Hewlett Packard Enterprise

A Path to Zero Trust Architecture in HPC and AI Using SPIFFE and Spire

Jeremy Duckworth
Distinguished Technologist

October 2023

Background

State of Cybersecurity in HPC

SPIFFE (Secure Production Identity for Everyone) and Spire Overview

SPIFFE and Spire use in HPC and AI Architectures

Current implementation in HPE Cray EX platform

Future implementation in HPE Denali platform

Background

State of Cybersecurity in HPC

NIST High-Performance Computing Security (HPCS) Project

- HPCS created in response to US Executive Order (EO) 14028, as a directive of the National Strategic Computing Initiative (NCSI)
- HPCS has held three workshops, the last in March 2023
  - Substantial peer-group intersection with Supercomputing Conference Cybersecurity Workshop Series
- Publications
  - (US) National Strategic Computing Initiative Update: Pioneering the Future of Computing (2019)
  - NIST SP (Special Publication) 800-223 - High-Performance Computing (HPC) Security: Architecture, Threat Analysis, and Security Posture (2023)
  - NIST IR (Interagency Report) 8476 - 3rd High-Performance Computing Security Workshop (2023)

Cybersecurity “Disruptors” in HPC

- United States Executive Order 14028 (2021) – Highlights
  - –Remove barriers to sharing of and response to threat intelligence
  - –Advance towards Zero Trust Architecture
  - –Mandates deployment of MFA and encryption within a specific time period
  - –Accelerate movement towards secure aaS models and providers
  - –Major emphasis on supply chain security and observable product security
- Workflow distribution over diverse, federated platforms and geolocations vs. HPC as an “island” or small ”archipelago”
- The convergence of HPC, AI, and ML in amalgamated workflows
- Emergent HPC aaS and cloud native operational models and technologies (e.g., containerization, virtualization)
- Expanded use of HPC for regulated processing (e.g., healthcare) and desires for operator/tenant isolation

"**Zero trust** focuses on protecting resources (assets, services, workflows, network accounts, etc.), not network segments, as the network location is no longer seen as the prime component to the security posture of the resource."

NIST SP 800-207 Zero Trust Architecture (Aug 2020)

