



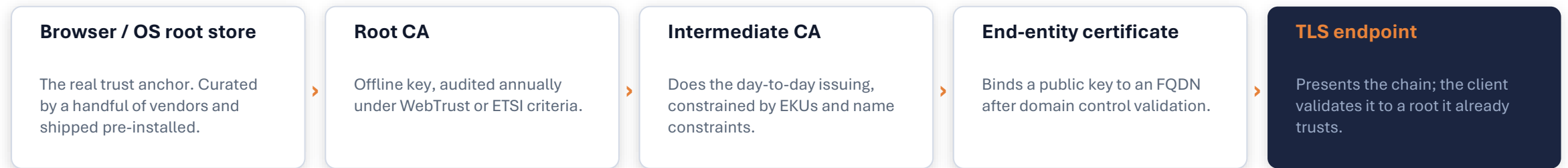
PKI in a World of Confidential Computing and AI Agents

Shalabh Jain
Center of Data for Public Good
Indian Institute of Science

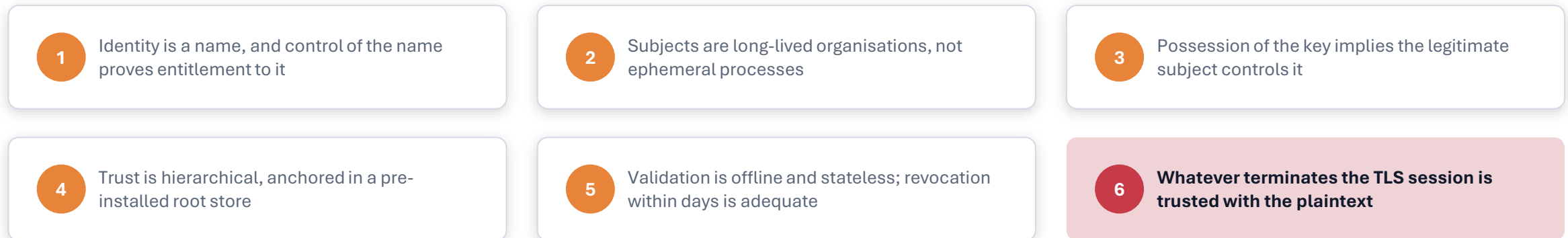
PKI in a World of CC and Agents

Introduction

The WebPKI works extremely well for the problem it was designed to solve.



Relies on six load bearing assumptions



Our computation paradigm is fundamentally changing...

Confidential Computing

Introduction



Confidential Computing

Introduction

Local compute for critical functions



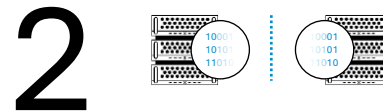
Outsourced compute on the cloud

Creates a need for data and compute sovereignty



SECURE DURING TRANSPORT

Components are protected during transfer from the client to the compute infrastructure



SECURITY AT REST

The security relevant components are protected while stored in storage-lakes



PROTECTED IN USE

The privacy relevant components of the system are protected during runtime

Provides the ability to exercise sovereignty

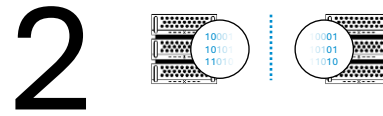
Confidential Computing

Introduction



SECURE DURING TRANSPORT

Components are protected during transfer from the client to the compute infrastructure



SECURITY AT REST

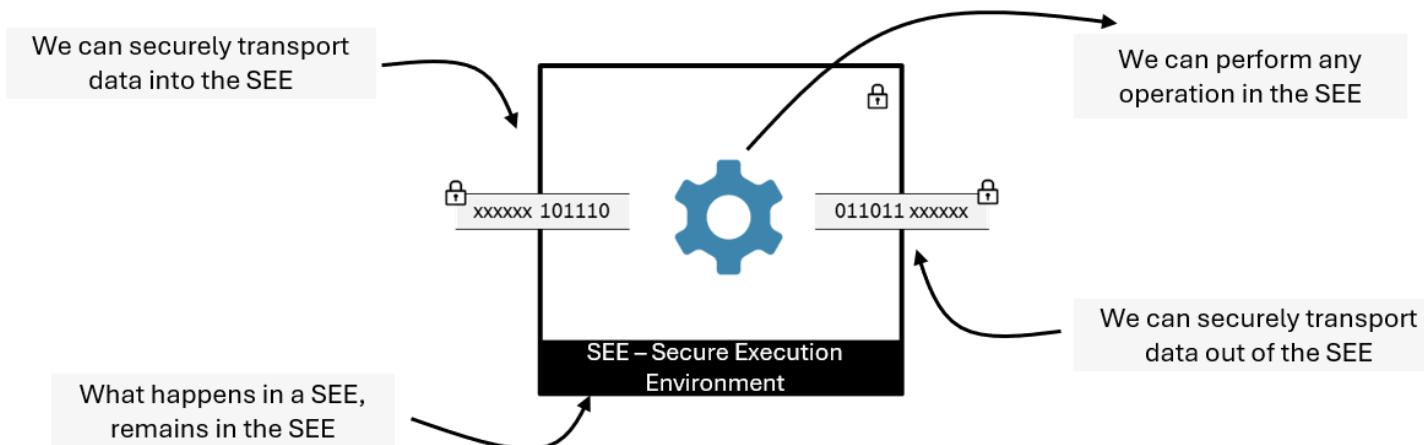
The security relevant components are protected while stored in storage-lakes



