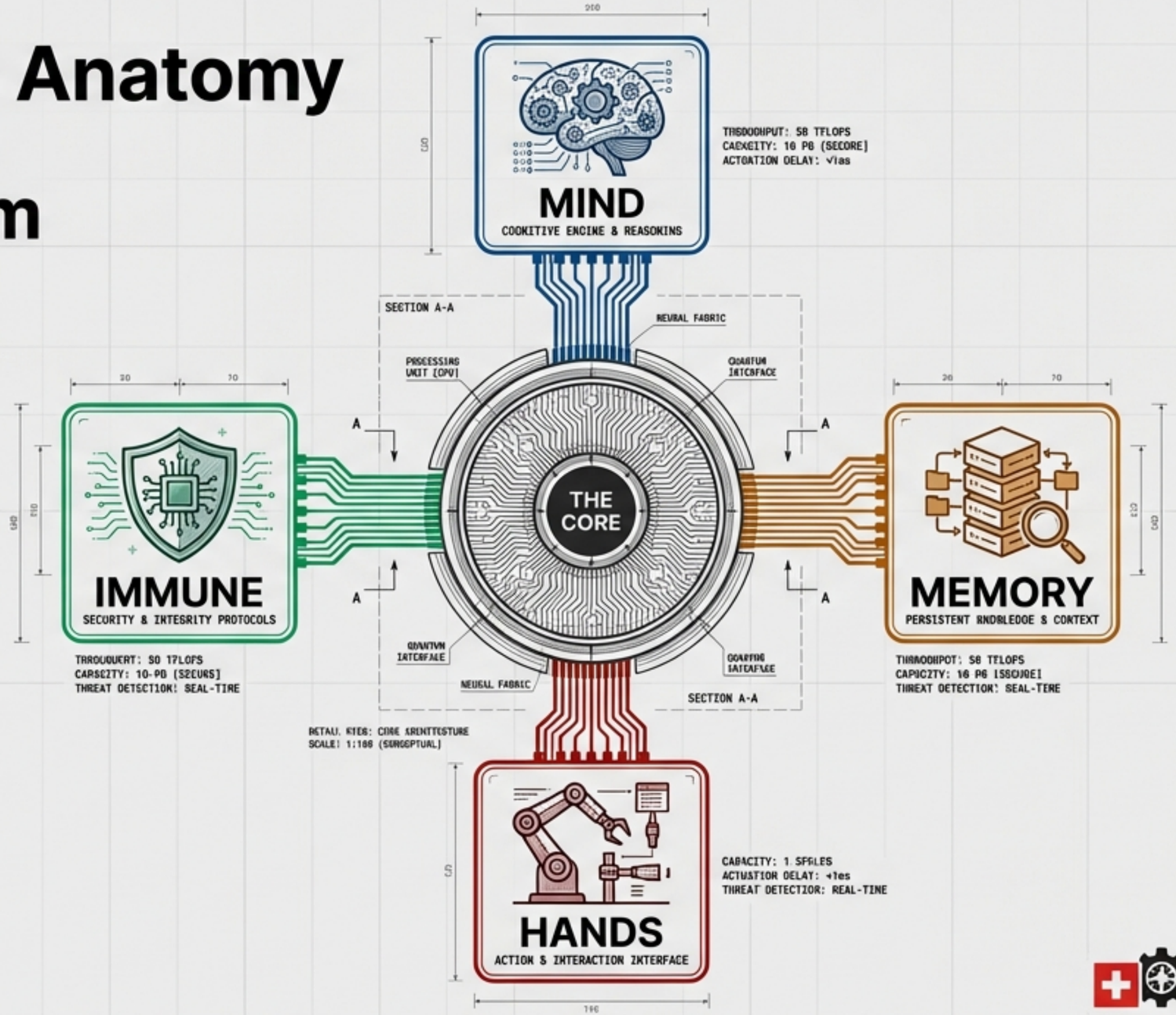
ContextUnity: The Anatomy of an Autonomous AI Operating System

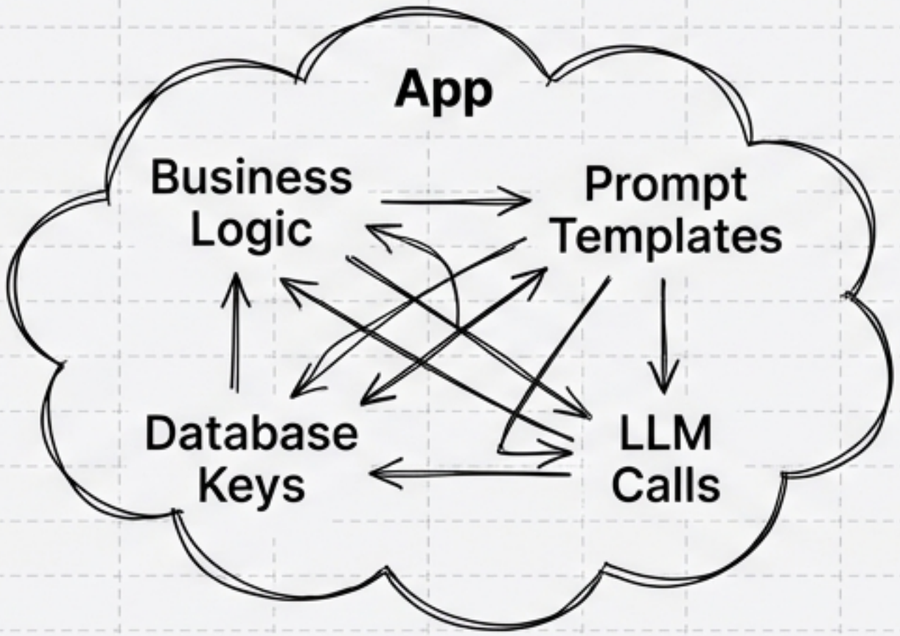
A Modular, Agnostic Infrastructure for Sovereign Digital Entities



[SOVEREIGN]: Tenant-Bound Data Ownership
[AGNOSTIC]: Decoupled Capability & Intent
[MODULAR]: Registry-Based Component Swapping

The Architectural Thesis: Inversion of Control

Legacy: Monolithic Chatbot



ContextUnity: Digital Entity Ecosystem

Domain / Project (Intent)

Intent is provided by the Project.

------ gRPC Boundary ------

Mind (Reasoning) & Hands (Execution)

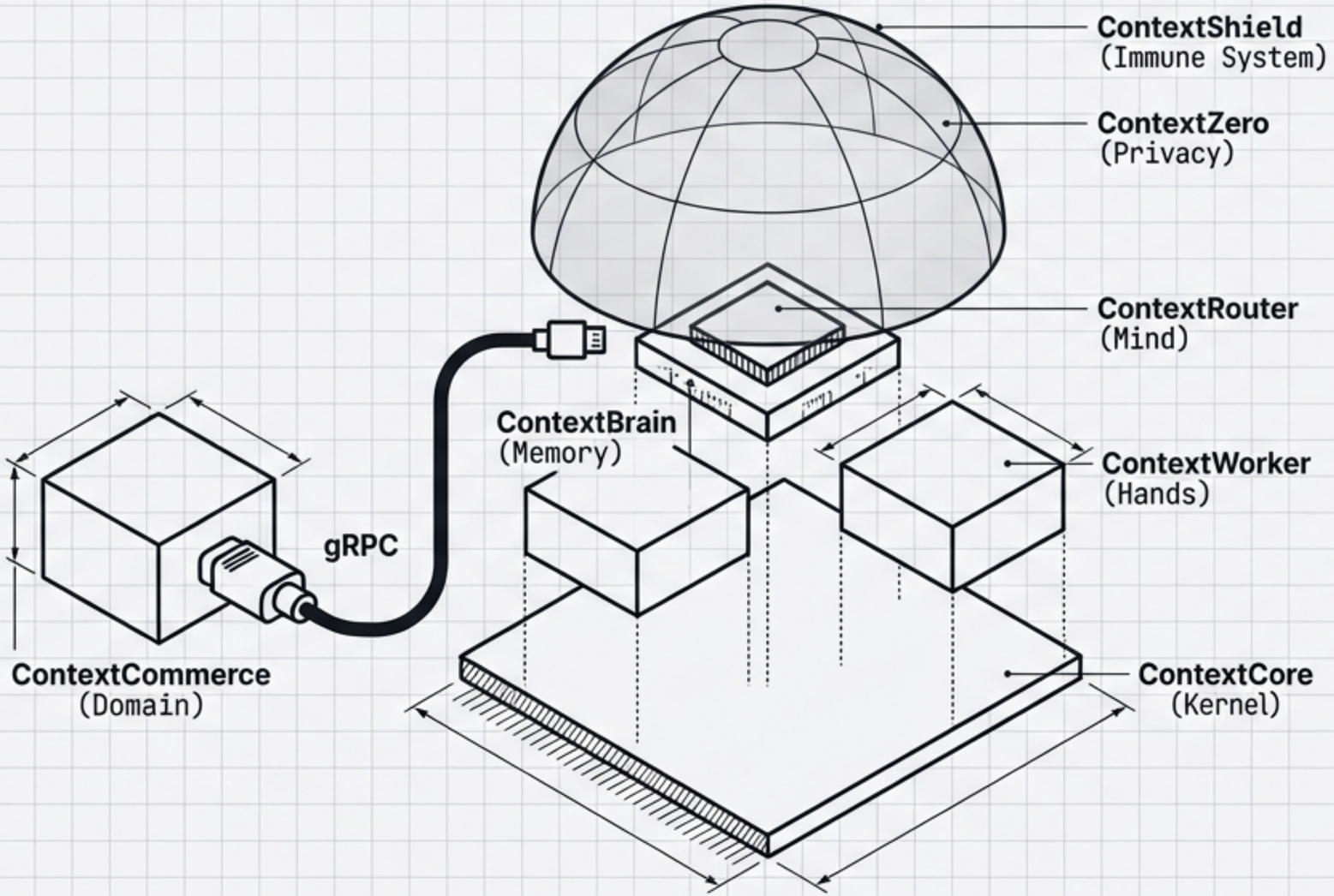
ContextCore (Kernel & Infrastructure)