Post-Exascale Vision for Leadership HPC

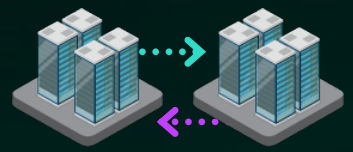
Today

Exascale Supercomputer

World's fastest
Supercomputer

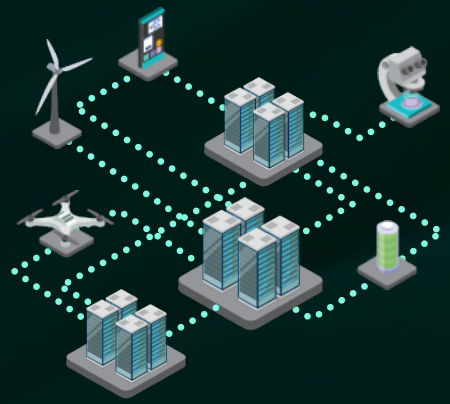
~10x Exascale

Productivity and agility for HPC and AI applications



Multi-faceted, complex workflows

For modeling, simulation, data analytics, and artificial intelligence

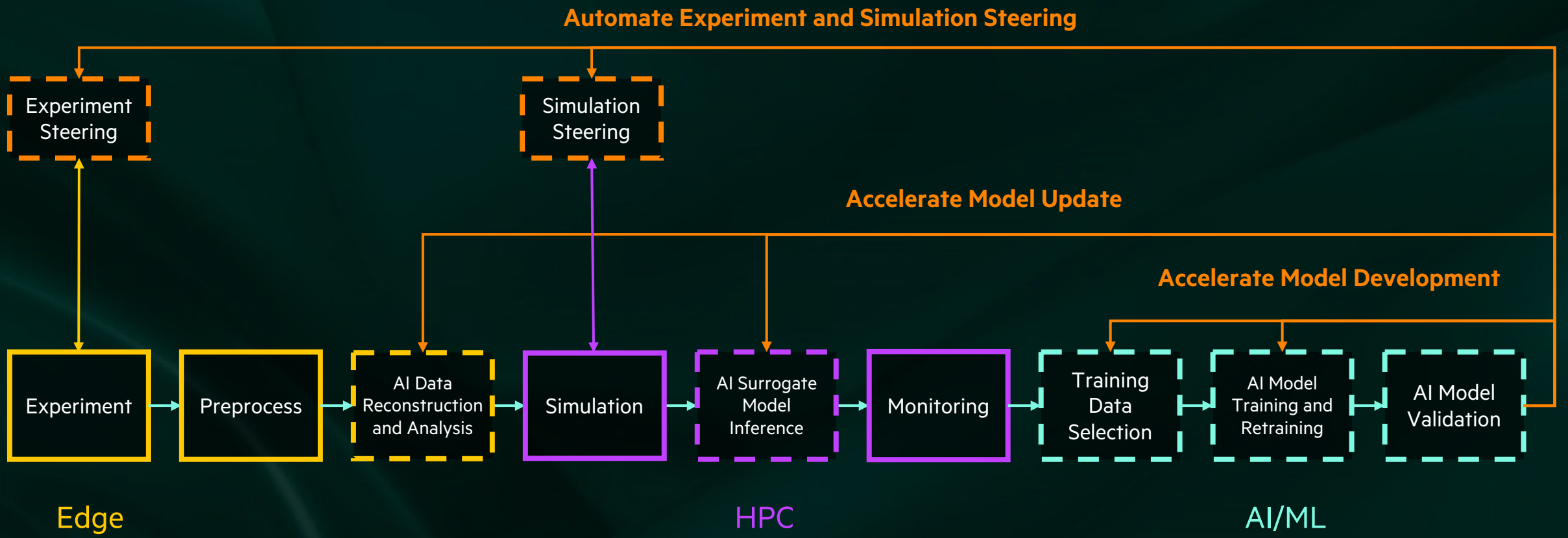


Federated Diverse Systems

Integrate, automate, and optimize workflows that span multiple locations, organizations, and vendors

World's fastest
Workflows

Edge-HPC-Analysis Amalgamated Workflows



SPIFFE and Spire Overview

Concepts and References

What are SPIFFE and Spire?

SPIFFE – Secure Production Identity for Everyone

- **SPIFFE** is a set of open-source specifications that provides a framework for bootstrapping and issuing strongly attested, **cryptographic identities to workloads** across heterogeneous environments and organizational boundaries. SPIFFE **is intended for identifying servers, services, and other nonhuman entities** communicating over a computer network.

- **Spire** is a production-ready implementation of the SPIFFE APIs (Application Programming Interfaces) that performs node and workload attestation in order to securely issue SVIDs (SPIFFE verifiable identity documents) to workloads, and verify the SVIDs of other workloads. Very mature on Linux platforms, experimental support on Windows.

- SPIFFE and Spire are ”graduate projects” of the Cloud Native Compute Foundation (CNCF)

What are SPIFFE and Spire?

SPIFFE and Spire Concepts

SPIFFE Identity

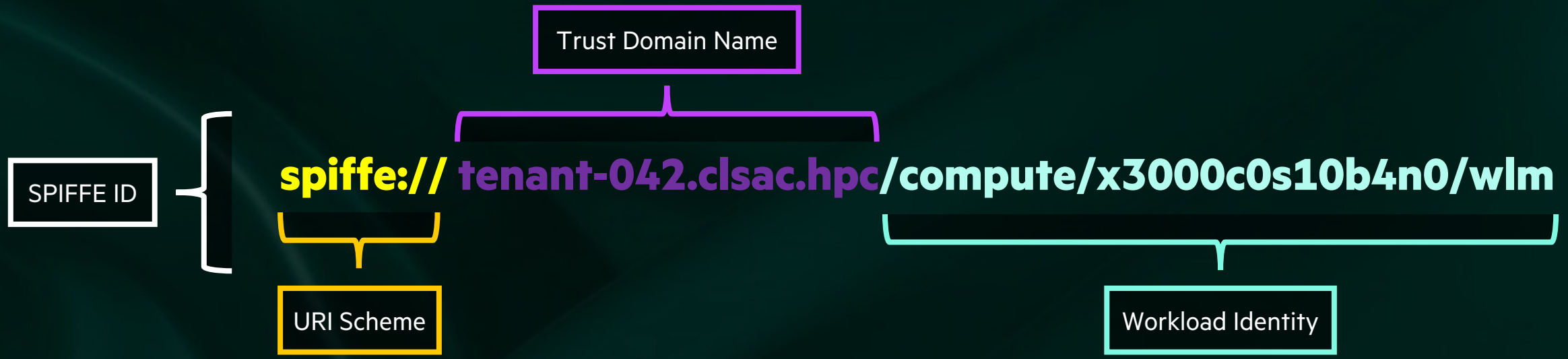
Is a URI (uniform resource locator) string that uniquely and specifically identifies a trust domain and workload.

Trust Domain

Corresponds to the trust root of the system. Can be used to federate multiple domains.

Workload Identity

Corresponds to the infrastructure and/or application upon which it runs.



What are SPIFFE and Spire?

SPIFFE and Spire Concepts (continued)

SPIFFE SVID

A SPIFFE Verifiable Identity Document, through which a workload proves its identity to a resource or caller. An SVID is considered valid if it has been signed by a CA authority within the SPIFFE ID's trust domain.

Trust Bundle

SPIFFE relies on Public Key Infrastructure (PKI) as a foundation for trust management. The trust bundle is the 'chain' of trusted certificate authority (CA), certificates that can be used to verify trust (of the infrastructure and workloads).

Spire Workload API

Issues cryptographic "passports," used by workloads, to prove their identity. Two types of identity formats are supported, an X509-SVID and a JWT-SVID. With the X509-SVID, the workload also has access to a private key which it can also use to sign data on behalf of the workload.

Spire Node Attestation

A process by which the node identity hosting a workload is verified. Multiple attestation options are available (AWS, GCP, Azure, K8S, Join Token (OTP), SSH (CA), x509, TPM LDevID).