PROTECTED IN USE

The privacy relevant components of the system are protected during runtime

Provides the ability to exercise sovereignty

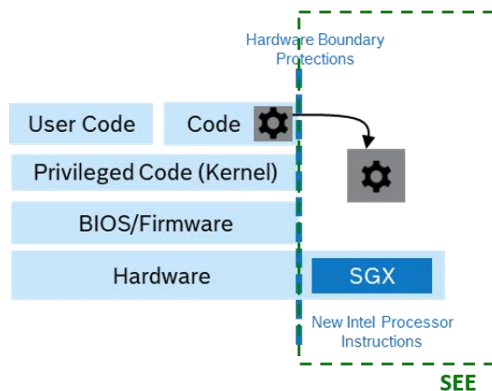


Confidential Computing

Flavours of PETs

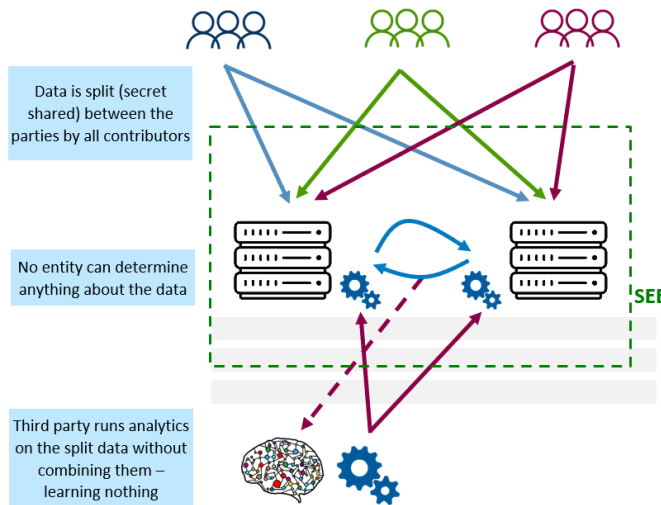
Trusted Execution Environments

TEE



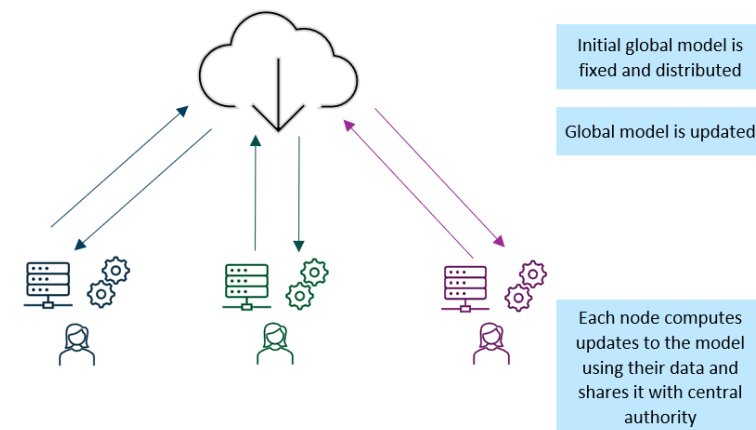
Secure Multiparty Computation

MPC



Federated Learning

FL

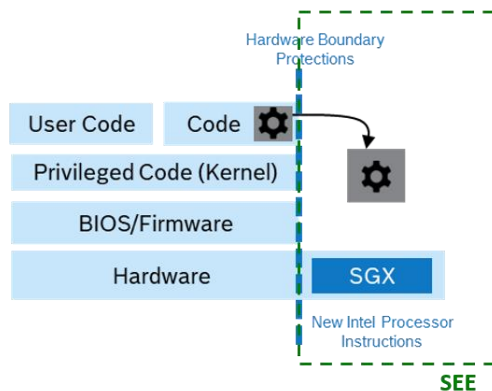


Confidential Computing

Flavours of PETs

Trusted Execution Environments

TEE



Secure Multiparty Computation

MPC

Federated Learning

FL

From the perspective of PKI, hardware based TEEs – Confidential Compute



Confidential Computing

TEE using AMD SEV-SNP

Hardware support for

