



TECHNICAL DEEP DIVE

Security Architecture for Financial Services

This document provides complete technical detail on every layer of the Alpheous security architecture. It is intended for CCOs, CIOs, operations teams, and due diligence evaluators who need to understand exactly how the system protects your firm's data, enforces regulatory compliance, and maintains fiduciary-grade operational integrity.

alpheous.com

Prepared for evaluation. © 2026 Alpheous, Inc.

Contents

01 Network Architecture

02 Agent Permission Matrix

03 Internal Data Boundaries

04 Identity Protection

05 Self-Healing Infrastructure

06 Alert System

07 Audit Trail and Compliance

08 Threat Mitigation

09 Inbound Defense

10 Credential Management

11 Regulatory Data Ingestion

Network Architecture

Your Alpheous infrastructure has no public ports, no open SSH, and no externally routable IP address. Advisor data sits behind infrastructure that does not exist on the public internet.

Tailscale Mesh VPN

Every connection is encrypted end-to-end and authenticated by device identity. There is no VPN concentrator to compromise. The mesh topology means no single point of failure. Access from anywhere without opening ports or configuring a traditional VPN. Critical for asset managers who need secure remote access to distribution operations.

macOS PF Firewall

A second layer of defense at the packet filter level. Even if Tailscale were bypassed, the firewall blocks all inbound connections that do not originate from the Tailscale network. Advisor data has defense in depth.

SSH Lockdown

Access restricted exclusively to Tailscale IP addresses. Password authentication is disabled. Only key-based authentication is accepted. No external party can reach the system.

The Result

There is no attack surface to scan, no port to probe, and no login page to brute-force. The infrastructure is invisible to the public internet. For asset managers handling sensitive advisor data, this eliminates the most common categories of cyber attack.

Internal Data Boundaries

Data separation inside the appliance is enforced at the database level, not through prompt instructions. Territory boundaries, knowledge tiers, and agent access controls are architectural.

Two-Tier Knowledge Separation

The knowledge system uses two separate database tiers to enforce access control at the database boundary:

Market & Public Knowledge

Detail. FOMC data, FRED economic indicators, general market research (accessible to all queries)

Firm-Level Knowledge

Detail. Investment philosophy, compliance frameworks, operational playbooks (accessible to firm-scoped queries only)

Cross-Agent Isolation Rules

No advisor data is shared between agents except through explicit, logged handoffs. The orchestrator coordinates workflows across agents but does not merge contexts. All data access is logged, making unauthorized cross-contamination both preventable and detectable.

3

Identity Protection

AI-generated communications must never reveal that AI was involved, internal operations, or that any other firm relationship exists.

Identity Injection. Every draft generation prompt includes a mandatory identity block. The AI is instructed that it is a team member at the specific firm, writing in the distribution leader's voice. It cannot reference AI tools, internal systems, other firms, or the fact that it is an AI system. Advisor-facing communications maintain institutional quality.

BLOCKLIST

Description. After marketing content is generated, it passes through a deterministic scanner. The scanner checks for AI tool references, competitor brand names, and content that could reveal AI involvement.

Enforcement. If any match is detected, the content is flagged. There is no override mechanism for the brand-guard scanner.

Rationale. This is a hard gate by design for marketing content. The cost of a false positive is dramatically lower than the cost of a brand violation.

Self-Healing Infrastructure

All agents and supporting services monitored continuously. When something breaks, the system fixes itself before it impacts distribution operations.

Watchdog Service

A dedicated monitor checks all agents and services continuously. When a service fails consecutive health checks, the watchdog auto-restarts it. If the restart fails, an alert escalates to the critical channel. Distribution operations are never silently degraded.

Scheduled Job Monitoring

Tracks whether all scheduled operations ran on time and completed successfully. Missing or late jobs trigger alerts. This includes compliance checks, research data ingestion, and filing deadline monitors.

Critical Escalation

If a core service is down for 10+ continuous minutes, or if 3+ services are simultaneously down, the system escalates to the dedicated critical alert channel. Distribution operations disruptions are never silent.

Boot Recovery

All services use process managers with KeepAlive and RunAtLoad settings. If the infrastructure reboots, every service comes back automatically. Each service pins a specific runtime version to prevent breakage from system updates.

Daily Encrypted Backup

Every database, knowledge base, and audit log is backed up daily with encryption. Combined with nightly snapshots, the system can be restored to any previous day's state. Advisor data is never at risk of loss.

Alert System

A dedicated alert channel surfaces the most important signals across all agents, all services, and all distribution operations.

