NornGate Trust Topology & DigitalOcean Infrastructure

SBS-AI-Security-Report

1. Document Control

Field	Value
Title	SBS AI Security Report: NornGate Trust Topology & DigitalOcean Infrastructure
Date	August 21, 2026
Version	21.0
Scope	`norngate-core` application, NornGate Light Paper topology, DigitalOcean Infrastructure, Legacy Data Ingestion, Business Unit Compartmentalization, and 3rd-Party Audit Manifest
Audience	SBS Executive Leadership, Board of Directors, Security/Compliance Teams, and Authorized 3rd-Party Auditors
Classification	Confidential — restricted distribution

Purpose: To establish NornGate's and SBS's underlying infrastructure security posture as a verifiable contract — what controls exist, what is missing, and what work is committed — ensuring every claim is understood by executive stakeholders and technically verifiable by external auditors.

Threat Model Summary: (a) External unauthenticated attacker against public endpoints; (b) Prompt-injected AI agent or hostile content processed by agents running on the host; (c) Malicious/compromised third-party npm plugin or adapter (supply chain); (d) Cross-tenant actor; (e) Malicious or compromised board operator (insider); (f) Attacker with host/LAN/backup access; (g) Lateral movement between Dev and Prod environments; (h) Lateral movement between internal business units (Compartments); (i) Malicious data injected via legacy system ingestion (Data Poisoning).

2. List of Acronyms

Acronym	Expansion
AI	Artificial Intelligence
API	Application Programming Interface
AuthN/Z	Authentication and Authorization
BCDR	Business Continuity and Disaster Recovery
CCPA	California Consumer Privacy Act
CORS	Cross-Origin Resource Sharing
CSP	Content Security Policy
CSRF	Cross-Site Request Forgery
CVE	Common Vulnerabilities and Exposures
DB	Database
DLQ	Dead-Letter Queue
DLP	Data Loss Prevention
DoS	Denial of Service
DoW	Denial of Wallet
DR	Disaster Recovery
GDPR	General Data Protection Regulation (EU)
HITL	Human-in-the-Loop
HSTS	HTTP Strict Transport Security
IaC	Infrastructure as Code
JWT	JSON Web Token
LHN	Ledger Hub Networks, LLC
LGPD	Lei Geral de Proteção de Dados (Brazil)
LLM	Large Language Model
MCP	Model Context Protocol
npm	Node Package Manager
OCR	Optical Character Recognition
OWASP	Open Web Application Security Project
PII	Personally Identifiable Information

Acronym	Expansion
RBAC	Role-Based Access Control
RCE	Remote Code Execution
RLS	Row Level Security
RPO	Recovery Point Objective
RTO	Recovery Time Objective
SBS	Spanish Broadcasting System, Inc.
SCA	Software Composition Analysis
SQL	Structured Query Language
SQLi	SQL Injection
SSRF	Server-Side Request Forgery
SOW	Scope of Work
TLS	Transport Layer Security
VPC	Virtual Private Cloud
XSS	Cross-Site Scripting

3. Executive Summary

NornGate is a multi-gated AI orchestrator with strong, well-implemented design invariants. The architecture is defined by the **5-Gate Taxonomy** (G0-G4) and the **Realm Trust Topology** (Asgard to Jotunheim). Core application strengths include:

- **Credential Starvation:** AI agents never hold raw credentials. G1 (Tyr) resolves the minimum-scope credential for an approved action, injects it at execution, and revokes it on completion. A prompt-injected agent has nothing durable to exfiltrate.
- **Origin Isolation:** Untrusted code and agent-rendered content execute in Jotunheim and are served from a separate registrable domain, preventing cookie theft against the control plane in Asgard.
- **Deterministic Gates:** Every consequential action transits an explicit, deterministic gate. *Deterministic* means decided by policy code, never LLM judgment. *Fail-closed* means an unreachable or erroring gate results in denial.
- **Telemetry Substrate (The Norns):** Every gate decision writes to **Urd** (audit log/past) before execution, monitored by **Verdandi** (real-time metrics/present), with **Skuld** (forecasting/future) ensuring system health.

Build Phase Context & Audit Commitment: It is critical to note that The Dash platform is currently in its **foundational deployment phase**. The gaps identified in this report—particularly

those related to DigitalOcean infrastructure and host access—are "Day 0" findings discovered during the active build process. These gaps are scheduled to be fully remediated before the platform processes live production data.

Upon completion of the foundational installation and gap remediation, Ledger Hub Networks will conduct rigorous internal and external security audits. Furthermore, SBS retains the option to engage an independent third-party security firm (such as Certek) to perform a comprehensive penetration test and compliance validation prior to production go-live.

Infrastructure Context (DigitalOcean As-Built): While NornGate's application-layer architecture is robust, the underlying DigitalOcean infrastructure is currently in a "Day 0" remediation state. Critical environmental gaps were identified during infrastructure auditing, including: plaintext credentials committed to operational changelogs, shared flat VPC networking between Dev and Prod, absent cloud/host firewalls on production droplets, unmitigated egress paths, and shared root SSH access. All infrastructure findings are tracked in the Build Register with active remediation workstreams.

---

4. The NornGate Trust Topology (Realms & Gates)

Subdomain = Realm. The marketing metaphor and the routing topology are the same diagram. Trust gradient runs from Asgard (highest privilege) to Jotunheim (zero standing privilege).