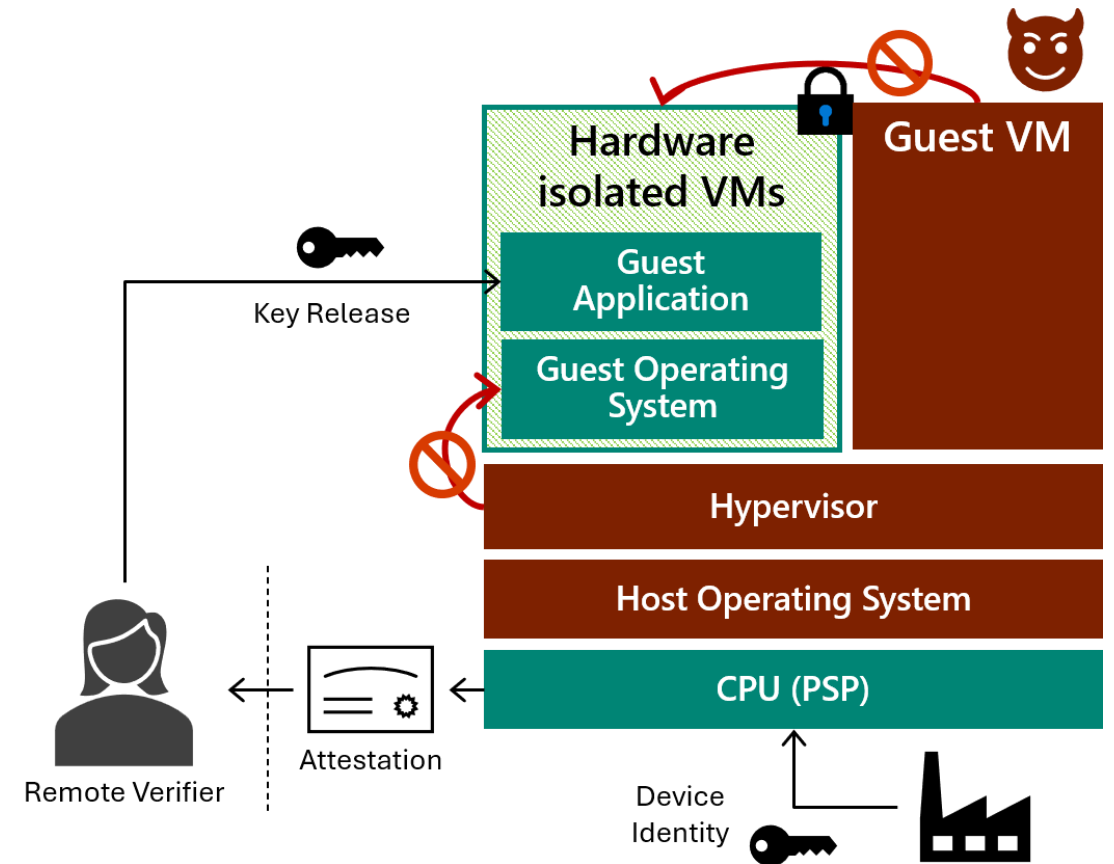
- » Security boundary at the VM level – aligns with existing programming model
- » Confidentiality and Integrity guarantees

Remotely verifiable state

- » Hardware supported process called remote attestation

Several challenges

- » Very large TCB
- » Freshness/replay guarantees lost



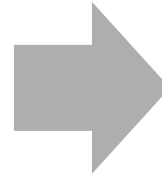
Confidential Computing

Trust stems from a different question

Classical PKI

“Who owns this key?”

- » Identity is a DNS name or a vetted legal entity
- » Trust is established once, then cached for months
- » The CA never inspects the code holding the key



Confidential Computing

What code, on what hardware, in what state, holds this key?”

