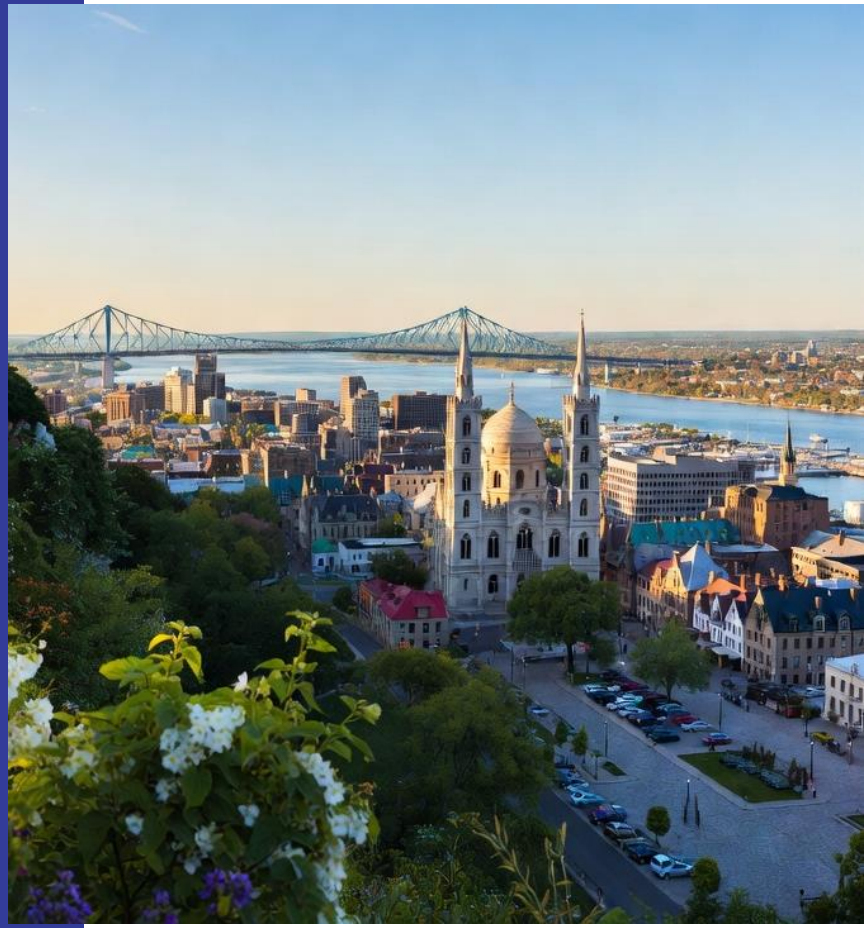
Aurora™ NETWORKS

AI on Both Sides of the Wire

Building Cryptographic Guardrails for Resilient Network Operations



Presenter: Xin Qiu
Authors: James Ni, Xin Qiu, Ting Yao, Lisa Yin, Jinsong Zheng

June 10, 2026

Xin Qiu, Ph.D.

Head of Security Solutions and PKI Center, Aurora Networks

Dr. Xin Qiu has over 25 years of experience in Public Key Infrastructure (PKI), device and software security, and has generated a patent portfolio of over 100 assets worldwide.

She leads a diverse team across R&D, security operations, product marketing and management, delivering security services to global device manufacturers and network operators.

Agenda

- AI-driven **asymmetry**: why security is harder now
- Where **trust breaks** in AI-driven systems
- Why **detection** alone **fails**
- Cryptographic **guardrails** for trust
- From trust signals to **enforcement**
- Enabling **resilient operations** at machine speed

The Shift: AI as an Operational Layer

- AI is now embedded in operations
  - AI-driven analytics, detection, automation
- The adversary has AI Too: attackers are operationalizing AI
  - Adaptive reconnaissance, targeting and penetration
  - Scalable impersonation and social engineering
  - Malware generation and mutation
  - Supply chain and dependency exploitation

