BONAFIDE™ SPECIFICATION V1.0 — PART 9

Network Security & Abuse Prevention

Threat model, transport security with mTLS, client attestation, DoS defense, validator network protection, relay abuse prevention, traffic analysis resistance, and incident response

1. Purpose

This part defines the network security architecture: how the Bonafide network defends against attackers ranging from script kiddies to nation-states, how transport is secured, how denial-of-service is mitigated, how the validator network is protected, how traffic analysis is resisted, and how incidents are detected and resolved.

2. Threat Model

2.1 Attacker Classes

Class	Capability	Goal	Defense Layer
Opportunistic	Automated scanning, known exploits, credential stuffing	Mass compromise for profit	Transport security, rate limiting, device attestation
Targeted	Custom tooling, social engineering, persistent access	Specific user's data	Per-quantum encryption, blind validation, key ephemerality
Insider	Privileged access to institutional infrastructure	Data exfiltration from a specific institution	API abstraction, enclave isolation, forensic watermarking
Network	Traffic interception, man-in-the-middle, DNS manipulation	Session hijack, metadata collection	mTLS, certificate pinning, traffic analysis resistance
Nation-state	Unlimited resources, legal coercion, infrastructure compromise	Bulk surveillance, targeted access	Content neutrality, stateless derivation, compliance profiles
Coercive	Physical access to user, legal or illegal compulsion	Force user to reveal vault contents	Duress personas, plausible deniability, silent alerts

2.2 Defense-in-Depth

No single layer defeats all attackers. Bonafide's network security is layered: transport security prevents interception, client authentication prevents unauthorized access, session integrity prevents replay and hijacking, DoS defense prevents resource exhaustion, validator protection prevents consensus manipulation, and traffic analysis resistance prevents metadata leakage. Each layer operates independently—failure of one does not compromise the others.

3. Transport Security

3.1 Mutual TLS Everywhere

All Bonafide network traffic uses mutual TLS (mTLS). Both client and server authenticate with certificates. There are no anonymous connections. Every participant—user device, institution, validator, relay—presents a certificate issued by the Bonafide certificate authority (ca.bonafide.id) or by a certified subordinate CA.

3.2 TLS Configuration

Parameter	Requirement
Minimum version	TLS 1.3
Cipher suites	AEAD only: AES-256-GCM with SHA-384, CHACHA20-POLY1305 with SHA-256
Key exchange	X25519 or X448 (ECDHE only, no static key exchange)
Certificate type	X.509v3 with Ed25519 or ECDSA P-384 signatures
Certificate pinning	Required for all long-lived connections (device-to-cloud, institution-to-cloud)
OCSP stapling	Required for all server certificates
Session resumption	TLS 1.3 PSK mode with 0-RTT disabled (prevents replay)

3.3 Certificate Hierarchy

Participant	Certificate Issuer	Lifetime	Revocation
User device	ca.bonafide.id (device enrollment)	1 year, auto-renewed	Immediate on device deauthorization
Institution	ca.bonafide.id (certification)	1 year, re-certification required	Immediate on certification revocation
Validator	ca.bonafide.id (validator certification)	1 year, re-certification required	Immediate on score drop below threshold
Relay operator	ca.bonafide.id (relay certification)	1 year, re-certification required	Immediate on certification revocation
Bonafide Cloud services	ca.bonafide.id (infrastructure)	90 days, auto-renewed	Standard CRL/OCSP

3.4 Hardware-Accelerated TLS

On ExaForge FPGA deployments and SmartNIC/DPU deployments, TLS termination occurs in hardware. Unauthorized connections are rejected at wire speed before reaching the host CPU.

This provides both performance (the host CPU handles only authenticated traffic) and security (the attack surface for DoS is the hardware TLS engine, which is purpose-built for rejection throughput).

4. Client Authentication

4.1 Device Attestation

Every client connection requires device attestation: cryptographic proof that the connecting device has a secure element of the claimed enclave tier, the secure element has not been tampered with, and the device's software environment is in a known-good state.