- » Identity is a measurement — an opaque hash of code and configuration
- » Trust is point-in-time and must be re-established continuously
- » The relying party must appraise hardware, firmware and TCB version

Confidential Computing

Four assumptions that break

1

Identity is a human-registered name

A TLS certificate binds a key to an FQDN after the CA checks domain control. A workload's meaningful identity is a hash of its code and configuration.

2

Subjects are long-lived

Certificates last up to 398 days today. Enclaves live seconds to minutes and are re-created constantly. Human-in-the-loop vetting cannot run at that cadence.

3

The CA vets the wrong property

CA/Browser Forum validation confirms organisational and domain control — not that the private key lives in genuine, up-to-date, unbroken TEE hardware running the expected binary.

4

Revocation is assumed to be slow but adequate

CRL and OCSP latency is measured in hours or days. An enclave that is compromised mid-session needs trust withdrawn in seconds.

THE ATTACK IT ENABLES

A cloud operator can hold a legitimate certificate for the domain and still terminate the TLS session outside the enclave — relaying traffic to code the client never approved.

Consequence

Attestation must be bound into the credential itself, not presented alongside it — and the binding has to survive an adversary who controls the network, the host, and a valid domain certificate.

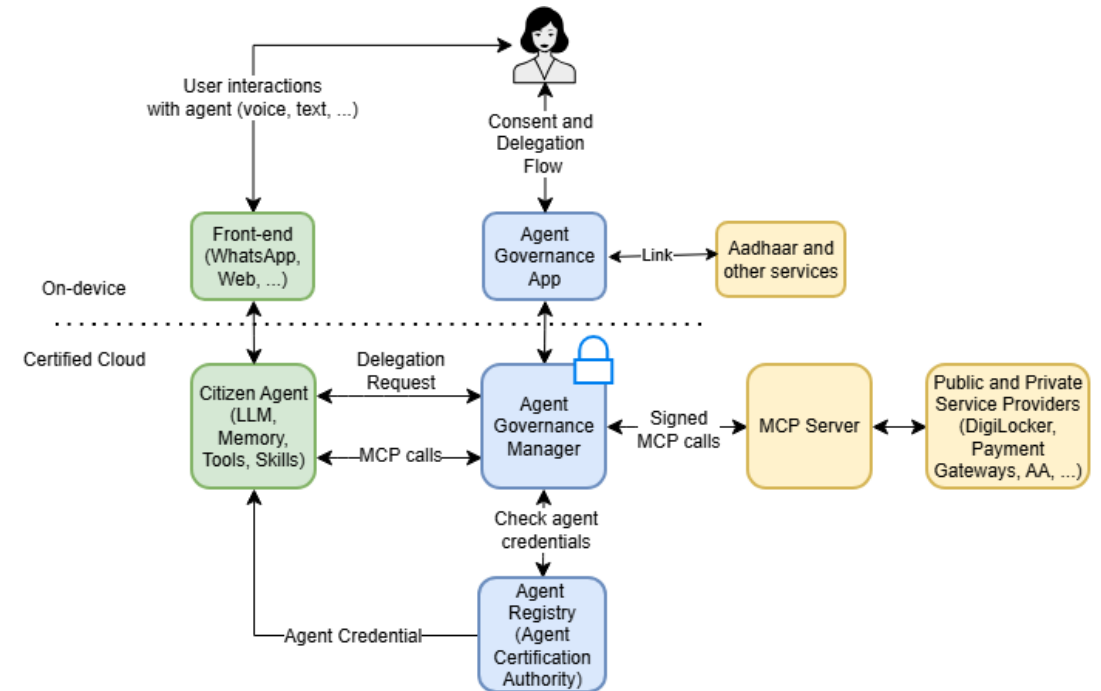
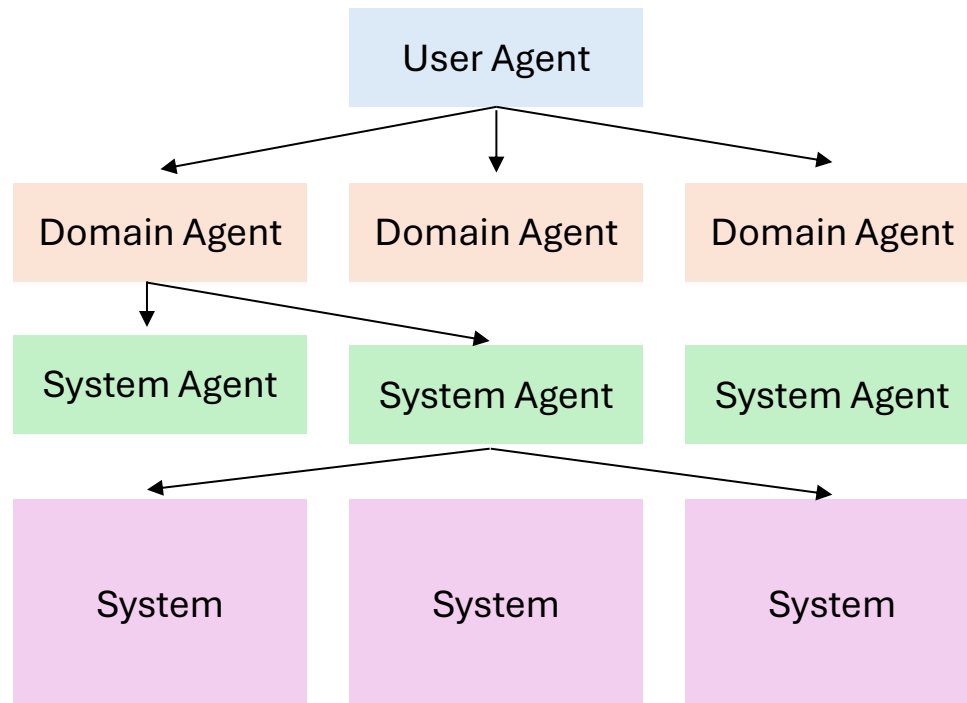
Our computation paradigm is fundamentally changing...