Capability is provided by the OS.

Logic Refactoring Threshold: >400 Lines

Engineering Title

System Overview: The Biological Metaphor



Kernel: Protocols & Contracts
Mind: Reasoning & Orchestration
Memory: Vector & Graph Storage
Hands: Durable Execution
Immune: Policy Enforcement

The Nervous System: ContextCore

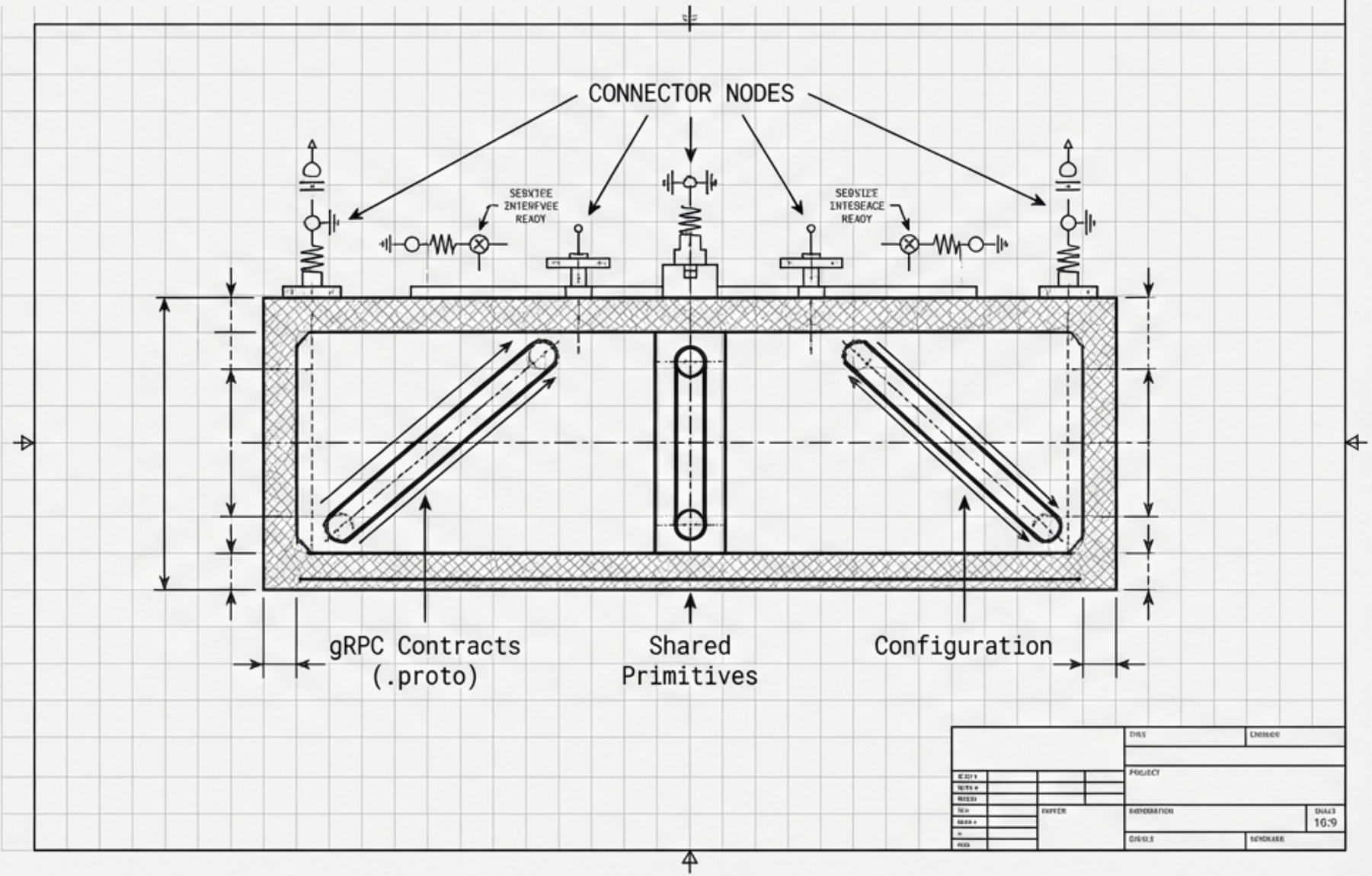
Infrastructure Zero

Rule 1: Zero Business Logic. Pure Infrastructure.

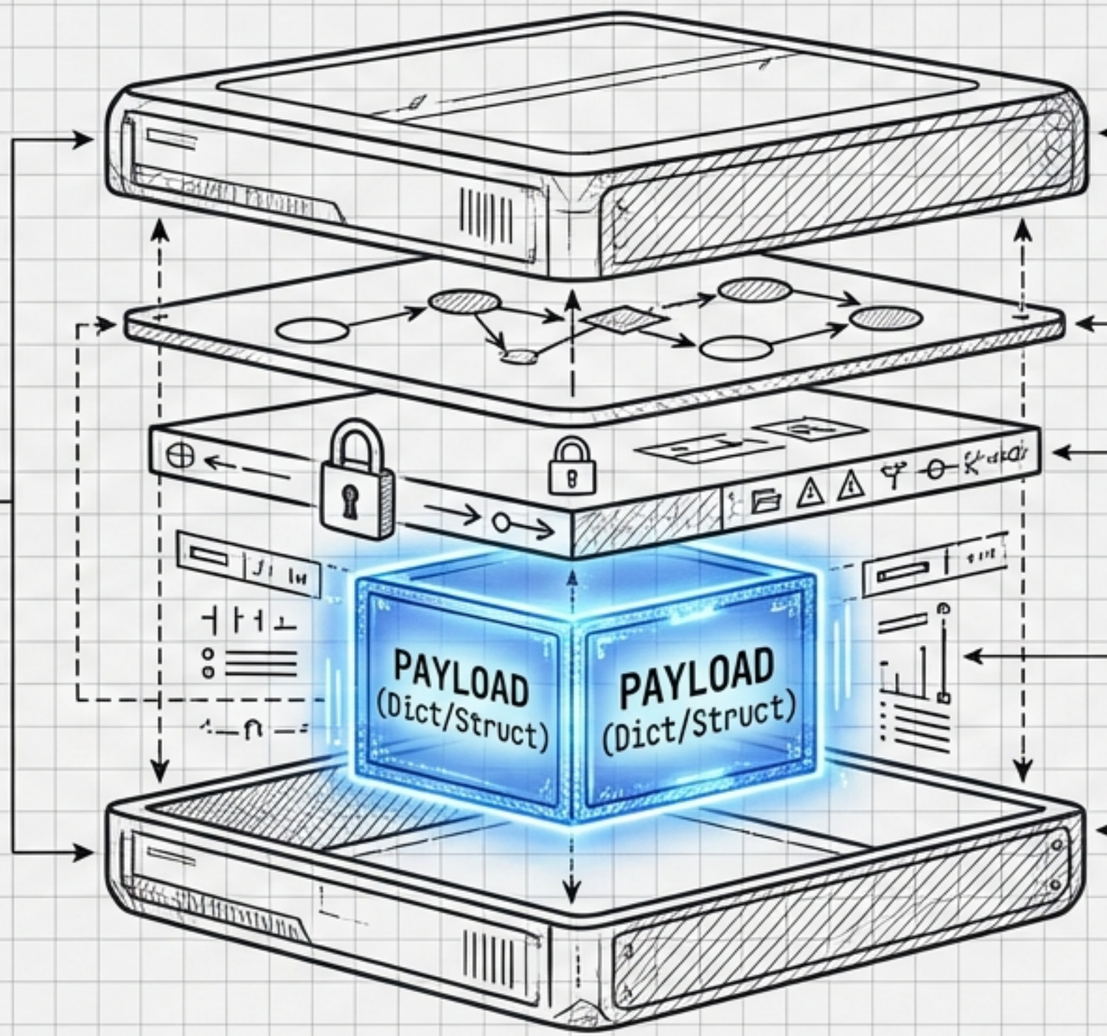
Rule 2: Strict Contracts via Protobuf.