4.1 The 5-Gate Execution Model

A **gate** is a deterministic, fail-closed checkpoint. No execution path produces an external side effect without transiting these gates:

Gate	Name / Owner	Enforces	Placement
G0	Ingress — The Gate / Warden	AuthN/Z, tenant resolution, rate limits, schema validation	System edge; nothing enters unauthenticated
G1	Policy — Tyr	Deterministic pre-execution policy: tool permissions, credential scope, RBAC	Before every tool call, every realm crossing
G2	Approval — HITL	Human approval for side effects above a configured risk class	Before external side effects (writes, spend, sends)

Gate	Name / Owner	Enforces	Placement
G3	Sandbox — Jotunheim Boundary	Egress control, no standing credentials, origin isolation	Around all untrusted code and tool calls
G4	Quality — Forseti	Consensus/judging on contradictory agent outputs; structured-output validation	Before results leave the system

4.2 Realms → Trust Zones

- **Midgard** (`norngate.com`): User-facing tier. Humans live here; sessions terminate here.
- **Asgard** (`asgard.`): Control plane / governance. Policy, scheduling, admin. Zero-trust access only.
- **Vanaheim** (`vanir.`): Federated external integrations. Third-party APIs with negotiated trust.
- **Jotunheim** (`jotun.`): Sandboxed / untrusted execution. Unvetted code + untrusted agent tool-calls. Outside Asgard control by design.
- **Valhalla** (`valhalla.`): Success archive. Immutable audit trail of completed jobs.
- **Helheim** (`hel.`): Dead-letter queue (DLQ). Jobs that did not "die gloriously."

4.3 Failure Semantics & Business Continuity (BCDR)

- **Two Afterlives:** Valhalla (immutable success archive) and Helheim (DLQ). A job is in exactly one afterlife, keeping the audit append-only and completion metrics honest.
- **Hermod (DLQ Reprocessor):** Classifies before retrying. Transient failures get capped exponential backoff. G1/G2 denials surface to a human queue — retrying a denial is an escalation attempt, not resilience.
- **Ragnarök (DR Drill):** The scheduled DR/chaos drill. The myth ends in rebirth — exit criterion is *verified recovery*: state restored from checkpoints, gates re-armed fail-closed, Norn telemetry confirming steady state.
- **BCDR Targets:** The infrastructure is designed for an **RPO (Recovery Point Objective) of < 24 hours** via nightly encrypted database backups to SFO3, and an **RTO (Recovery Time Objective) of < 4 hours** for critical control plane restoration via infrastructure-as-code (IaC) redeployment.

---

5. Operational Context — Closed System

The organization declares NornGate operates as a **Closed System**:

- **System Closed:** NornGate is not exposed directly to the internet; all access is mediated through gateway-type connections.
- **Data Non-Exposed:** Data managed by the system is never exposed outside the operational perimeter.
- **Risk Modulation:** This deployment model reduces the probability of exposure for most breaches. Realistic exploitation scenarios concentrate on insider threats, supply chain (npm packages), and lateral movement, rather than direct external attack.

Technical Honesty Note: The closed system context is a declared risk factor, not a code-implemented control. It modulates probability but does not alter the intrinsic severity of the documented gaps. Critical and High findings remain in active remediation as part of the foundation build-out.

---

6. Layered Security Architecture

"*[committed]*" = *exists in* `doc/plans` *or spec;* "*[recommendation]*" = *proposal from this report.*

Layer	Name	Status	Dominant Risk
0	Host & Deployment (DigitalOcean)	◫ Gap (Day 0)	Shared root SSH, plaintext creds in changelog (SBS-33, SBS-34)
1	Network & Transport (VPC & Firewalls)	◫ Gap (Day 0)	Shared flat VPC, missing firewalls, open egress (SBS-35, SBS-36, SBS-39)
2	Auth & Authz (G0/G1)	◫ Partial	No rate limiting; shared dev secret (SBS-05, SBS-07)
3	Multi-Tenant & Compartment Isolation	◫ Partial	No RLS: second line of defense absent (SBS-11)
4	Execution & Containment (G3/G4)	◫ Partial	Untrusted code without real sandbox; open host egress; hallucination risk (SBS-01, SBS-02, SBS-03)

Layer	Name	Status	Dominant Risk
5	Secrets, Data & PII (G1)	◫ Partial	Leakage to `server.log`; key co-located with backups (SBS-04, SBS-10)
6	Audit & Observability (The Norns)	◫ Partial	Falsifiable log, no integrity/retention (SBS-15)

Layer 0 — Host & Deployment (DigitalOcean) — ◫ Gap (Day 0)

- **Present:** Dedicated SBS-owned DigitalOcean account; phishing-resistant MFA enforced at account root; Docker 29.1.3 / Compose 2.40.3 deployed; Prod application server isolated via jump-host/bastion pattern through monitor-servers.
- **Gaps:** Root SSH login enabled fleet-wide with a single shared SSH keypair (SBS-33); shared `sbsRoot` account with `NOPASSWD:ALL` and docker group equivalence (SBS-33); plaintext credentials (shared root, GitHub, Wiki, Grafana) discovered in `SERVER_CHANGELOG.md` (SBS-34); production OS standardization pending (mix of Ubuntu 24.04 and 26.04).
- **Roadmap:** [committed] Rotate all compromised credentials and migrate to Vault; strip credentials from changelog history; create per-person named SSH users with scoped sudo; disable direct root login.

Layer 1 — Network & Transport (VPC & Firewalls) — ◫ Gap (Day 0)

- **Present:** DigitalOcean VPCs utilized for private networking; Prometheus/Grafana internal services bind to `127.0.0.1` behind nginx TLS; Wiki TLS live via Let's Encrypt.
- **Gaps:** Dev and Prod environments share a flat VPC (`default-nyc2`), breaking the Dev↔Prod isolation rule (SBS-35); **No Cloud Firewalls or host firewalls (ufw/nftables) exist on Prod or Dev droplets**—public interfaces are protected solely by service bind addresses (SBS-36); Wiki droplet is in NYC1 (cross-region), forcing unencrypted metric scraping over the public internet (SBS-37); Production apex DNS (`sbsdash.com`) currently resolves to an undocumented Dev-Gitlab-server, not Prod (SBS-38).
- **Roadmap:** [committed] Rebuild fleet into dedicated `prod-vpc-nyc2` and `dev-vpc-nyc2`; deploy 4 tag-scoped Cloud Firewalls immediately; repoint apex DNS to `sbsdash-server-prod`; rebuild Wiki in NYC2.