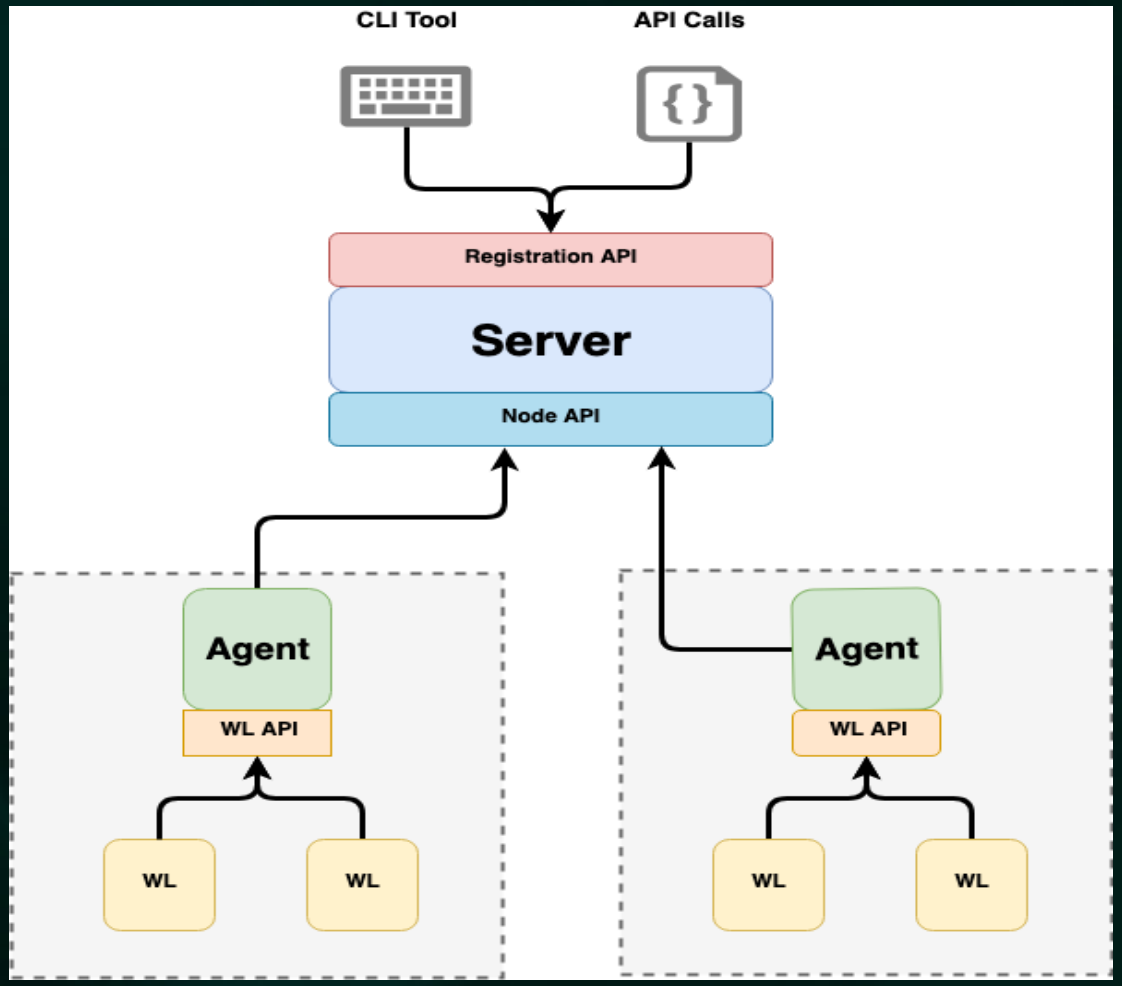
Spire Workload Attestation

A process by which a workload identity is verified. Multiple workload attestors are available (Unix, K8S, Docker).