Compliance Risk **CRITICAL**

Marketing Rule violations detected in drafts, missing disclaimers, unsubstantiated performance claims, filing deadline warnings

Security Violations **CRITICAL**

Cross-firm data references in drafts, credential exposure, unauthorized access attempts, configuration drift

Operational Failures **CRITICAL**

Core service down 10+ minutes, 3+ services simultaneously down, crash-looping agents

Advisor Relationship Signals **HIGH**

Unanswered advisor inquiries, warm-gone-cold detection, negative sentiment in communications

Distribution Operations **HIGH**

Stalled outreach sequences, missed follow-up windows, coverage gaps detected

Data Integrity **HIGH**

Audit log integrity violations, backup failures, credential drift detection, dependency vulnerabilities

30-minute deduplication prevents alert fatigue. All agents feed into the alert system. Graceful fallback ensures that if the alert system itself is unavailable, all agents continue running normally.

6

Autonomous Security Agent

A dedicated security agent monitors the entire infrastructure around the clock, diagnoses issues, and fixes them automatically with safety guardrails designed for financial services operations.

Continuous Security Checks

On a continuous cycle, the security agent validates file permissions on all credential-bearing files, scans for unexpected open ports, checks service logs for credential leaks and prompt injection attempts, verifies configuration integrity, detects configuration drift via SHA-256 baselines, and monitors log health across all agents.

Weekly Deep Security Audit

Every week, a comprehensive audit adds additional checks: data access control verification across all firm boundaries, cross-agent permission validation, dependency vulnerability scanning, and threat intelligence monitoring. A full report with risk scoring is produced automatically.

Monthly Compliance Posture Report

Monthly automated assessment of the platform's compliance posture: SEC recordkeeping adherence, audit trail integrity, data isolation verification, access control review, and documentation completeness. Designed for CCO review and regulatory examination readiness.

Autonomous Self-Healing

When the security agent finds an issue it cannot auto-fix with standard rules, it escalates to a multi-stage diagnostics pipeline: gather context, diagnose with a reasoning model, execute the fix with safety guardrails, and verify the result. Dangerous commands are blocked by an explicit allowlist and blocklist system.

Credential and Config Drift Monitoring

All configuration files and credential stores are baselined with SHA-256 hashes. Any unauthorized modification triggers an immediate alert and investigation. Configuration drift is detected before it becomes a security exposure. Critical for maintaining the security posture financial regulators expect.

Log Integrity Monitoring

Integrity is verified at two levels: per-entry SHA-256 hash chains detect modification, insertion, or deletion of individual log entries, each day's record is closed with a cryptographically signed root hash, while file-level hashes detect bulk truncation or replacement. Stale services trigger alerts. Tamper-evident logging is essential for SEC recordkeeping compliance.

7

Audit Trail and Compliance Documentation

LEFT

Title. SEC-Ready Audit Trail

Comprehensive Action Logs

Every agent action records: which agent, which advisor data accessed, what action taken, what output produced, timestamp, compliance status, and whether a human approved it. Designed for SEC recordkeeping requirements.

Cryptographic Hash Chain

Every log entry carries a SHA-256 hash of its contents and a reference to the previous entry's hash, forming a tamper-evident chain. Any modification, insertion, or deletion of entries breaks the chain. A single verification command confirms integrity back to the first entry. Essential for regulatory examination defense. Each day's chain then closes with a root hash signed with the appliance's Ed25519 audit key, so a rewrite of an entire day's record is detectable as well.

Write Safety

Log writes protected by file locking with retry timeout. Hash chain computation happens under the exclusive lock. No data corruption from concurrent agent operations.

Log Rotation with Chain Continuity

Rotating file handlers prevent unbounded disk growth. Hash chains bridge rotation boundaries automatically. Nightly cleanup deletes temp logs older than 30 days while preserving audit-relevant records.

Version Control

Daily auto-commit captures the full state of all non-secret files to a private repository. Roll back to any previous day's state. Combined with the hash-chained audit logs, this provides complete historical documentation for regulatory inquiries.

RIGHT

Title. Code Quality as Security

Zero bare except clauses

in production code. Every exception handler catches specific types. No silent failures that could mask compliance issues.

Explicit timeouts

on all external API calls. No call hangs indefinitely. Distribution operations are never blocked by an unresponsive external service.

Parameterized queries

throughout all database operations. Formal audit identified and fixed all injection patterns. Advisor data is never exposed through query manipulation.

Atomic file writes

using the tmp-plus-rename pattern. No corruption from interrupted writes. Audit logs and firm data are never partially written.

Formal code audits

with documented, repeatable process. All critical issues resolved. Audit documentation maintained for due diligence review.