Layer 2 — Auth & Authz (G0 / G1) — ◫ Partial

- **Present:** Better Auth with required secret; `boardMutationGuard` anti-CSRF; API keys 192 bits hashed (SHA-256); JWT HS256 with strict `alg`/`exp` checks; timing-safe comparisons.
- **Gaps:** No rate limiting on auth/keys/invites (SBS-05); sign-up open by default in private (SBS-16); shared dev secret signs JWTs (SBS-07); agent keys lack expiration (SBS-19).
- **Roadmap:** [committed] Tier 2/3 agent auth; [recommendation] Global rate limiting; independent JWT secret with entropy validation.

Layer 3 — Multi-Tenant & Compartment Isolation — ◫ Partial

- **Present:** `company_id` and `compartment_id` NOT NULL with `assertCompanyAccess` / `assertCompartmentAccess` enforced at application layer; cross-tenant boundary tests; object storage prefixed with `${companyId}/${compartmentId}/` anti-traversal. SBS internal business units (HR, Finance, Sales) are isolated into distinct Compartments.
- **Gaps:** Isolation *only* at application layer — no RLS (Row Level Security) (SBS-11). A single missed filter or SQLi exposes the whole instance.
- **Roadmap:** [recommendation] Postgres RLS or scoped DB roles as a second line of defense. (See Appendix D for Business Unit mapping).

Layer 4 — Execution & Containment (G3 Jotunheim Boundary) — ◫ Partial

- **Present:** `low_trust_review` preset with fail-closed resolution; runtime containment requiring `sandbox` driver + isolated workspace; quarantine of output with source-trust tagging; sandbox callback bridge with allowlist.
- **Gaps:** Local CLI agents run **unsandboxed** on host inheriting server env, using `--dangerously-skip-permissions` (SBS-02); plugin workers lack real OS isolation (SBS-03); external adapters load in-process via `import()` (SBS-01); built-in sandbox provider `fake` is a no-op but accepted by containment check (SBS-25); **No egress control exists on any droplet**, allowing potential LLM01 exfiltration (SBS-39).
- **Hallucination & Liability Containment:** All outward-facing AI generation (Sol Vega scripts, automated client communications) must transit G4 (Forseti - Output Arbitration). G4 enforces structured-output validation and fact-checking against the 40-year model baseline. Any AI output lacking sufficient grounding confidence is routed to the Helheim DLQ for human review, preventing brand-damaging hallucinations from executing.

- **Roadmap:** [committed] Centralized low-trust enforcement; `adapter_managed`/ `cloud_sandbox` modes; deploy nftables default-deny outbound with strict allowlists. [recommendation] Block `fake` provider for low-trust; run workers in microVMs (gVisor/Firecracker).

Layer 5 — Secrets, Data, & PII (G1 Credential Starvation) — □□ Partial

- **Present:** AES-256-GCM encryption at rest; master key 0600 with health check; AWS provider with SigV4; **no reveal endpoint** (100% board-only management); just-in-time secret injection as env vars before spawn; redaction in logs/transcripts.
- **PII & Privacy Controls:** Listener data ingested via the LaMusica app and geofence engine is subject to strict data minimization. NornGate enforces CCPA/LGPD/GDPR compliance by design through anonymized cohort targeting (no targeting an audience of one) and just-in-time data resolution, ensuring personal identifiers are never persistently stored alongside behavioral profiles.
- **Gaps:** `reqBody` with secrets leaks to `server.log` on HTTP ≥400 errors (SBS-04); master key co-located with DB/backups (SBS-10); plugin configs stored in clear while secret-refs disabled (SBS-22); backups unencrypted (SBS-23).
- **Roadmap:** [committed] Company-scoped plugin secret-refs; [recommendation] Structural logger redaction; master key rotation; encrypted backups.

Layer 6 — Audit & Observability (The Norns) — □□ Partial

- **Present:** Activity log company-scoped with ~239 instrumentation points and structural sanitization; transactional approvals; budget hard-stop; run logs confined with sha256; telemetry opt-out with hashed private refs.
- **Gaps:** Activity log falsifiable by board (manual POST, no hash-chain) (SBS-15); no retention/rotation for `activity_log` or `server.log` (SBS-18); unaudited read access to run logs.
- **Roadmap:** [committed] "Immutable audit" for escalation routes; pluggable `RunLogStore`. [recommendation] Hash-chain/sign log entries + end-to-end request-ID; restrict manual activity POST; configurable retention.

---

7. Prioritized Findings Matrix

Remediation States: □□ Active Remediation · □□ Planned · □□ Under Evaluation. *All Critical and High findings have active remediation work in progress to be resolved before production go-live.*

ID	Sev.	State	Domain	Description	Recommendation
SBS-01	Critical	□□ Active	Supply Chain	External adapters: `import()` in-process + `npm install` without `--ignore-scripts` → RCE	`--ignore-scripts`, out-of-process worker, integrity pinning
SBS-02	Critical	□□ Active	Execution	Local CLI agents unsandboxed on host, inherit env, use `--dangerously-skip-permissions`	Dedicated user/seccomp/sandbox; env denylist; restrict `--allowedTools`
SBS-03	Critical	□□ Active	Plugins	Plugin workers lack real sandbox (fs/net open); VM sandbox dead code; `http.outbound` doesn't limit process net	OS isolation/microVM; document as "trusted code" if not sandboxed
SBS-04	High	□□ Active	Secrets/Logging	`reqBody` (with secrets) logged to `server.log` on ≥400 errors; only `authorization` redacted	Structural redaction of `reqBody`/`reqQuery` in logger
SBS-05	High	□□ Active	Auth/API	No rate limiting on login, invites, claim, webhooks (SPEC §16 requires it)	Global + endpoint-specific rate limiting
SBS-06	High	□□ Active	Deploy/DB	Fixed PG credentials `norngate/norngate` + port 5432 published to host	Random credentials per install; do not publish 5432
SBS-07	High	□□ Active	Secrets/Auth	Dev secret `norngate-dev-secret` signs sessions and JWTs (shared fallback)	Mandatory independent JWT secret; entropy validation

