The storage dilemma

Our personal data, once centralised on a local PC, is getting increasingly federated across many devices and online services. This, amongst many benefits, leads to a number of issues:

- challenging to maintain an overview over an archive of digital assets,
- personal data may be controlled by intermediating parties (platform providers),
- data provenance is uncontrolled and unmonitored,
- uncontrolled sharing can lead to privacy violations.

A solution: Nimbus

- Provides a unified way of accessing federated data in a transparent and automatic system.
- Exposes plugin API that front ends and actuators can be built upon.

Existing solution attempts

- Personal databases [1, 2] – Schema problem
- Synchronisation services [3, 4] – No cross-platform standards
- Controlled environments – Unrealistic outside high-security settings
- Secure wide-area filesystems [5, 6] – Require global deployment

Fountains of trust

A simple, flat trust architecture

Individuals or organisations run trust fountains

- permanently active servers
- providers of cryptographic hashes that can be used for attestation.

Unlike PGP signatures, no trust hierarchy required.

- Single, not multi-hop trust paths.
- Can choose appropriate signatories as appropriate.
- No acknowledgement or action required from signatory.

Unidirectional operation

HTTP GET

\[ h(K_t, t) \]

e.g. to attest communication.

Bidirectional operation

HTTP POST

\[ h(K_t, t, h(F_t, t)) \]

e.g. to attest data item ownership.

Verifying an attestation

- Log of hashes on append-only storage with fountain owner.
- To verify, ask the fountain owner to confirm a hash stored with a data item (N.B. deniability for the fountain owner).

Fuzzy provenance

and how it is useful

Relaxed scope: no requirement for full forward provenance.

- Traditional provenance requires us to know where every data item is (or has been) at any point in time.
- Assume that copies are only dangerous if used.

Can always find out where a data item came from, but not necessarily where it went.

Combining the two... we can get provenance tracking based on trust fountain attestation primitives.

Advantages

- Decouples provenance and data: can still prove origin even if provenance information removed or lost.
- Makes provenance tractable for uncontrolled systems.