UCN WP1

Lead UCAM w/ Notts ++
Jon.crowcroft, amir.chaudhry, thomas.gazagnaire,
anil.madhavapeddy, carlos.molina, david.sheets @cl.cam.ac.uk

PII&PIH

PII: Personal (Identifying) Information

- Explicit (usual list in data protection *law*
- Implicit (anything you can infer things about a person from)
- Data mined to target ads and for analytics

PIH – repository under control of subject –

- Privacy preserving – various techniques mostly decentralization & cryptographic, but more subtle too
- Talks Tuesday about system, storage, privacy & business models
- (n.b. have other new project on cloud/legal MCCRC)

T1.2: PIH Architecture

Provide **secure, privacy-preserving service domains** within which application code can execute against users' data.

Provide **private provenance-aware storage** where data is encrypted and selectively exposed with verifiable privacy properties.

T1.2: PIH Architecture

Hardware prototype: Cubieboard2

- ARM Cortex A20, costs £49.

Software for private service domains:

- Xen 4.4 released with ARM support.
- Mirage/ARM works in userspace, and almost kernel mode (a few weeks).

T1.2: PIH Architecture

Software libraries Mirage for UCN being developed by Cambridge and Nottingham:

- Web social network hooks for data retrieval (e.g. Facebook and Twitter)
- IMAP and SMTP services for e-mail use-cases, released as open-source. XMPP ongoing.
- Selected for Google Summer of Code 2014.

T1.2: PIH Architecture

Provide **secure, privacy-preserving service domains** within which application code can execute against users' data.

Provide **private provenance-aware storage** where data is encrypted and selectively exposed with verifiable privacy properties.

T1.2: PIH Architecture

Storage efforts have been focussed on **Irminsule**, the distributed branch-consistent database.

- *Tutorial:* https://samoht.github.com/irminsule
- *Status:* interops with Git, memory/disk backends.
- *Next steps:* convergent encryption, integrate with real applications, compile to JavaScript.

T1.2: Deliverables

T1.1: PIH Requirements M1-M9

Partners involved: CAM, NOTT, NICTA, TSA, EURC, PTIN

Through interaction with WP2, in which data will be gathered concerning existing user experience, understanding and conception of privacy, requirements for the PIH architecture will be articulated and developed. Formally linking this work to that of WP2 is required due to people's (users') complex and shifting conceptions of privacy. As has been argued [Brundell11] it is all too easy to build infrastructure technology that unwittingly prevents subsequent development of suitable user-facing abstractions. By incorporating user considerations from the start we will ensure that these problems are avoided, and that the systems we build are designed to be fit for purpose from the start.

D1.2 @ M18 PIH Architecture, D1.3 @ M33 PIH Implementation

Possible Business Case Requirements ☺

1. P2p. mutual, no money at all:)
2. government pays
3. subscription, paid by users
4. two-sided market paid by adverts, not users
5. three sided market, paid by adverts and service providers analytics
6. ...add yours here

Therefore need…

Crypted storage

Possibly decentralized storage

Capability for access

Either diff priv or FHC or GC operations

=> create Provenance trail & FHC/GC/Summaries
at update time + update provenance at access time
->see Irminsule later…

Thus...

. a fully decentralized system (pushed to the very edge of the ne
nterconnected by secure social channels, perhaps using signpost for setu
nd key distribution etc with distributed differential privacy as search/analytic
PI

. another extreme is fully centralized personal data, but partitioned purely b
rypto techniques

. 1+2 hybrid, a model where the user pushes some data to some partial
entralized data stores (perhaps scrubbed in a fuzzing/differential privac
anner) so that a) its backed up/ b) has more availability/performance and c
an have central analytics run on it...

. central storage but decentralized processing one can take the full edg
ystem, and relax slightly by having a per-user proxy for every edge user in
cloud" as a personal mirror....this is really a no-op --well, client side crypt
as risks

More tomorrow…

Or over dinner ☺

Oh, as well as MCCRC, also have HAT as use case