Project Horizon

Next-Generation AI-Driven Security & Operations Platform

Project Title: Project Horizon - Unified Zero Trust Architecture

Project Type: Prototype - Free Tier and Open Source Tools

1. Executive Summary

Project Horizon was initiated to design and implement a **comprehensive Zero Trust security architecture**, addressing modern enterprise challenges across identity, data, workloads, and detection. The core objective was to build a **fully functional open-source MVP** that demonstrates technical feasibility, integration depth, and alignment with strategic cybersecurity priorities including **Zero Trust Network Access (ZTNA)**, **Secure Access Service Edge (SASE)**, **Data Security Posture Management (DSPM)**, **Cloud-Native Application Protection (CNAPP)**, and **Extended Detection & Response (XDR)**.

Over the course of the project, the team successfully delivered **five integrated solution pillars** that collectively establish a robust security foundation:

- **ZTNA & SASE** - Implemented identity-aware secure access using Keycloak, Pomerium, and Suricata to enforce per-user authentication, TLS encryption, and network-layer inspection.
- **Customer Data Platform (CDP)** - Developed a synthetic data analytics and engagement pipeline that simulates user behavior, performs churn scoring, and triggers targeted campaigns with full telemetry.
- **DSPM** - Deployed open-source DSPM capabilities for sensitive data discovery, classification, flow mapping, and policy enforcement using MinIO, NiFi, Presidio, Custodian, and ELK.
- **CNAPP** - Secured workloads and container environments through image scanning, runtime policy enforcement, and observability, laying the groundwork for Zero Trust workload protection.
- **XDR & Insider Risk** - Integrated network, identity, and behavior telemetry for advanced detection engineering, threat correlation, and insider risk analytics.

These components were deployed in a **controlled lab environment** using **open-source and free-tier tools**, and were fully integrated through a **centralized ELK-based telemetry backbone**. Together, they deliver a cohesive end-to-end security architecture that enforces identity-driven access, monitors data flows, detects abnormal behaviors, and enables foundational response automation.

The **business value** of this MVP is threefold:

1. **Strategic Alignment** - The architecture directly supports priorities of Zero Trust enforcement, regulatory compliance, data security, and SOC modernization.
2. **Cost-Effective Feasibility** - By leveraging open-source components, the team demonstrated enterprise-grade capabilities without commercial licensing overheads, making it ideal for rapid prototyping.
3. **Scalable Foundation** - The modular design allows each solution to operate independently while integrating seamlessly, enabling future scaling to production environments.

Looking ahead, the MVP will evolve into a **production-grade enterprise platform** by introducing federated identity, automated PKI, high-availability deployments, enriched telemetry pipelines, granular policy enforcement, and compliance automation. This phased roadmap ensures that Project Horizon will mature from a functional prototype into a **fully operational, scalable, and compliant Zero Trust architecture** that meets enterprise security requirements.

2. Introduction

2.1 Objectives

Strategic cybersecurity vision is centered on **building a unified Zero Trust security architecture** that ensures **secure access, data visibility, regulatory compliance**, and **advanced threat detection** across its enterprise ecosystem.
The key objectives driving **Project Horizon** were:

- **Zero Trust Enforcement**
  Implement identity-driven access control for applications, data, and workloads, eliminating implicit trust and adopting a “never trust, always verify” posture.

- **Centralized Visibility & Compliance**
  Gain unified visibility into authentication, data flow, and security events while aligning with regulatory frameworks such as GDPR and PCI-DSS.

- **Data Protection & Security Posture Management**
  Discover, classify, and secure sensitive data across storage and processing systems, enforcing security policies and remediating misconfigurations in real time.

- **SOC Modernization & Detection Fabric**
  Establish a foundation for modern SOC operations through telemetry integration, UEBA, XDR capabilities, and workflow automation to reduce detection and response times.

- **Scalable, Modular Architecture**
  Adopt an open, extensible design that supports incremental scaling from lab prototypes to enterprise production deployments.

2.2 Project Mandate

The **Project Horizon mandate** was to design and implement a **functional, open-source MVP** demonstrating how Zero Trust objectives could be realized across five critical solution pillars:

1. **Zero Trust Network Access (ZTNA) & SASE** – Secure, identity-aware access perimeter

2. **Customer Data Platform (CDP)** – Behavioral analytics & engagement pipeline

3. **Data Security Posture Management (DSPM)** – Sensitive data discovery, policy enforcement, compliance visibility

4. **Cloud-Native Application Protection Platform (CNAPP)** – Workload protection, container security, runtime visibility

5. **Extended Detection & Response (XDR) & Insider Risk** – Centralized telemetry, threat detection, and behavioral analytics

The team was tasked with proving **end-to-end feasibility**, demonstrating **integration across solutions**, and aligning the MVP implementation with **requirements and operational goals** all within the constraints of open-source tooling and a controlled lab setup.

2.3 Timeline & Deliverables

The MVP was developed and deployed in a **phased manner** over the project period, covering solution design, integration, and validation:

Phase	Timeline	Key Activities	Deliverables
Phase 1	Solution Design	Architecture definition, tool selection, topology planning	Design blueprints, integration plan
Phase 2	Deployment & Configuration	Component setup, customizations, baseline telemetry pipelines	Working lab deployments for each solution pillar
Phase 3	Integration & Testing	Telemetry centralization, identity and data flow integration, security validation	Fully integrated MVP environment
Phase 4	Evidence & Reporting	Validation of workflows, log capture, dashboarding, documentation	Solution reports, dashboards, configuration records

Key Deliverables:

- Five working solution pillars (ZTNA, CDP, DSPM, CNAPP, XDR)
- Centralized ELK telemetry pipeline
- End-to-end detection and response workflows
- Evidence (logs, dashboards, configurations)
- Strategic roadmap for production scaling

3. Scope

3.1 Current Prototype Scope

The current implementation of **Project Horizon** represents a **fully functional open-source MVP** deployed within a **controlled lab environment**. Its primary purpose is to validate the **feasibility and integration** of core Zero Trust security concepts across multiple security domains, using lightweight and cost-effective tooling.

Key characteristics of the current scope include:

- **Deployment Environment:**
  All components are hosted on **single-node virtual machines** within a controlled lab

network. TLS is configured using self-signed certificates to simulate secure communication channels.

- **Implemented Security Pillars:**
  The MVP covers **five core solutions** forming the foundation of Zero Trust strategy:

  1. **ZTNA & SASE** – Identity-aware secure access perimeter
  2. **CDP** – Behavioral analytics & engagement pipeline
  3. **DSPM** – Sensitive data discovery, policy enforcement, and compliance dashboards
  4. **CNAPP** – Workload protection and container runtime visibility
  5. **XDR & Insider Risk** – Integrated telemetry, UEBA, and detection fabric

- **Functional Demonstrations:**
  End-to-end workflows have been validated, including identity-based access control, data flow analysis, telemetry ingestion, detection alerts, and basic response actions.

- **Centralized Telemetry:**
  All solutions forward structured logs to a **unified ELK stack**, serving as the backbone for visibility, dashboarding, and detection engineering.

This scope successfully demonstrates how **open-source tools can be integrated to build a cohesive Zero Trust architecture** that spans identity, data, network, workloads, and detection layers.

3.2 Constraints

The MVP intentionally leverages **open-source tooling and free-tier components**, focusing on functional coverage rather than production-grade scale. As a result, several constraints apply:

- **Open-Source / Community Editions:**
  All components (Keycloak, Pomerium, ELK, MinIO, NiFi, Suricata, etc.) are deployed using their community versions, with no commercial licenses or enterprise support.

- **Single-Node Deployments:**
  Services are hosted on individual virtual machines without clustering, load balancing, or automatic failover, limiting scalability and resilience.

- **TLS & PKI Limitations:**
  Self-signed certificates are used for secure communication, requiring manual trust configuration. There is no enterprise PKI integration or automated certificate lifecycle management.
- **Synthetic Data & Test Scenarios:**
  Synthetic telemetry and demo datasets are used to simulate user behavior, network traffic, and data flows. No live production data is processed.
- **Policy & Detection Simplification:**
  Access and security policies are basic (e.g., route-based ZTNA, bucket-level DSPM), and detection rules are limited to MVP scenarios.
- **No High Availability or Federation:**
  There is no integration with enterprise IdPs (e.g., Azure AD/Okta), and no multi-node or multi-region deployment.

These constraints reflect a deliberate **MVP focus on rapid prototyping** rather than enterprise hardening, ensuring functional demonstration of all core security capabilities within project timelines.