Result: machine-speed offense

AI on Both Sides: The Asymmetry Problem

- Defenders must stop EVERY attack attempt
- Attackers need only ONE successful attempt
- AI accelerates both sides
- Detection alone cannot keep pace with AI-driven adaptive attacks
- Without strong guardrails, attackers gain the advantage

Where This Asymmetry Becomes Critical

- AI-driven systems span many use cases
- The asymmetry exists broadly across AI-enabled environments
- Agentic AI systems amplify this problem
  - Autonomous decision-making
  - Dynamic interaction across components
  - Invocation of external tools and services

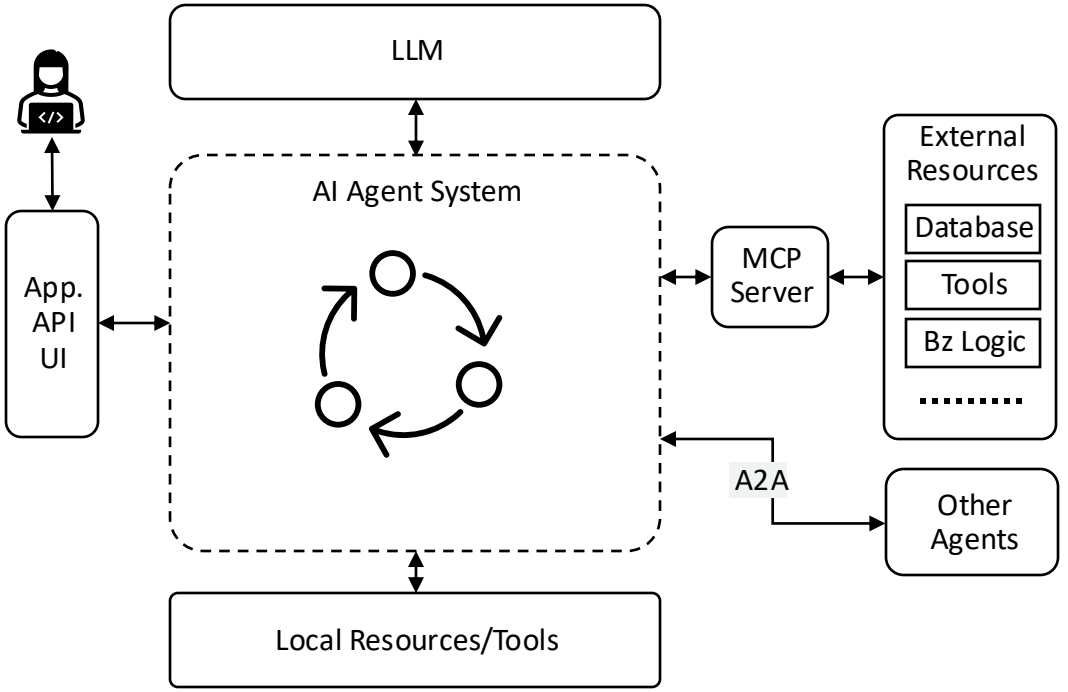
Focus of this talk: Agentic AI as a representative, emerging operational model

An Active and Evolving Security Landscape

- NIST driving AI Agent Identity & Authorization standards (NCCoE)
- Industry building workload and agent identity frameworks (e.g., SPIFFE / SPIRE, verifiable credentials for delegated authority)
- Emerging interaction models (e.g. agent-to-agent, tool & API invocation, language model context)
- Expanding attack surface (e.g. prompt & tool abuse)