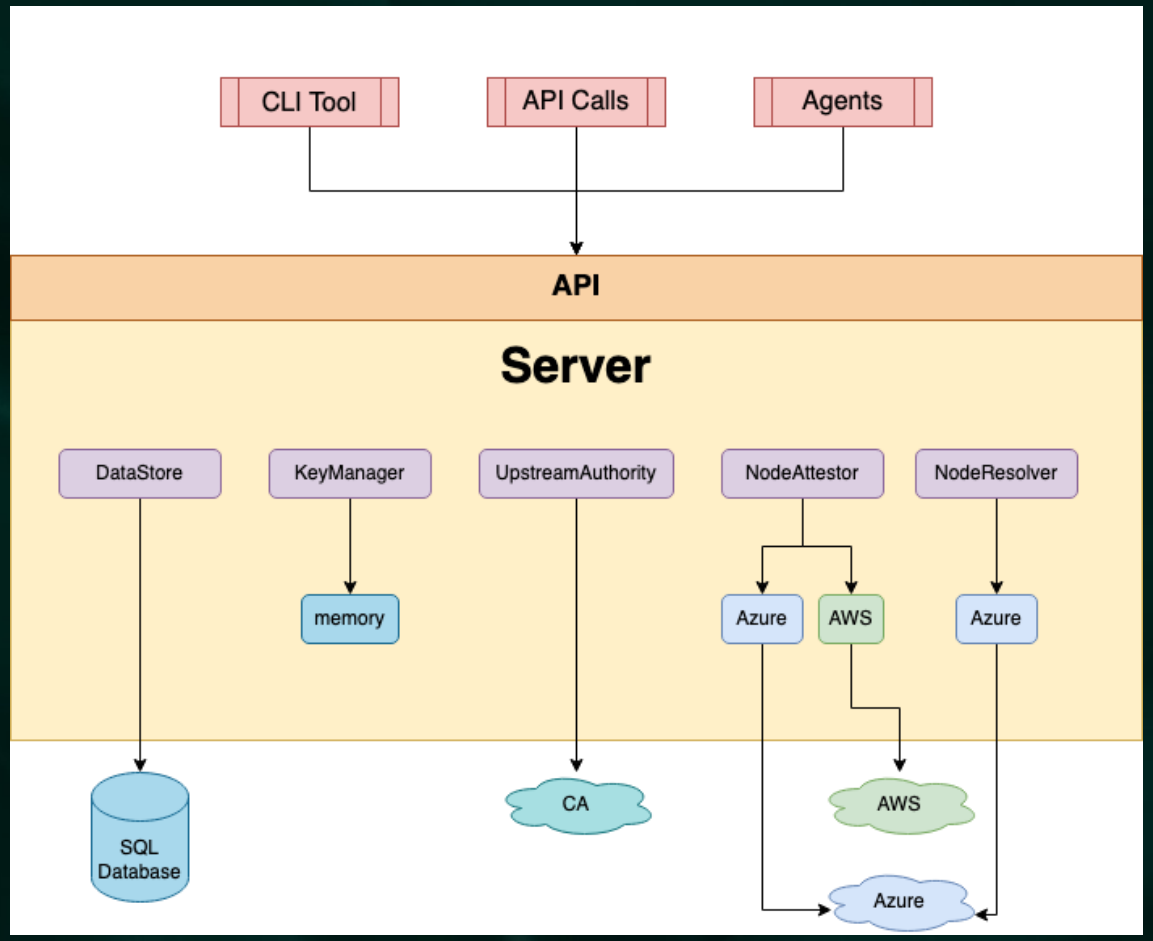
Spire Workload Registration

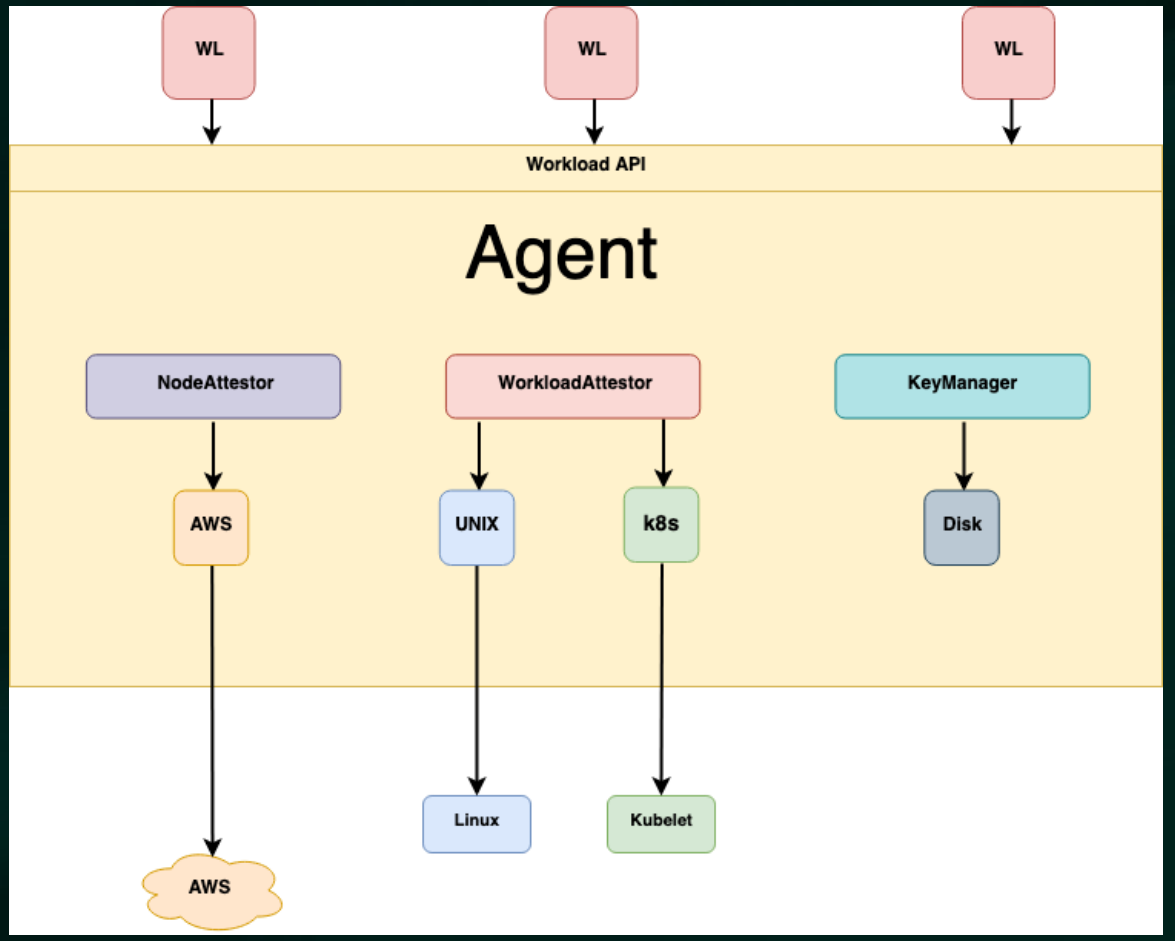
A mapping of workloads to nodes based on selectors that a workload must possess to be issued a particular identity. Mapping can be dynamic based on underlying infrastructure (e.g., AWS EC2 instance tags) or static. The mapping can be applied to a single node or multiple.