3.3 Gaps & Future Enhancements (Requirements Mapping)

The table below maps **requirements** to **identified gaps** in the MVP and outlines **future enhancements** needed to achieve full production and enterprise readiness:

Requirement	Current Gap (MVP)	Future Enhancement
Federated Identity & SSO	Standalone Keycloak, no Azure AD/Okta integration	Enterprise IdP federation, MFA, SCIM provisioning, conditional access policies
High Availability & Scaling	Single-node deployments, no clustering or LB	Multi-node clusters, auto-scaling, load balancers, failover mechanisms
Certificate & Key Management	Self-signed certs, manual trust setup	Integration with enterprise PKI, automated cert issuance and renewal
Policy Enforcement Depth	Basic route and bucket policies	Context-aware ZTNA, fine-grained DSPM rules, CNAPP runtime protections

Telemetry & SIEM Integration	Raw JSON logs, basic dashboards	Enriched telemetry pipelines, SIEM integration, UEBA models, advanced correlation
Compliance & Reporting	Manual dashboards, no automated compliance mapping	GDPR/PCI-DSS framework mapping, compliance dashboards, reporting & audit integrations
Production Rollout Readiness	Lab setup, synthetic data, no multi-region coverage	Cloud-native deployments, multi-region rollout, real datasets, SOC integration

These enhancements form the foundation of the **scaling strategy and enterprise roadmap** detailed in later sections of this report.

4. Solution Portfolio Overview

Project Horizon delivers a **modular yet integrated portfolio of five security solutions**, each addressing a critical layer of Zero Trust cybersecurity strategy. Together, these solutions form a **defense-in-depth architecture** that spans identity, network, data, workloads, and detection layers.

The implementation focused on using **open-source components** to demonstrate technical feasibility and integration across these layers in a controlled lab environment.

4.1 Solution Pillars and Strategic Roles

S.no.	Solution Pillar	Strategic Role	Key Technologies
1	**Zero Trust Network Access (ZTNA) & SASE**	Establishes **identity-aware secure access** to internal applications with strong authentication, TLS encryption, and network-level inspection.	Keycloak, Pomerium, Suricata, pfSense
2	**Customer Data Platform (CDP)**	Simulates **user behavior analytics and churn scoring**, generates alerts, and triggers campaigns, while maintaining full telemetry and observability.	PostgreSQL, Python, Mailgun, ELK

3	**Data Security Posture Management (DSPM)**	Provides **sensitive data discovery**, classification, policy enforcement, and compliance dashboards, ensuring visibility and protection of data flows.	MinIO, Apache NiFi, Microsoft Presidio, Custodian, ELK
4	**Cloud-Native Application Protection Platform (CNAPP)**	Implements **workload protection** by scanning container images, applying runtime security policies, and monitoring cloud-native environments for threats and misconfigs.	Trivy, Falco, K3s, Kyverno, Cosign, Checkov
5	**Extended Detection & Response (XDR) & Insider Risk**	Integrates identity, network, and behavior telemetry to enable **threat detection, correlation, and insider risk analytics**. Forms the detection & response fabric.	ELK Stack, Suricata, UEBA modules, detection rules

Each solution is **independently functional**, yet designed to **integrate seamlessly** through a **central telemetry backbone (ELK)**, enabling coordinated security visibility and response.

4.2 Mapping Requirements to Solution Pillars

The table below shows how each solution contributes to fulfilling **key cybersecurity requirements**:

Requirement	ZTNA & SASE	CDP	DSPM	CNAPP	XDR & Insider Risk
Zero Trust Enforcement	Identity-based access	–	–	Workload enforcement	Detection telemetry
Centralized Visibility & Logging	Auth + proxy logs	Engagement telemetry	Data flow & DSPM logs	Runtime logs	Integrated in ELK
Data Protection & Compliance	–	–	Data discovery,	Misconfig detection	Compliance alerting

			classification, policy		
SOC Modernization & Threat Detection	Suricata alerts	Behavior signals	Policy violations	Runtime detection	UEBA + Detection rules
Scalable, Modular Architecture	Modular proxy + IdP	Modular pipeline	Policy-driven stack	Container-native	SIEM-agnostic detection

4.3 Integration Highlights

The five solution pillars are **not siloed** they are deliberately **interconnected** to reflect a **realistic enterprise SOC architecture**:

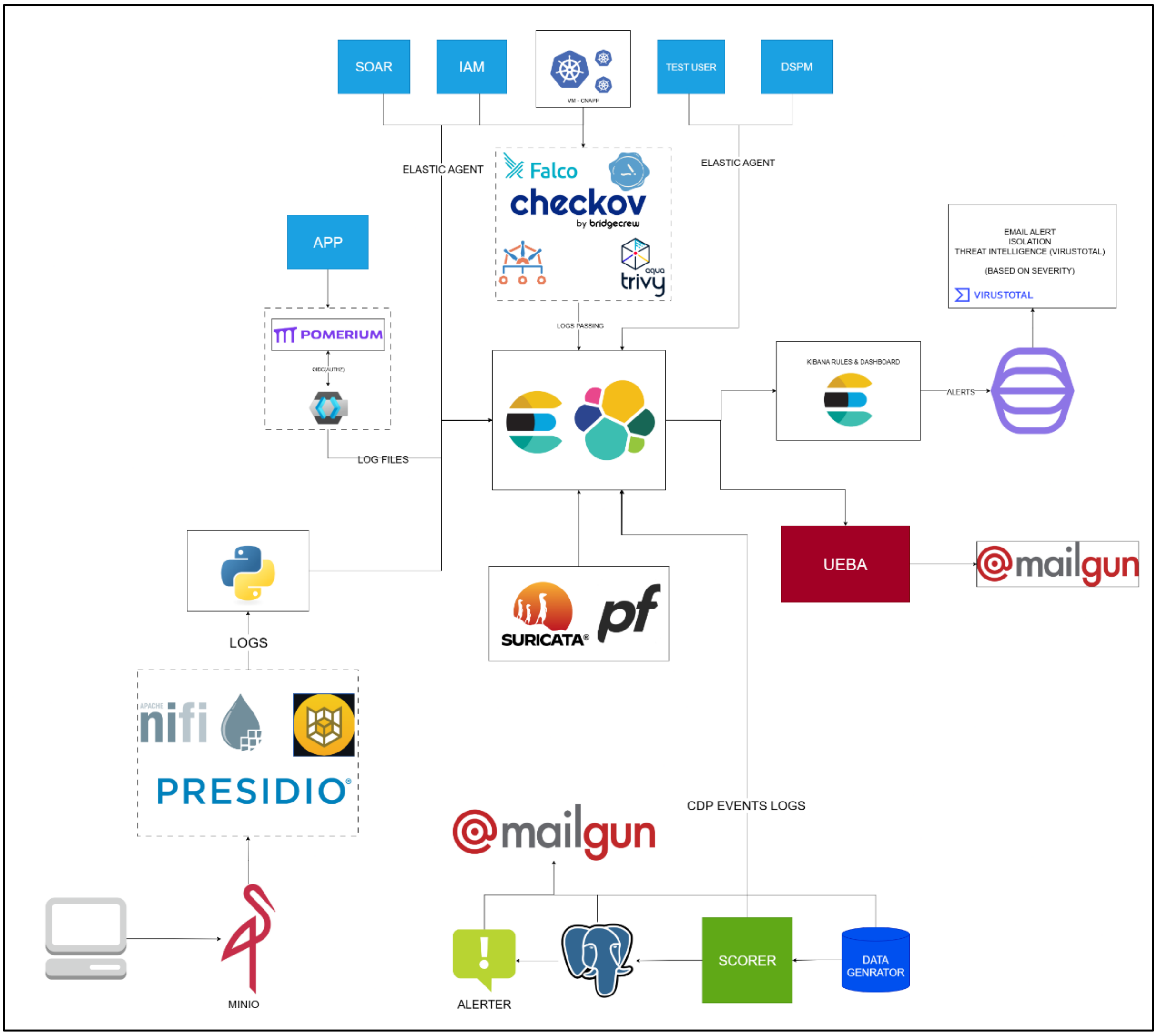
- **Centralized Telemetry:**
  All solutions forward structured logs to **Elasticsearch** for indexing, visualization, and detection rule application. Kibana dashboards provide unified visibility across identity, data, and workload layers.
- **Identity-Centric Access Control:**
  Keycloak + Pomerium handle authentication and access enforcement for internal applications. Logs from these components feed into ELK and correlate with network telemetry from Suricata.
- **Data Flow & Policy Enforcement:**
  DSPM pipelines ingest data from MinIO, classify using Presidio, enforce policies via Custodian, and visualize compliance posture in Kibana.
- **Workload & Runtime Security:**
  CNAPP components (Trivy, Falco) provide scanning and runtime monitoring of containerized environments. Alerts feed into the central ELK pipeline for correlation.
- **Detection Fabric:**
  XDR logic leverages logs from ZTNA, DSPM, and CNAPP, applying UEBA and detection rules to surface insider threats and multi-stage attack behaviors.

