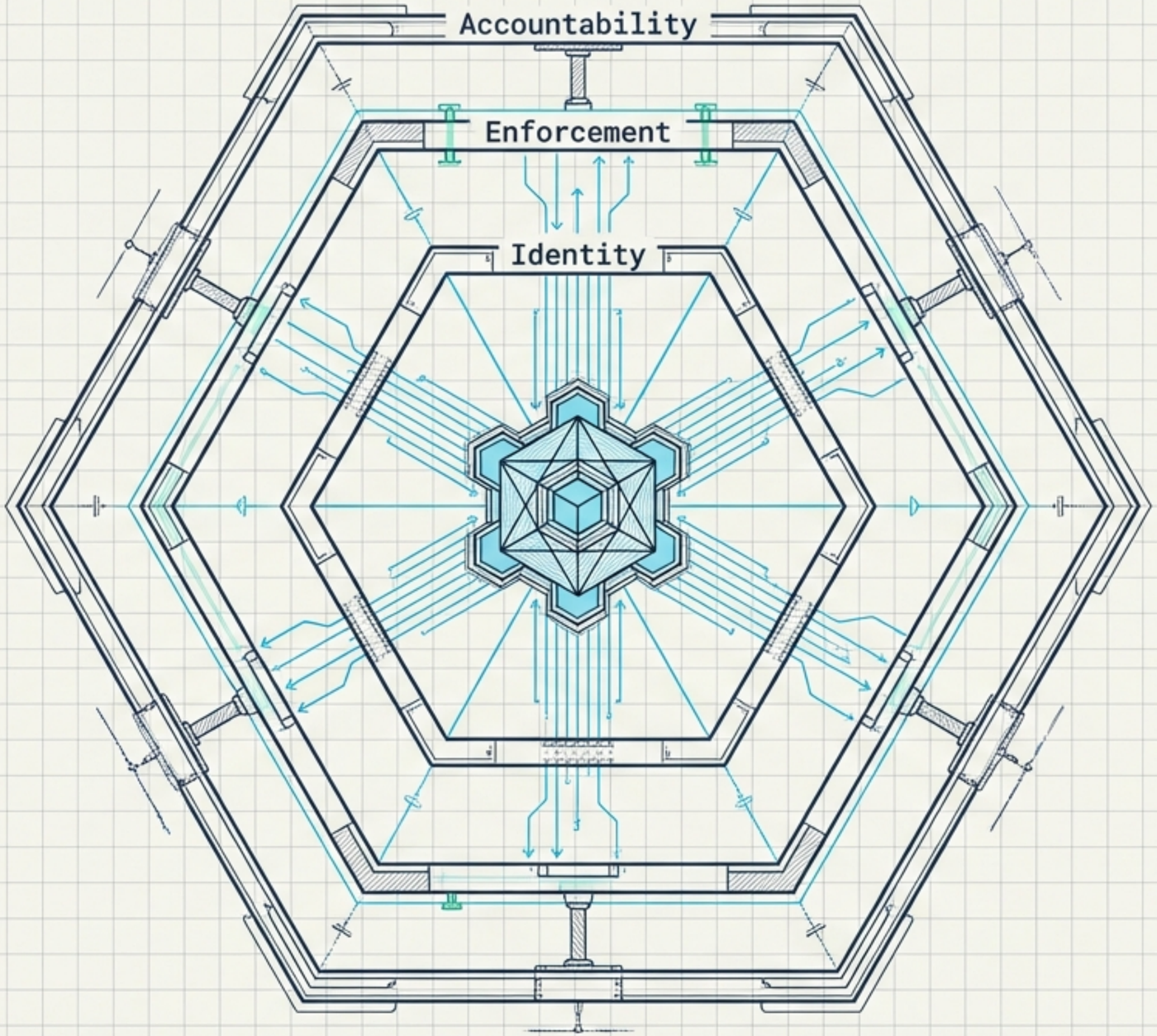
CC-401

Agent Governance Framework

Identity, Enforcement, and Accountability for Autonomous Systems

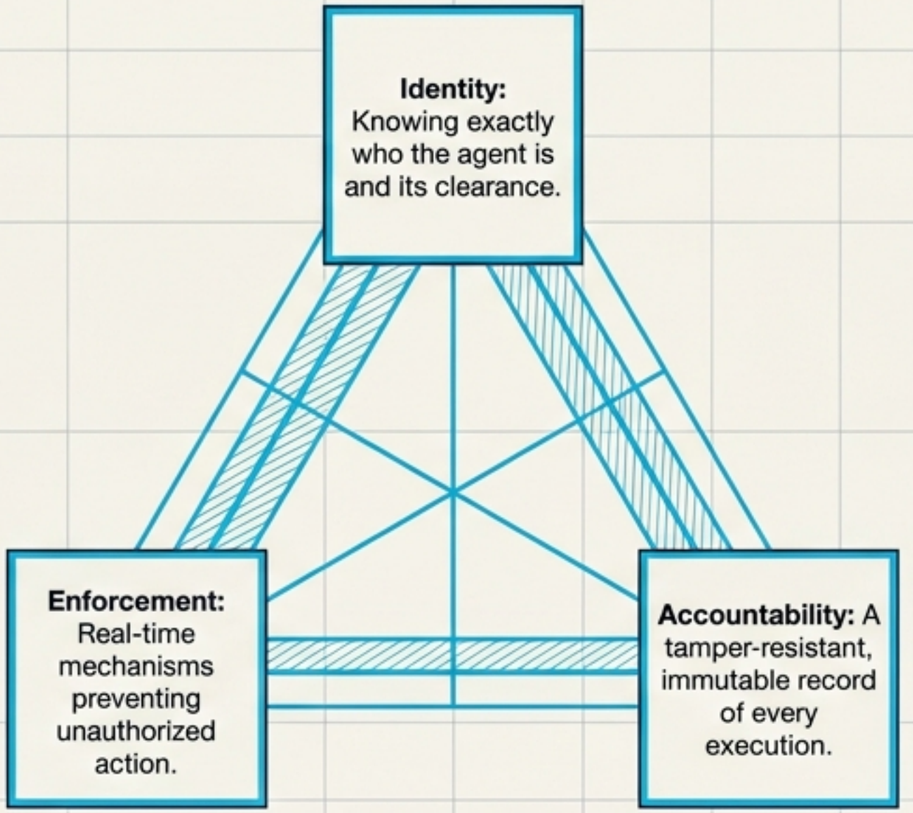


The Ungoverned Risk

```
> rm -rf .env
> git push --force
```

Without governance, tool-equipped agents have unbounded access to modify files, execute shell commands, and exfiltrate data.