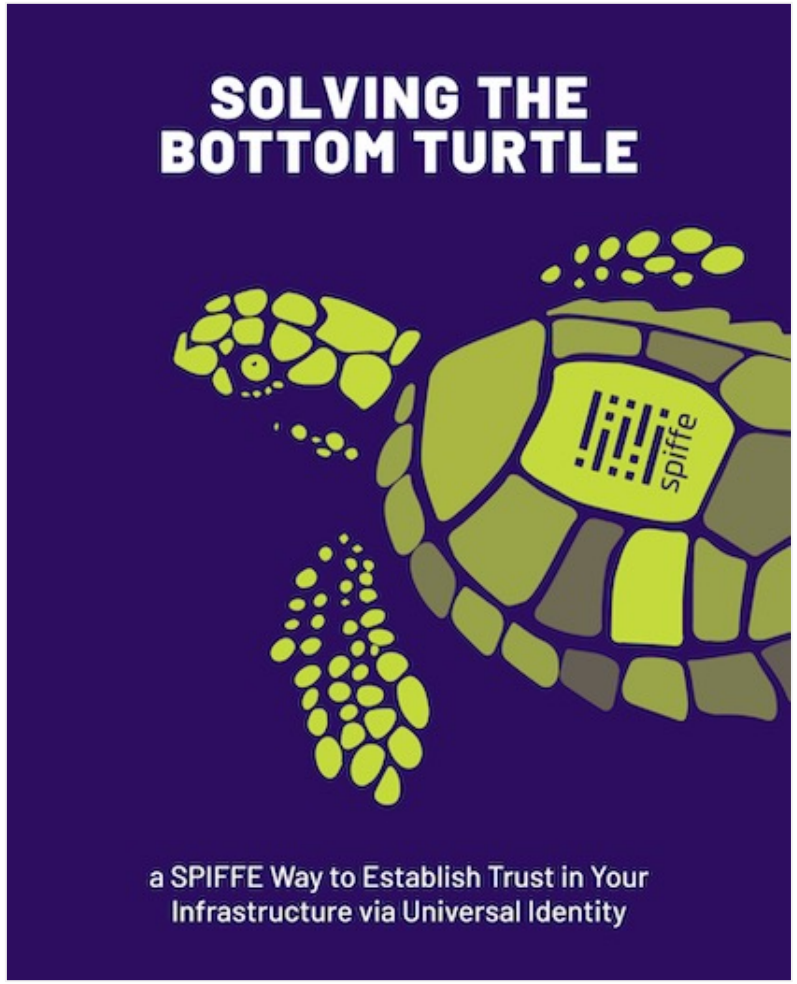
What are SPIFFE and Spire?
Spire Architecture



What are SPIFFE and Spire?
Spire Architecture







This book presents the SPIFFE standard for service identity, and SPIRE, the reference implementation for SPIFFE. These projects provide a uniform identity control plane across modern, heterogeneous infrastructure.