This **interoperability** ensures that **threats can be detected across multiple layers simultaneously**, mirroring how modern SOC platforms function in production environments.

5. Architecture & Integration

5.1 High-Level Architecture

The Project Horizon MVP has been designed using a **layered Zero Trust security model**, integrating multiple open-source components to cover **identity, data, network, workload, and detection layers**.



Architecture Layers:

- **Identity & Access Layer:**
  Keycloak (IdP) and Pomerium (proxy) enforce strong authentication, TLS encryption, and identity-based routing for internal applications.

- **Data & DSPM Layer:**
  MinIO acts as object storage, with NiFi handling ingestion pipelines, Presidio classifying sensitive data, and Custodian applying policy enforcement. This layer provides visibility and control over data flows.

- **Network Layer:**
  Suricata monitors ingress and egress traffic to detect anomalies, policy violations, and suspicious behavior, feeding alerts to ELK.

- **Telemetry & Detection Layer:**
  All components send structured logs to Logstash → Elasticsearch for indexing. Kibana dashboards provide visibility, and detection rules + UEBA models enable threat detection and insider risk analytics.

- **Workload & Runtime Security:**
  CNAPP tools (Trivy, Falco, K3s) secure containerized workloads, scan for vulnerabilities, and monitor runtime activity.

This layered structure enables **policy enforcement and monitoring at every control point**, ensuring no single point of trust exists aligning strongly with Zero Trust principles.

5.2 Integration Highlights

The strength of Project Horizon lies in the **tight integration between these components**, transforming them from isolated tools into a **cohesive security fabric**.

5.2.1 Identity & Access Integration

- Keycloak serves as the **central Identity Provider**, managing users, groups, and authentication flows.

- Pomerium integrates with Keycloak to **enforce identity-aware routing**, providing ZTNA-style secure access to internal applications.

- Authentication and access logs are shipped to ELK, where they are correlated with Suricata network logs for visibility and detection.

5.2.2 Data & DSPM Integration

- NiFi pipelines pull data from simulated sources and route it to MinIO object storage.
- Presidio scans data in motion and at rest to detect sensitive fields (PII/PCI).
- Custodian applies remediation policies (e.g., blocking, redaction, tagging).
- All DSPM activities generate structured logs, ingested into ELK for compliance dashboards and detection triggers.

5.2.3 Telemetry & Detection Integration

- Suricata generates real-time network telemetry and security alerts.
- Keycloak, Pomerium, NiFi, Custodian, Trivy, and Falco all forward logs to Logstash.
- Logstash enriches and normalizes data before indexing into Elasticsearch.
- Kibana dashboards provide unified visibility across identity, network, data, and workload dimensions.
- Detection rules and UEBA analytics correlate multi-domain events to detect insider threats, lateral movement, and policy violations.

5.2.4 CNAPP & Runtime Integration

- Container images are scanned with Trivy for vulnerabilities before deployment.
- Falco monitors runtime activity, generating alerts on anomalous container behavior.
- Runtime telemetry is integrated into the ELK pipeline, enabling detection rules to correlate workload activity with network and identity events.

5.2.5 End-to-End Workflow Example

Scenario: An external user attempts to access a sensitive internal dashboard.

1. **Authentication:** Keycloak authenticates the user; Pomerium enforces route-based access.
2. **Network Monitoring:** Suricata inspects the traffic for anomalies.
3. **Data Classification:** DSPM detects sensitive data within accessed resources.

4. **Telemetry Centralization:** All logs flow to ELK; detection rules check for unusual behavior.
5. **Alert Generation:** A correlated alert is triggered if abnormal access + sensitive data + suspicious traffic are observed simultaneously.
6. **Response:** SOC can investigate via Kibana dashboards, and in future phases, SOAR workflows will automate containment.

6. Per-Solution Summaries

6.1 Zero Trust Network Access (ZTNA) & SASE

6.1.1 Objective & Requirement Mapping

The goal of the **ZTNA & SASE** pillar was to establish a **secure, identity-aware access perimeter** for internal services, aligning with Zero Trust enforcement requirements. Key objectives included:

- Enforce **strong user authentication** and **per-route authorization** for all internal web applications.
- Secure communications end-to-end using **TLS encryption**.
- Implement basic network-layer inspection to detect suspicious activity at the perimeter.
- Forward all authentication and traffic logs to ELK for centralized visibility and detection.

Requirement	**ZTNA & SASE Coverage**
Identity-based access control	Keycloak + Pomerium authentication & routing
Encrypted communications	TLS (self-signed) between clients, proxy, and backend
Secure perimeter enforcement	Reverse proxy with policy-based routing
Network telemetry & threat visibility	Suricata IDS monitoring of ingress traffic
Centralized logging & detection	Auth logs + network logs forwarded to ELK

6.1.2 Tools & Components

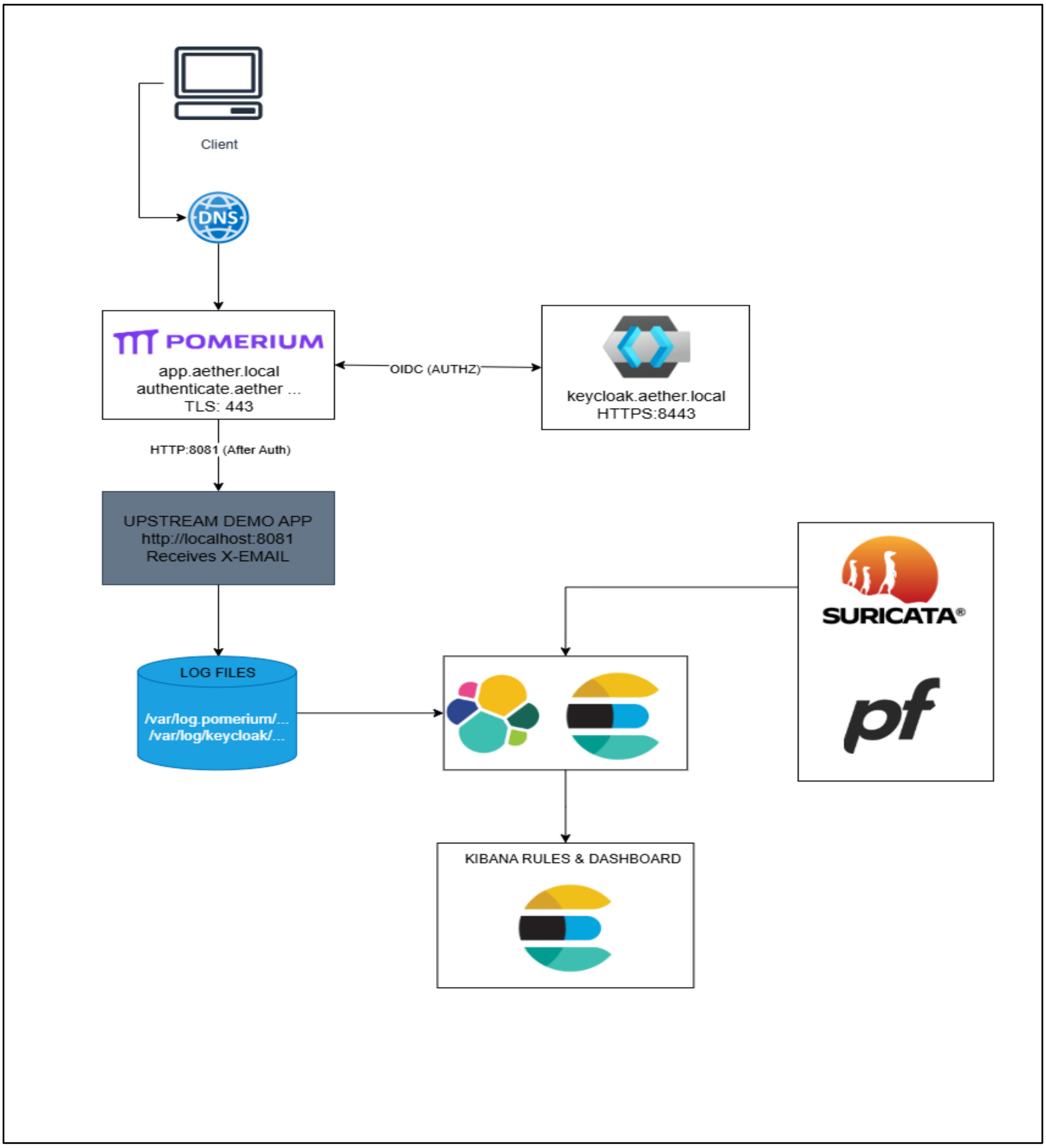
Component	**Purpose**
Keycloak	Acts as the Identity Provider (IdP), handling user authentication and token issuance.
Pomerium	Functions as an identity-aware reverse proxy, enforcing per-route access policies based on Keycloak authentication.
Suricata	Provides network intrusion detection at the perimeter, monitoring ingress traffic for anomalies.
pfSense	Used in some flows for basic routing and firewall controls.
ELK Stack	Receives and indexes logs from Keycloak, Pomerium, and Suricata for visibility and detection.