8

Threat Mitigation for Financial Services

Not theoretical risks. These are documented attack vectors relevant to AI systems handling firm data and advisor communications.

Prompt Injection

Risk. Inbound emails, document uploads, and ingested content could contain adversarial instructions designed to manipulate AI behavior: causing an agent to leak advisor data, fabricate data, or bypass compliance gates.

Mitigation. A dedicated multi-category sanitization engine strips adversarial content from all untrusted input before it enters any AI model. It detects instruction overrides, role markers, boundary injection, jailbreak patterns, role spoofing, and unicode control character attacks. Deployed across every agent that processes external content. The compliance gate and human approval provide additional defense layers.

Data Exfiltration

Risk. AI agents with access to advisor data could be manipulated into including sensitive information in outputs: advisor contact details in public content, fund terms in marketing materials, or performance data in non-compliant contexts.

Mitigation. Per-firm data isolation at the database level prevents cross-firm exposure. The identity protection scanner checks all outbound content for sensitive information. The Compliance Review agent validates that no non-public fund information appears in public-facing content. No agent can access data outside its defined firm boundary.

Excessive Agency

Risk. The OWASP Agentic AI guidelines identify this as a primary failure mode: an AI system taking actions beyond its intended scope. For asset managers, this could mean unauthorized advisor communications or unapproved performance disclosures.

Mitigation. The trust escalation model, agent permission matrix, compliance gate, and human approval framework are direct mitigations. No agent can expand its own permissions. Every new capability follows a defined deployment process with compliance review. Advisor-facing approval gates are permanent and cannot be disabled.

9

Inbound Defense

Every piece of external content passes through multiple detection layers before it reaches any AI model or distribution operations workflow.

Prompt Injection Sanitization

A dedicated multi-category sanitization engine processes all untrusted input before it enters any AI model. It detects and strips instruction overrides, role markers, boundary injection, jailbreak patterns, role spoofing, and unicode control character attacks. The sanitizer never raises exceptions: it logs warnings and returns cleaned text. Deployed across every agent that processes external content.

Phishing Detection

Inbound communications are scanned for brand spoofing, display name impersonation, leet-speak evasion, and homoglyph typosquatting across financial institution domain mappings. Detected phishing attempts are flagged and quarantined before any agent processes them. Critical for protecting distribution operations from social engineering.

Document Screening

Uploaded documents and data feeds are screened for embedded adversarial content before entering agent workflows. Malicious payloads hidden in advisor documents, research reports, or data files are neutralized before they can influence any AI model.

10

Credential Management

Credentials are treated as high-value targets at every layer. Financial services data requires financial services security.

File Permissions

All credential-bearing files set to owner read/write only. No credential file is world-readable. All service configurations locked to restrictive permissions.

No Hardcoded Keys

All API keys externalized to configuration files. Source code loads secrets at runtime via dedicated loader functions. Zero inline API keys in production files, validated through formal security audit.

Git Prevention

Comprehensive exclusion rules block all auth files, credential configs, and backup files from version control. History scrubbed to remove any previously committed secrets from every historical commit.

Centralized Config

A single configuration file with restrictive permissions serves as the source of truth for primary credentials. Service-specific configs stored separately, each excluded from version control and permission-locked. Compromise of one service does not expose others.

Credential Audit

Formal audit documents every credential: where stored, what scopes it has, minimum required scopes, and remediation needed. All actionable items resolved including over-scoped permissions. Credential drift monitoring detects unauthorized changes to credential files in real time.

11

Regulatory Data Ingestion

Public filings, ingested on-appliance, with no advisor PII in the pipeline.

SEC EDGAR 13F Ingest

Daily pull from the SEC EDGAR FULL-INDEX. 13F-HR filings are parsed, holdings extracted, and mapped to your fund lineup via OpenFIGI public API for CUSIP resolution. All processing happens on-appliance. No advisor personally identifiable information flows in this pipeline; the data is institutional holdings reported by registered investment advisors.

IAPD Form ADV Amendment Watch

Daily scrape of the IAPD public dataset. 411 fields tracked per amendment: ownership structure, AUM, custodian relationships, disciplinary history, personnel changes, and more. Amendments are diffed against the prior filing and scored for materiality. Results feed the FIT Score and surface in the morning brief and the work queue when a filing affects an advisor in your coverage universe.

OpenFIGI CUSIP Resolution

CUSIPs from 13F filings are resolved through the OpenFIGI public API to map institutional holdings to your fund lineup. No proprietary data leaves the appliance in this process; only public CUSIP identifiers are sent to the API.