Where Trust Breaks in AI-Driven Systems

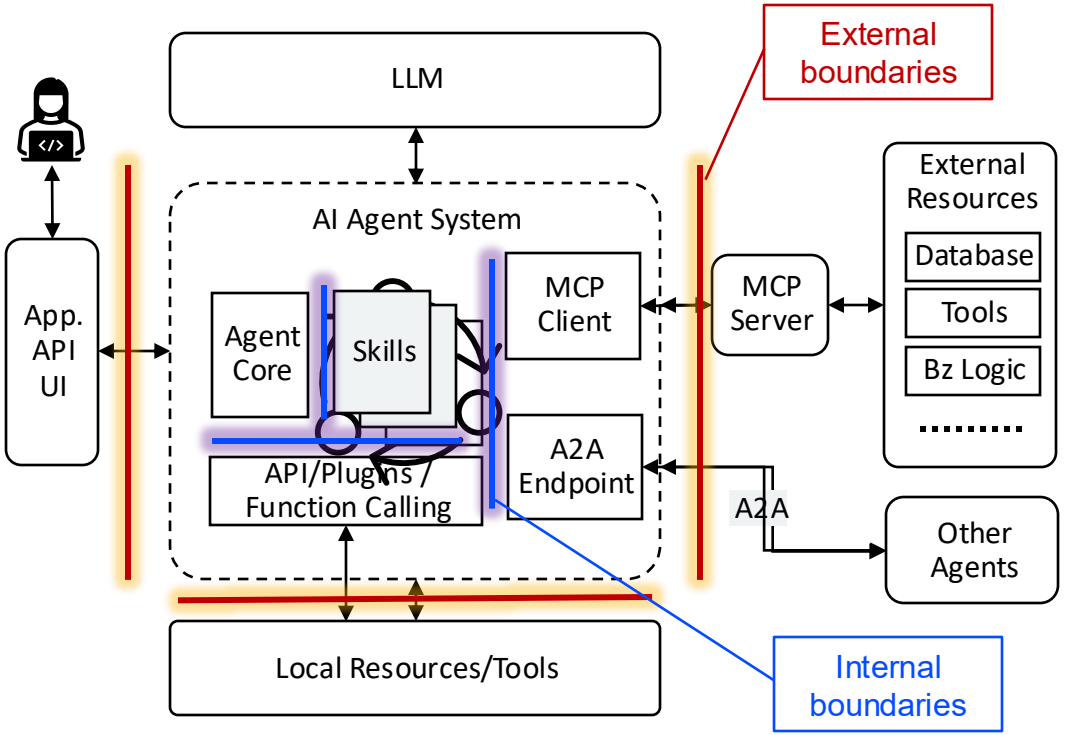
- AI systems run autonomously
- AI systems composed of distributed, multi-party components
- Components interact across trust boundaries
- External inputs and dependencies are not inherently trusted
- No built-in enforcement of identity or integrity