6.1.3 Configurations & Customizations

- **Keycloak Realm Configuration:**
  - A dedicated realm was created for Project Horizon.
  - User roles and groups were defined to represent internal access tiers.
  - Token lifetimes were adjusted for testing secure session expiry.
- **Pomerium Policy Rules:**
  - Route-based access rules were defined to restrict access to internal dashboards and APIs.
  - Pomerium was configured to use Keycloak as its OIDC provider with TLS enabled.
  - Custom certificates were used to secure communication between proxy and backends.
- **Suricata:**
  - Deployed in inline/monitor mode to observe ZTNA ingress traffic.
  - Rulesets were customized to alert on protocol anomalies, brute force attempts, and suspicious URLs.
  - Output was configured to JSON format for ELK ingestion.

- **Log Forwarding:**
    - Keycloak and Pomerium logs were shipped via Filebeat to Logstash → Elasticsearch.
    - Suricata alerts were ingested directly into ELK with custom index templates.

6.1.4 Mini Architecture Diagram



6.1.5 Achievements & Gaps

Achievements:

- Functional ZTNA perimeter with identity-aware routing and TLS.
- Centralized telemetry covering identity and network layers.
- Basic detection rules operational for authentication anomalies.
- Integration between Keycloak, Pomerium, Suricata, and ELK verified end-to-end.

Gaps & Enhancements:

- No enterprise federation (e.g., Azure AD, Okta).
- TLS uses self-signed certs; no automated PKI integration.
- Single-node deployment without HA.
- Limited policy depth (e.g., no device posture or geo rules).

6.2 Customer Data Platform (CDP)

6.2.1 Objective & Requirement Mapping

The **Customer Data Platform (CDP)** component was developed to **simulate user behavior analytics and engagement pipelines** as part of broader **data visibility and detection strategy**. While not a production marketing system, this CDP prototype acts as a **behavioral telemetry generator**, allowing us to model **user activity**, **data processing**, and **alert generation** in a controlled lab environment.

Key objectives included:

- Ingest and process **synthetic user interaction data** to mimic real-world customer behaviors.
- Perform **basic churn scoring** and generate engagement events.
- Forward all telemetry to ELK for centralized logging, visualization, and detection rule testing.
- Provide a **data source for DSPM** and XDR components to analyze downstream flows and policy enforcement.

Requirement	CDP Coverage

Data visibility and telemetry	Synthetic user events processed and logged
Behavior analytics	Basic churn scoring and event enrichment
Integration with detection & DSPM	CDP telemetry consumed by DSPM classification and XDR detection
Centralized logging	Structured logs forwarded to ELK
Simulated engagement triggers	Campaign events generated for testing pipelines

6.2.2 Tools & Components

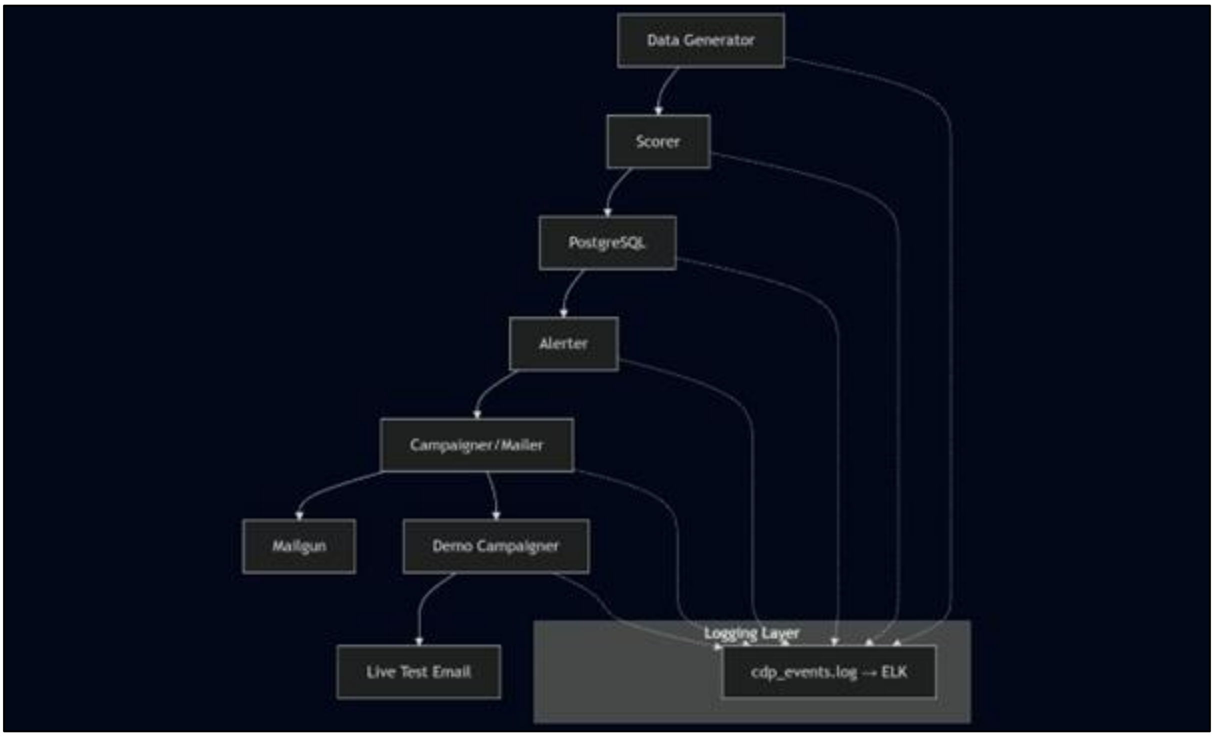
Component	**Purpose**
PostgreSQL	Stores synthetic user and interaction data used for churn scoring and event generation.
Python Scripts	Process data, calculate churn scores, and generate synthetic engagement events.
Mailgun API	Simulates outbound campaign triggers (email alerts) from engagement events.
ELK Stack	Central telemetry sink for user events, engagement logs, and processing pipeline outputs.

6.2.3 Configurations & Customizations

- **Database Initialization:**
  - PostgreSQL was populated with synthetic user profiles and interaction histories.
  - Data included session timestamps, activity types, frequency scores, and churn likelihood fields.
- **Behavioral Analytics Scripts:**
  - Custom Python scripts computed churn scores based on inactivity periods, session drops, and frequency anomalies.
  - Engagement events were generated for high-risk users (e.g., retention campaigns).
- **Campaign Simulation:**

  - Mailgun API was integrated to simulate real outbound campaign triggers.
  - Each engagement trigger was logged with metadata (user ID, campaign type, timestamp).
- **Log Forwarding:**
  - All CDP outputs (raw events, churn scores, campaign triggers) were logged in JSON format and shipped to ELK.
  - Logstash pipelines parsed the fields for dashboarding and correlation.

6.2.4 Mini Architecture Diagram



6.2.5 Achievements & Gaps

Achievements:

- Functional behavioral telemetry generator aligned with enterprise data visibility use cases.
- Realistic churn scoring pipeline implemented in lab environment.
- Centralized logging and visualization of user behavior patterns.

- Successful integration with downstream DSPM and XDR components for data flow and detection testing.

Gaps & Enhancements:

- Synthetic data only; no real production datasets.
- No advanced ML-based analytics basic heuristics only.
- Mailgun integration was simulated; not production-compliant.
- No scaling or high availability for analytics pipeline.

6.3 Data Security Posture Management (DSPM)

6.3.1 Objective & Requirement Mapping

The **DSPM** solution was implemented to provide **visibility, classification, and protection of sensitive data** across storage and processing systems, while enabling **policy enforcement** and **compliance reporting**.

Its purpose is to answer the key Zero Trust question:

“Where is our sensitive data, who is accessing it, and how is it being protected?”

Key objectives included:

- **Discover and classify** sensitive data (e.g., PII, PCI) in storage systems.
- **Enforce data security policies** such as redaction, blocking, or tagging.
- **Visualize compliance posture** through ELK dashboards.
- **Integrate DSPM telemetry** with other pillars (CDP, CNAPP, XDR) to enable detection and data-centric analytics.
- Provide a **foundation for compliance mapping** (e.g., GDPR, PCI-DSS) in future phases.