ID	Sev.	State	Domain	Description	Recommendation
SBS-08	High	◫ Active	Webhooks	Plugin webhooks public, no host-level signature verification or rate limit	Optional HMAC at host + ingestion throttling
SBS-09	High	◫ Active	SSRF	Native HTTP adapter: `fetch` arbitrary URLs without private IP filtering; cloud metadata reachable	Reuse plugin SSRF guard in HTTP adapter
SBS-10	High	◫ Active	Secrets	Master key co-located with DB and backups	Key outside data volume; encrypted backups; fix guide
SBS-11	High	◫ Active	Multi-tenant	Isolation only at app layer: no RLS, single DB user with total access	Postgres RLS or scoped DB roles
SBS-12	High	◫ Active	Auth	Cloud tenant token provisions `instance_admin` to any user; replayable	Per-request nonce/timestamp; declared roles
SBS-13	Medium	◫ Planned	HTTP	Guards trust `X-Forwarded-Host` without `trust proxy`	Configure `trust proxy` or validate real `Host`
SBS-14	Medium	◫ Planned	HTTP	No CORS or global security headers; SPA lacks CSP/HSTS	helmet + restrictive CSP
SBS-15	Medium	◫ Planned	Audit	Activity log falsifiable by board; no hash-chain or request-ID	Integrity (hash-chain), request-ID, restrict manual POST
SBS-16	Medium	◫ Planned	Auth	Sign-up open by default; browser claim available in `authenticated/private`	`DISABLE_SIGN_UP=true` by default; claim notification

ID	Sev.	State	Domain	Description	Recommendation
SBS-17	Medium	◫ Planned	SSRF	Portability import accepts arbitrary HTTPS hostnames	Host allowlist + private IP blocking
SBS-18	Medium	◫ Planned	Logging	Workspace-operation and plugin logs lack secret redaction	Apply `redactSensitiveText` to both pipelines
SBS-19	Medium	◫ Planned	Creds	Agent API keys never expire; JWT lacks `jti` /revocation	Key expiration/rotation; short TTL + denylist
SBS-20	Medium	◫ Planned	Containers	Main container starts as root; no `cap_drop`; image not reproducible	`USER` in image, hardening, digest pinning
SBS-21	Medium	◫ Planned	Invites	Invite tokens ~41 bits vs 192 bits elsewhere	Increase entropy to ≥128 bits
SBS-22	Medium	◫ Planned	Secrets/Plugins	Sensitive plugin config in clear while secret-refs disabled	Complete company-scope redesign
SBS-23	Medium	◫ Planned	Data	Sessions/OAuth tokens in clear in DB; backups unencrypted	Hash session tokens; encrypt backups
SBS-24	Medium	◫ Planned	Telemetry	Telemetry allows arbitrary dimensions from plugins (exfil channel)	Sanitize/limit plugin dimensions
SBS-25	Medium	◫ Planned	Low-trust	Containment check accepts `fake` (no-op) sandbox provider	Reject `fake` for low-trust runs
SBS-26	Medium	◫ Planned	Low-trust	Possible boundary escape via PATCH `executionPolicy`	Deny `executionPolicy` mutation to low-trust actors
SBS-27	Low	◫ Eval	Realtime	WS accepts `?token=` in URL; credential bypass in `local_trusted`	Header/Sec-WebSocket-Protocol only

ID	Sev.	State	Domain	Description	Recommendation
SBS-28	Low	◫ Eval	CLI/MCP	CLI tokens in plain JSON; MCP static bearer in env	OS keyring; short-lived tokens
SBS-29	Low	◫ Eval	Audit	`x-norngate-run-id` trusted without validation → attribution forgery	Validate runId or ignore for `none` actors
SBS-30	Low	◫ Eval	Disclosure	`openapi.json` public; `/api/health` exposes mode/status	Require auth for openapi; minimize public health
SBS-31	Low	◫ Eval	Plugins	Plugin manifest executed in-process during discovery	Declarative JSON manifest or isolated discovery
SBS-32	Low	◫ Eval	Integrations	openclaw-gateway allows `ws://`; private key in clear config	Require `wss://` outside loopback; move key to secret-ref
SBS-33	Critical	◫ Active	Host Access	Root SSH enabled fleet-wide; single shared SSH keypair for root/sbsRoot; `sbsRoot` has `NOPASSWD:ALL` and docker group	Per-person named users; disable root SSH; scoped sudo; restrict `sbsRoot` to break-glass
SBS-34	Critical	◫ Active	Secrets/Ops	Plaintext credentials (root, GitHub, Wiki, Grafana) committed to `SERVER_CHANGELOG.md`	Rotate all credentials; migrate to Vault; purge git history; enforce distinct per-service creds
SBS-35	High	◫ Active	Network	Dev and Prod share flat VPC (`default-nyc2`), breaking required Dev↔Prod isolation	Rebuild droplets into dedicated `prod-vpc-nyc2` and `dev-vpc-nyc2`

ID	Sev.	State	Domain	Description	Recommendation
SBS-36	Critical	□□ Active	Firewalls	No Cloud Firewalls or host firewalls (ufw/nftables) active on Prod or Dev droplets	Deploy 4 tag-scoped DO Cloud Firewalls; activate host firewalls mirroring admit set
SBS-37	Medium	□□ Active	Network	Wiki droplet in NYC1 cannot join NYC2 VPC; scraped over public internet unencrypted	Rebuild Wiki in NYC2 to rebind to private VPC networking
SBS-38	High	□□ Active	DNS	Prod apex DNS (`sbsdash.com`) points to undocumented Dev-Gitlab-server, not Prod	Repoint A record to `sbsdash-server-prod` (107.170.72.145) upon proxy/TLS readiness
SBS-39	High	□□ Active	Egress	No egress control on any droplet; all outbound traffic open (LLM01 exfiltration risk)	Deploy nftables default-deny outbound with strict allowlists (Anthropic, GitHub, OS mirrors)

---

8. Security Roadmap & Incident Response

Committed in Project (Sources: `doc/plans/`, specs, LH-SBS-INF-001; state □□):

- **Agent Auth Tier 2/3:** JWT signing key rotation, invite expirations, long-lived credential renewal.
- **Low-Trust Enforcement:** Centralized enforcement layer, content-trust tagging, denial regression tests.
- **Plugin Secret-Refs:** Company-scoped end-to-end (post-PAP-2394 redesign).
- **Immutable Audit:** Hash-chained, signed action records for escalation routes (Valhalla archive integrity).
- **Ragnarök DR Drills:** Scheduled chaos drills with verified recovery (Trickster running destruction in Jotunheim under G3).