Highly distributed, loosely trusted system

Where Trust Breaks in AI-Driven Systems

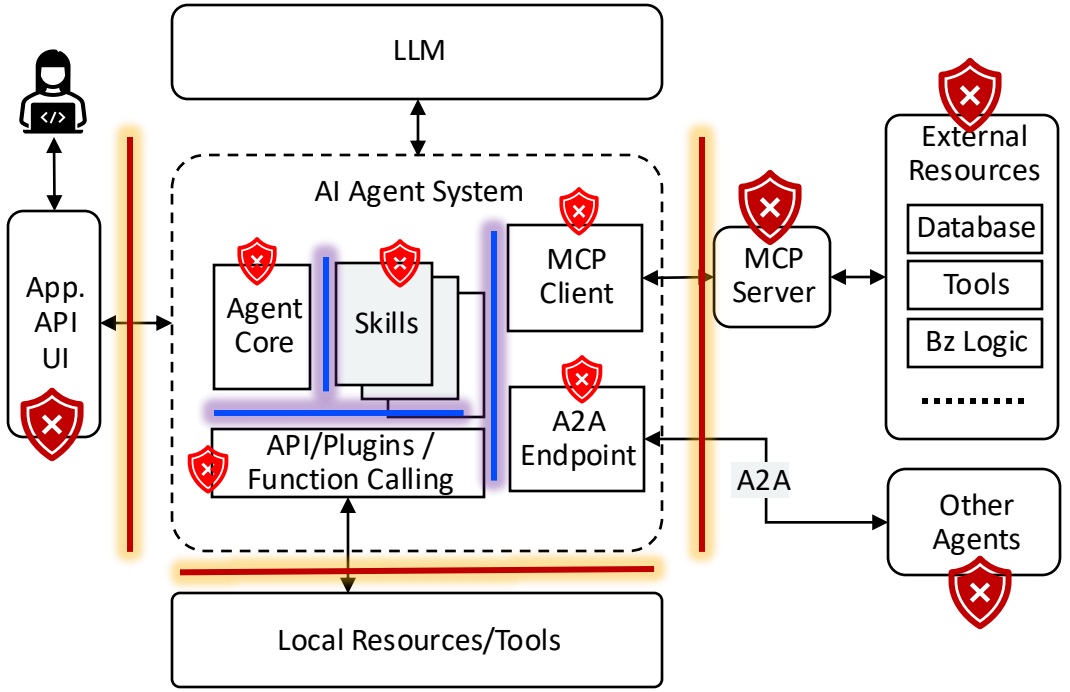
- AI systems run autonomously
- AI systems composed of distributed, multi-party components
- Components interact across trust boundaries
- External inputs and dependencies are not inherently trusted
- No built-in enforcement of identity or integrity



Highly distributed, loosely trusted system

Where Trust Breaks in AI-Driven Systems

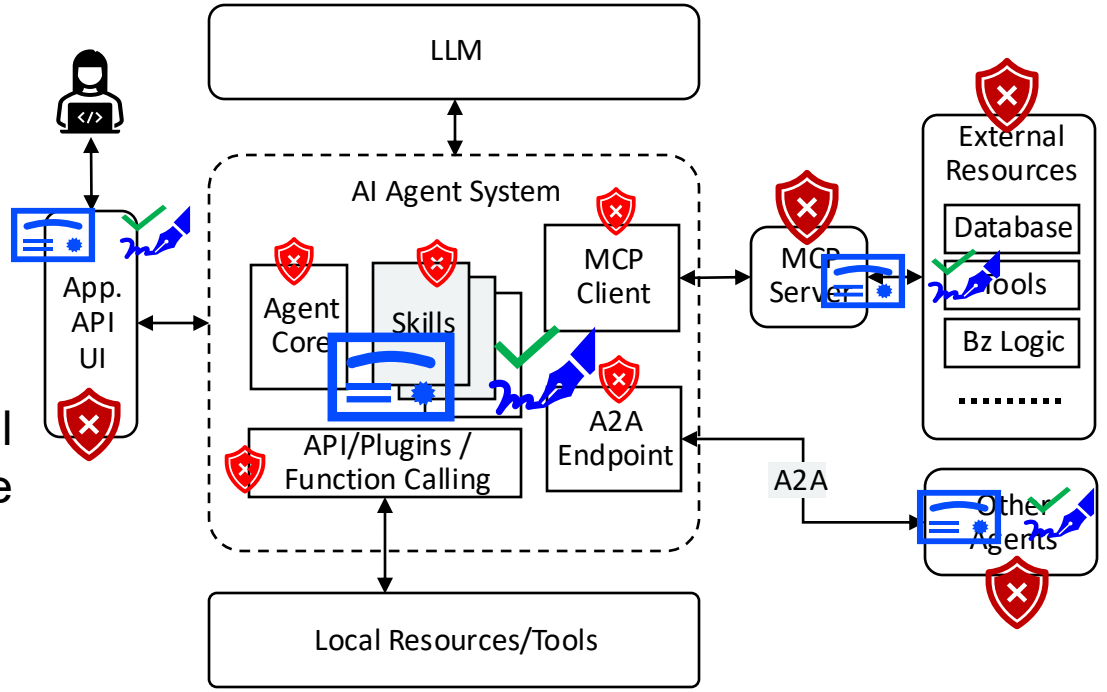
- AI systems run autonomously
- AI systems composed of distributed, multi-party components
- Components interact across trust boundaries
- Internal interactions, external inputs and dependencies are not inherently trusted
- No built-in enforcement of identity or integrity