Requirement	**DSPM Coverage**
Data discovery & classification	Presidio-based classification on data flows and storage
Policy enforcement	Custodian rules for redaction, blocking, tagging

Compliance visibility	DSPM dashboards in Kibana with policy violation logs
Data flow telemetry	Integration with ELK for monitoring and XDR correlation
Scalable architecture foundation	Modular pipeline using MinIO + NiFi + Presidio + Custodian

6.3.2 Tools & Components

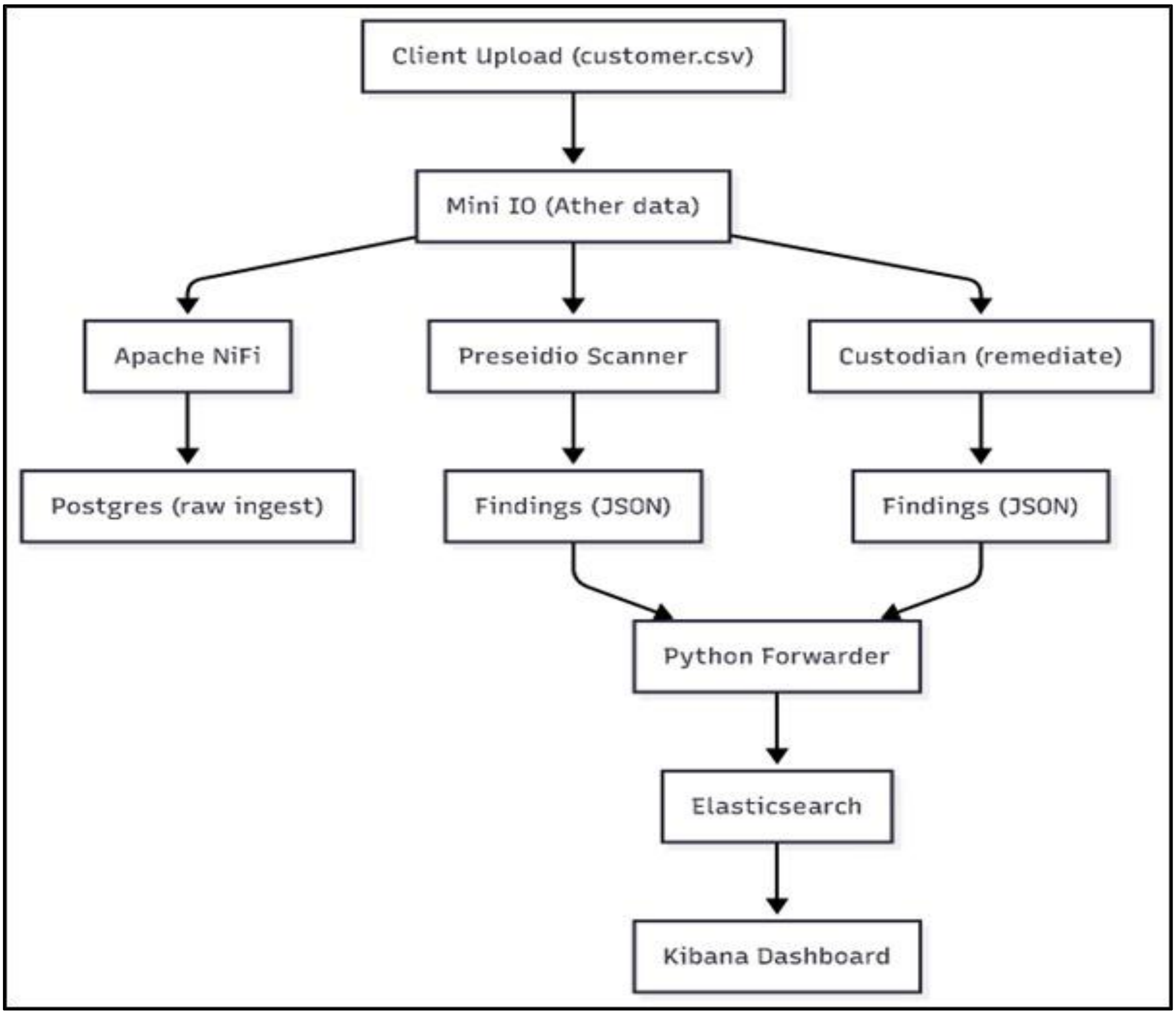
Component	**Purpose**
MinIO	Acts as S3-compatible object storage for synthetic datasets.
Apache NiFi	Manages data ingestion pipelines and routing between storage and processors.
Microsoft Presidio	Performs entity recognition and classification (e.g., credit cards, emails, PII).
Cloud Custodian	Applies policy actions on classified data (e.g., redact, delete, tag).
ELK Stack	Aggregates and visualizes DSPM telemetry (flows, classifications, violations).

6.3.3 Configurations & Customizations

- **MinIO Storage Setup:**
  - Configured MinIO buckets (customer-data, logs, classified) with proper access keys.
  - Synthetic datasets containing PII-like fields were uploaded for testing.
- **NiFi Pipeline:**
  - Designed flow to pull data from MinIO → pass to Presidio → re-ingest into classified bucket.
  - NiFi processors were tuned for JSON parsing, error handling, and tagging.
- **Presidio Classification:**
  - Configured recognizers for credit card numbers, email addresses, phone numbers, and national IDs.

  - Output included entity type, confidence score, and location in the dataset.
- **Custodian Policies:**
  - Wrote custom YAML policies to **redact sensitive fields** and **log violations** when encountering classified data.
  - Policies included tagging non-compliant objects and deleting high-risk entries in test scenarios.
- **Logging & Visualization:**
  - NiFi, Presidio, and Custodian logs were shipped to Logstash for normalization.
  - Elasticsearch indexed logs in dspm-* indices, and Kibana dashboards were created to show:
    - Classification activity over time
    - Policy violations by type
    - Data flow lineage and transformations

6.3.4 Mini Architecture Diagram



6.3.5 Achievements & Gaps

Achievements:

- End-to-end DSPM pipeline implemented using open-source components.
- Successful classification of PII/PCI entities in synthetic datasets.
- Policy enforcement via Custodian with logged violations.
- DSPM telemetry integrated with ELK for dashboards and detection.
- Demonstrated data flow lineage and compliance reporting potential.

Gaps & Enhancements:

- Synthetic data only; lacks real production coverage.
- No automatic remediation workflows beyond basic redaction.

- No integration with enterprise DLP or classification taxonomies.
- Single-node setup without redundancy or scale-out processing.

6.4 Cloud-Native Application Protection Platform (CNAPP)

6.4.1 Objective & Requirement Mapping

The **CNAPP** solution was implemented to **secure containerized workloads**, **detect vulnerabilities**, and **enforce runtime security** within the Project Horizon lab environment.

Its purpose is to protect the **workload and application layer** of the Zero Trust model ensuring that even if identity or network layers are compromised, the workloads themselves are hardened and monitored for suspicious activity.

Key objectives included:

- **Scan container images** for known vulnerabilities and misconfigurations prior to deployment.
- **Monitor container runtime activity** to detect anomalous or malicious behavior.
- **Forward security events and telemetry** to ELK for centralized visibility.
- **Integrate CNAPP alerts** with XDR and DSPM components for correlated detection.
- Establish a **baseline pipeline for secure cloud-native deployments** using open-source tools.