Attestation methods by enclave tier:

Tier	Attestation Method	Verification
Tier S (FPGA)	FPGA bitstream attestation + PUF challenge-response	Verified by validator network
Tier 1 (Dedicated SE)	Apple SEP / Google Titan M2 / Samsung Knox attestation	Verified against manufacturer root of trust
Tier 2 (TEE)	SGX/TDX/SEV-SNP/TrustZone remote attestation	Verified against platform attestation service
Tier 3 (TPM)	TPM 2.0 remote attestation (EK + AIK)	Verified against TPM manufacturer CA
Tier 4 (Software)	Software environment report (limited trust)	Verified by cloud service, restricted capabilities

4.2 Guest Enclave Restrictions

Devices without a recognized secure element can connect through the guest enclave with restricted capabilities: limited to Level 0–1 data access, limited session duration (24 hours), no key caching, and no institutional peering. Guest enclave access is the onramp for new users and the fallback for unrecognized devices.

5. DoS Defense

5.1 Tiered Capacity Architecture

Bonafide isolates traffic into five capacity tiers with independent resource pools. Exhausting one tier does not affect the others:

Tier	Traffic Type	Protection	Priority
Tier 1	Authenticated vault operations from enrolled devices	Full resources, no throttle unless abuse detected	Highest
Tier 2	Institutional runtime operations with valid certificates	Dedicated pool, rate-limited per institution	High
Tier 3	Validator consensus traffic	Isolated network, priority routing	High
Tier 4	Third-party access (share links, verifications)	Rate-limited per authorization, resource-capped	Medium
Tier 5	Unauthenticated (guest enclave, discovery, public endpoints)	Proof-of-work required, heavily rate-limited	Lowest

5.2 Connection-Level Defenses

- **SYN cookies:** TCP SYN flood protection without state allocation for incomplete handshakes.
- **TLS handshake timeout:** Incomplete TLS handshakes are terminated after 5 seconds, preventing slowloris-style attacks.
- **Certificate validation before resource allocation:** No application-layer resources are allocated until the client presents a valid certificate. Invalid or missing certificates result in immediate connection termination.
- **Anycast distribution:** API endpoints are anycast across multiple regions, distributing traffic geographically and providing automatic failover.

5.3 Application-Level Defenses

- **Per-identity rate limiting:** Each authenticated identity has rate limits appropriate to their participation model (user, institution, validator).
- **Proof-of-work for unauthenticated requests:** Unauthenticated requests must include a proof-of-work solution. The difficulty adjusts dynamically based on current load.
- **Backpressure propagation:** When a downstream service approaches capacity, it signals upstream services to reduce request rates. The system degrades gracefully rather than failing catastrophically.

- **Request prioritization:** During high load, vault operations (Tier 1) take precedence over share link access (Tier 4) and unauthenticated traffic (Tier 5).

6. Validator Network Protection

6.1 Cost Tokens

Every validation request includes a cost token: a proof-of-work or signed authorization that prevents garbage flooding. Submitting validation requests costs computation or reputation. An attacker who floods the validator network with invalid requests burns resources without affecting legitimate traffic.

6.2 Selection Diversity

Validator selection enforces organizational, geographic, and network diversity (Part 5, Section 5.2). An attacker who controls validators in one organization, region, or network cannot dominate any committee.

6.3 Consensus Communication Security

- Consensus messages are signed with per-round ephemeral keys (forward secrecy).
- Inter-validator communication uses mTLS through the federation layer.
- Validators do not communicate directly peer-to-peer, preventing network topology mapping.
- Consensus results are signed by all agreeing validators (non-repudiation).

6.4 Sybil Resistance Summary

Certification cost, probationary period, organizational diversity limits, diminishing returns for same-organization validators, and mentorship risk create a defense-in-depth against Sybil attacks. The full analysis is in Part 5, Section 7.

7. Relay Abuse Prevention

7.1 Per-Proxy Rate Limiting

Each proxy address has independent rate limits for inbound and outbound messages. Limits are configurable per proxy and per relay operator. Default limits prevent a single proxy from being used for spam relay while allowing normal communication volumes.

7.2 Sender Reputation

The relay network tracks sender reputation across proxy addresses. Known-bad senders (spam sources, abuse sources) are blocked at the relay before reaching the user’s proxy. Reputation data is shared across relay operators through the federation layer.