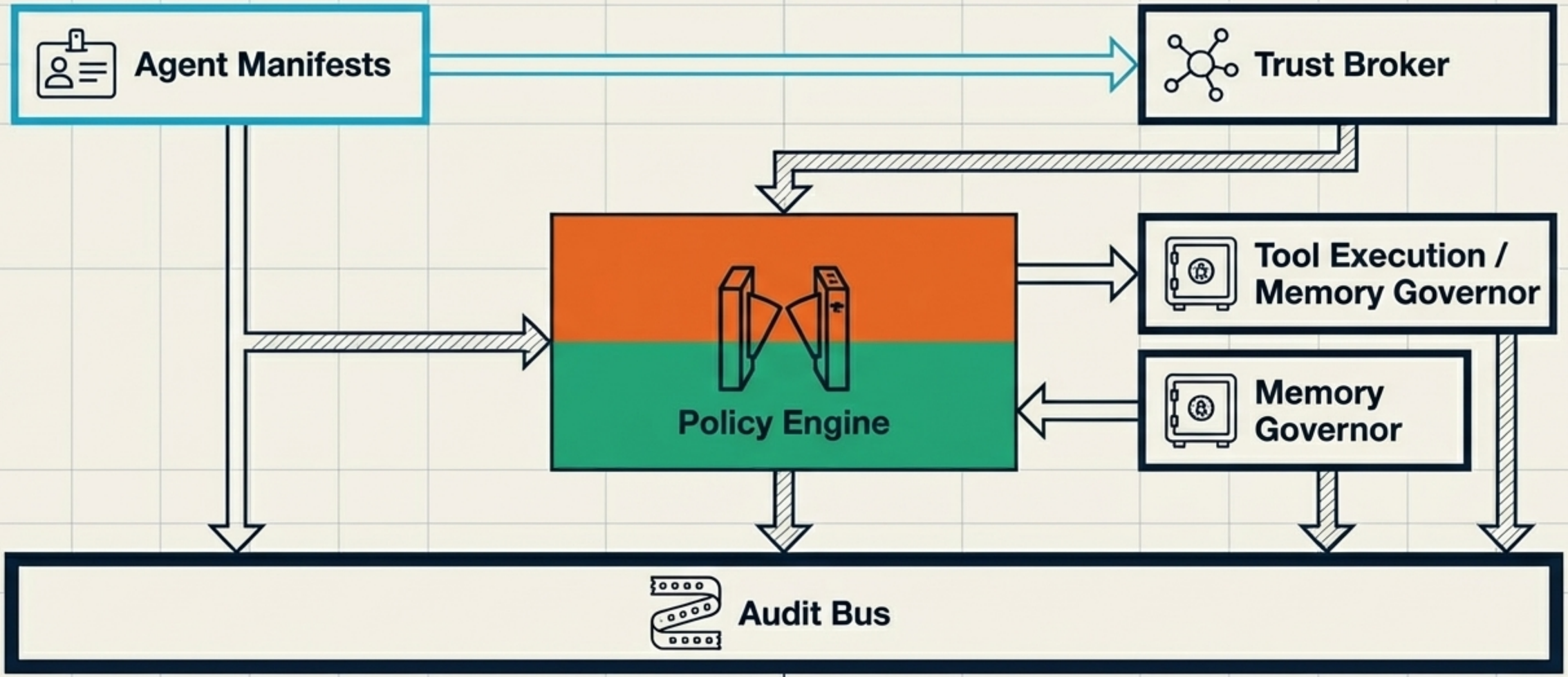
The CC-401 Solution



Governance ensures agents do only what they are authorized to do.

