Identity Management, Aadhar and Replicated State Machines

Emergent Trust in a Networked World inspired by Tinker Bell

A VARA Perspective
Dr. Partha Das Chowdhury

Background and Threat Model

Authentication, authorisation and audit are three traditional concerns in building a privilege management infrastructure. Traditionally, authentication is strong and is based on fixed credentials linked to long term stable identity; and audit is linked to authorization via the same fixed credential. Security literature, traditionally has the good principals Alice and Bob with Carol as a system and/or service owner. The conventional approach by Alice or Bob to authenticate Carol using fixed credentials linked to long term stable identity led to compulsive trust relationships between the legitimate users and large parts of the system infrastructure. Trust became a substitute rather a way out or a cover for certain knowledge, which is difficult to gather or share, thus forcing participants involved in any such interaction (with an unknown) to a compulsive relationship called Trust relationship.

Furthermore in shared systems we have no control over the participants who would be using the system and the network, threats can arise from insiders as we have seen in numerous instances; moreover system domains cannot also neatly map into administrative domains and there would be shared resources between internal and even with external domains which are otherwise in a different system/security boundary. The obvious consequences of this information asymmetry led to the manifestation of threats like

- **Identity Theft** – Stealing identities to access privileges is a billion dollar industry now. We have seen consequences where it has been nearly impossible to restore stolen identity to the legitimate holder of the long term stable identity.
- **Threats to Individual Privacy** – We now live in a world which is the ultimate Panopticon where each and every individual can be observed without them knowing when they are being observed and when not.
- **Non-repudiation** – It is indeed hard to gather evidence in the electronic world, a defendant can ensure that an instance of semantic communication between computer systems leaves behind no unequivocal evidence of its having taken place. Litigation are lengthy and expensive and often without any consequence.
- **Loss of Confidentiality** – Large corporations store and process information and are bound by regulations and laws within which they operate, for example GDPR in the Europe. However, massive data breaches are common as is sharing of data by corporations with external entities. Such sharing is often without consent and is not as per the regulatory framework.

(We make a distinction between confidentiality and privacy here; by confidentiality we mean protecting someone else's data where as by privacy we mean our own security policy w.r.t. our own data.)

For authorization we ascertain 'unambiguous and verfiable' bindings between bits (such as a name, a cryptographic key, or a program text) and a real-world entity (such as a person, a smart card, or a process running on a particular machine.) A bit pattern can be freely copied and modified in cyberspace; whereas entities of the second type have provenance in the real world. Our experience of security failures of the last few decades tell us that it is indeed a hard problem to uniquely identify a real world entity in the cyber space and then ensure that Alice is indeed speaking to Bob when she thinks she is speaking to Bob. Conventional mechanisms to prevent unauthorized users of the system from masquerading as another range from passwords, certificates, tickets and more recently the use of biometrics.

Aadhar

Aadhar, is supposedly the largest program of its kind to identify around a billion population uniquely and then ensure that delivery of subsidies, benefits and other services reach in a transparent and fair manner. The program involves capturing the biometric data of the subscribers and storing them in a secured manner and the subsequently allow authentication over public networks while subsidies and benefits are being disbursed to prevent Eve from claiming Alice's subsidy.

The scheme was subsequently expanded to other areas as well like telecommunications, driving licenses and other areas which did not involve benefits or subsidies.

There were public interest litigation filed at the Supreme Court of India and subsequently the Supreme Court while exempting Aadhar from non welfare schemes have made verification of Aadhar mandatory for welfare schemes, subsidies and benefits. However Aadhar is not mandatory for availing a SIM card or opening a bank account or appearing for exams. The Supreme Court is of the opinion that with minimalistic biometric authentication it will be difficult to profile an individual; so Aadhar will not violate the right to privacy. Aadhar is perceived to be the foundation of a privilege management infrastructure in Indian governance and hence a critical enabling factor.

The singular most important reason underlying this perception is that biometric data cannot be copied or modified; thus Alice can be sure that she is speaking to Bob and not Eve. There are other underlying mathematical assumptions like the Birthday paradox and other social engineering attacks about which the Supreme Court of India with all its wisdom was 'technically' satisfied before mandating Aadhar for access to all welfare and benefits.

Distributed Ledgers and Replicated State Machines

The idea of append only ledgers existed around 25 years back where as a distributed ledger existed around the same time for PKI systems.

However what is of interest from a security and integrity point of view are the programmable replicated state machines which is roughly three years old; Block chain brings in a consensus mechanism that cannot be tampered with. Though a fair decentralised consensus is difficult in an open environment, however for closed environments one can achieve decentralised fair consensus.

It is worth mentioning here that in the original paper Nakamoto mentions the Bitcoin consensus as a weaker form of consensus because his contention was that the incentives are more for the attacker to keep the system going rather than creating a fork.

A number of propositions to leverage the property of consensus in the domain of identity management has evolved since. The primary reasons behind such efforts are reduction of costs with repeated customer/user on-boarding, introducing fairness in the system and preventing any unauthorized modification; while at the same time allowing cross domain verification of credentials.

The Way Forward – Tinker Bell Model of Trust for Identity Management

The computer science research community have always advocated for a localization of trust relationships where the user being the basic modelling unit of systems plays a critical role in establishment of trust relationship in a bottom up approach.

In the play Peter Pan, the fairy Tinker Bell was about to die since nobody believed in her any longer, but is saved by the belief of the audience. Gods in ancient Greece drew their power from how many mortals sacrificed to them. This is more democratic and follows a social consensus.

The rationale for block chain is exactly this, truth arrives through a consensus and not through favoured pawns as has been the case with distributed systems even with a decentralised authority. Thus an identity management platform to replace traditional CAs where identity is verified through a consensus ensures trust flows from below rather top down.

The theoretical foundation as well as the practical reasoning behind the adoption of consensus for identity management is that membership could be granted and revoked by a consensus of the existing members and can be stored in the block chain once that consensus is reached.

So the membership is visible and as long as the starting membership set is greater than one(so that no one is special) a fair consensus is possible.

Moreover the consensus on the possible set of members is integral to the future extension of the members making adding new members through malice difficult. Even then mathematicians might argue that since all distributed consensus algorithms share the old Byzantine general's problem; one can delay progress by partitioning the network at just the right time (Arrow's impossibility theorem).

However in practice we have become exceedingly good these days at being on the internet. Regular long term episodes of non connectedness of the internet does not happen in practice but only in text books on consensus. Given the theoretical understanding accrued over decades over research into state machines and consensus algorithms one can safely bet on the applicability of consensus on identity on-boarding, storage and subsequent verification.

Our contention is to Aadhar within a consensus framework will do better in meeting the legitimate needs of the various stakeholders like the user, regulators, welfare providers, existing block members and others.

- Aadhar will continue to be the enabler for welfare and subsidies. However onboarding based on consensus of various stakeholders (issuers of various fixed credentials) will add to the security and prevent against masquerading.
- It will not be possible to introduce inconsistencies by partitioning sets. Any future propagation will be dependent on the consensus. The consensus threshold can be slided based on the security requirements and policies.
- There is no unauthorized exposure to sensitive information. Fingerprints and unique information are not exposed through the chain thus preventing aggregating and inference attacks.
- Right to Access as mandated by the regulators for example GDPR guidelines can be implemented in a transparent manner.