Digital Identity

Identity, Security, and Data Provenance

IARIA Annual Congress 2022
Frontiers in Science, Technology, Services, and Applications
Nice, France, July 22-28, 2022
(https://www.iaria.org/conferences2022/IARIACongress22.html)

Stan McClellan, PhD
Texas State University, San Marcos TX, USA
Abstract

• Digital Identity is an abiding problem

• Data-Centric Protection:
  - Can augment conventional AAA
  - Can provide context-sensitive policy
  - Is compatible with Zero-Trust Architectures
  - Can provide Digital Identity
Stan McClellan

Professional Experience
- Co-Director, Connected Infrastructure Initiative (CIEDAR), Texas State University
- Professor, Ingram School of Engineering, Texas State University (2008 – Present)
- Director, Ingram School of Engineering, Texas State University (2013 – 2018)
- CTO & co-Founder, Power Tagging Technologies (2008-2010)

Publications & Activities
- Smart Cities in Application: Healthcare, Policy, and Innovation. Springer. 2019
Basic Agenda

• Background
  – Classical AAA
  – Contemporary Approaches

• Problems
  – High Profile
  – Constant Failure

• (re)Definition
  – Phases & Principles
  – Use Cases & Comparison

• Possible Outcome
  – Protected Data
  – Zero-Trust Architecture
Classical Authentication

- **What I know**
  - Password, Challenge/Response, etc.
- **What I have**
  - Access Card, USB Dongle, etc.
- **What I am**
  - Fingerprint, Retina Scan, etc.
Classical AAA

- **Authentication**
  - Are you who you say you are?
  - Exercises “Know / Have / Am” of classical authentication

- **Authorization**
  - Should you have access to this data?
  - Typically via access control lists (ACL) and user databases

- **Accounting**
  - Access for how long, and in what fashion?
  - Most often used for billing purposes and audit trails
Approaches: Technologies

- **Network-based**
  - TACACS/+ 
  - RADIUS (RFC-2865 et.al.) 
  - DIAMETER (RFC-6733, et.al.)

- **Person-based**
  - Self-Sovereign Identity (SSI) 
  - Decentralized Identifiers (DID) 
  - e.g. EU “ESSIF” per eIDAS

- **Application-based**
  - SSL/TLS (RFC-8446): encryption 
  - OAuth2: constrained delegation of access to applications 
  - UMA: user-managed access, extensions of OAuth 
  - FIDO2 (WebAuthn, CTAP2.x): client-to-authenticator protocol 
  - OpenID/FAPI: decentralized attestation
Approaches: Companies

Large Companies

- Okta (https://www.okta.com)
- Docusign (https://www.docusign.com/products/identify)

Small Companies

- ImageWare (https://imageware.io)
- Mitek (https://www.miteksystems.com)
- Vouched (https://www.vouched.id)
- Trulioo (https://www.trulioo.com)
- iComply (https://icomplyis.com)
- InCode (https://info.incode.com)
- TeleSign (https://www.telesign.com)

There are a bunch of them ... The market is crowded and growing
By 2023, at least 80% of government services that require citizen authentication will support access through multiple digital identity providers.

By 2024, at least a third of national governments and half of U.S. states will offer citizens mobile-based identity wallets.

Only a minority will be interoperable across sectors and jurisdictions.

This is a problem
Defining a “Digital Identity”

• Bundle of identifying attributes and data
  - Discrete, secure, self-contained, extensible ("atomic")

• Authentication + Authorization
  - Uniquely identifies the entity to which it belongs

• Portable
  - Can be sent to insecure location via insecure network
Not Digital Identity

• National/Civil Identity
  - Passport, Driver License, Social Security, etc.

• Online Identity
  - Breadcrumbs, purchase history, public information, etc.

• Computer Identity
  - Usernames, passwords, encryption keys, etc.

• Encryption (!)
## Digital ID vs. Encryption

<table>
<thead>
<tr>
<th>Function</th>
<th>Encryption</th>
<th>Digital ID</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Support multiple algorithms (e.g. AES-128)</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Support multiple keys per user or instance</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Partial decryption / partial disclosure</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td><strong>Detection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interval (dates, times)</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Locations (geo, network, system)</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Attempt tracking (number, lockout)</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td><strong>Countermeasure</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Notification (of owner – email, text, etc.)</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Escalate (new &amp; stricter challenges, etc.)</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Self-Destruct</td>
<td>N</td>
<td>Y</td>
</tr>
</tbody>
</table>
Phases of Identity (C.Allen)