Requirement	**CNAPP Coverage**
Container image vulnerability scanning	Trivy scanning integrated into CI/lab workflows
Runtime threat detection	Falco runtime monitoring for container anomalies
Centralized logging & visibility	All CNAPP telemetry forwarded to ELK
Integration with detection fabric	CNAPP alerts correlated with XDR rules and Suricata telemetry
Secure workload posture	Baseline lab configuration for container protection using K3s + Falco + Trivy

6.4.2 Tools & Components

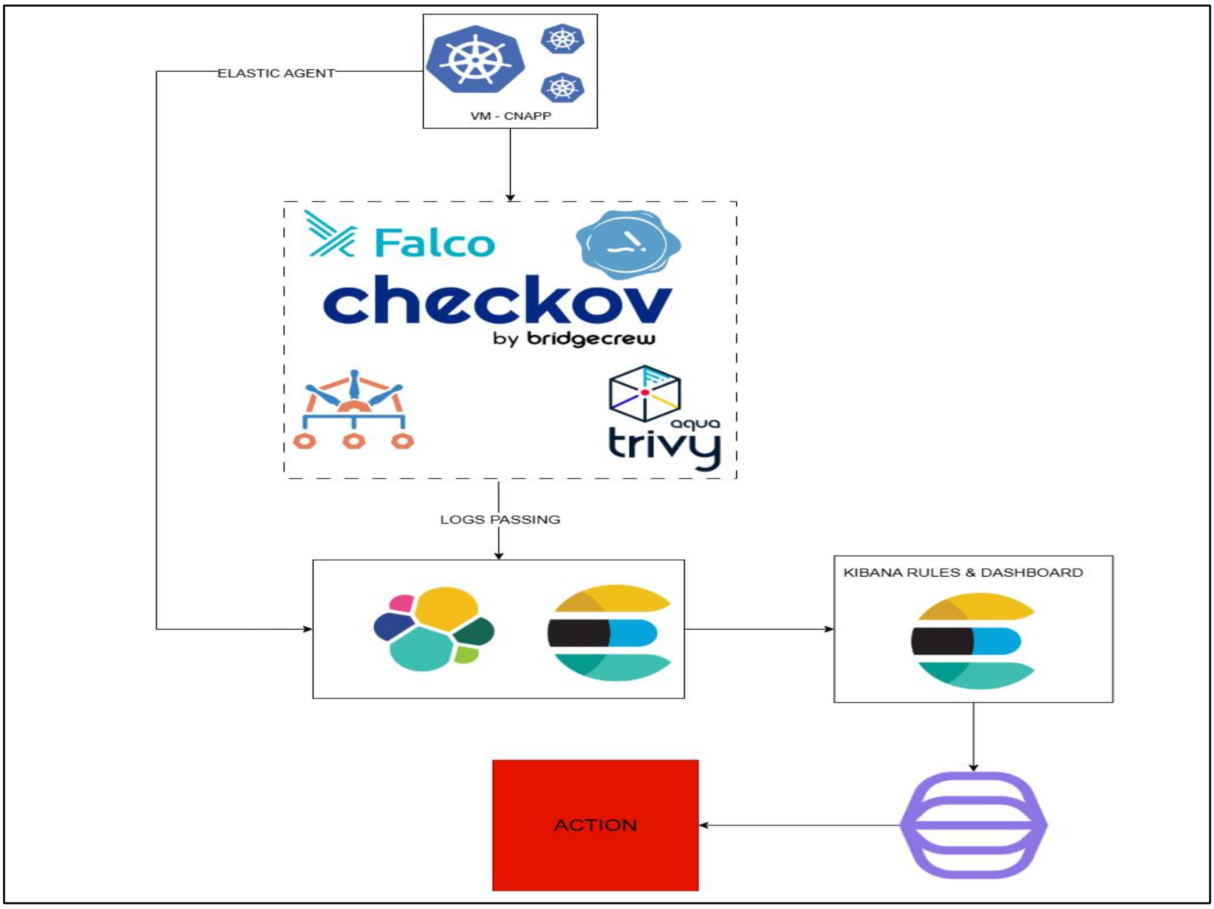
Component	Purpose
K3s	Lightweight Kubernetes distribution used to deploy and run containerized applications in the lab.
Trivy	Performs vulnerability scanning of container images to detect CVEs and misconfigurations.
Falco	Monitors container runtime activity for abnormal behavior, system call anomalies, and suspicious patterns.
ELK Stack	Central logging and detection platform for all CNAPP alerts and telemetry.

6.4.3 Configurations & Customizations

- **K3s Cluster Setup:**
  - Deployed single-node K3s cluster hosting multiple lab containers (e.g., NGINX, telemetry apps).
  - Enabled necessary CNI and API access for Falco and Prometheus integration.
- **Trivy Scanning:**
  - Configured Trivy to scan container images both **pre-deployment** and periodically on deployed workloads.
  - Reports were generated in JSON format and shipped to ELK for indexing under cnapp-trivy-*.
- **Falco Runtime Security:**
  - Deployed Falco as a DaemonSet within K3s to monitor all container activity.
  - Custom Falco rules were added to detect:
    - Unexpected process execution inside containers (e.g., /bin/bash in NGINX pods)
    - File system modifications in restricted directories
    - Outbound network connections from non-network pods
  - Alerts were exported via Falcosidekick to ELK.

- **Log Forwarding:**
  - All CNAPP logs (Trivy, Falco, K3s audit) were routed through Logstash pipelines and indexed in Elasticsearch.
  - Detection rules were created to trigger on specific Falco event types combined with network or identity anomalies.

6.4.4 Mini Architecture Diagram



6.4.5 Achievements & Gaps

Achievements:

- Full CNAPP pipeline deployed in a lab K3s environment.
- Vulnerability scanning operational with Trivy, feeding ELK.
- Real-time runtime detection achieved with Falco + Falcosidekick.
- Integration with ELK, XDR, and DSPM demonstrated.
- Detection of simulated attacker behaviors inside containers verified.

Gaps & Enhancements:

- No image signing or admission controller enforcement yet.
- Single-node cluster; no HA or scaling.
- Vulnerability management not integrated with patching workflows.
- Runtime rules were limited to basic attack behaviors.

6.5 Extended Detection & Response (XDR) & Insider Risk

6.5.1 Objective & Requirement Mapping

The **XDR & Insider Risk** pillar was designed to act as the **central detection, analytics, and response layer** of Project Horizon. Its role is to **aggregate telemetry** from all other solution pillars, **detect multi-stage or insider attack patterns**, and **provide SOC operators with a unified view** of security events.

Key objectives included:

- Ingest and normalize telemetry from **identity**, **network**, **data**, and **workload** layers.
- Implement **detection rules** and **behavioral analytics** to surface actionable security alerts.
- Build **dashboards and investigation views** to help SOC analysts detect insider threats and complex attack chains.
- Validate detection coverage through controlled attack simulations.
- Lay the foundation for future **SOAR (Security Orchestration, Automation & Response)** workflows.

Requirement	**XDR & Insider Risk Coverage**
Centralized telemetry ingestion	Identity (Keycloak), Network (Suricata), Data (DSPM), Workload (Falco), Behavior (CDP)
Threat detection & correlation	Multi-source detection rules using ELK and UEBA
Insider risk detection	Behavioral anomalies and unusual access patterns analyzed through UEBA logic

Investigation & response dashboards	Kibana visualizations and detection tables built for SOC analysis
SOC modernization foundation	XDR layer acts as the detection and analytics fabric across all pillars

6.5.2 Tools & Components

Component	**Purpose**
ELK Stack	Core telemetry pipeline and detection engine for all integrated sources.
Logstash	Ingestion and normalization of structured logs from all solution components.
Kibana	Dashboards, detection views, and investigation panels for SOC analysts.
UEBA Modules	Basic user and entity behavior analytics logic applied on identity, network, and CDP logs.
Detection Rules	Custom correlation rules in ESQL for attack pattern detection.

