovation/rii-les-donnees-17-18-octobre-2017/programme

Rencontre Inria - Industrie, Les données et leurs applications, Démonstration du projet Domycile, Nicolas Anciaux, , 6 Juin 2017. https://www.inria.fr/centre/paris/innovation/rii-les-donnees-17-18-octobre-2017/demos/domycile

Rencontre Inria - Industrie, Les données et leurs applications, Démonstration du projet Domycile, Paul Tran Van, , 6 Juin 2017.

Round table: Accès aux données et aux compétences, Convention Systematic Paris-Région, Luc Bouganim, June 2017.

Demonstration: PlugDB and the Secure Mobile Laboratory, Turing building inauguration, Luc Bouganim, Iulian Sandu Popa, Riad Ladjel, February 2017.

10. Bibliography

Major publications by the team in recent years

[1] T. Allard, B. Nguyen, P. Pucheral. *MetaP: Revisiting Privacy-Preserving Data Publishing using Secure Devices*, in "Distributed and Parallel Databases", June 2014, vol. 32, n^{o} 1, pp. 191-244 [*DOI :* 10.1007/s10619-013-7122-x], https://hal.archives-ouvertes.fr/hal-00934586

[2] N. Anciaux, P. Bonnet, L. Bouganim, B. Nguyen, P. Pucheral, I. Sandu Popa. *Trusted Cells : A Sea Change for Personnal Data Services*, in "CIDR 2013 - 6th Biennal Conference on Innovative Database Research", Asilomar, United States, 2013, 4 p. , http://hal.inria.fr/hal-00768379

[3] N. Anciaux, L. Bouganim, T. Delot, S. Ilarri, L. Kloul, N. Mitton, P. Pucheral. *Folk-IS: Opportunistic Data Services in Least Developed Countries*, in "40th International Conference on Very Large Data Bases (VLDB)", Hangzhou, China, Zhejiang University, September 2014, https://hal.inria.fr/hal-00906204

[4] N. Anciaux, L. Bouganim, P. Pucheral, Y. Guo, L. Le Folgoc, S. Yin. *MILo-DB: a personal, secure and portable database machine*, in "Distributed and Parallel Databases", March 2014, vol. 32, n^{o} 1, pp. 37-63 [*DOI :* 10.1007/s10619-012-7119-x], https://hal.archives-ouvertes.fr/hal-00768355

[5] N. Anciaux, S. Lallali, I. Sandu Popa, P. Pucheral. *A Scalable Search Engine for Mass Storage Smart Objects*, in "Proceedings of the 41th International Conference on Very Large Databases (VLDB)", Kohala Coast, Hawaii, United States, August 2015, vol. 8, n^{o} 9, pp. 910-921 [*DOI :* 10.14778/2777598.2777600], https://hal.inria.fr/hal-01176458

[6] N. Anciaux, B. Nguyen, I. Sandu Popa. *Tutorial: Managing Personal Data with Strong Privacy Guarantees*, March 2014, pp. 672-673 [*DOI :* 10.5441/002/edbt.2014.71], https://hal.inria.fr/hal-01096633

[7] G. Scerri, B. Warinschi, M. Barbosa, B. Portela. *Foundations of Hardware-Based Attested Computation and Application to SGX*, March 2016, pp. 245-260 [*DOI :* 10.1109/EuroSP.2016.28], https://hal.inria.fr/hal-01417137

[8] C. Q. To, B. Nguyen, P. Pucheral. *TrustedMR: A Trusted MapReduce System based on Tamper Resistance Hardware* , in "Proceedings of the 23rd International Conference on Cooperative Information Systems (COOPIS)", Rhodes, Greece, October 2015, pp. 38-56 [*DOI :* 10.1007/978-3-319-26148-5_3], https://hal.inria.fr/hal-01254951

[9] C. Q. To, B. Nguyen, P. Pucheral. *Private and Scalable Execution of SQL Aggregates on a Secure Decentralized Architecture*, in "ACM Transactions on Database Systems", 2016, vol. 41, n^o 3, pp. 16:1-16:43, https://hal.archives-ouvertes.fr/hal-01296432

[10] D. H. Ton That, I. Sandu Popa, K. Zeitouni. *TRIFL: A Generic Trajectory Index for Flash Storage*, in "ACM Transactions on Algorithms", July 2015, vol. 1, n^o 2, 44 p. [*DOI :* 10.1145/2786758], https://hal.inria.fr/hal-01176563

Publications of the year

Articles in International Peer-Reviewed Journals

[11] S. Lallali, N. Anciaux, I. Sandu Popa, P. Pucheral. *Supporting secure keyword search in the personal cloud*, in "Information Systems", December 2017, vol. 72, pp. 1 - 26 [*DOI :* 10.1016/j.is.2017.09.003], https://hal.inria.fr/hal-01660599