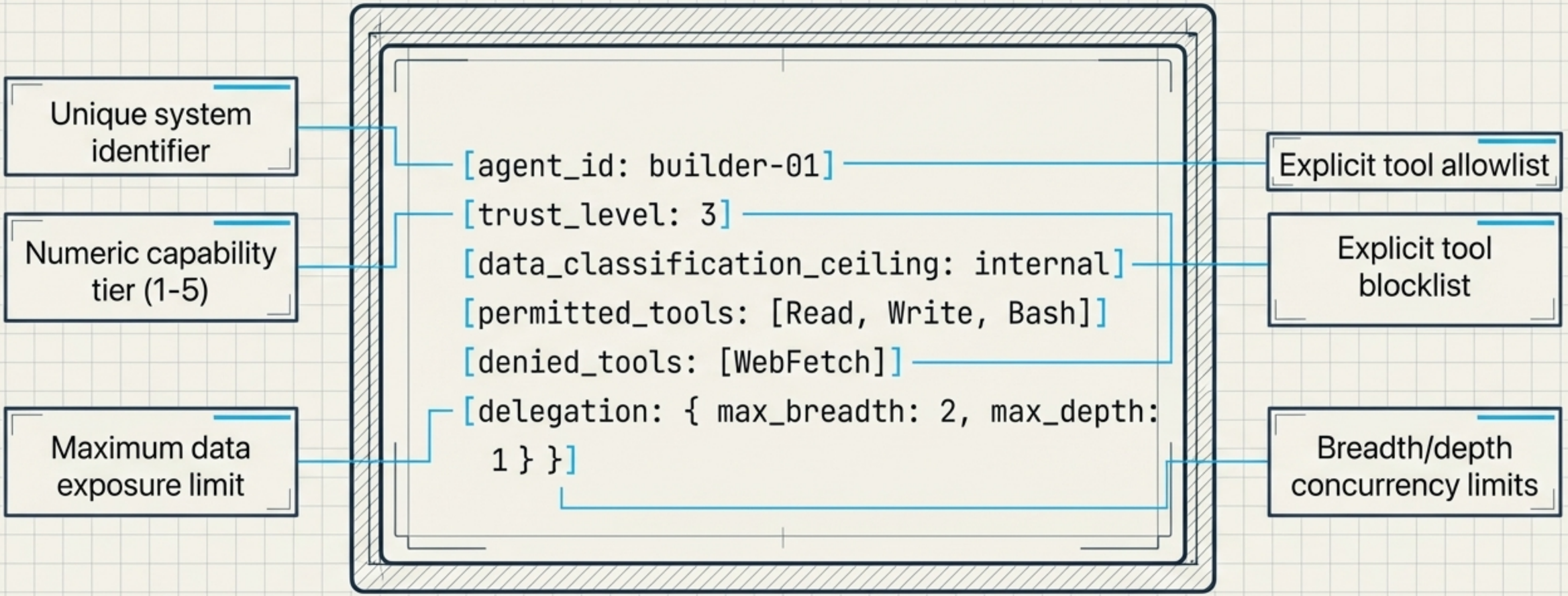
The Five Pillars of System Architecture

Every request must navigate this exact sequence. Identity is verified, boundaries are negotiated, actions are authorized, memory is sanitized, and everything is permanently recorded.



Agent Identity Manifests

Every agent requires a registered YAML identity document. It serves as both a security clearance and a job description.



The Trust Level Hierarchy

Capabilities are inherited upwards. Agents operate strictly within their numeric tier to prevent over-privileged execution.

Level 5: Orchestrator
Full system, Tier classification.

Level 4: Integrator
Network access, WebFetch.

Level 3: Builder
Unrestricted local Bash.