Rule 3: Limited Dependencies: grpcio, protobuf, pydantic.

Output: Unified Structured Logging & Telemetry.



The Universal Protocol: ContextUnit



Envelope
(UUID: Trace_ID)

Provenance Chain
[router:ingest -> shield:scan -> brain:store]

Security Context
[Scopes: Read/Write]

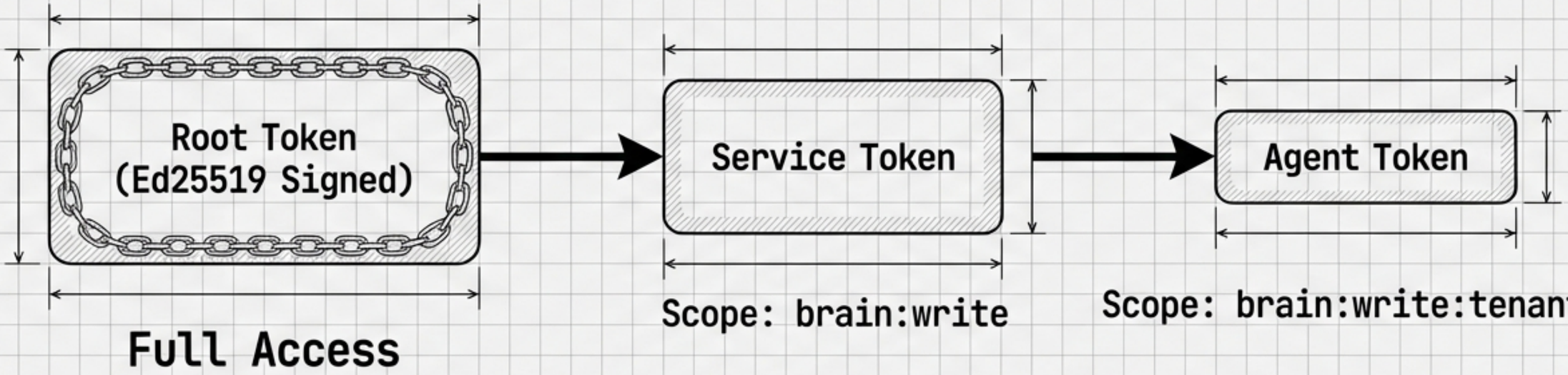
Metrics
[Latency, Cost, Tokens]

Envelope
(UUID: Trace_ID)

Modality Agnostic:
Text / Audio / Spatial

Every atomic unit of data is traceable, measurable, and secure.

Identity & Access: The ContextToken

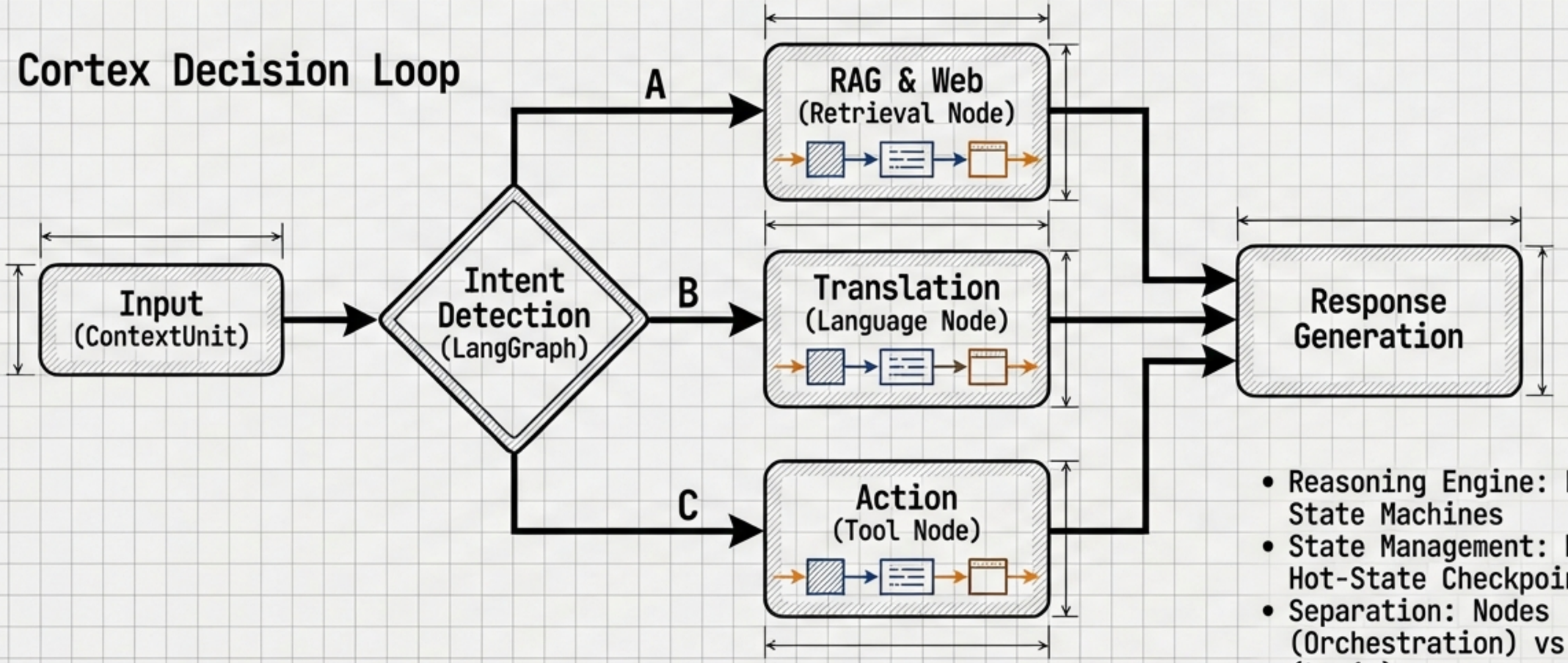


Key Concepts

- Capability-Based Security (Not Identity-Based)
- **Verification:** Public Key Only (No DB Lookup)
- **Attenuation:** Permissions can only shrink, never grow.
- **Scope Syntax:** {domain}:{action}:{resource}