- **Infrastructure Remediation (LH-SBS-INF-001):** Rebuild fleet into dedicated VPCs; deploy Cloud Firewalls; migrate to per-person SSH users; deploy Vault; deploy nftables egress control.
- **Pre-Launch Security Audits:** Upon completion of the foundational installation and gap remediation, conduct rigorous internal and external security audits. SBS retains the option to engage an independent third-party security firm (such as Certek) to perform a comprehensive penetration test and compliance validation prior to production go-live.

Incident Response (IR) & AI Containment: In the event of a confirmed breach or agent compromise, NornGate supports an instant "Kill Switch" revocation of all agent JWTs and API keys. The IR playbook mandates immediate isolation of the Jotunheim sandbox realm, followed by forensic extraction of the Urd (audit log) to determine blast radius. SBS IT Leadership will be notified within 24 hours of any Critical-severity security event.

Recommended by this Report (Mapped to findings):

- **Immediate:** Global rate limiting (SBS-05); structural logger redaction (SBS-04); `--ignore-scripts` + out-of-process adapter loading (SBS-01).
- **High Priority:** helmet/CSP/HSTS (SBS-14); independent JWT secret (SBS-07); random PG credentials (SBS-06); block `fake` sandbox in low-trust (SBS-25); anti-SSRF in HTTP adapter (SBS-09).
- **Medium Priority:** Audit log hash-chain (SBS-15); Postgres RLS (SBS-11); master key rotation & encrypted backups (SBS-10, SBS-23); CI security scanning.

---

9. Shared Responsibility Model

Security and compliance are jointly shared between Ledger Hub Networks (LHN) and Spanish Broadcasting System (SBS) IT.

Entity	Responsibilities
Ledger Hub Networks (LHN)	NornGate application architecture (G0-G4 gates); IaC code generation and maintenance; AI agent sandboxing and containment logic; audit logging (Urd) generation; cryptographic secret vault architecture.
SBS IT Infrastructure	DigitalOcean account-level security (MFA); execution of IaC firewalls and VPC segmentation; physical backup retention validation; enforcement of per-person SSH access; network-level egress enforcement.

---

10. Appendix A — Audit Scope & Technical Manifest

(This appendix is provided specifically for third-party security auditors and AI red-team firms to define the technical attack surface, rules of engagement, and testing scope.)

10.1 Technology Stack & Attack Surface

The Dash and NornGate platform are composed of the following authorized technologies. Auditors should focus vulnerability scanning and exploitation attempts against these specific frameworks and versions.

- **Cloud Infrastructure:** DigitalOcean (Droplets, VPCs, Cloud Firewalls).
- **Operating Systems:** Ubuntu 24.04 LTS / 26.04 LTS.
- **Containerization:** Docker 29.1.3, Docker Compose 2.40.3.
- **Web Servers / Proxies:** Nginx 1.28.3, Caddy 2.11.
- **Backend Application Stack:** Node.js (Express), Rust (Cargo), Python (scripts/adapters).
- **Frontend Application Stack:** React (NornGate Console), Vue 3 / Vite (SBS Intranet).
- **Databases:** PostgreSQL, MongoDB, Redis.
- **Authentication:** Better Auth (Session management), JWT HS256 (Agent auth).
- **AI/LLM Integration:** Anthropic API (High-end reasoning), Open-Weight Models via OpenRouter / DigitalOcean GPUs (Kimi/Z.ai for routine orchestration).
- **AI Tooling:** Model Context Protocol (MCP) for agent-tool communication.
- **Legacy Data Ingestion:** WideOrbit, Oracle, WideOrbit Automation, MusicMaster, SIMS TV, CounterPoint, ADP, and unknown Digital Systems. Ingestion methods include MCP, direct API, Raw Data Dumps (SFTP/DB exports), and Report OCR Scraping.

10.2 Target Environment & Access Points

The audit scope is limited to the following SBS-owned operational surfaces.

In-Scope URLs:

- `sbsdash.com` / `admin.sbsdash.com` / `clients.sbsdash.com` (Primary Tenant Surfaces)
- `solvega.sbsdash.com` (Public Sol Vega splash site / video delivery)
- `monitor.sbsdash.com` (Grafana / Prometheus observability stack)
- `wiki.sbsdash.com` (BookStack documentation portal)

In-Scope Infrastructure:

- DigitalOcean Droplets residing in the NYC2 region (Prod, Dev, Monitor).
- The NornGate application codebase (White-Box access to be provided via private GitHub repository invitation).

10.3 Rules of Engagement (RoE)

The auditor agrees to the following constraints to protect SBS's production environment and live broadcast capabilities:

1. **Testing Methodology:** White-Box testing is authorized. Source code, IaC configurations, and infrastructure access (via jump host) will be provided.
2. **No Denial of Service (DoS):** Auditors are prohibited from executing volumetric DDoS attacks or aggressive fuzzing that degrades the availability of `solvega.sbsdash.com` or the underlying DigitalOcean droplets. Resource exhaustion tests must be coordinated and throttled.
3. **No Cloud Provider Attacks:** Attempts to compromise the underlying DigitalOcean hypervisor, control plane, or physical infrastructure are strictly out of scope.
4. **Third-Party API Safety:** Auditors must not attempt to breach, overwhelm, or extract data from third-party LLM providers (e.g., Anthropic, OpenRouter). Prompt injection tests must be self-contained and not designed to elicit harmful content from the foundational LLM itself, but rather to test NornGate's G1/G3 containment boundaries.
5. **Data Handling:** Any PII, API keys, or secrets discovered during the audit must be handled under strict confidentiality, not exfiltrated, and securely destroyed at the conclusion of the engagement.