https://spiffe.io/book/

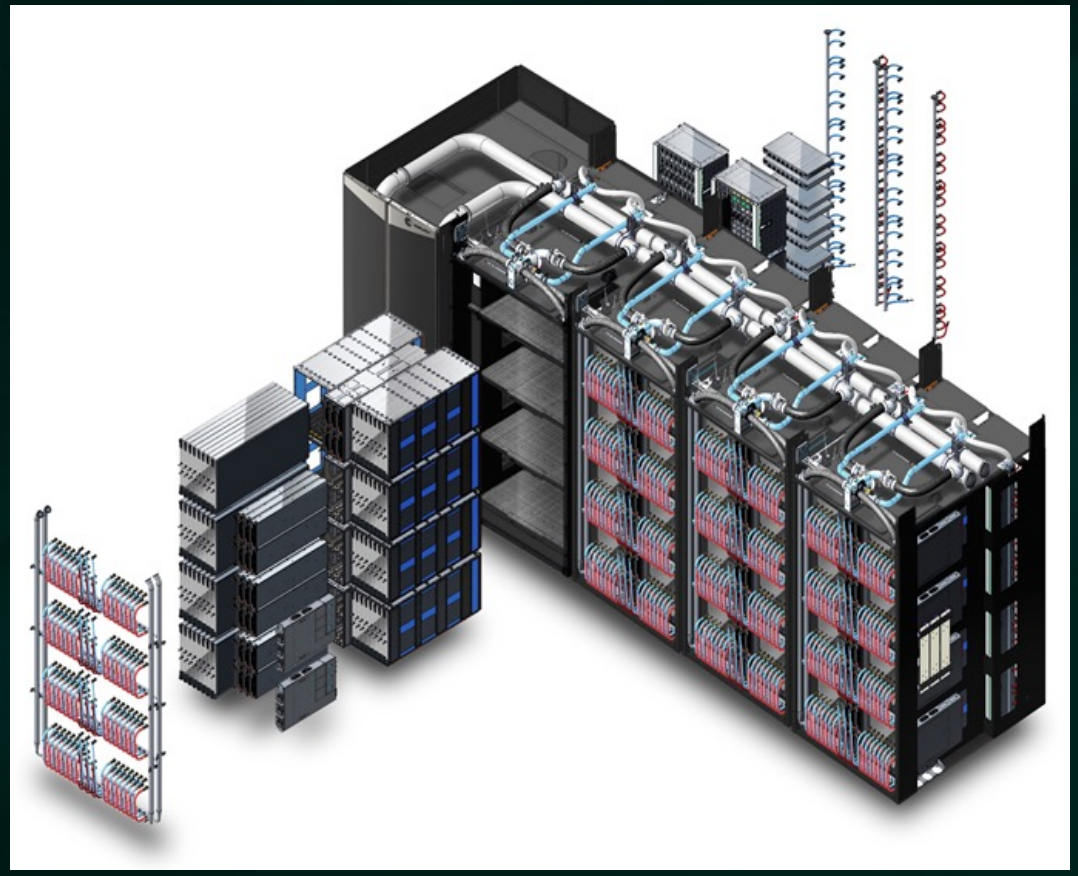
SPIFFE and Spire use in HPC and AI Architectures

Robust non-person entity (NPE) identities and supply chain attestations

HPE HPC and AI

Application of SPIFFE and Spire in the HPE Cray Ex Platform

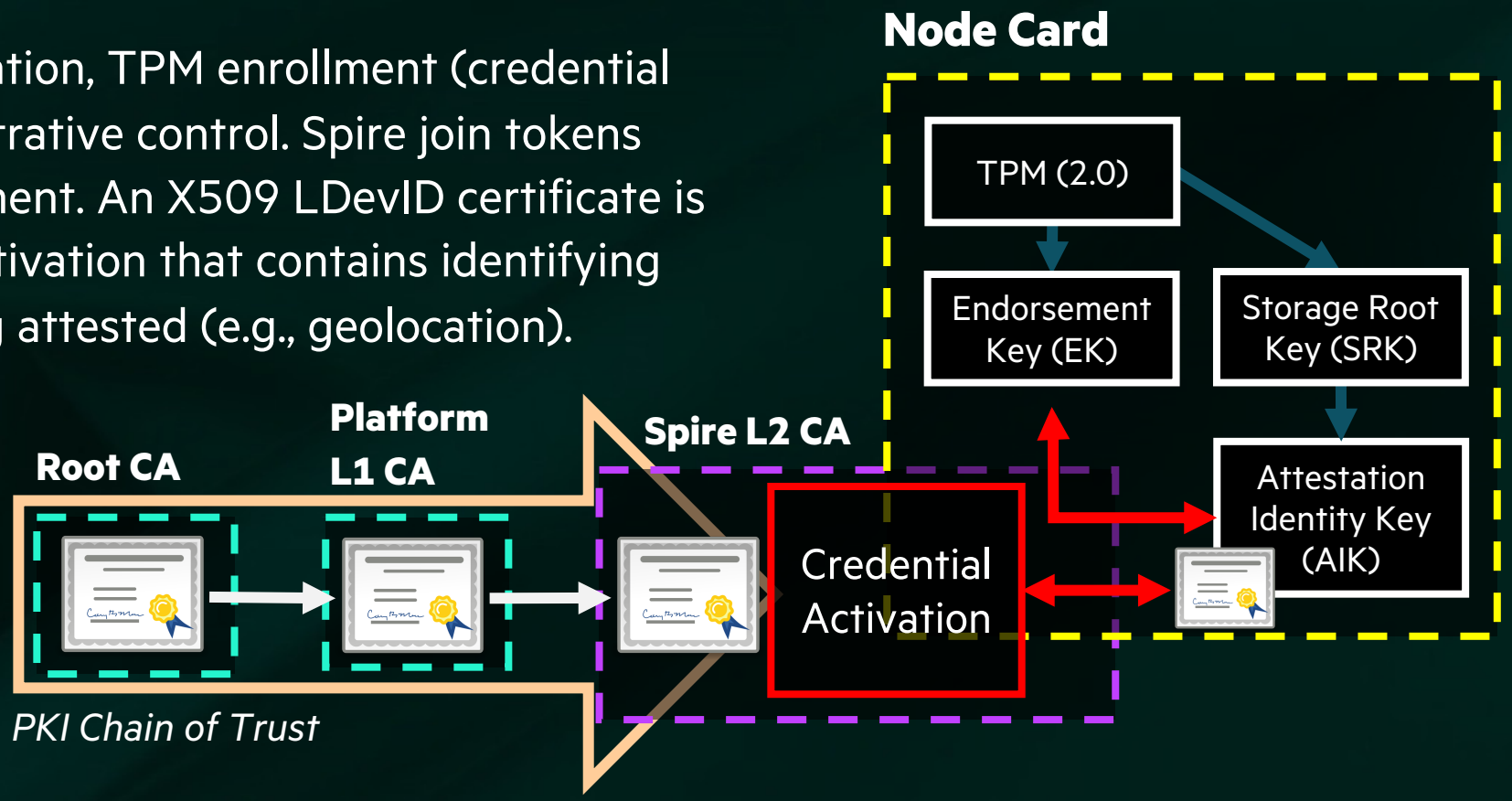
- Flat trust domain scheme, dormant tenancy, no federation; Spire servers and components deployed per-physical cluster
- Integrated with platform and customer upstream PKI
- Parent SPIFFE IDs (nodes) based on node type (e.g., management, compute), geolocation, and application.
- Node attestation via either Join Tokens (OTPs) or TPM proof-of-possession using node-based discrete TPMs
- Workloads mapped to cluster control plane functions: boot and file system content projection, orchestrated boot, post-boot customization, diagnostic sub-systems, node heartbeat, WLM (workload manager) orchestration, LDevID enrollment (TPM).
- Use of JWT-SVIDs to authenticate against cluster APIs using least-privilege layer 7 authorization and intra-cluster geofencing to limit blast radius should a breach occur.