Highly distributed, loosely trusted system

Where Trust Breaks in AI-Driven Systems

- AI systems run autonomously
- AI systems composed of distributed, multi-party components
- Components interact across trust boundaries
- Internal interactions, external inputs and dependencies are not inherently trusted
- Need built-in identity for integrity and enforcement



Highly distributed, loosely trusted system

What Happens Without Trusted Guardrails

Examples:

- Compromised components drive destructive actions
- AI outputs can be manipulated or weaponized
- Sensitive data leaks across APIs & system boundaries
- Components can manipulate or abuse each other
- Internal and external services can be misused

Root problem: Lack of enforceable trust

Why Detection Alone Fails

- Attacks adapt faster than manual or static defenses
- Alert volume exceeds response capacity
- No trusted signal to trigger enforcement decisions
- Detection ≠ containment

Cryptographic Guardrails for Trust

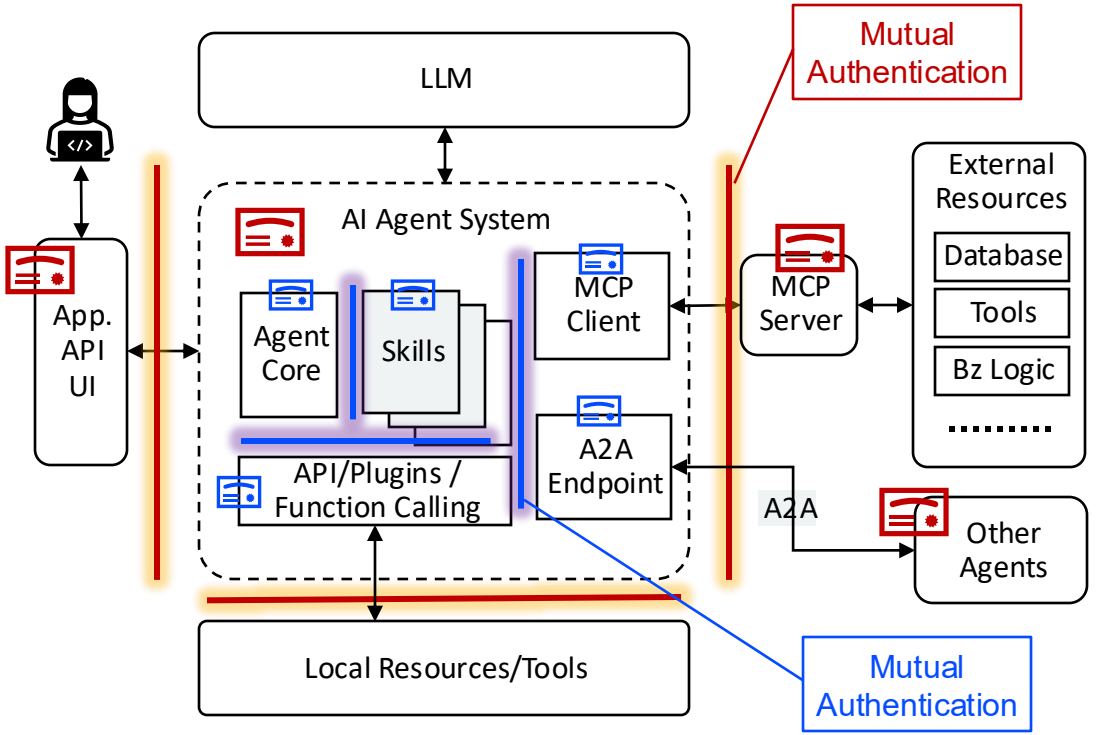
Move from detection to enforcement

- **Guardrail 1**: Cryptographic Identity (across trust boundaries)
  - Identity issued, verified, and anchored in trusted roots
  - Mutual authentication before interaction
- **Guardrail 2:** Component Security (installation + runtime)
  - Signing (Skills, MCP Clients, A2A endpoints, configurations, etc.) with keys protected in hardware (HSM /TPM secure enclave)
  - Integrity verification at install
  - Secure startup/invocation at runtime