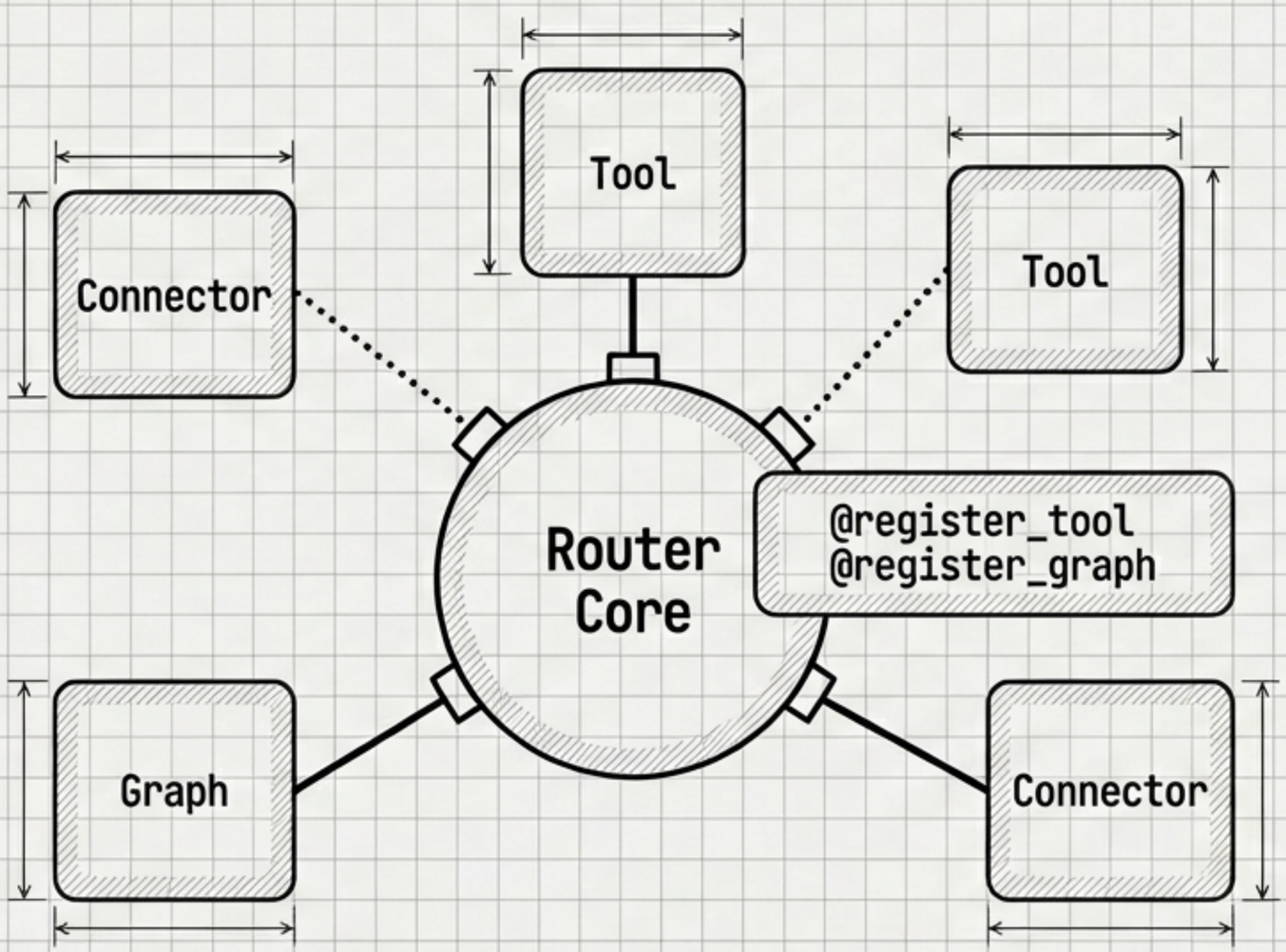
The Mind: ContextRouter & Cortex

Cortex Decision Loop



- Reasoning Engine: LangGraph State Machines
- State Management: Redis Hot-State Checkpoints
- Separation: Nodes (Orchestration) vs. Steps (Logic)

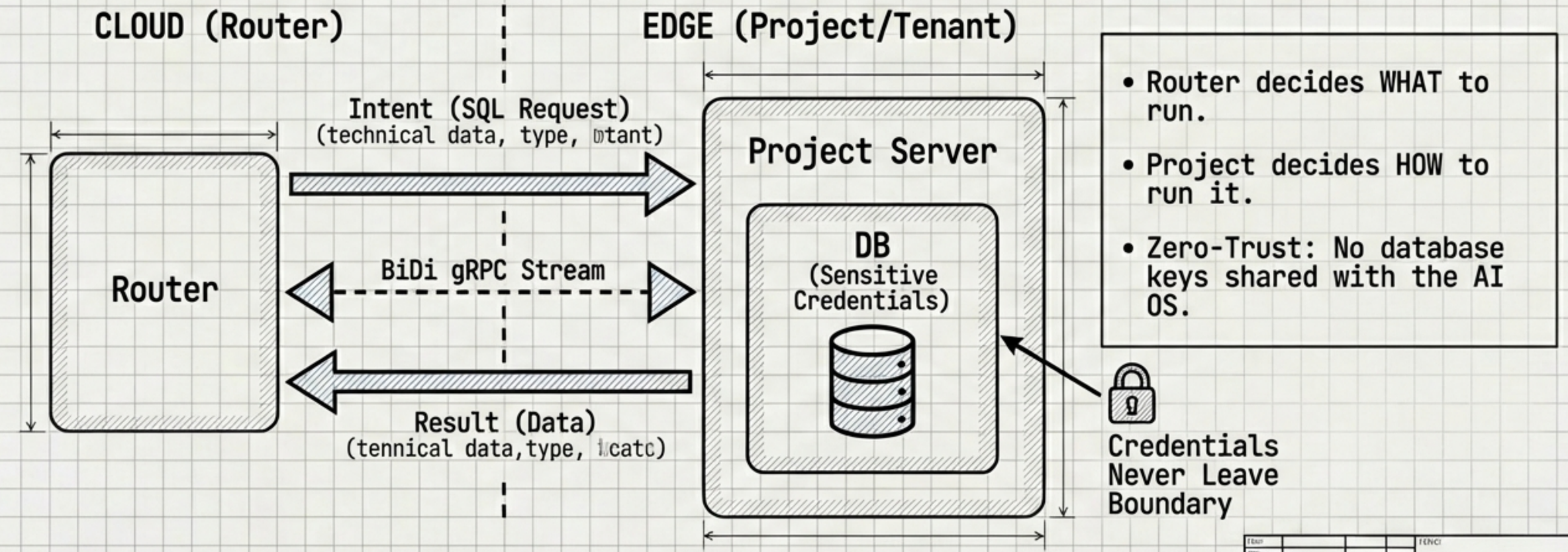
Dynamic Capabilities: The Registry System



- **Runtime Discovery:** Automatic loading from 'modules/tools'
- **Hot-Swapping:** No restart required for new logic.
- **Universal Model Registry:** OpenAI / Vertex / Anthropic / Ollama (Local)
- **Fallback Strategies:** Sequential -> Parallel -> Cost-Priority