10.4 AI Red-Team Mandate

While traditional web application testing (OWASP Top 10) is required, SBS mandates that a minimum of 40% of the auditor's effort be focused on AI-specific threat vectors (OWASP LLM Top 10). The auditor must specifically attempt to:

1. **G1 Policy Bypass (Tyr):** Attempt to craft prompts that trick Workflow Agents into executing tool calls or data mutations outside their assigned RBAC scope.
2. **G3 Sandbox Escape (Jotunheim):** Attempt to break out of the agent execution environment to access the host server's filesystem, environment variables, or the Asgard control plane.
3. **Credential Starvation Circumvention:** Attempt to force the AI to leak injected just-in-time credentials or API keys to the console, logs, or external endpoints.
4. **Denial of Wallet (DoW):** Attempt to trigger infinite agent loops or excessive token consumption to test the effectiveness of the NornGate `budget hard-stop` auto-pause mechanism.
5. **Data Poisoning (Skuld/Mimir):** Analyze how historical data is ingested. Attempt to inject anomalous or malicious data via legacy API ingestion to see if it degrades the 40-Year Model's targeting intelligence.

6. **Hallucination Exploitation (G4 Forseti):** Attempt to generate outward-facing content (e.g., Sol Vega scripts, automated emails) that bypasses structured-output validation to produce brand-damaging or legally liable statements.

---

11. Appendix B — Data Flow & Trust Boundary Architecture

(This appendix maps the logical flow of data across the NornGate Realms, explicitly identifying trust boundaries and the gates that enforce them. Auditors should use this to plan cross-boundary penetration tests.)

11.1 Trust Boundaries & Data Flow Map

Boundary 1: The Edge (Internet → Midgard)

- **Flow:** User/Client Request (HTTP/HTTPS) → `api.norngate.com` / `sbsdash.com`.
- **Gate Enforced: G0 (The Gate / Warden)** — AuthN/Z, tenant resolution, rate limits, schema validation.
- **Data State:** Untrusted input. No access to internal services until G0 passes.

Boundary 2: Control Plane Entry (Midgard → Asgard)

- **Flow:** Authenticated User Request → Yggdrasil (Orchestrator Core).
- **Gate Enforced: G1 (Tyr) / G0 (Warden)** — RBAC verification, session validation.
- **Data State:** Trusted session context. User can view dashboards and trigger workflows, but cannot execute code or access raw data directly.

Boundary 3: Agent Execution (Asgard → Jotunheim)

- **Flow:** Orchestrator dispatches task to AI Agent (Odin/Thor) → Agent spawns in Sandbox.
- **Gate Enforced: G1 (Tyr) / G3 (Jotunheim Boundary)** — Pre-execution policy check, tool permission validation, sandbox initialization.
- **Data State:** Agent is provided with minimum-scope, just-in-time credentials (Credential Starvation). Agent output is tagged as `quarantined` if `low_trust_review` is active.

Boundary 4: External LLM Communication (Jotunheim → Vanaheim)

- **Flow:** AI Agent requires reasoning → API call to Anthropic / OpenRouter.
- **Gate Enforced: G3 (Egress Control)** — nftables/firewall allowlist permits traffic *only* to approved LLM endpoints (e.g., `api.anthropic.com`).

- **Data State:** Data leaves the SBS perimeter. PII is minimized/redacted prior to payload construction by G1. Return payload is treated as untrusted hostile content.

Boundary 5: Data Persistence (Jotunheim/Asgard → Niflheim/Database)

- **Flow:** Agent or Orchestrator writes result to PostgreSQL/MongoDB → Audit event written to Valhalla/Urd.
- **Gate Enforced: G1 (Tyr) / G4 (Forseti)** — Structured output validation, SQL sandbox keyword banlist, `company_id` enforcement.
- **Data State:** Data is encrypted at rest (AES-256-GCM for secrets). Audit log entry is append-only.

Boundary 6: External Side Effects (Jotunheim → Vanaheim/Internet)

- **Flow:** AI Agent attempts to send email, post to social media, or modify a third-party API (e.g., Audacy invoice generation).
- **Gate Enforced: G2 (HITL Approval)** — Action is parked. Human Board member must explicitly approve the prepared action.
- **Data State:** Zero execution occurs until human approval token is provided.

11.2 Legacy System Ingestion Boundary (SBS Internal Systems → Jotunheim/Asgard)

The Dash aggregates 40 years of SBS historical and operational data. This data lives in disparate, legacy systems (WideOrbit, Oracle, WideOrbit Automation, MusicMaster, SIMS TV, CounterPoint, ADP, and unknown Digital Systems). Data is collected via four distinct ingestion paths, all strictly governed by G1/G3:

1. **MCP / API Integration:** For modern systems (WideOrbit, Oracle, ADP), NornGate utilizes Model Context Protocol (MCP) and direct APIs. G1 enforces read-only, minimum-scope credentials just-in-time. The API connection is terminated in Jotunheim.
2. **Raw Data Dump:** For systems that only support scheduled exports (e.g., MusicMaster logs, SIMS TV traffic), CSV/SQL dumps are pushed via SFTP to an isolated staging directory in Jotunheim. A Low-Trust Agent parses the file, strips PII, and validates the schema before writing to Asgard.
3. **Report OCR Scraping:** For legacy systems with no API or export capability (e.g., unknown Digital Systems, legacy CounterPoint reports), PDF/visual reports are ingested. OCR processing is executed by a Low-Trust Agent in Jotunheim. The raw report is quarantined, and only the extracted, structured JSON data passes G4 validation to enter the Asgard data lake.

4. **Data Poisoning Prevention:** Because legacy ingestion is a primary vector for Data Poisoning (Threat AI-T4), all ingested data is compared against historical baselines (Skuld). Anomalous data (e.g., a sudden 500% spike in a specific demographic) triggers a G1 denial, quarantining the data in Helheim for human review before it can degrade the 40-Year Model.

---

12. Appendix C — Role-Based Access Control (RBAC) Matrix

(This matrix defines the strict least-privilege access model for all human and non-human actors within NornGate. Auditors should use this to write automated tests attempting to escalate privileges or cross tenant boundaries.)