Together: generate authoritative, enforceable trust signals

Guardrail 1: Cryptographic Identity

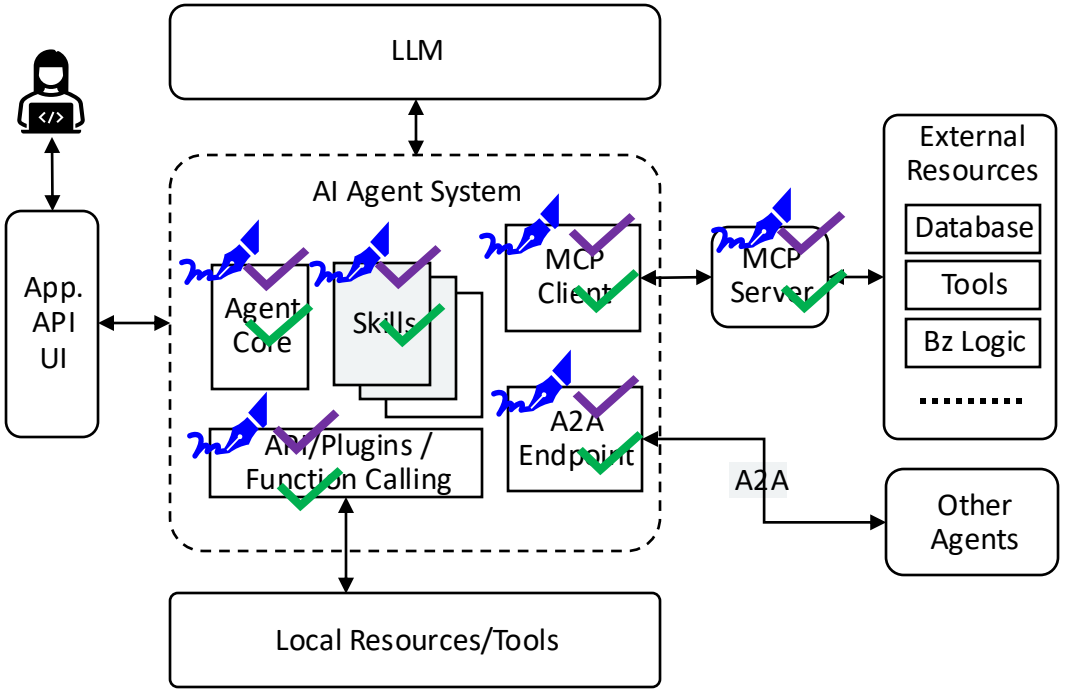
- Across trust boundaries
- Component / device / workload identity
- Verifiable identity anchored in a trust root protected by HW
- Mutual authentication before interaction



Highly distributed, loosely trusted system

Guardrail 2: Component Software Security

- Signing components with keys protected in hardware
- Integrity verification at install
- Secure startup/invocation at runtime
- Verification keys have to be immutable



Highly distributed, loosely trusted system

Core Trust Signals for Enforcement

- **Identity**: trustable device and workload credentials
- **Integrity**: signed code and verified execution
- **Provenance**: trusted origin of software and models
- **Attestation**: runtime state validation
- **Revocation**: rapid removal of compromised trust

Outcome: Enforce at cryptographic layer
(not application logic)