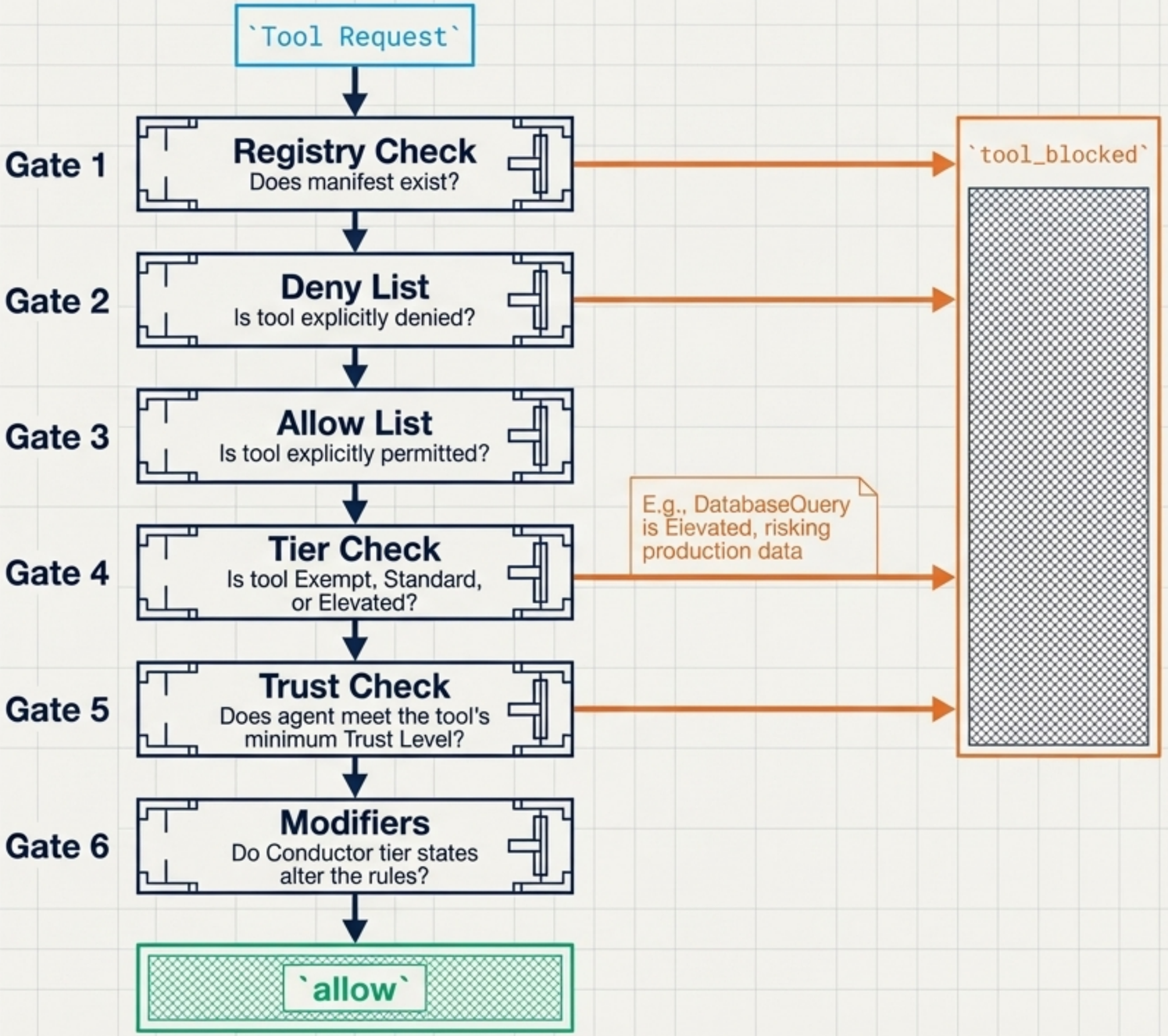
Level 2: Contributor
Write, Edit, limited Bash.

Level 1: Observer
Read, Glob, Grep. (Read-only).

Elevated tools require Level 4+. Agent receives a denial message and emits `tool_blocked` to the audit bus.