Actor / Role	Environment / Realm	Permissions & Capabilities	Explicit Denials (Fail-Closed)
SBS Board / Admin	Midgard / Asgard	Read dashboards, approve G2 actions, manage user invites, rotate secrets metadata (no view), configure budget limits.	Cannot view raw secret values. Cannot bypass G1 policy. Cannot execute code directly.
SBS IT / System Admin	Asgard / Host Layer	SSH access (per-person named users), Docker container management, firewall rule deployment via IaC, database backup restoration.	No access to tenant data payloads. No access to NornGate application root without break-glass.
Human User (Employee)	Midgard	Read assigned compartment data, submit tasks to Workflow Agents, view personal audit history.	Cannot create new agents. Cannot alter G1 policies. Cannot access other compartments' data.
Workflow Agent	Jotunheim	Execute approved tool calls, read/write to assigned DB schema, call allowlisted external APIs (via G3 egress).	No standing credentials. Cannot mutate `executionPolicy`. Cannot access host filesystem. Cannot bypass G4 output validation.
Guardian Agent	Jotunheim / Asgard	Enforce policy (Tyr), tag output trust levels, quarantine low-trust content.	Cannot execute external side effects. Cannot approve its own generated content.

Actor / Role	Environment / Realm	Permissions & Capabilities	Explicit Denials (Fail-Closed)
Low-Trust Agent	Jotunheim (Isolated)	Process untrusted external content (e.g., support tickets, social media scraping, OCR report ingestion).	Output strictly quarantined. No direct DB write access. No access to secrets vault. Egress restricted to specific endpoints.

---

13. Appendix D — Business Unit Compartmentalization & Data Movement Policy

(This appendix details how SBS internal departments and external end-users are contained, how cross-department data movement is restricted by an Orchestrator traffic blacklist, and the specific Read/Write mappings enforced by G1 (Tyr).)

13.1 Business Unit Containment (Compartments)

SBS does not operate on a flat internal network. Internal business units (HR, Finance, Sales, Traffic, Marketing) are isolated into distinct **Compartments** within the NornGate control plane (Asgard).

- **Network Isolation:** Each Compartment operates on its own dedicated Docker Compose network. Cross-Compartment traffic is denied by network non-membership. An agent in the Sales Compartment cannot natively route traffic to the HR Compartment's database container.
- **Schema Isolation:** Data services (PostgreSQL, MongoDB, Redis) enforce per-Compartment separation. A Compartment can only read/write to its explicitly assigned database schema.
- **Agent Containment:** Workflow Agents are scoped strictly to their assigned Compartment. An HR Agent processing onboarding cannot execute tool calls against the Traffic/Continuity ad log.

13.2 Orchestrator Traffic Blacklist (Data Loss Prevention)

To prevent lateral movement of sensitive data between departments, the Orchestrator Core (Yggdrasil), enforced by G1 (Tyr), operates a strict **Data Movement Blacklist**. This acts as an internal Data Loss Prevention (DLP) layer.

- **Policy-as-Code:** G1 evaluates every attempted data transfer or tool call between Compartments against a deterministic policy file.
- **Denied Transfers (Examples):**
  - *HR → Marketing:* PII (Employee SSNs, home addresses) is blacklisted from being passed into Marketing campaign generation agents.
  - *Finance → Sales:* Raw GL/banking credentials are blacklisted from being passed into Sales media deck generation.
  - *Traffic → Marketing:* Unrevealed advertising rate cards are blacklisted from Marketing's promotional output.
- **Enforcement:** If an agent attempts to cross a Compartment boundary with blacklisted data payloads, G1 denies the action, logs the violation to Urd, and routes the job to Helheim (DLQ) for administrative review.

13.3 Business Unit Read/Write Access Matrix

Access to data within Compartments is strictly governed by RBAC. The following matrix maps the Read (R) and Read/Write (RW) permissions for SBS core business units.

Business Unit	HR Data	Finance/GL	Sales/CRM	Traffic/Inventory	Marketing Ops
HR	**RW**	R (Budget codes only)	Deny	Deny	Deny
Finance	R (Payroll metadata)	**RW**	R (Revenue tracking)	R (Billing reconciliation)	Deny
Sales	Deny	R (Client billing status)	**RW**	R (Avails/Inventory)	R (Campaign performance)
Traffic	Deny	R (Invoicing)	R (Client spots)	**RW**	Deny
Marketing	Deny	Deny	R (Audience targets)	Deny	**RW**

(Note: "R" requires G1 policy validation. "RW" requires G1 validation for reads, and G2 (Human-in-the-Loop) approval for writes that impact external systems or financial state.)

13.4 End-User Containment (Credential Rights)

External end-users (LaMusica app listeners, SBS Intranet viewers, Sol Vega interactors) are strictly contained by scoped credential rights enforced at G0 (The Gate).

- **Authentication:** End-users authenticate via Better Auth (OAuth/Email) or receive scoped, short-lived JWTs.
- **Realm Restriction:** End-users exist exclusively in **Midgard**. Their credentials explicitly deny access to the Asgard control plane, Jotunheim agent execution environments, and all internal Business Unit Compartments.
- **Data Scope:** An end-user JWT only permits access to their own individual behavioral profile and interaction history (e.g., their chat history with Sol Vega, their location fence data). They cannot query global audience data, advertiser metrics, or other users' PII.
- **Action Constraint:** End-users can submit input (e.g., song requests, chat messages) which are routed to Low-Trust Agents in Jotunheim. End-users possess zero direct execution credentials; all side effects (ad insertion, content generation) are handled by agents acting on their behalf, strictly governed by G1-G4.

---

14. Appendix E — AI Compute & Inference Strategy

(This appendix outlines the tiered AI compute architecture for auditors. Understanding this 80/20 split is critical for testing Denial of Wallet (DoW) and egress control mechanisms.)

14.1 The 80/20 Inference Split

To balance high-end AI reasoning with strict cost control (AI Unit Economics), NornGate does not route all traffic to premium external APIs. Compute is split across two distinct tiers:

- **80% Routine Orchestration (Open-Weight Models):** The vast majority of agent execution—including log reconciliation, DAI assembly, traffic routing, and standard content generation—is handled by open-weight LLMs (specifically **GLM / Kimi**). These models are hosted in-house, meaning inference costs drop to near-zero (just compute electricity).
- **20% Complex Reasoning (Premium API):** High-stakes, nuanced tasks—such as Sol Vega scriptwriting, complex contextual ad arbitration, and meta-reasoning (Odin

agent)—are routed via API to **Anthropic (Claude)**.