second big change...

AI Agents in Computing

Introduction

Agentic AI ecosystems – Computation is no longer performed by humans, rather by agents



AI Agents in Computing

Four assumptions that break



Subject = human, org, or domain

Certificates bind a key to a domain (DV) or a vetted legal entity (OV/EV). There is no notion of an "agent instance."



Long-lived, stable identities

Certs historically lasted 1–2 years; the subject (a server, a company) is expected to persist and rarely change.



Hierarchical, transitive trust

Trust chains from a root CA in a trust store down to the leaf. Relying parties trust anything that chains to a root.



Revocation via CRL / OCSP

Compromise is handled by publishing revocation — a model that already scales poorly and is being deprecated (OCSP).

AI Agents in Computing

Identity is insufficient

» Authentication

Proving this workload / model / agent is what it claims. Best served today by attested workload identity (SPIFFE SVIDs).

Addressed by current
PKI certificates

» Delegation

Proving the chain of authority and that each hop only narrows it. PKI has no first-class delegation primitive.

» Authorization

Fine-grained, contextual, real-time-revocable permission — not coarse static roles baked into a long-lived cert.

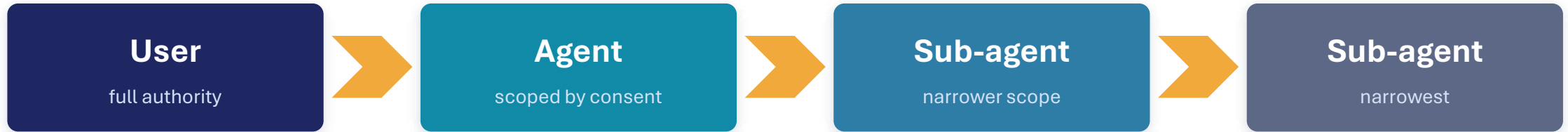
» Attestation

Binding the cryptographic identity to the actual running code, model, and policy (RATS, TPM/TEE, supply-chain provenance).

Brings in Confidential
Compute issues

AI Agents in Computing

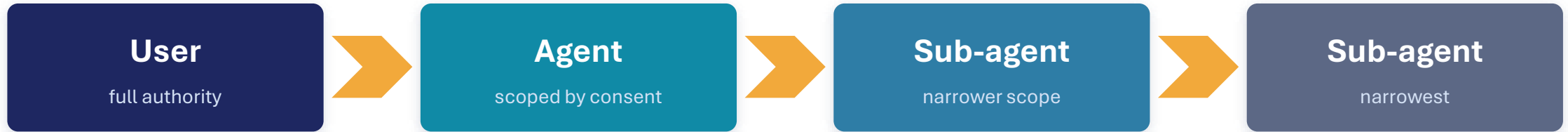
Chained Delegation



Scope changes at every hop – no way of achieving this at a granular level under current PKI