HPE HPC and AI

TPM-based Spire Attestation, Proof-of-Possession, and LDevID (802.1 AR-2018)

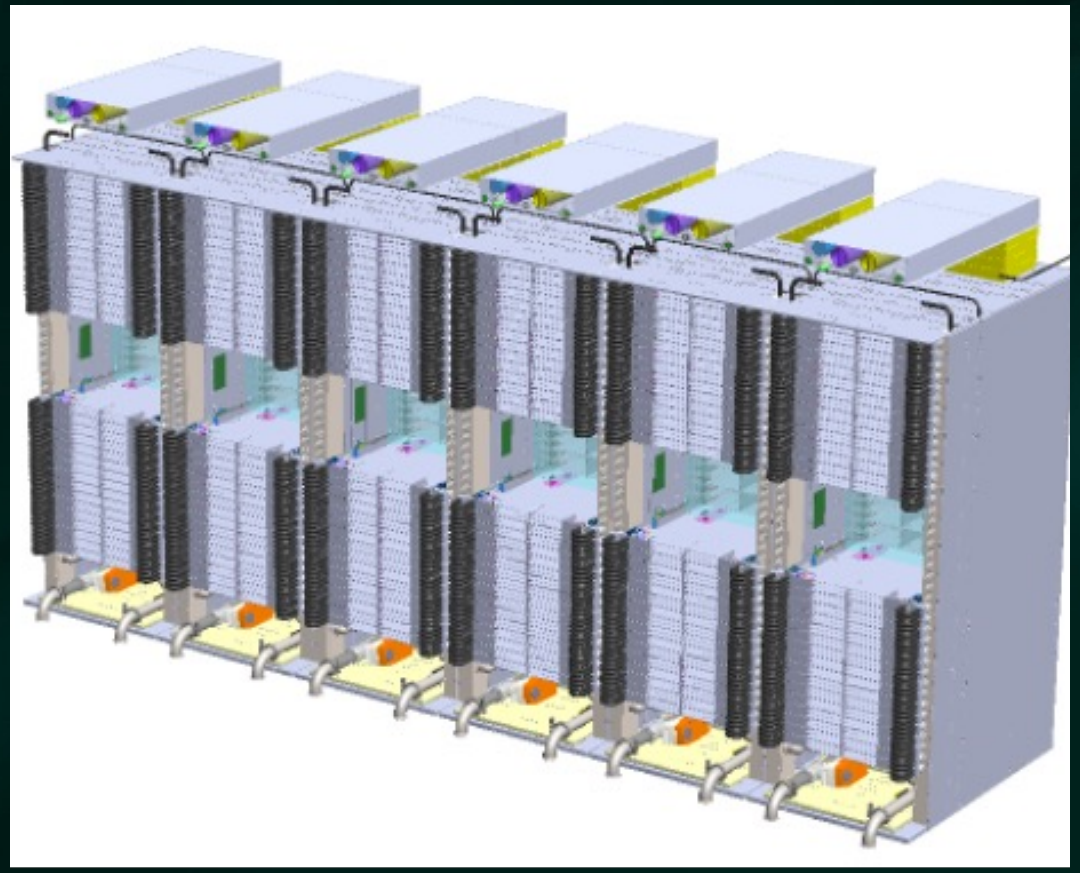
- Spire TPM-based attestation is implemented using TPM credential activation, which chains an attestation key back to the TPM manufacturer endorsement key (EK). This establishes proof-of-possession.
- In the HPE Cray Ex implementation, TPM enrollment (credential activation) is gated by administrative control. Spire join tokens (OTPs) are used during enrollment. An X509 LDevID certificate is generated during credential activation that contains identifying attributes of the instance being attested (e.g., geolocation).