The Policy Engine Check Flow

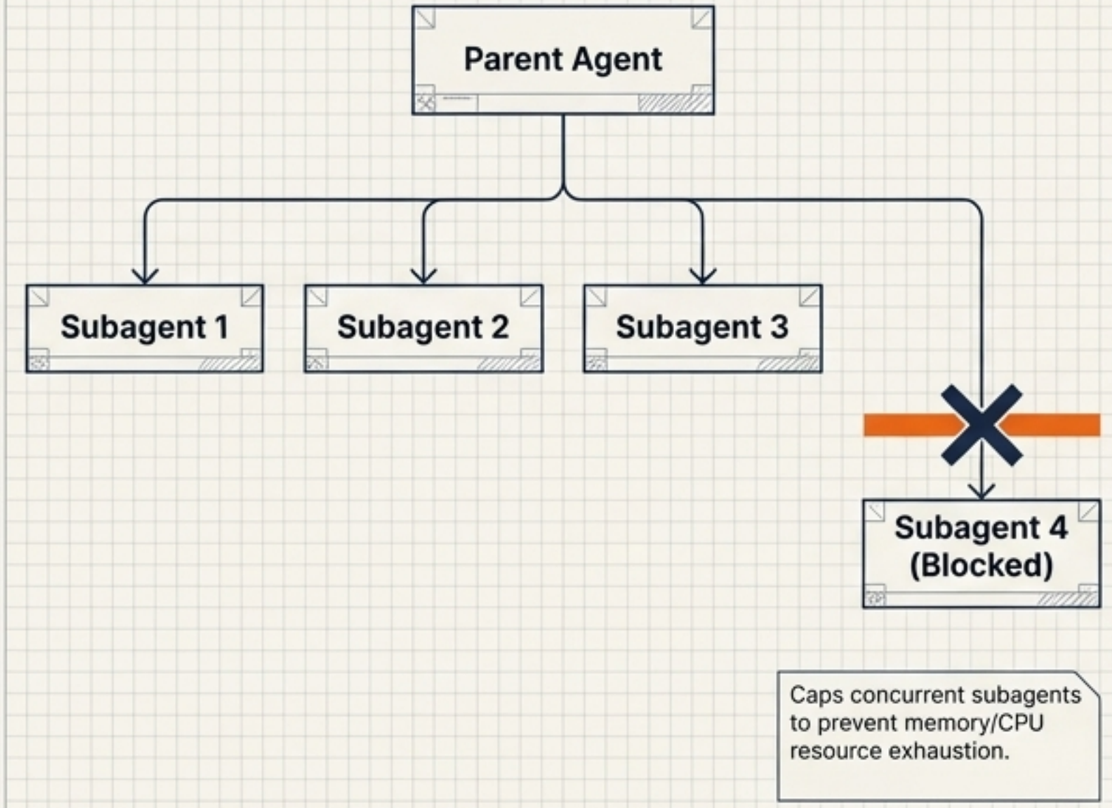
The runtime gatekeeper evaluates every tool request sequentially before execution.