AI Agents in Computing

Chained Delegation



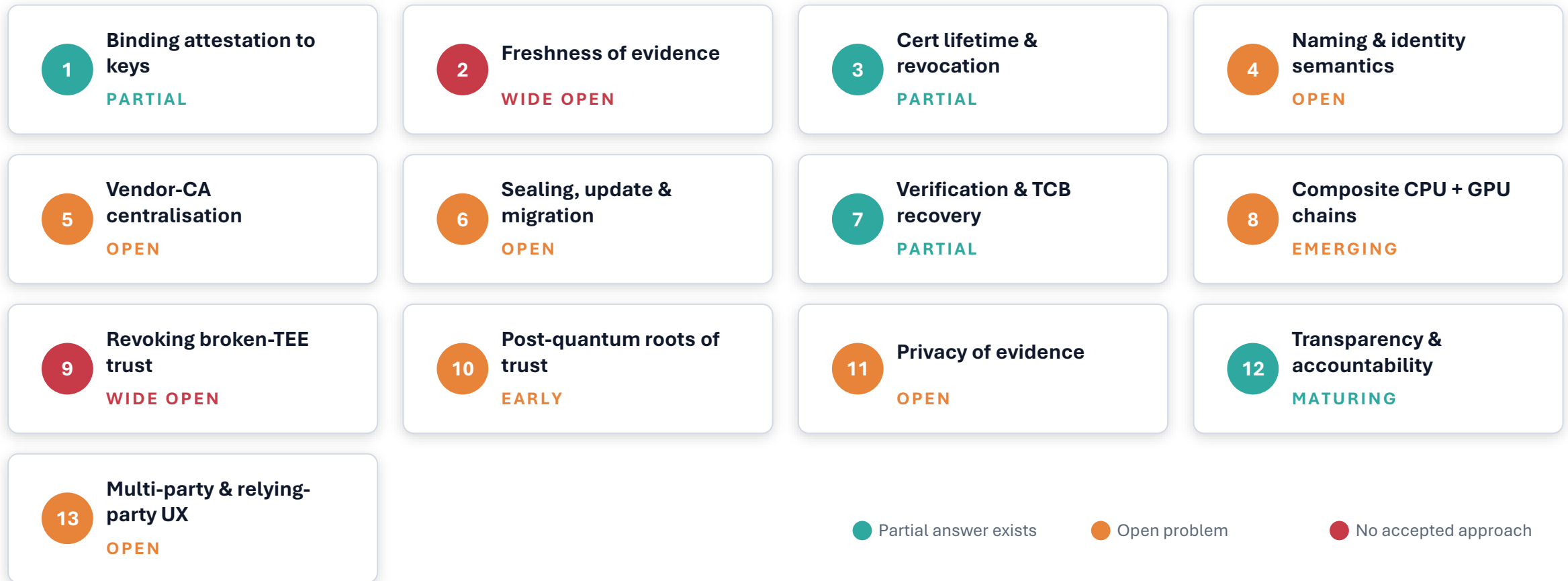
Scope changes at every hop – no way of achieving this at a granular level under current PKI

Strong need of auditability and non-repudiation

Going back to confidential computing..