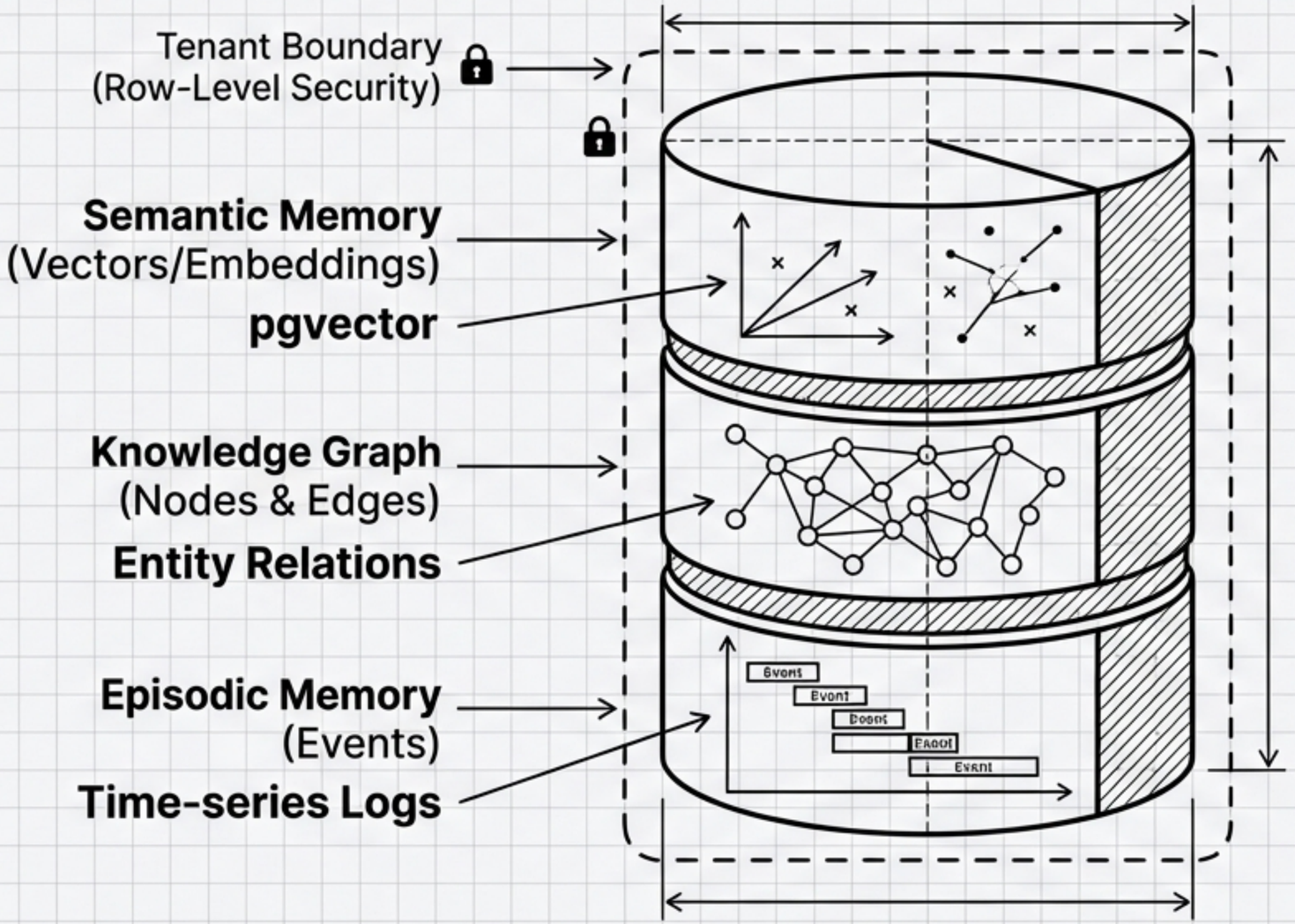
Federated Execution: BiDi Binding

Inversion of Control / Zero-Knowledge Execution



- Router decides WHAT to run.
- Project decides HOW to run it.
- Zero-Trust: No database keys shared with the AI OS.

The Memory: ContextBrain



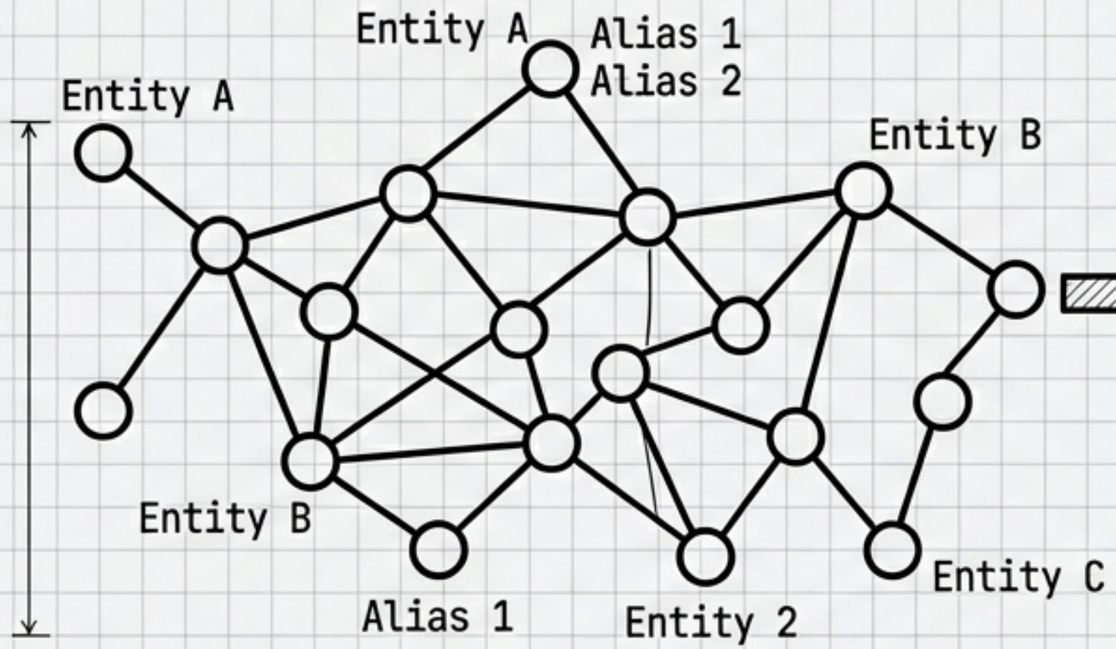
Unified Storage:
`PostgreSQL + Extensions`

Caching:
`Redis-backed Embedding Cache (emb:{sha256})`

Procedural Memory:
`Stored Skills`

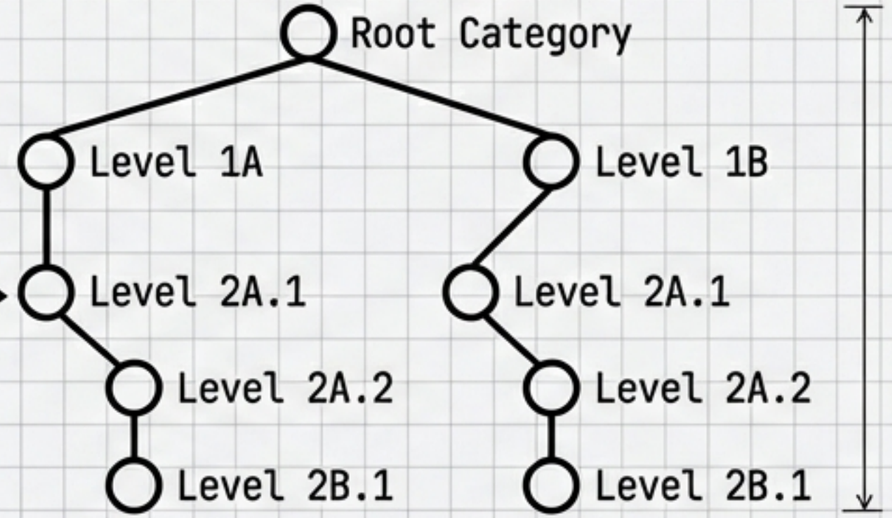
Structured Knowledge & Taxonomy

Knowledge Graph (Relations)



Entity Aliases & Synonyms

Taxonomy (ltree)

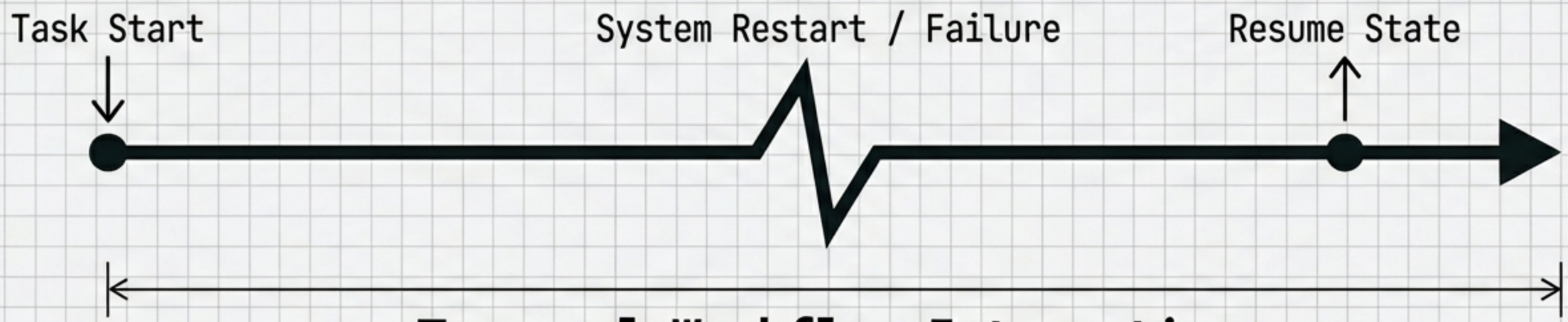


Category Paths

PostgreSQL Extensions: pgvector, ltree, uuid-ossp
Isolation: SET LOCAL app.current_tenant

