Shield *for LLM* Runtime

Continuous runtime defense that runs *inside every live inference.*
Three engines. One coordinated response.

LCAC, LBF, and CBE operate in concert — controlling context access, detecting zero-day jailbreaks through behavioral fingerprinting, and enforcing hard execution limits. Shared intelligence. Policy evaluation in under 10 ms.

● real-time enforcement · sub-10ms policy evaluation

<10ms	3	Real-Time
POLICY EVALUATION LATENCY PER INFERENCE	COORDINATED ENGINES · LCAC · LBF · CBE	CROSS-ENGINE INTELLIGENCE SHARING

RUNTIME ENGINES

Three engines. One coordinated shield.

Each engine covers a distinct attack surface. All three share threat signals in real time, creating a defense posture no single-layer solution can match.

LCAC

LLM CONTEXT ACCESS CONTROLLER

Context Layer Defense

Controls precisely what context the model can access at inference time. Enforces scope boundaries, redacts out-of-policy knowledge, and blocks retrieval paths that exceed the model's permitted information aperture.

- Dynamic context window scoping per session role
- Retrieval path allowlist enforcement
- Out-of-policy knowledge redaction at token level
- Cross-request context bleed prevention
- Privileged document access gating
- Real-time scope violation alerting

LBF

LLM BEHAVIORAL FINGERPRINTER

Behavioral Layer Defense

Monitors how the model behaves across requests — not just what it says. Detects zero-day jailbreaks, persona drift, and emergent unsafe behaviors through statistical fingerprinting of output patterns and response trajectories.

- Zero-day jailbreak detection via behavioral delta
- Persona drift and role-override fingerprinting
- Response trajectory anomaly scoring
- Cross-session behavioral baseline comparison
- Emergent unsafe pattern flagging
- Adaptive threshold tuning per model family

CBE

CAPABILITY BOUNDARY ENFORCER

Execution Layer Defense

Enforces hard limits on what the model can do at execution time. Caps tool chaining depth, sets execution ceilings for agentic loops, and prevents capability escalation attempts from reaching connected infrastructure.

- Configurable tool chaining depth caps
- Agentic loop execution ceiling enforcement
- Capability escalation attempt interception
- Per-tool permission scope validation
- Hard kill-switch for runaway agent chains
- Execution audit trail with replay support

SHARED INTELLIGENCE

All three engines. One threat picture.

Shield's engines do not operate in silos. Every signal from every engine is shared in real time, enabling coordinated responses that no individual layer can trigger alone.

EVALUATION LATENCY

<10ms

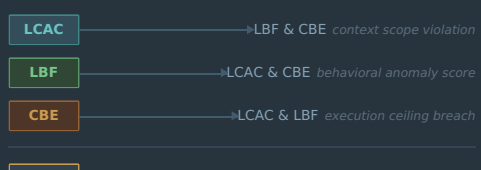
Cross-engine policy evaluation time per inference — including signal aggregation and enforcement decision.

SIGNAL PROPAGATION

3-way

Every engine broadcasts its threat state to the other two on every inference cycle, in real time.

LIVE SIGNAL EXCHANGE



ENFORCEMENT PIPELINE

Four stages. Every inference.

Shield's pipeline runs synchronously on every request — with no perceptible latency impact at production throughput.

01

Intercept & Parse

Incoming request and context payload captured. Role, session state, tool config, and retrieval scope extracted into a canonical inference record.

02

Engine Evaluation

LCAC, LBF, and CBE evaluate the inference record in parallel. Each engine emits a threat signal and enforcement recommendation within 10 ms.

03

Signal Aggregation

Cross-engine signals aggregated into a unified policy decision. Escalation, throttle, redact, or block actions resolved by severity weighting.

04

Enforce & Log

Enforcement action applied before the model output is returned. Full audit record written locally with provenance, severity, and replay support.

DEPLOYMENT OPTIONS

Your environment. Your rules.

No architecture changes. No migration. Shield integrates at the inference layer — wherever your models run.

DEVELOPER · SIDECAR

SDK & Middleware

- Python & Node.js drop-in wrappers
- OpenAI & Anthropic API compatible
- Configurable engine thresholds per env
- Local policy file support
- Async & streaming inference support

PLATFORM · GATEWAY

Inference Proxy

- Transparent HTTPS inference gateway
- Zero client-side code changes
- Multi-model & multi-provider routing
- Per-route engine config profiles
- gRPC & REST control plane API

ENTERPRISE · DASHBOARD

Command Centre

- Unified threat dashboard — LCAC, LBF, CBE
- RBAC & SSO via SAML 2.0 / OIDC
- Real-time alert routing — Slack, PagerDuty
- Compliance export & audit log trail
- Custom policy authoring interface

ENGINE QUICK REFERENCE

ENGINE	FULL NAME	ATTACK SURFACE	PRIMARY DEFENSE	RESPONSE
LCAC	LLM Context Access Controller	Context window & retrieval	Scope enforcement, redaction of out-of-policy knowledge	**Redact / Block**
LBF	LLM Behavioral Fingerprinter	Output behavior & model drift	Zero-day jailbreak detection via behavioral delta scoring	**Alert / Halt**
CBE	Capability Boundary Enforcer	Tool use & agentic execution	Execution ceiling & tool chain depth hard limits	**Cap / Kill**

COMPLIANCE ALIGNMENT: SOC 2 Type II · ISO 27001 · NIST AI RMF · EU AI Act · OWASP LLM Top 10 · GDPR

"Runtime is the most hostile phase of an LLM's lifecycle. Context can be poisoned, behavior can drift, and agents can chain tools in ways nobody intended. Shield is the only product that defends all three attack surfaces simultaneously — with shared threat intelligence and sub-10ms enforcement."

CORELAYER · SECURITY RESEARCH TEAM · 2026

Request an enterprise evaluation.

A solutions engineer will deploy Shield against your production inference environment, tune LCAC, LBF, and CBE thresholds for your model stack, and deliver a runtime risk assessment within five business days.

REQUEST A DEMO →

READ THE DOCS