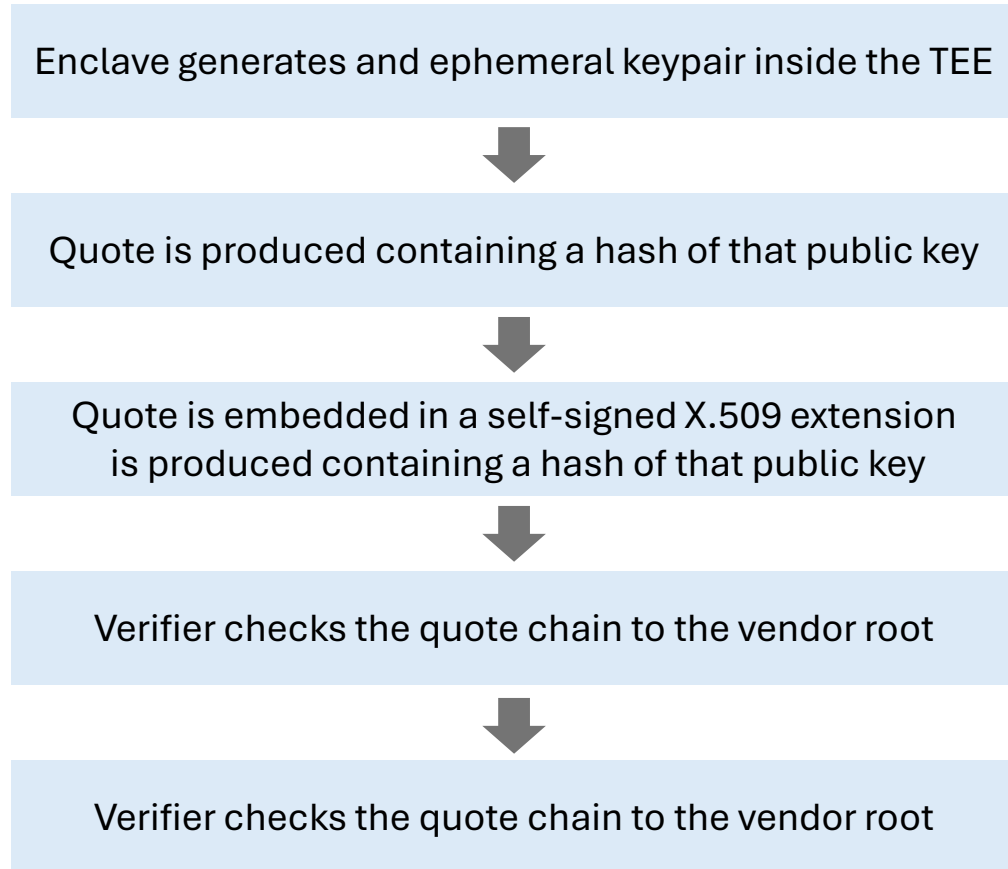
PKI in a World of CC and Agents

Gaps to be addressed



PKI in a World of CC and Agents

Binding attestation evidence to keys

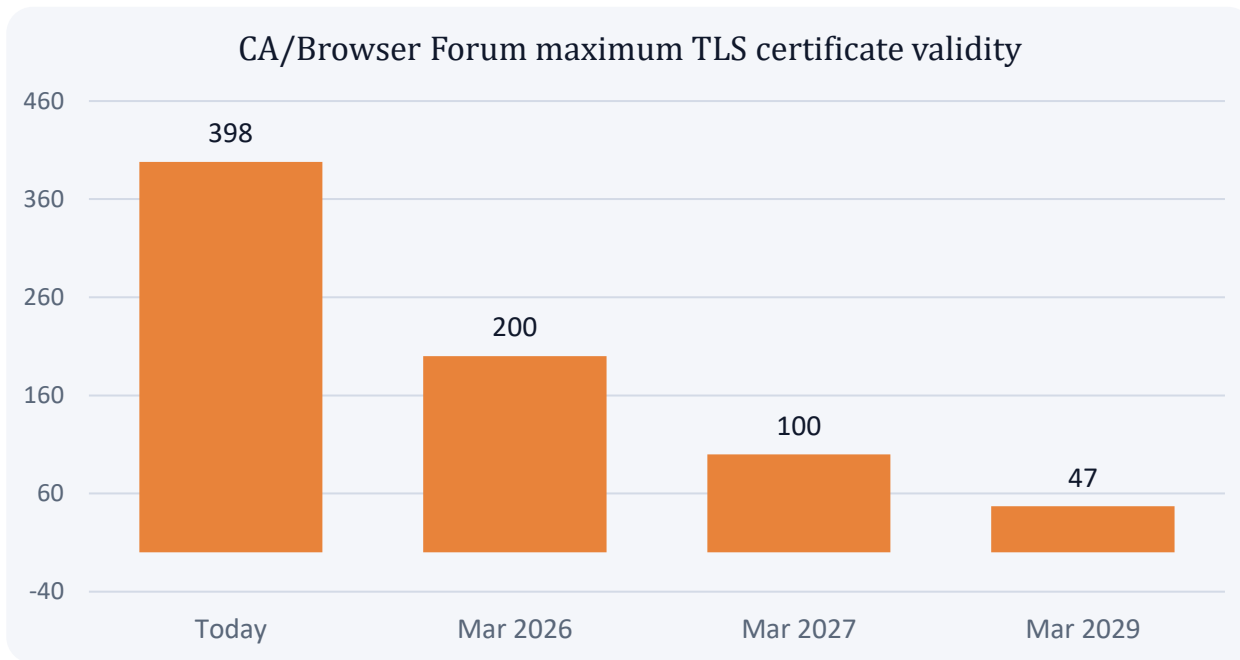


Two incompatible families

- » Modify the TLS handshake to carry evidence - **breaks legacy clients**
- » Embed evidence in certificates - **breaks compatibility with existing CAs**, so also breaks legacy clients

PKI in a World of CC and Agents

Freshness, lifetime and revocation



Even 47 days is too long for an enclave

- » Confidential computing goes beyond past short-lived certificates toward attestation-gated, single-use credentials
- » Credentials issued at machine speed - expiry is so fast that revocation becomes moot,
- » Backed by continuous re-attestation.