The Hands: ContextWorker

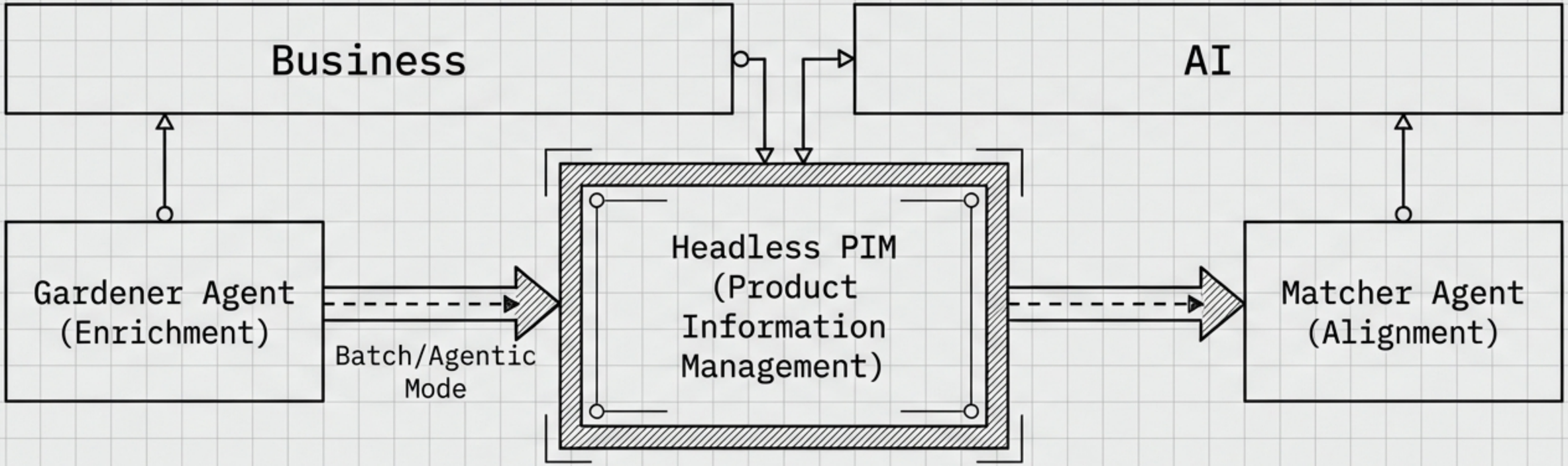
Durable Execution



Temporal Workflow Integration

Mode A: Harvester (Heavy Data Ingestion)
Mode B: Agent Registry (Background Polling)

The Domain: ContextCommerce



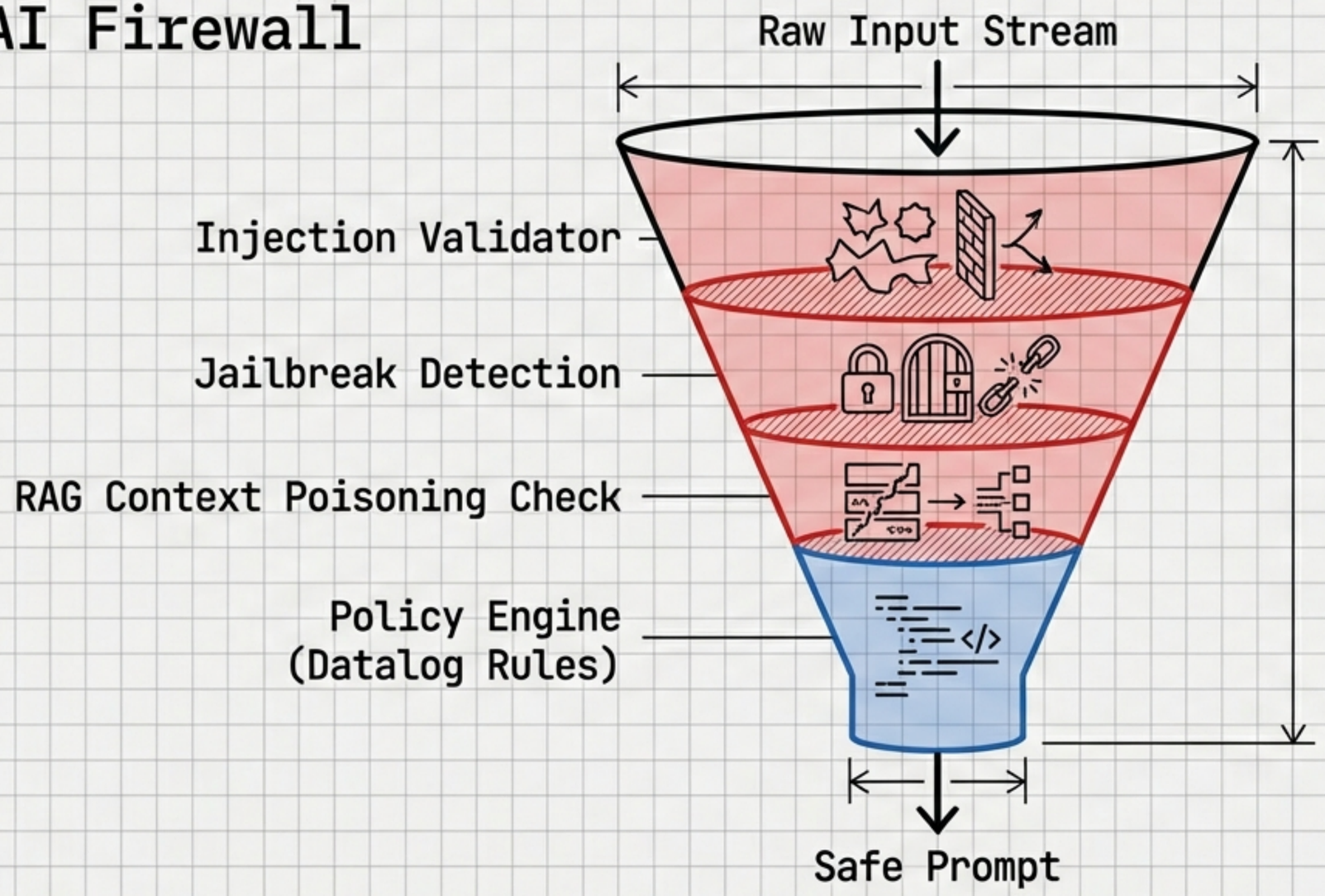
Process List:

- Batch Mode: High Volume (2-5s/item)
- Agentic Mode: Complex Reasoning (10-30s/item)
- Manual Mode: Human-in-the-Loop Review

The Platform provides the Tools.
The Domain provides the Taxonomy.

The Immune System: ContextShield

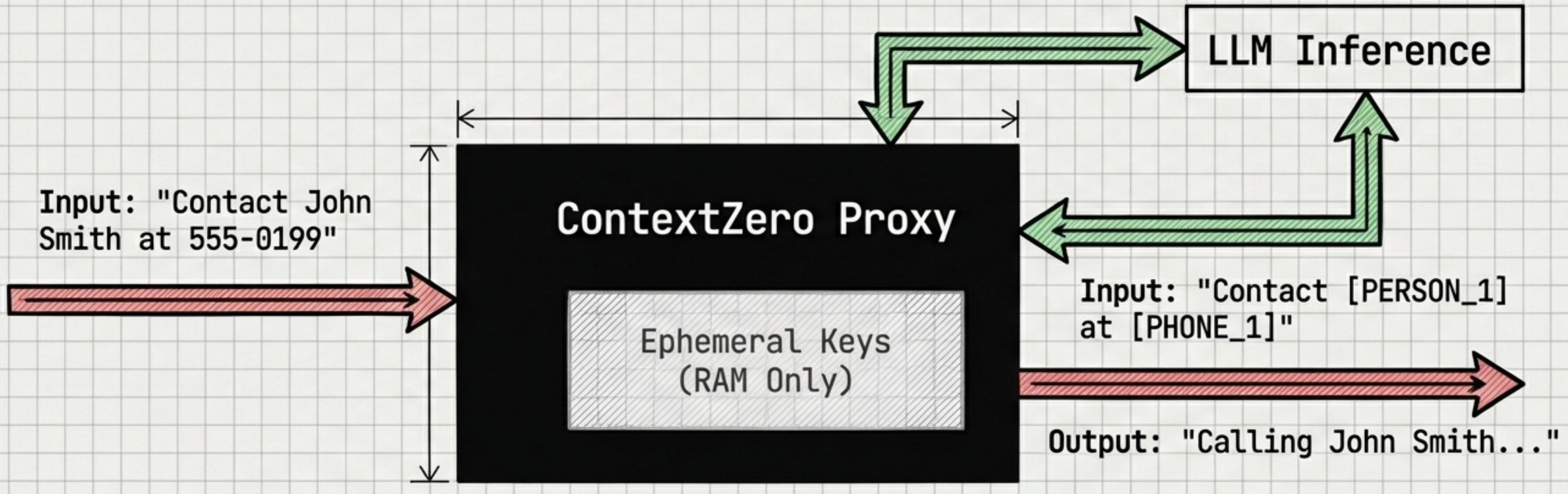
AI Firewall



Audit Trail:

Immutable JSON-lines logging to ClickHouse/S3.