HPE HPC and AI

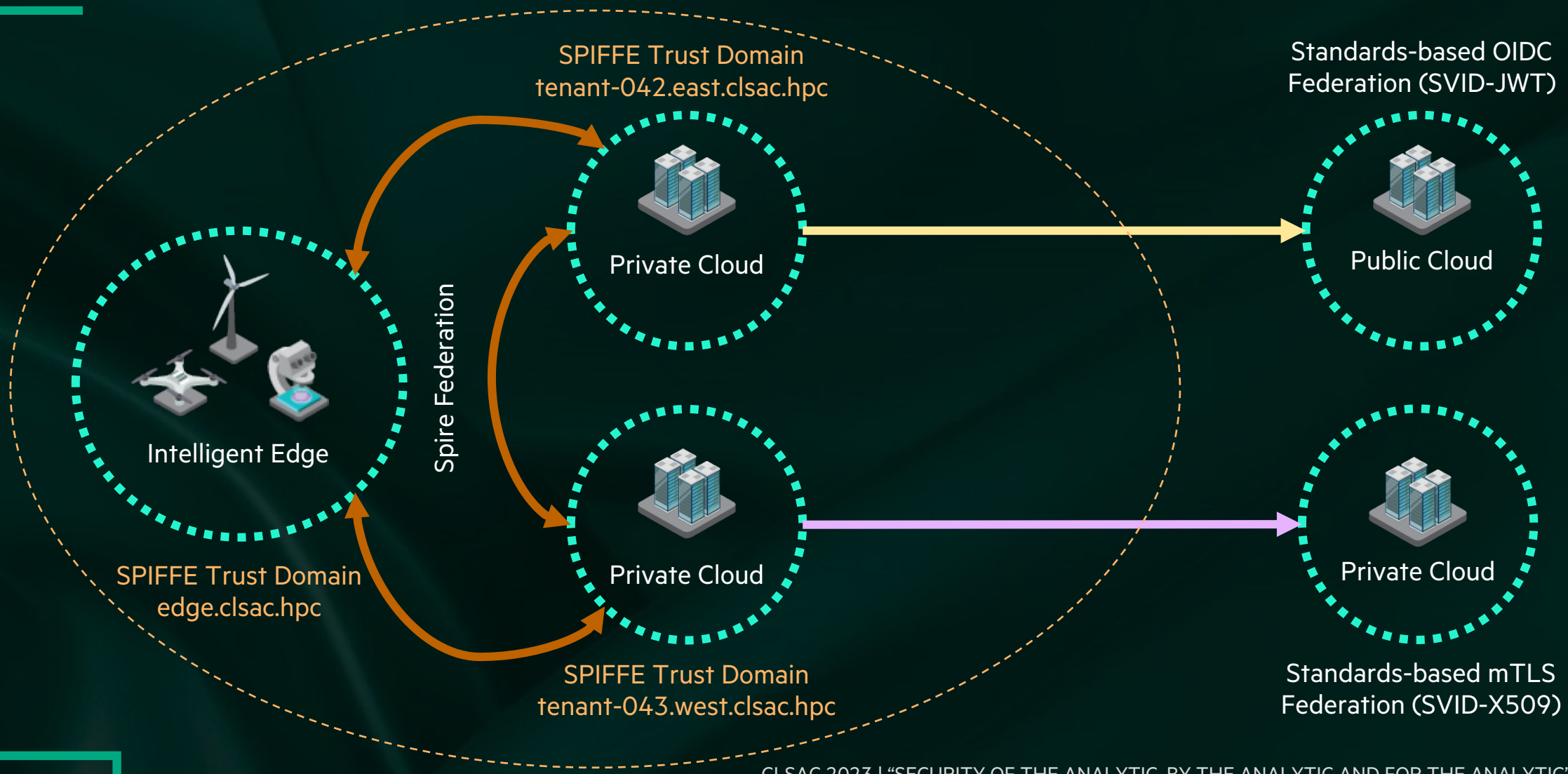
Evolution of SPIFFE and Spire in HPE Denali Platform (Future Product)

- Multiple trust domains, explicit tenancy; Spire server and federation hub capabilities in siloed security infrastructure
- Federation of composable infrastructure enclaves and external clouds, through Spire federation
- Support of isolated, OOB enclave with Hardware Root of Trust and instance metadata for Spire Node Attestation, that is suitable for nested software identities (bare metal, virtual machines, containers)
- HPE factory provisioning of IDevID (Initial Device ID) to chain the LDevID to secure HPE manufacturing, in addition to upstream TPM manufacturer or equiv.
- Combined use of JWT-SVID and X509-SVID depending upon use case; continued push away from long-lived secrets
- Support for instance-based metadata and other dynamic workload registration capabilities.



HPE HPC and AI

Evolution of SPIFFE and Spire in HPE Denali (Next Generation Leadership Compute)



Questions?