[12] S. J. Pan, I. Sandu Popa, C. Borcea. *DIVERT: A Distributed Vehicular Traffic Re-Routing System for Congestion Avoidance*, in "IEEE Transactions on Mobile Computing", January 2017, vol. 16, n^o 1, pp. 58-72 [*DOI :* 10.1109/TMC.2016.2538226], https://hal.inria.fr/hal-01426424

Articles in National Peer-Reviewed Journals

[13] N. Anciaux, P. Pucheral, M. BEHAR-TOUCHAIS, V.-L. Benabou, G. BRUNAUX, A. LEFEVRE, N. MARTIAL-BRAZ, J. ROCHFELD, N. SAUPHANOR-BROUILLAUD, B. SCHULZ, J. SENECHAL, C. Zolynski. *Dossier Contenus Numériques Revue Contrats, Concurrence, Consommation - Contenus Numériques*, in "Contrats concurrence consommation", February 2017, https://hal.archives-ouvertes.fr/hal-01432544

[14] C. Berthet, C. Zolynski, N. Anciaux, P. Pucheral. *" Contenus numériques et récupération des données : un nouvel outil d' 'empouvoirement' du consommateur ? "*, in "Dalloz IP/IT", January 2017, vol. IP IT / 10, https://hal.inria.fr/hal-01429939

[15] P. Pucheral, N. Anciaux, M. BEHAR-TOUCHAIS, V.-L. Benabou, N. MARTIAL-BRAZ, J. ROCHFELD, N. SAUPHANOR-BROUILLAUD, B. SCHULZ, J. SENECHAL, C. Zolynski. *Dossier Contenus Numériques / Données*, in "Contrats concurrence consommation", February 2017, https://hal.inria.fr/hal-01429951

International Conferences with Proceedings

[16] C. Jacomme, S. Kremer, G. Scerri. *Symbolic Models for Isolated Execution Environments*, in "2nd IEEE European Symposium on Security and Privacy (EuroS&P'17)", Paris, France, C. Hriţcu (editor), Proceedings of the 2nd IEEE European Symposium on Security and Privacy, Springer, April 2017, https://hal.inria.fr/hal-01396291

[17] I. L. Picoli, C. V. Pasco, B. Þ. Jónsson, L. Bouganim, P. Bonnet. *uFLIP-OC: Understanding Flash I/O Patterns on Open-Channel Solid-State Drives*, in "APSys'17", Mumbai, India, September 2017, pp. 1-7 [*DOI :* 10.1145/3124680.3124741], https://hal.archives-ouvertes.fr/hal-01654985

[18] P. Tran-Van, N. Anciaux, P. Pucheral. *A new Sharing Paradigm for the Personal Cloud*, in "TrustBus 2017 - 14th International Conference on Trust, Privacy and Security in Digital Business", Lyon, France, August 2017, https://hal.inria.fr/hal-01675092

[19] P. Tran-Van, N. Anciaux, P. Pucheral. *SWYSWYK: a new Sharing Paradigm for the Personal Cloud*, in "ADMA 2017 - International Conference on Advanced Data Mining and Applications", Singapore, Indonesia, November 2017, https://hal.inria.fr/hal-01675091

[20] P. Tran-Van, N. Anciaux, P. Pucheral. *SWYSWYK: A Privacy-by-Design Paradigm for Personal Information Management Systems*, in "International Conference on Information Systems Development (ISD)", Cyprus, Cyprus, September 2017, https://hal.inria.fr/hal-01675090

[21] P. Tran-Van, N. Anciaux, P. Pucheral. *Reconciling Privacy and Data Sharing in a Smart and Connected Surrounding*, in "International Conference on Extending Database Technology (EDBT)", Vienna, Austria, March 2018, https://hal.inria.fr/hal-01675093

Conferences without Proceedings

[22] A. Michel, B. Nguyen, P. Pucheral. *Managing Distributed Queries under Personalized Anonymity Constraints*, in "DATA", MADRID, Spain, 2017, https://hal.archives-ouvertes.fr/hal-01682316

Scientific Books (or Scientific Book chapters)

[23] L. Bouganim. *Data Skew*, in "Encyclopedia of Database Systems (2nd edition)", L. Liu, M. O¨zsu (editors), Springer, 2017 [*DOI :* 10.1007/978-1-4899-7993-3_1088-2], https://hal.archives-ouvertes.fr/hal-01656691

[24] L. Bouganim. *Query Load Balancing in Parallel Database Systems*, in "Encyclopedia of Database Systems (2nd edition)", L. Liu, M. O¨zsu (editors), Springer, 2017, pp. 1-6 [*DOI :* 10.1007/978-1-4899-7993-3_1080-2], https://hal.inria.fr/hal-01660649

References in notes

[25] S. Lallali, N. Anciaux, I. Sandu Popa, P. Pucheral. *A Secure Search Engine for the Personal Cloud*, in "Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data (SIGMOD '15). Demo paper", Melbourne, Australia, 2015, pp. 1445-1450 [*DOI :* 10.1145/2723372.2735376], https://hal.inria.fr/hal-01176473