14.2 Infrastructure Scaling Roadmap

The hardware hosting the 80% open-weight tier scales deliberately across three phases to protect SBS's capital expenditure:

1. **Phase 0/1 (Agility - Current):** Initial deployment routes traffic through **OpenRouter**. This allows rapid integration and testing of multiple models without heavy upfront capital expenditure.
2. **Phase 2/3 (Scaling):** As inference volume grows, compute migrates to **rented DigitalOcean GPU clusters**. This lowers marginal costs, secures data processing environments within the VPC, and tests sustained load before hardware is purchased.
3. **Phase 4+ (Maturity):** Once usage patterns are proven and optimized, SBS will acquire **dedicated H200 GPU infrastructure**. This drives inference costs to the floor and ensures complete sovereign control of the 40-Year Model and the 80% open-weight tier.

14.3 Auditor Implications

- **Denial of Wallet (DoW) Testing:** Auditors testing DoW (Threat AI-T2) should understand that infinite loops targeting the Anthropic API will trigger the `budget hard-stop` (Hermod). However, loops targeting the open-weight tier (GLM/Kimi) will primarily consume local DigitalOcean GPU resources, testing the container resource limits (CPU/RAM) rather than token budget.
- **Egress Testing:** Egress allowlists (G3) must explicitly permit traffic to OpenRouter, Anthropic, and DigitalOcean internal GPU endpoints, but deny all other external AI APIs.

---

15. Appendix F — AI-Specific Threat Vector Matrix

(While NornGate addresses the OWASP LLM Top 10, this matrix maps the specific AI attack vectors relevant to SBS's deployment of The Dash. Auditors must specifically target these vectors during the Red-Team engagement.)

Threat ID	OWASP LLM Mapping	Attack Vector Description	NornGate Mitigation (Gate)	Auditor Testing Goal

AI-T1	LLM01: Prompt Injection	Attacker submits malicious input via LaMusica app chat or social media scraper, instructing the agent to ignore previous rules and exfiltrate data.	**G1 (Tyr) / G3 (Jotunheim):** Agents have zero standing credentials (Credential Starvation). Even if tricked, they lack API keys to exfiltrate data. Egress is firewalled.	Attempt to trick Sol Vega into returning DB connection strings or API keys via the chat interface.
AI-T2	LLM08: Excessive Agency / DoW	Attacker triggers a recursive loop in a Workflow Agent (e.g., infinite podcast generation), draining the Anthropic API token budget (Denial of Wallet).	**G1 (Tyr) / Hermod:** `budget hard-stop auto-pause`. Scheduler skips invocation if agent is paused or over budget.	Attempt to trigger an infinite agent loop and verify that the budget hard-stop terminates the process before exceeding the $12k/mo cap.
AI-T3	LLM02: Insecure Output Handling	LLM returns a payload containing malicious script (XSS) or SQL injection, which the system executes without sanitization.	**G4 (Forseti):** Structured-output validation and consensus/judging. Output must match strict JSON schema before leaving Jotunheim.	Inject a payload via LLM response designed to drop a database table or execute XSS in the admin console.
AI-T4	LLM04: Data Poisoning	Attacker slowly feeds anomalous data into the 40-Year Model via legacy API ingestion (e.g., WideOrbit, SIMS TV) or OCR scraping, skewing ad targeting to benefit a competitor.	**G1 (Tyr) / Skuld:** Continuous data-integrity monitoring. Anomalous data fed via legacy APIs or OCR is flagged and quarantined before model degradation.	Attempt to push malformed audience demographic data via the WideOrbit API adapter or a crafted OCR report and verify it is quarantined.
AI-T5	LLM05: Supply Chain Vulnerability	A third-party npm plugin or adapter is compromised (typosquatting) and attempts to execute malicious code inside the NornGate host.	**G0/G1 / Host Hardening:** `npm install --ignore-scripts` enforcement. External adapters load out-of-process. Plugin capability gating denies unknown operations.	Submit a mock malicious plugin with a `postinstall` script and attempt to execute it via the adapter loader.
AI-T6	LLM06: Sensitive Info Disclosure	LLM accidentally includes PII or secrets from its training context/prompt in its generated response to a user.	**G1 (Tyr) / Layer 5:** PII data minimization. Just-in-time credential injection. Redaction pipelines in run logs and transcripts.	Attempt to trick the agent into repeating the system prompt or revealing the just-in-time Anthropic API key.

16. Closing & Next Steps

The transition of Spanish Broadcasting System into an AI-native media ecosystem is a monumental undertaking that requires an uncompromising commitment to security. This report outlines the robust architectural design of NornGate—anchored by the 5-Gate Taxonomy, Credential Starvation, and strict Business Unit Compartmentalization—which provides a fundamentally secure foundation for The Dash.

While the "Day 0" infrastructure and application gaps identified herein are inherent to an active foundational deployment, they are explicitly tracked, actively remediated, and will be fully closed before any live production data transits the platform. The closed-system deployment model further mitigates exposure during this critical build phase.

Security is not a static achievement but a continuous operational discipline. Upon completion of the foundational installation, Ledger Hub Networks will execute comprehensive internal and external security audits. We fully support and welcome SBS's option to engage an independent third-party security firm (such as Certek) to perform final penetration testing, AI Red-Team validation, and compliance certification prior to production go-live.

The Dash is designed to be the strongest amplifier SBS has ever had; NornGate ensures that this amplification never comes at the cost of data integrity, listener privacy, or operational security. We are committed to delivering a platform that is not only intelligent and profitable but inherently trustworthy.

Document End.*Confidential — Ledger Hub Networks, LLC × Spanish Broadcasting System, Inc.*

Revision #2
Created 2026-07-31 15:37:40 UTC by SBS Admin
Updated 2026-07-31 16:04:44 UTC by SBS Admin