Architecture Decisions

Foundational Platform Service v2.0

Version: 1.0 **Author:** John Sasser **Status:** Draft for Review

Evaluation Criteria

1. **Portability:** Does it work across AWS, Azure, and air-gapped environments?
2. **Compliance:** Does it accelerate ATO and reduce manual evidence burden?
3. **Operational Simplicity:** Can a small team maintain it without specialized knowledge?

Table of Contents

1. Key Architecture Decisions

SPIFFE/SPIRE for Workload Identity

Use SPIFFE/SPIRE as the universal workload identity layer with federation to cloud-native IAM

IDENTITY

CONTEXT

Zero Trust requires cryptographically verifiable workload identity. Options considered: cloud-native only, SPIFFE/SPIRE everywhere, or a hybrid approach.

DECISION MATRIX

APPROACH	PROS	CONS
Cloud-native only	Simpler, managed service	No portability, no air-gap support
SPIFFE/SPIRE only	Full control, portable	Operational complexity
Hybrid (Chosen)	**Best of both, portable**	**Moderate complexity**

RATIONALE

- SPIFFE provides vendor-neutral identity that works in air-gapped K3s
- Cloud-native federation reduces operational burden in connected environments
- OIDC federation to AWS STS/Azure AD enables seamless cloud API access
- Pattern proven in DoD environments (Platform One Big Bang uses SPIRE)

OPA/Gatekeeper for Policy Enforcement

Use Open Policy Agent with Gatekeeper for Kubernetes admission control and Conftest for IaC validation

POLICY

CONTEXT

Policy enforcement needed at multiple layers: Terraform planning, Kubernetes admission, runtime evaluation.

DECISION MATRIX

TOOL	PROS	CONS
OPA/Gatekeeper (Chosen)	**CNCF standard, Rego flexibility**	**Rego learning curve**
Kyverno	YAML policies (easier)	Less flexible, newer
Cloud-native (SCPs)	No additional tooling	Not portable

RATIONALE

- OPA is CNCF graduated with broad ecosystem
- Single policy language (Rego) across all enforcement points
- Conftest uses same OPA engine for Terraform validation
- Better audit trail than cloud-native alternatives

Istio Ambient for Service Mesh

Use Istio Ambient mesh (sidecarless) for connected environments, traditional sidecar for air-gapped

NETWORKING

CONTEXT

Service mesh required for mTLS, traffic management, and observability. Need consistent behavior across environments.

DECISION MATRIX

MESH	PROS	CONS
Istio Ambient (Chosen)	**Sidecarless, lower overhead**	**Newer, less production history**
Istio Sidecar	Proven, full features	Resource overhead (~128MB/pod)
Linkerd	Lightweight, simple	Fewer features, smaller community
Consul Connect	HashiCorp ecosystem	Different from Istio in air-gap

RATIONALE

- Istio Ambient reduces resource overhead by 40–60%
- Native SPIRE integration for SPIFFE identity
- L4 ztunnel handles most security requirements
- L7 waypoint available when needed (routing, authorization)

Flux CD for GitOps

Use Flux CD for GitOps-based configuration management

GITOPS

CONTEXT

Need declarative, auditable configuration management that works in air-gapped environments.

DECISION MATRIX

TOOL	PROS	CONS
Flux CD (Chosen)	**Multi-tenancy, lightweight**	**Smaller community than Argo**
ArgoCD	Larger community, better UI	Heavier, complex RBAC
Raw kubectl + CI/CD	Simple	No drift detection, no reconciliation

RATIONALE

- Flux's multi-tenancy model aligns with platform tenant isolation
- Lighter resource footprint than ArgoCD
- Better integration with Zarf for air-gapped deployments
- Kustomization controller handles environment-specific overlays

Zarf for Air-Gap Packaging

Use Zarf for air-gapped package management and deployment

AIR-GAP

CONTEXT

Air-gapped environments need a secure, verifiable way to receive updates.