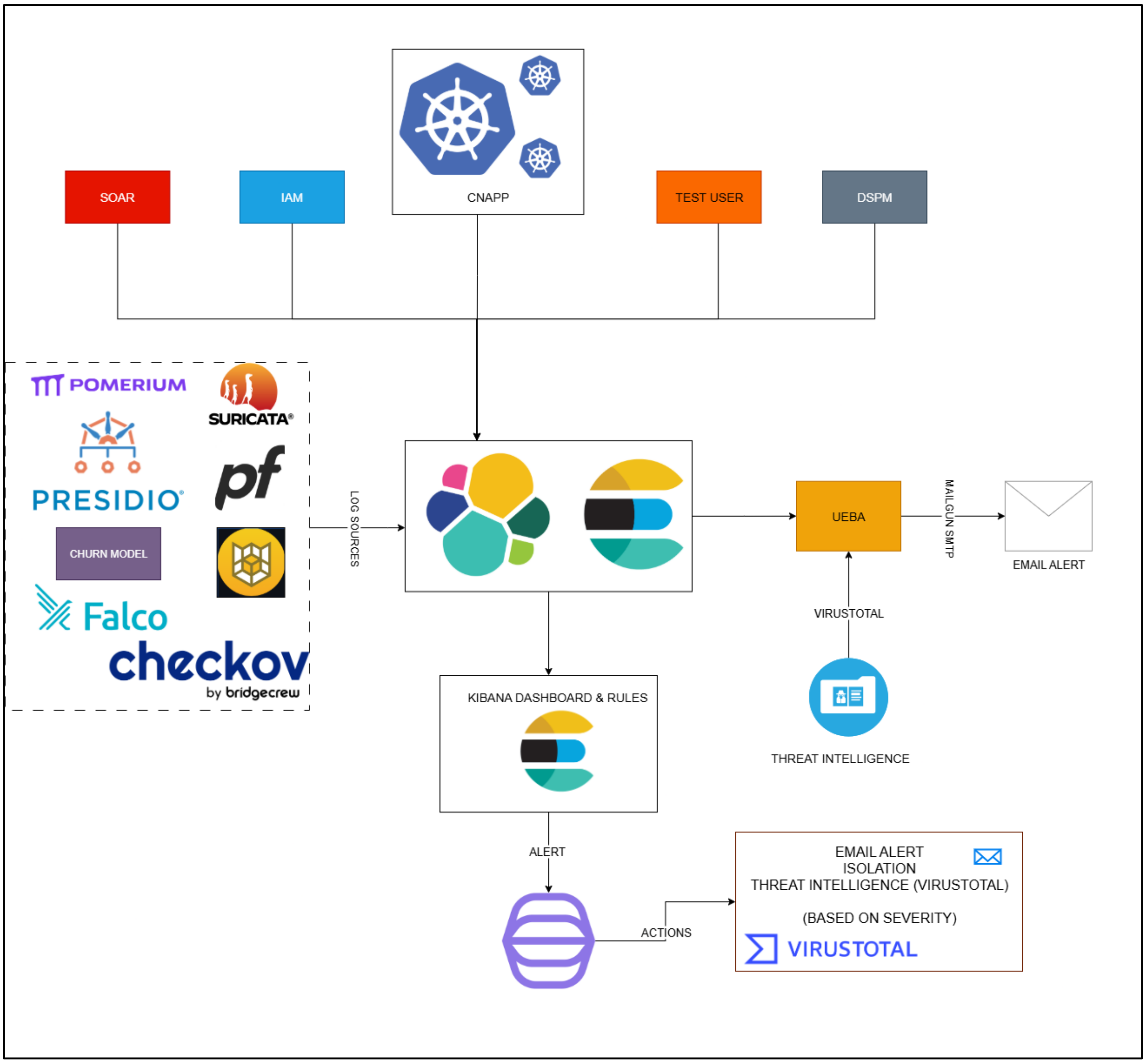
6.5.3 Configurations & Customizations

- **Log Ingestion Pipelines:**
  - Logstash pipelines were configured to accept logs from:
    - Keycloak (auth events),
    - Pomerium (access logs),
    - Suricata (network alerts),
    - DSPM (classification & policy violations),
    - Falco (runtime security),
    - CDP (behavioral telemetry).
  - Data was normalized into structured Elasticsearch indices by source type (auth-*, net-*, dspm-*, falco-*, cdp-*).
- **Detection Rule Development:**

- ESQL correlation rules were created to detect multi-stage activities. Examples:
    - **Brute Force + Suspicious Access**: multiple failed logins from Suricata anomaly IP → successful access within 5 min.
    - **Sensitive Data + Suspicious Network Activity**: DSPM violation followed by outbound traffic.
    - **Runtime Anomaly + Lateral Movement**: Falco alert combined with Suricata scanning activity.
- Detection rules were tested using controlled lab simulations.

- **UEBA Logic:**
    - Basic behavior baselines were generated from CDP churn telemetry and Keycloak auth logs.
    - UEBA flagged unusual login times, access location deviations, and churn anomalies used in detection context.
- **Dashboards & Visualizations:**
    - Multiple Kibana dashboards were created, including:
        - **Attack Timeline View** (time-ordered events across pillars)
        - **Top Targeted Assets** (based on Suricata + Falco telemetry)
        - **Policy Violation Heatmaps** (from DSPM)
        - **Insider Risk Behavior Dashboard** (CDP + Keycloak UEBA signals)
        - **Detection Evidence Table** (rule matches, IPs, threat IDs, timestamps)

6.5.4 Mini Architecture Diagram



6.5.5 Achievements & Gaps

Achievements:

- Centralized telemetry and detection fabric successfully built on ELK.
- Detection rules operational and validated via MITRE ATT&CK technique simulations.
- UEBA logic enriched detection contexts with behavioral anomalies.
- Dashboards provided unified visibility and investigation views.
- Full integration achieved across identity, network, data, workload, and behavioral telemetry.

Gaps & Enhancements:

- UEBA is basic - lacks advanced baselining and ML capabilities.
- Detection logic is handcrafted; no automated tuning or rule lifecycle management.
- No SOAR response automation implemented yet.
- Lab-only scale needs distributed deployment for production SOC use.

7. Roadmap & Scaling Strategy

The **Project Horizon MVP** successfully demonstrates the **technical feasibility and integration** of a Zero Trust security architecture using open-source components. The **next phase** focuses on evolving this MVP into a **scalable, production-grade enterprise platform** capable of meeting operational, compliance, and resilience requirements.

This section outlines a **phased roadmap** structured across **short-term**, **medium-term**, and **long-term** milestones, ensuring a controlled and strategic transition from lab prototype to full enterprise deployment.

7.1 Strategic Goals

The scaling strategy is designed to meet the following overarching goals:

- **Production-Grade Resilience** - introduce clustering, HA, and automated PKI for secure, always-on operations.
- **Federated Identity & Access** - integrate enterprise IdPs and MFA for centralized, conditional access.
- **Telemetry & Detection Maturity** - move from basic MVP logs to enriched, correlated telemetry and advanced detection analytics.
- **Compliance Alignment** - implement GDPR/PCI reporting, automated compliance checks, and auditable pipelines.
- **SOC Modernization** - establish automated workflows, SOAR integrations, and advanced threat detection models.

7.2 Phased Roadmap

Phase	Key Initiatives	Expected Outcomes

Phase 1 Short-Term	- Integrate Keycloak with Azure AD / Okta - Automate TLS certificate issuance via ACME/PKI - Enhance Logstash pipelines with enrichment - Create detection content packs based on MVP rules - Stabilize DSPM pipelines with real datasets	Enterprise SSO + MFA Secure perimeter with managed PKI Richer telemetry Improved data protection accuracy
Phase 2 Medium-Term	- Deploy clustered ELK and Keycloak setups - Introduce Kubernetes HA for CNAPP - Expand DSPM policy coverage (PII, PCI, PHI) - Add anomaly detection models for UEBA - Implement compliance dashboards (GDPR/PCI)	Scalable core services Advanced data security posture Behavioral analytics at scale Regulatory visibility
Phase 3 Long-Term	- Multi-region deployment for resiliency - Deploy automated SOAR workflows (e.g., Tines, Cortex) - Integrate real-time threat intelligence feeds - Mature XDR detection with ML/AI-assisted triage - Align architecture with SOC2 / ISO 27001 compliance frameworks	Enterprise-grade resilience Automated response & threat hunting Predictive detection capabilities Full compliance alignment

7.3 Key Enhancement Areas

Capability Area	**MVP State**	**Target State**
Identity & Access	Standalone Keycloak, no MFA, self-signed TLS	Federated IdP (Azure AD/Okta), MFA, PKI automation, conditional access
Telemetry & SIEM	Raw logs, basic detection rules	Enriched pipelines, advanced detection content packs, cross-layer correlation
DSPM	Synthetic data, basic rules, manual dashboards	Real datasets, expanded entity types, automated remediation, compliance dashboards

CNAPP	Single-node K3s, manual scanning	HA clusters, admission controllers, image signing, automated patching workflows
XDR & UEBA	Handcrafted rules, basic behavior analysis	Anomaly detection models, ML-assisted detection, UEBA enrichment from real user behavior
Compliance & Governance	Manual dashboards, no reporting	GDPR/PCI compliance mapping, automated reporting, SOC2/ISO alignment
SOC Operations	Manual investigation, no SOAR	Automated incident response, enrichment pipelines, threat intelligence integration, playbook execution

7.4 Integration Priorities

To ensure maximum impact with minimal disruption, the following integration priorities are recommended:

1. **Identity Federation First**
   - Connect Keycloak to Azure AD/Okta with SAML/OIDC.
   - Enforce MFA and SCIM provisioning.
   - Replace lab self-signed certificates with automated PKI (e.g., Let’s Encrypt or enterprise CA).
2. **Telemetry Enrichment Second**
   - Expand Logstash ingestion pipelines with enrichment (GeoIP, threat intel, user context).
   - Standardize log schemas for all pillars to support correlation.
3. **Detection & DSPM Maturity Third**
   - Convert MVP detection rules into reusable content packs.
   - Expand DSPM classification & policy scope, integrate real data flows.
4. **Infrastructure Scaling Fourth**
   - Cluster ELK, Keycloak, and CNAPP components for HA and resiliency.

    - Enable centralized secrets & configuration management (e.g., Vault, Sealed Secrets).

5. **SOAR & Compliance Final**
    - Integrate SOAR workflows to automate containment and investigation.
    - Deploy compliance dashboards mapped to GDPR/PCI/SOC2 frameworks.