PKI in a World of CC and Agents

Naming and identity semantics

Even aspects such as the name of the subject in a cert requires rework

Measurements

- » SGX MRENCLAVE / MRSIGNER
- » TDX MRTD plus RTMRs
- » SEV-SNP launch measurement
- » NVIDIA firmware and VBIOS digests

Precise, but opaque
and brittle — every
legitimate update
changes the identity.

DNS names

- » Legacy interoperability
- » Meaningful to humans
- » Consumable by any TLS client
- » Backed by existing CA processes

Understandable, but
says nothing
whatsoever about the
code holding the key.

SPIFFE IDs

- » Two-step node then workload attestation
- » TDX node attestor: MRTD plus UUID
- » Hybrid attestor: SEV-SNP plus cloud selectors
- » Keylime plugin ties SVIDs to TPM state

PKI in a World of CC and Agents

Hardware vendors are de facto certificate authorities

Certificates are rooted in Hardware Vendor

Intel

PCS / PCCS distribute PCK certificates and TCB info. Trust Authority is an operator-independent verifier SaaS (GA September 2023).

AMD

KDS issues VCEK (per-chip, per-TCB) and VLEK (per-CSP) certificates, chaining ARK → ASK → VCEK.

Arm & NVIDIA

CCA platform endorsements; NVIDIA per-device keys fused at manufacture, chaining to an NVIDIA CA and NRAS.

Significant risks – Availability, Jurisdiction, Offline USe

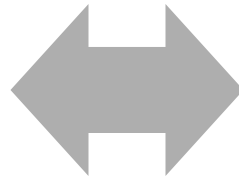
PKI in a World of CC and Agents

Privacy and Transparency

Privacy can be a concern

Unlinkability is desired

- » SGX's original EPID used group signatures for unlinkability; the shift to ECDSA/DCAP traded anonymity for manageability and exposed per-platform PCK and VCEK identities
- » Direct Anonymous Attestation and shared attestation keys anonymise individual devices — but you lose the ability to revoke a single key
- » LAMPS warns that publishing evidence in issued certificates leaks hardware detail about the end entity, and is NOT RECOMMENDED



Transparency

Public verifiability is also desired

- » Certificate Transparency's insight — make misbehaviour visible rather than impossible — is being ported to attestation and to binaries
- » Apple Private Cloud Compute publishes every production image's measurements to a public append-only log, and devices refuse to encrypt to unlogged software
- » Sigstore/Rekor, in-toto, SLSA and IETF SCITT extend the same idea to build provenance and supply-chain claims

Zero-knowledge attestation: prove that a valid quote exists for an approved measurement without revealing which platform produced it.

PKI in a World of CC and Agents

Summary

Attestation-aware PKI profiles

- » Standard X.509 profiles and CSR flows that carry — or reference — evidence and attestation results, with explicit privacy rules. Open question: how to reference rather than embed, to avoid leaking hardware detail while keeping legacy clients working.

Freshness protocols

- » Reconcile point-in-time evidence with certificate-lifetime trust. Nonce-based challenge/response, epoch and heartbeat re-attestation, and binding freshness to key lifetime. Formalise how fresh is fresh enough.

Vendor-independent and threshold verification

- » Multi-verifier quorum, threshold attestation verification and independent reference-value providers, so no single silicon vendor is a global single point of trust or failure.

Open-source silicon roots of trust

- » OpenTitan — production silicon fabrication began February 2025 — and Caliptra as transparent, auditable anchors that reduce the “trust me” content of fused vendor keys.

Privacy-preserving attestation with ZK proofs

- » Prove a valid quote exists for an approved measurement without revealing platform identity. zkVM-based DCAP verification is an early instantiation; combine with DAA-style unlinkability that still permits targeted revocation.

Thank You

Shalabh Jain
shalabh.jain@cdpg.org.in