DECISION MATRIX

APPROACH	PROS	CONS
Zarf (Chosen)	**DoD standard, signed packages, SBOM**	**Zarf-specific packaging**
Custom tarball scripts	Full control	No verification, error-prone
Replicated/Helm charts	Commercial support	Vendor dependency

RATIONALE

- Zarf is the de facto standard in DoD Kubernetes deployments
- Single binary, single tarball simplifies transfer process
- Cryptographic verification ensures package integrity
- Built-in SBOM generation for supply chain compliance

K3s for Air-Gapped Kubernetes

Use K3s for air-gapped/edge Kubernetes deployments

COMPUTE

CONTEXT

Need lightweight Kubernetes for resource-constrained, disconnected environments.

DECISION MATRIX

DISTRIBUTION	PROS	CONS
K3s (Chosen)	**Single binary, low resources, proven air-gap**	**Not "enterprise" K8s**
RKE2	More enterprise features	Heavier than K3s
Kubeadm	Standard K8s	Complex setup, many dependencies
OpenShift	Enterprise support	Massive footprint

RATIONALE

- K3s is ~50MB single binary vs. 500MB+ for kubeadm clusters
- Proven air-gap deployment pattern with Zarf
- Rancher (SUSE) provides commercial support if needed
- Platform One Big Bang uses K3s for edge

Four-Tier Network Architecture

Implement four subnet tiers: Public, Private, Isolated, Data

NETWORK

CONTEXT

Need defense-in-depth network segmentation aligned with SCCA requirements.

DESIGN

TIER	PURPOSE	EGRESS	INGRESS
Public	NAT Gateways, ALB	Internet	Internet (filtered)
Private	EKS nodes, applications	Via NAT	Via ALB only
Isolated	Internal services	None	Private tier only
Data	Databases	None	Private/Isolated only

RATIONALE

- Aligns with DoD SCCA boundary patterns (CAP, VDSS, VDMS)
- Isolated tier for sensitive workloads without egress
- Data tier provides additional layer for databases
- VPC endpoints eliminate need for data tier egress

Security Lake for Log Aggregation

Use Amazon Security Lake with OCSF normalization for centralized logging

LOGGING

CONTEXT

Need centralized, compliant log aggregation across multiple accounts and clouds.

DECISION MATRIX

SOLUTION	PROS	CONS
Security Lake (Chosen)	**Native AWS, OCSF, cost-effective**	**AWS-centric**
Splunk Cloud	Powerful, mature	Expensive at scale
ELK/OpenSearch	Open source, flexible	Operational burden
Sumo Logic	Cloud-native, good gov	Yet another vendor

RATIONALE

- Security Lake is FedRAMP High authorized
- OCSF normalization enables cross-cloud correlation
- Cost-effective for high-volume log storage
- Native integration with Security Hub, GuardDuty

2. Compliance Decisions

80/20 Control Inheritance Model

Target 80% of FedRAMP High controls satisfied at Platform layer, 20% at Application layer

COMPLIANCE

RATIONALE

- CSP inherits ~140 controls (AWS GovCloud FedRAMP authorization)
- Platform provides additional ~80 controls via standardized configuration
- Applications only need ~50–80 unique controls
- Dramatically reduces ATO effort for tenant applications

EVIDENCE STRATEGY

- Automated evidence generation via Terraform state + AWS Config
- AI-assisted SSP section drafting from infrastructure code
- Continuous compliance dashboard via Security Hub

Continuous ATO (cATO) Architecture

Design for continuous authorization rather than point-in-time ATO

COMPLIANCE

RATIONALE

- Traditional ATO is a 12–18 month point-in-time snapshot
- cATO provides real-time security posture visibility
- Aligns with DoD DevSecOps Reference Design
- Enables faster feature delivery without re-authorization

IMPLEMENTATION

- Security Hub as continuous compliance dashboard
- AWS Config rules for drift detection
- OPA for policy enforcement in CI/CD
- Automated evidence bundle generation