7.5 Expected Impact

By following this phased roadmap, Project Horizon will transform from a **functional MVP** to a **production-grade Zero Trust platform** with the following impacts:

- **Stronger Identity Security:** Enterprise SSO, MFA, and conditional access significantly reduce credential compromise risk.
- **Improved Data Posture:** Expanded DSPM ensures sensitive data is classified, protected, and monitored at scale.
- **Smarter Detection:** Advanced correlation and UEBA enable earlier detection of multi-stage and insider threats.
- **Automated Response:** SOAR workflows reduce analyst workload and response times.
- **Enterprise Resilience:** Multi-region deployments ensure availability and compliance with business continuity objectives.
- **Regulatory Confidence:** GDPR, PCI-DSS, and SOC2 alignment provides auditable security posture to stakeholders and regulators.

8. Meeting Requirements at Production & Enterprise Scale

The Project Horizon MVP establishes a **functional Zero Trust security architecture** using open-source tools in a controlled lab environment.
To meet **enterprise requirements**, the next phase focuses on **eliminating lab constraints**, **introducing production-grade resilience**, and **aligning with enterprise security and compliance frameworks**.

This section maps how each major requirement will be addressed through targeted enhancements, resulting in a **secure, scalable, and auditable enterprise platform**.

8.1 Strategic Objective

The strategic goal is to **evolve the MVP into a production-ready security platform** that delivers:

- **High availability and fault tolerance** through clustering, load balancing, and multi-region deployments.
- **Federated Identity and Access** with enterprise IdPs, MFA, and conditional access.
- **Scalable telemetry pipelines** with enriched analytics for advanced threat detection.
- **Granular policy enforcement** across identity, data, and workloads.
- **Compliance-aligned observability and reporting** for GDPR, PCI-DSS, and SOC2.
- **Robust security hardening** and automated response workflows.

8.2 Key Enhancement Areas

Capability Area	MVP State	Production & Enterprise State
Identity & Access	Standalone Keycloak, local users, self-signed TLS	Federated SSO (Azure AD / Okta), MFA, SCIM provisioning, enterprise PKI, conditional access
Infrastructure	Single-node deployments, manual scaling	Clustered deployments, load balancers, auto-scaling groups, failover, centralized secrets management
Network & TLS	Static routes, self-signed certificates	Managed PKI, automated TLS issuance, private DNS zones, service mesh for east-west control
Telemetry & SIEM	Raw logs, basic dashboards, limited detection	Enriched pipelines, threat intel integration, cross-layer correlation, advanced UEBA, SOC dashboards
Policy & DSPM	Basic bucket-level DSPM, manual enforcement	Expanded entity coverage (PII/PCI/PHI), automated remediation, policy orchestration, compliance dashboards

Workload Protection	Single-node K3s, manual scanning, basic Falco rules	HA clusters, image signing, admission control, automated patch workflows, expanded runtime rule sets
Compliance & Reporting	Manual dashboards, no formal framework mapping	GDPR/PCI/SOC2 compliance mapping, automated audit reports, long-term retention, integration with GRC systems
SOC Operations	Manual investigation, no SOAR	Automated response workflows, playbook execution, threat intel feeds, SOC process integration

8.3 Production-Grade Deployment Blueprint

The transformation involves a **layered upgrade** across all solution components:

1. **Identity & Access Layer**
    - Integrate Keycloak with **Azure AD or Okta** using OIDC/SAML for enterprise SSO.
    - Enforce **MFA** and **conditional access** policies.
    - Replace self-signed certs with **automated PKI** (Let's Encrypt or corporate CA).
    - Implement SCIM for user/group provisioning automation.
2. **Data & DSPM Layer**
    - Expand classification coverage to include PII, PCI, PHI with automated tagging and remediation.
    - Integrate DSPM dashboards with compliance frameworks (GDPR, PCI-DSS).
    - Deploy NiFi and Custodian in HA mode with real-time processing.
3. **Telemetry & Detection Layer**
    - Migrate to **clustered ELK** for scale and resilience.
    - Enrich telemetry with GeoIP, threat intel, and user context.
    - Expand detection rules and UEBA logic for real-world behavior analytics.

    - Integrate real-time threat feeds.

4. **Workload & CNAPP Layer**
    - Deploy K3s/Kubernetes clusters with **multi-node HA**.
    - Enforce image signing and admission controller policies (e.g., Kyverno, Cosign).
    - Integrate Falco with SOAR for automated containment.
5. **Compliance & SOC Layer**
    - Build **compliance dashboards** mapped to GDPR/PCI/SOC2 frameworks.
    - Automate reporting and audit logging.
    - Integrate SOAR (e.g., Tines, Cortex) to orchestrate response workflows.
    - Align with SOC2 / ISO 27001 controls and documentation requirements.

8.4 Expected Outcomes

By applying these enhancements, the Project Horizon architecture will achieve:

- **Stronger Identity Assurance** - MFA, federation, and PKI will ensure secure, centralized access control.
- **Resilient Infrastructure** - Clustering, scaling, and HA eliminate single points of failure.
- **Advanced Detection & Response** - Enriched telemetry and UEBA provide early detection of complex threats.
- **Comprehensive Data Security** - Expanded DSPM ensures sensitive data is continuously classified, protected, and monitored.
- **Compliance Readiness** - GDPR/PCI/SOC2 mapping with automated dashboards and reports.
- **Operational Efficiency** - SOAR automation reduces analyst workload and improves mean time to response (MTTR).
- **Scalability & Flexibility** - The architecture becomes adaptable to production workloads, multi-region environments, and regulatory requirements.

8.5 Alignment with Requirements

The table below summarizes how the enterprise-scale enhancements directly address strategic security objectives:

Objective	MVP Coverage	Enterprise Enhancement Impact
Zero Trust Enforcement	ZTNA perimeter, identity-aware routing, TLS	MFA, conditional access, mesh security, federated SSO
Centralized Visibility & Detection	ELK ingestion from all pillars, dashboards, basic rules	Enriched pipelines, advanced correlation, UEBA, SOAR
Data Protection & Compliance	DSPM pipeline for classification & policy enforcement	Real datasets, expanded entity types, compliance dashboards, automated reporting
SOC Modernization & Threat Hunting	XDR detection rules, UEBA, dashboards	SOAR workflows, threat intel integration, advanced detection, automated investigations
Scalable Architecture	Single-node lab environment	Clustered, multi-region deployments, enterprise HA, scalable pipelines
Regulatory Readiness	Manual dashboards	GDPR/PCI/SOC2 mapping, automated reports, long-term retention, auditable processes

9. Conclusion, Final Remarks & Acknowledgment

9.1 Conclusion

The **Project Horizon MVP** demonstrates the **feasibility, effectiveness, and strategic value** of implementing a **Zero Trust security architecture** using **open-source technologies** within a controlled lab environment.

Across five integrated solution pillars **ZTNA & SASE**, **CDP**, **DSPM**, **CNAPP**, and **XDR & Insider Risk** the project successfully showcased:

- **End-to-end Zero Trust Enforcement** through identity-aware access control, TLS encryption, workload hardening, and policy enforcement.
- **Centralized Telemetry & Detection** via a unified ELK-based pipeline ingesting identity, network, data, behavioral, and runtime logs.
- **Data Discovery & Protection** using DSPM pipelines with classification, policy actions, and compliance dashboards.
- **Workload Protection** through vulnerability scanning and runtime anomaly detection using Trivy and Falco.
- **Behavioral Analytics & Detection Engineering** leveraging synthetic CDP telemetry, UEBA logic, and custom detection rules mapped to MITRE ATT&CK techniques.
- **Integrated SOC Visibility** with functional dashboards, timelines, and evidence supporting security investigations.

The MVP fulfills its mandate of validating key Zero Trust objectives and lays a **solid technical foundation** for enterprise-scale security transformation.

9.2 Final Remarks

While the MVP was intentionally built with **lab constraints** such as single-node deployments, self-signed TLS, synthetic data, and basic detection logic it successfully **proved the architectural soundness and interoperability** of the entire stack.

The **phased scaling roadmap** (Section 7) and **enterprise requirement alignment** (Section 8) clearly outline how these constraints will be systematically eliminated to achieve a **production-grade, federated, compliant, and resilient Zero Trust platform**.

Key takeaways:

- The **architecture is modular and extensible**, enabling seamless evolution without redesigning core components.
- **Open-source tools are viable** for enterprise security when combined thoughtfully with layered integration and detection engineering.
- **Telemetry correlation across multiple layers** dramatically improves detection quality, especially for multi-stage and insider attack scenarios.

- **Operationalization through SOAR and compliance automation** will be the critical next steps to transform this MVP into a production SOC platform.

Project Horizon not only meets the immediate goals of demonstrating technical feasibility but also provides with a **clear, actionable path** toward **full enterprise adoption** of Zero Trust security principles.