7.3 Enumeration Defense

An attacker attempting to discover valid proxy addresses receives uniform responses for all addresses—valid and invalid. Messages to non-existent proxies are accepted and silently discarded. Timing is uniform (tarpitting). There is no observable difference between sending to a real proxy and sending to a non-existent one.

7.4 User-Configurable Filtering

Users can configure per-proxy filtering policies: whitelist-only (accept messages only from known senders), domain filtering (accept only from specific domains), content-type filtering (reject attachments), and time-based policies (proxy active only during business hours). Filtering is applied by the relay operator before forwarding.

8. Traffic Analysis Resistance

8.1 Threat

Even with encrypted transport, an observer can learn information from traffic patterns: message timing, sizes, frequency, and source/destination. A nation-state adversary monitoring network traffic could correlate vault operations with real-world events, identify communication patterns, or detect when a user accesses specific branches.

8.2 Countermeasures

- **Constant-rate padding:** Vault holder devices maintain a constant rate of network traffic to the Bonafide network, regardless of actual operation volume. Real operations are interleaved with dummy operations that are indistinguishable to an observer.
- **Uniform request/response sizing:** Requests and responses are padded to power-of-two size classes (256B, 512B, 1KB, 2KB, ...). An observer cannot distinguish a small query from a large document transfer if both fall in the same size class.
- **Timing jitter:** Random delays are added to request/response timing to prevent timing correlation between operations and external events.
- **Optional ORAM:** For high-security deployments, the specification supports Oblivious RAM patterns for vault access. ORAM ensures that the access pattern to stored quanta reveals nothing about which quanta are being accessed. ORAM has significant performance overhead and is recommended only for Classification S deployments with Level 7+ data.

8.3 Traffic Analysis and Privacy Classification

Traffic analysis resistance is a component of the privacy classification assessment. Classification S deployments are expected to implement constant-rate padding and ORAM for high-security channels. Classification A deployments implement padding and uniform sizing. Classification B and C deployments implement uniform sizing at minimum.

9. Incident Response

9.1 Trust Score as Immune System

The trust scoring system (Part 5, Section 4) acts as an automated immune system. Misbehaving participants—validators that produce incorrect results, institutions with attestation failures, relay operators with content inspection—see their trust scores degrade automatically. Scores below thresholds trigger automatic capability restrictions. The system self-heals without requiring manual intervention for most incidents.

9.2 Certificate Revocation

When a participant's certificate must be revoked (compromise detected, certification revoked, trust score below threshold):

- The certificate is added to the CRL (Certificate Revocation List) published by ca.bonafide.id.
- OCSP responders are updated within 60 seconds.
- All active sessions using the revoked certificate are terminated.
- The revocation event is recorded in the ledger.

9.3 Incident Types and Response

Incident	Detection	Automated Response	Manual Response
Validator compromise	Consensus disagreement, attestation failure	Score reduction, suspension at threshold, committee exclusion	Investigation, re-certification, or permanent ban
Institutional key compromise	Attestation failure, anomalous access patterns	Certificate revocation, runtime lockdown, user notification	Forensic investigation, re-keying, breach notification
Relay content inspection	Canary detection in relay layer, audit finding	Certification revocation, relay defederation	Investigation, potential legal action
Device compromise	Attestation failure, anomalous location/behavior	Device deauthorization, session termination, user alert	Device re-enrollment or replacement
DDoS attack	Traffic volume anomaly, backpressure triggers	Proof-of-work difficulty increase, tier isolation, anycast absorption	Traffic analysis, upstream filtering coordination
Data breach at institution	Canary detection, external report, anomalous access	Privacy score impact, user notification, enhanced monitoring	Forensic investigation, regulatory reporting, remediation

9.4 Forensic Audit Trail

Every security-relevant event is recorded in the immutable ledger: certificate issuance and revocation, trust score changes, attestation results, consensus outcomes, revocation events, and incident responses. The ledger provides a permanent, tamper-proof record for forensic investigation and regulatory compliance.

Bonafide™ — Privacy by architecture, not by promise.

An open specification by Sly Technologies Inc. | bonafide.id | bonafideid.org
V1.0 — February 2026