The Privacy Proxy: ContextZero

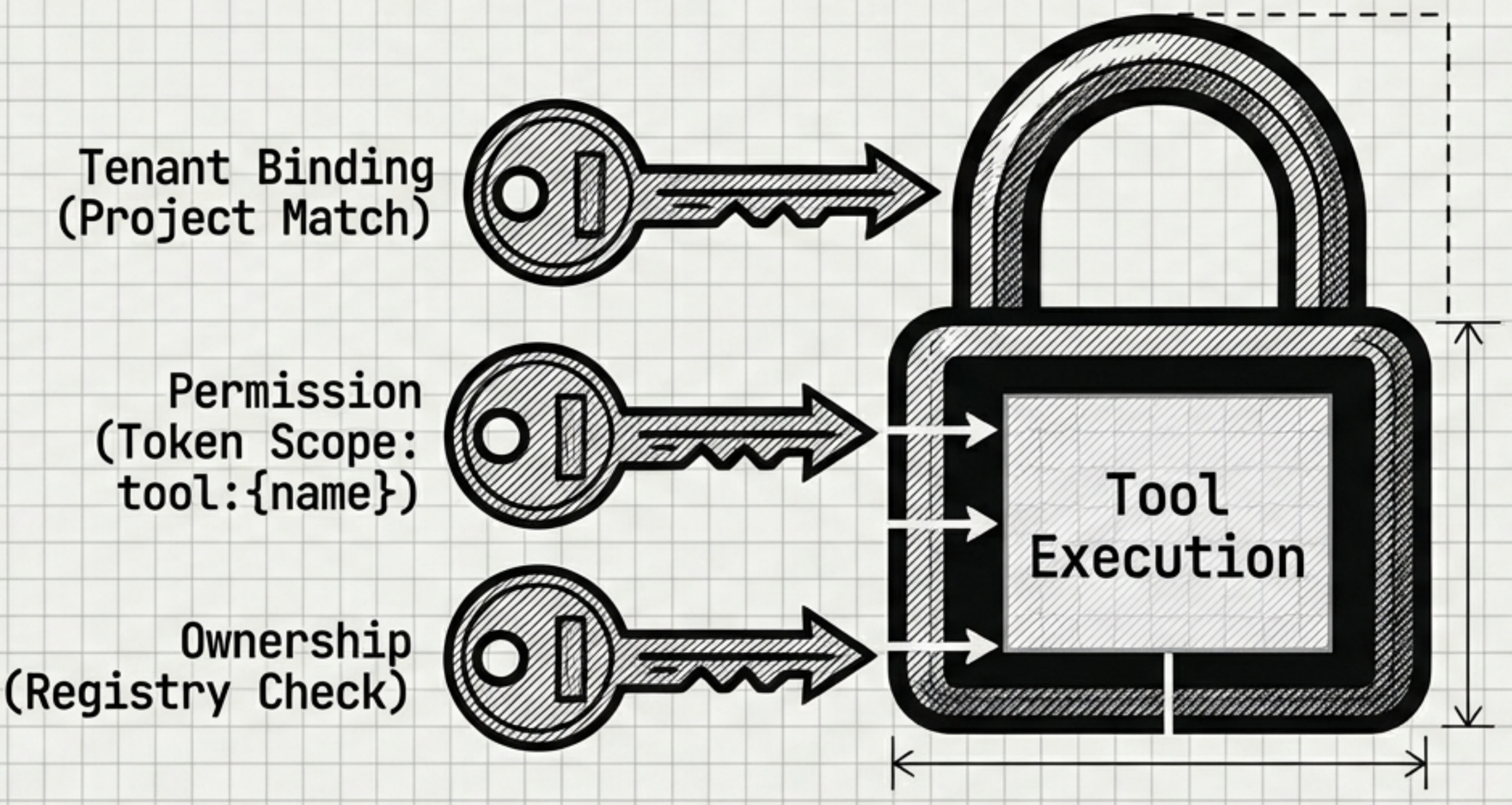


Key Features:

- PII Masking: Regex + Named Entity Recognition
- Deanonymization: Automatic restoration on return
- No Persistence: Keys destroyed post-session

The ContextZero Proxy sanitizes sensitive data before it reaches the AI and restores it in the output, ensuring privacy without compromising functionality.

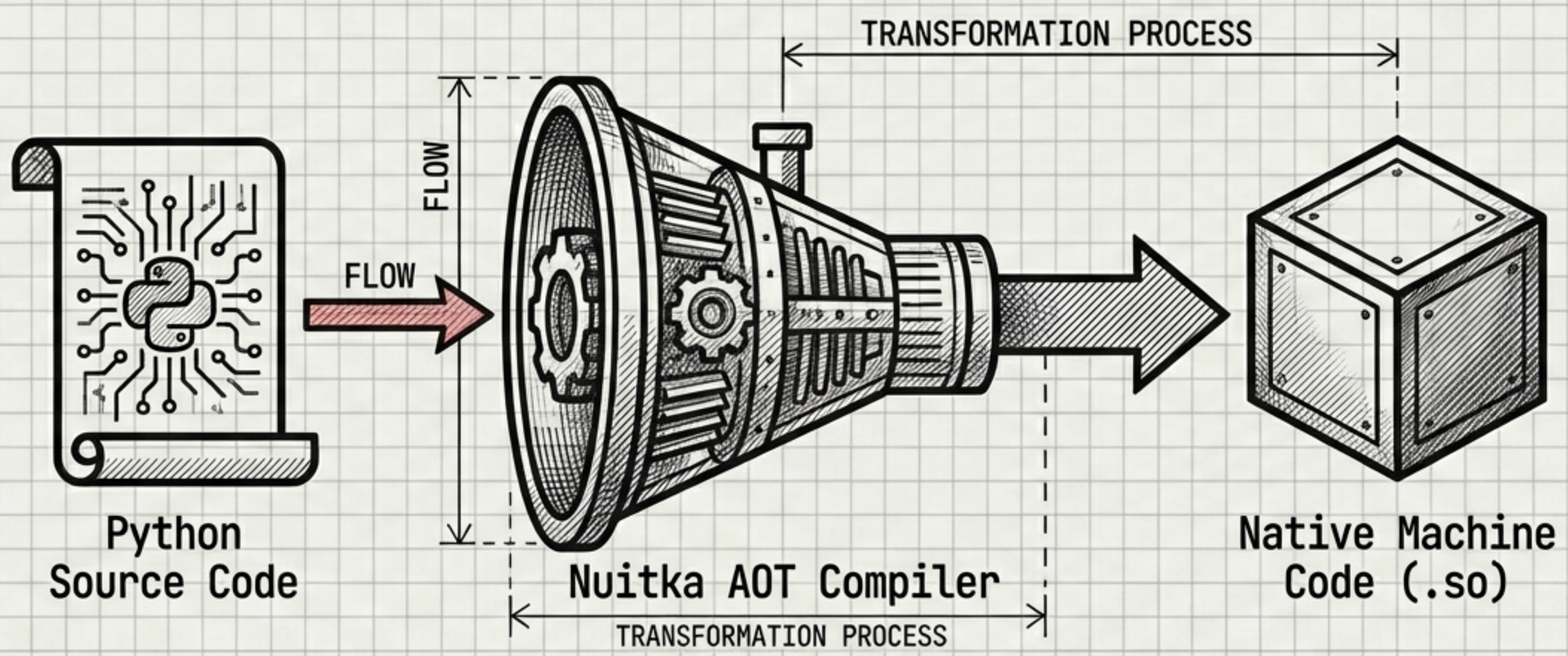
Secure Tool Enforcement



Fail-Closed Architecture: No Token = No Execution.

Zero-Trust Stream: Authenticated via Shield.