- Centralized (unitary)
  - Single administrative authority
- Federated (multi-central / oligarchy)
  - Multiple administrative authorities, federated
- User-Centric (multi-central / individual)
  - Multiple administrative authorities, federated
- Self-Sovereign (non-central)
  - Individual control regardless of authorities
Ten Principles of Identity (C.Allen)

1. Exist Independently
2. User Control
3. Self-Owned Data
4. System Transparency
5. Persistence
6. Transportable
7. Widely Used
8. User Consent
9. Minimal Disclosure
10. Protection of Rights
EcoSystem is Mandatory

• Creation
  - Created and owned by the entity identified
  - More than one ID per entity (many to one)
  - Identifying data provided at creation (schema)
  - Requires secure, validated “writer” to ingest data, create bundle

• Usage
  - Network needed to share and for some countermeasures
  - May require centralized management (ala PKI?)
  - Requires secure, validated “reader” to ingest bundle, validate access
    Must be cross-sector and cross-jurisdiction, and linked to valuable use-cases
Blockchain is not Identity

• Myths
  – Use Blockchain as a database to store personally identifying information (PII)
  – Use Blockchain as a distributed hash table (DHT) for PII data that is stored off-chain

• Reality
  – Blockchain is transparent, immutable, reliable and auditable
  – It can be used in the secure exchange of cryptographic keys … e.g. PKI not PII
  – This may be a step toward decentralized public key infrastructure (PKI) which can lead to management of PII
High Profile, High Cost

• April 2021
  - UN Data Breach
    • Fraudulent credentials allow access to sensitive data

• January 2022
  - Okta Identity Management compromised by Lapsus$
    • 2.5% of customers data “may have been viewed or acted upon”
    • https://www.wired.com/story/okta-hack-customers-lapsus-breach/
  - Tesla cars compromised by German researcher
    • Bug in open source logging tool exposed cars directly to the internet
    • https://techcrunch.com/2022/01/24/teslamate-bug-teslas-exposed-remote/

• July 2022
  - MICODUS GPS Tracker compromised by Bitsight
    • Exploit tracks and remotely manipulates “at least a million vehicles”
    • https://techcrunch.com/2022/07/19/micodus-gps-tracker-exposing-vehicle-locations

• [T]raditional protections just aren’t working …
• [T]he solution is actually quite simple: Protect the data itself
Key: Zero Trust Architecture

- Everything is a resource. All resources can present a threat.
- All communication is secured, regardless of location.
- Access to a resource is on a per-person basis, with minimal privilege granted.
- Access policies are dynamic, and based on telemetry.
- All assets are monitored. No asset is inherently trusted.
- Authentication and authorization are enforced per-resource, requiring identity, credential, access, and asset management.
- Telemetry of access requests and asset state is used for continual improvement

(per NIST SP 800-207)
Use Case: Supply Chain

• Problem
  - Layered security model doesn’t work
  - Boundary, Network, System, File … easily exploited

• Approach
  - *Data-Centric Protection*
  - Augments the layered enterprise security model
  - Built-in policy-based tracking and protection

• Results
  - Intelligent data self-enforces protection policies
  - Self-destruct, invoke different access procedures, call-home, honeypot, etc.
Use-Case: IT/OT Convergence

• Problem
  – Endpoints are small, remote, with limited CPU and memory
  – Battery-powered devices conserve energy by “sleeping”
  – Data may transit unknown networks from insecure locations

• Approach
  – Secure the data at the source before transmission
  – Track the data when it is received and utilized via enclosed policies

• Result
  – Independence from incompetent device manufacturers
  – Independence from insecure intervening networks and paths
  – Policy-driven visibility for all activities, states, and locations of the data itself
Intelligent Data

- Time Based
  - Unauthorized Scheduled Period
  - Unauthorized Time of Day
  - Unauthorized Day of Week
  - Unauthorized Geo Location
  - Unauthorized Network
  - Unauthorized Hardware System
  - Exceeded Allowed Access Count
  - Multiple Failed Access Attempts

- Environment Based

- Action Based

Access violations trigger self-managed mitigating actions by UXP Policy Engine

- Policy Violations
  - Text/Email Home
  - Self-Destruct (Delete)
  - Digital Shred
  - Present Honeypot
  - Escalate Multi-Factor Auth. Level
  - Deny Access

UXP

July 2022
Data-Centric Protection

• Adheres to:
  – Principles of Zero-Trust Architecture
  – Conventional AAA principles

• Is not:
  – BlockChain, but can be co-implemented
  – Encryption, but depends on it

• Provides:
  – Use-Case-Aware security
  – Context-sensitive policy

• Implements:
  – Digital Identity
Thank You!

• Useful References
  - Decentralized Identifiers (DIDs) v1.0. W3C Recommendation. July 2022. https://www.w3.org/TR/did-core/