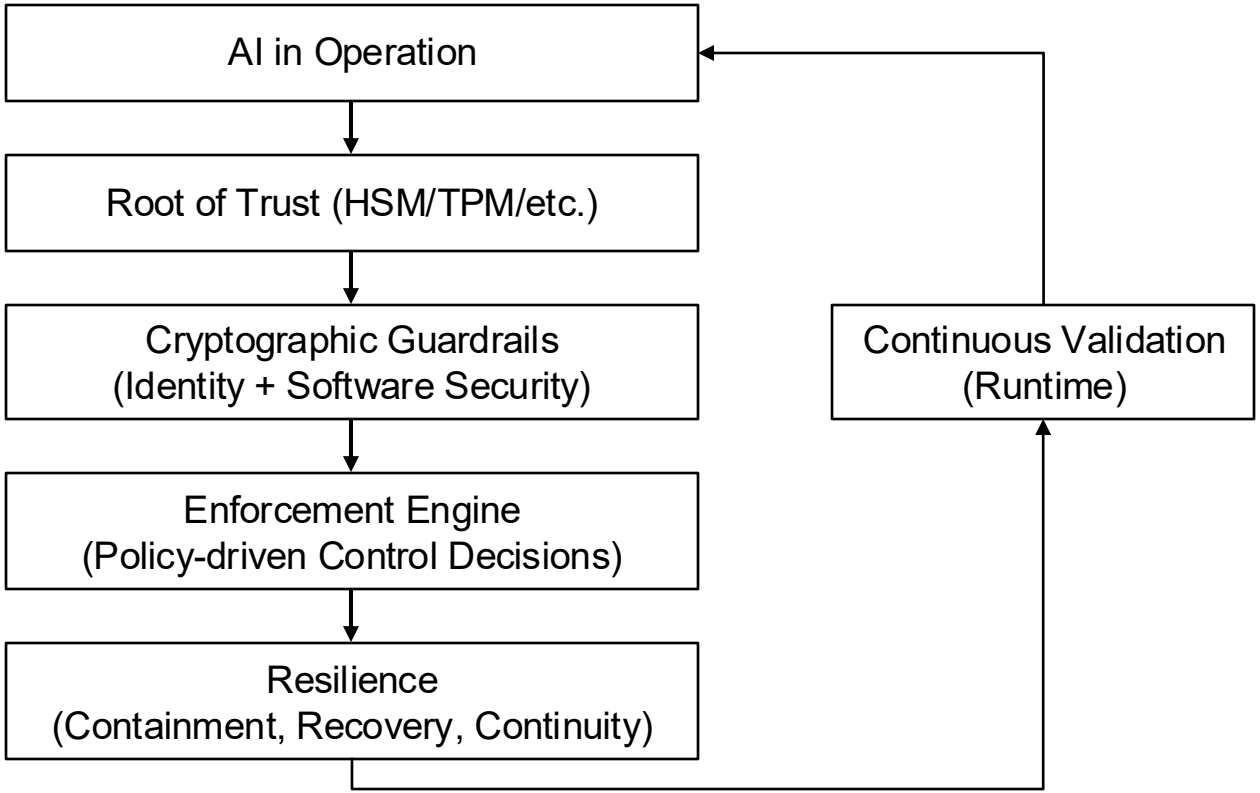
Containment driven by trust, not alert volume

Operational Resilience
(Containment + Recovery)

- Automated, trust-based enforcement reduces response latency
- Continuous validation via attestation (runtime state confidence)
- Rapid recovery when compromise is confirmed:
  - Revocation of identity/trust
  - Secure re-provisioning / re-admission to service

Detection → Trusted Enforcement → Resilient Operations

Loop Enforcement for AI-Driven Operations



Key Takeaways

- AI amplifies both attack and defense
- Detection alone fails in machine-speed environments
- Cryptographic guardrails enable trusted enforcement
- AI must be bound to cryptographic signals

→ **Resilient, abuse-resistant operations**

Xin Qiu

Head of Security Solutions and PKI Center™

Aurora Networks

Email: xin.qiu@auroranetworks.com

LinkedIn: https://shorturl.at/QMDkN

pki-center.com