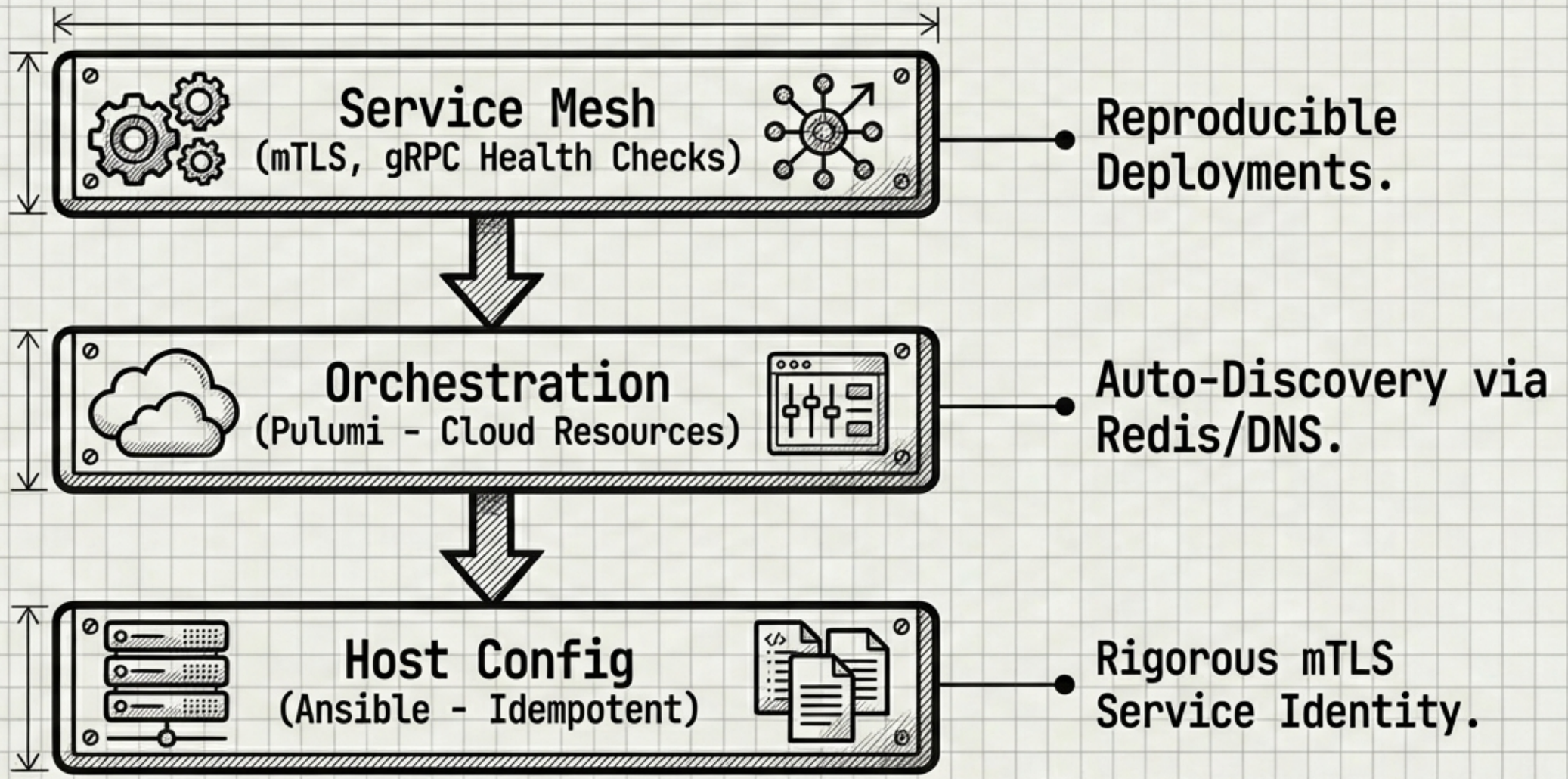
Platform Protection & Licensing



Key Features:

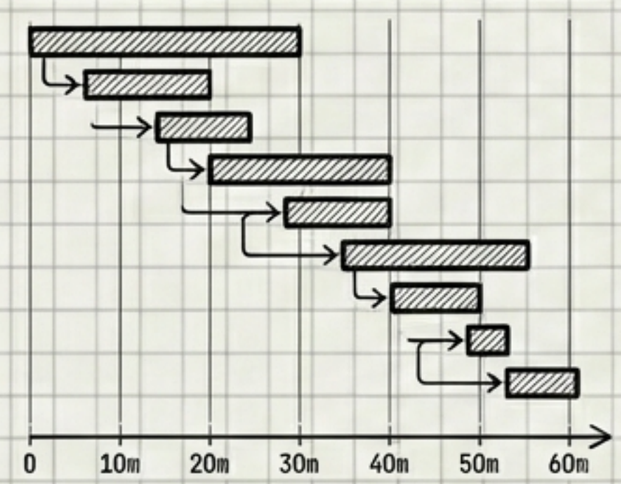
- Binary Distribution: Python bytecode stripped.
- Hardened Architecture: Critical modules (Shield, Zero) compiled.
- Licensing: Ed25519 Signed JSON Verification.

Infrastructure as Code



Observability & Telemetry

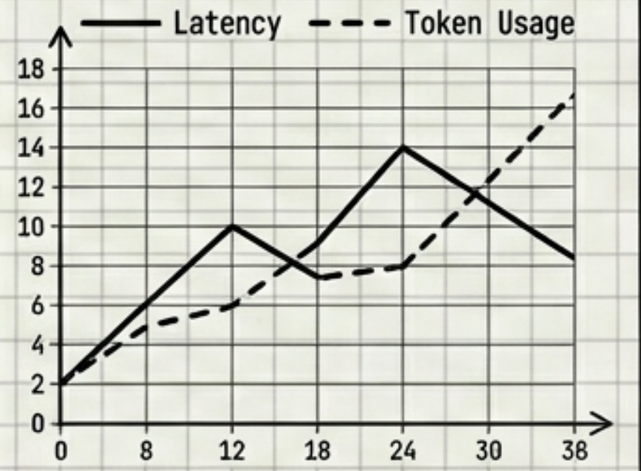
OpenTelemetry / Langfuse



Logs

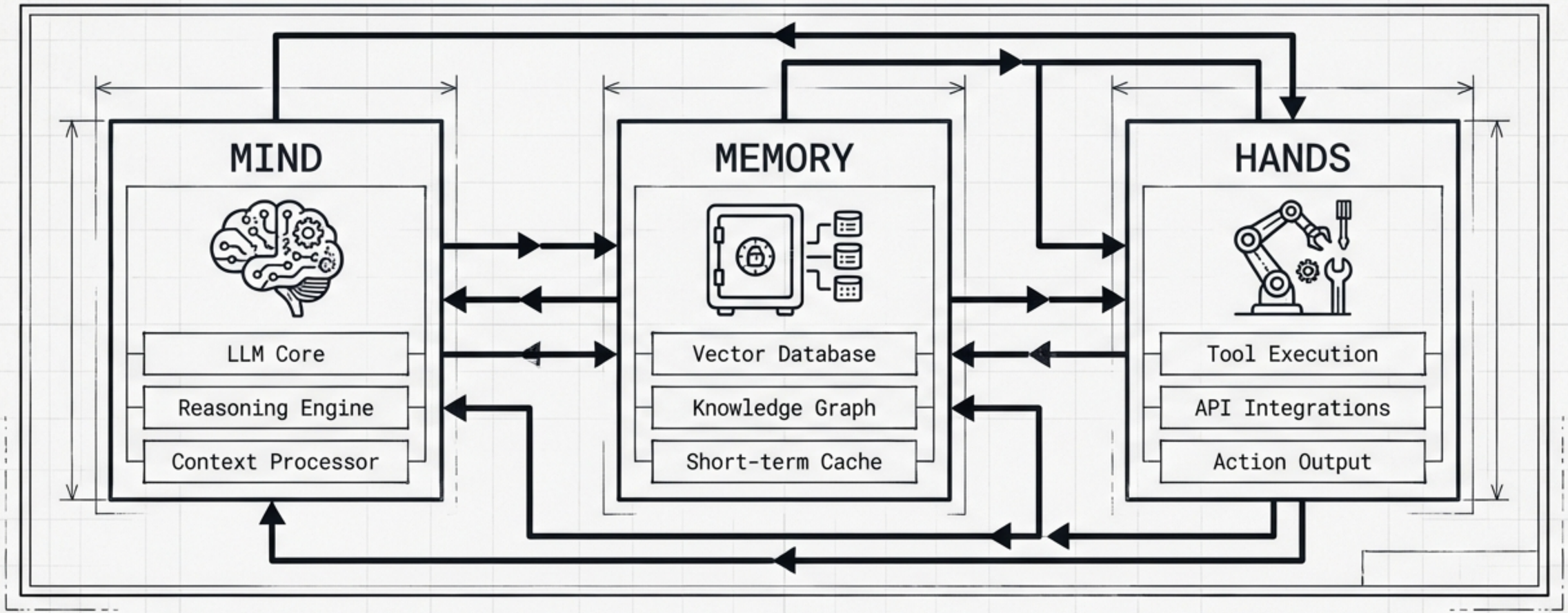

```
{
  "timestamp": "2024-01-01T10:00:00Z",
  "level": "INFO",
  "service": "llm-engine",
  "trace_id": "abcdef12345",
  "user_id": "user_001",
  "action": "completion_request",
  "model": "gpt-4",
  "parameters": {
    "temperature": 0.7,
    "max_tokens": 100
  },
  "payload": "{\"prompt\": \"Translate
            'Hello' to French.\"}",
  "password": "[REDACTED]"
}
```


Metrics



- Deep Introspection of LLM Reasoning.
- Distributed Tracing across Service Boundaries.
- Structured JSON Logging with Secret Redaction.

The Complete Digital Entity



UNIVERSAL. SECURE. SOVEREIGN.

A unified, programmable AI Operating System.