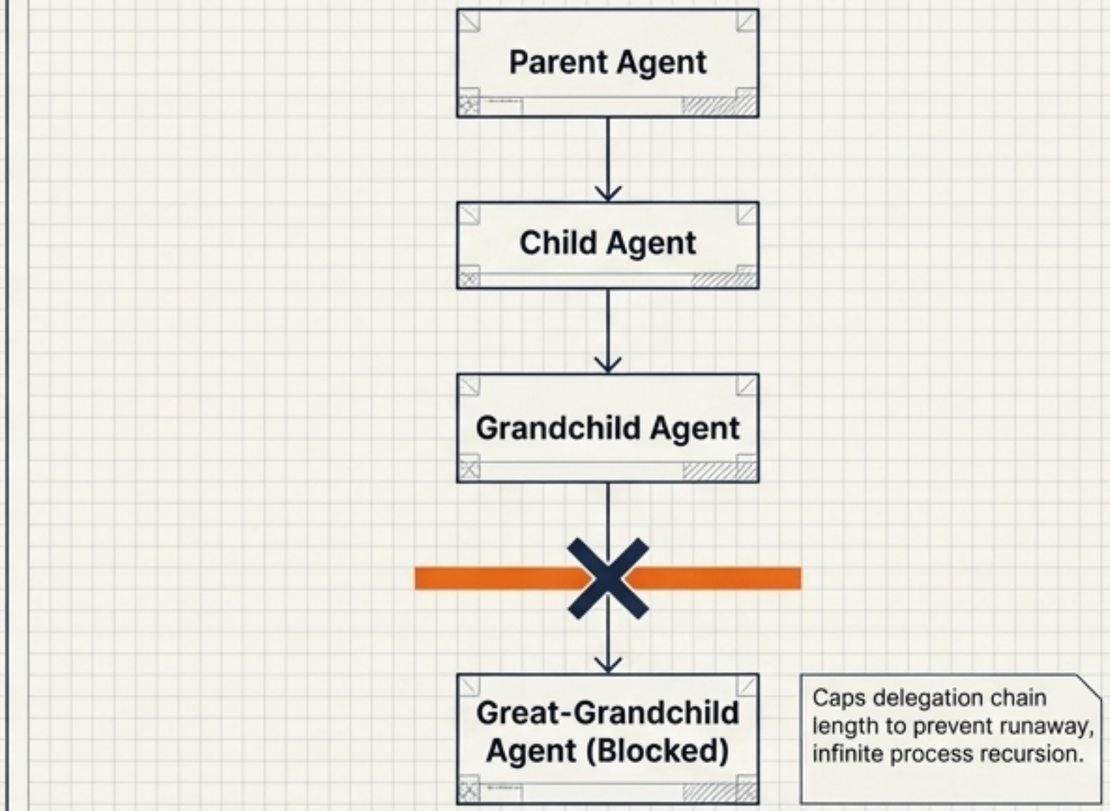
Trust Broker: Boundaries of Delegation

The Trust Broker issues **TTL** (Time-To-Live) **tokens** for **subagents** to prevent unbounded system behavior.

Horizontal Axis (Breadth Limits)

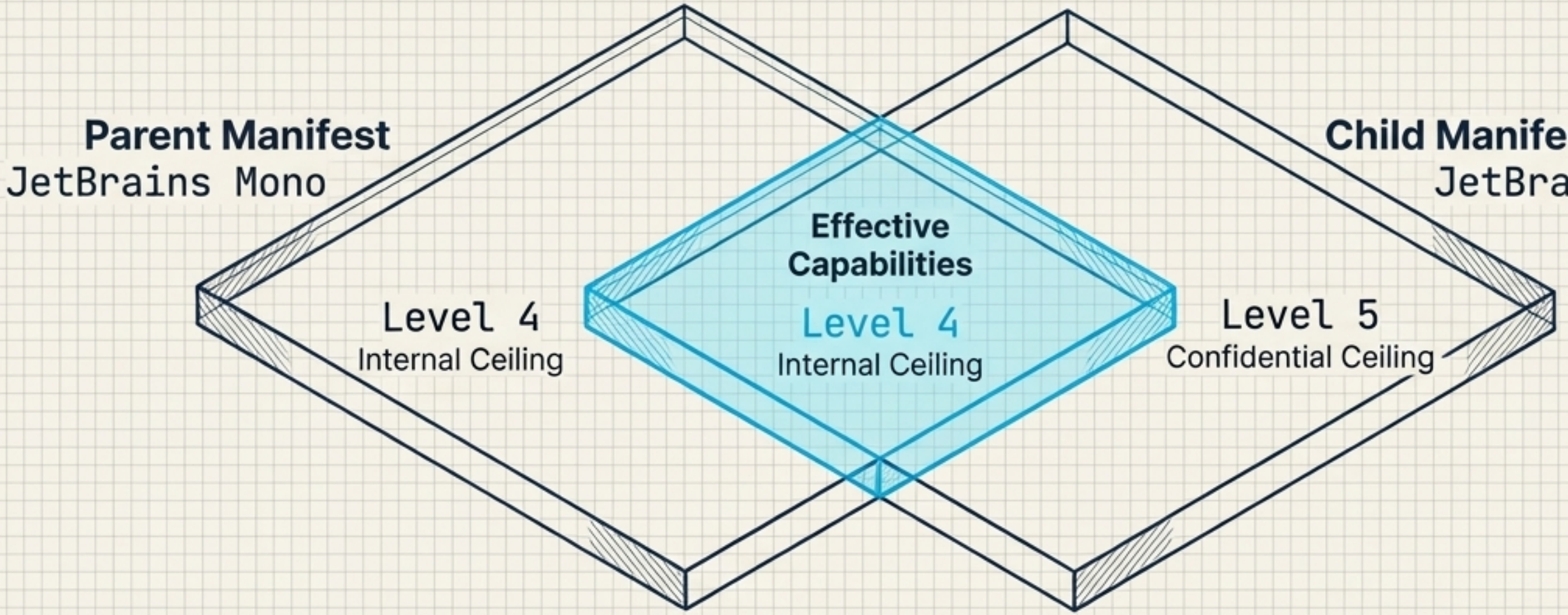


Vertical Axis (Depth Limits)



Parent Ceiling Enforcement

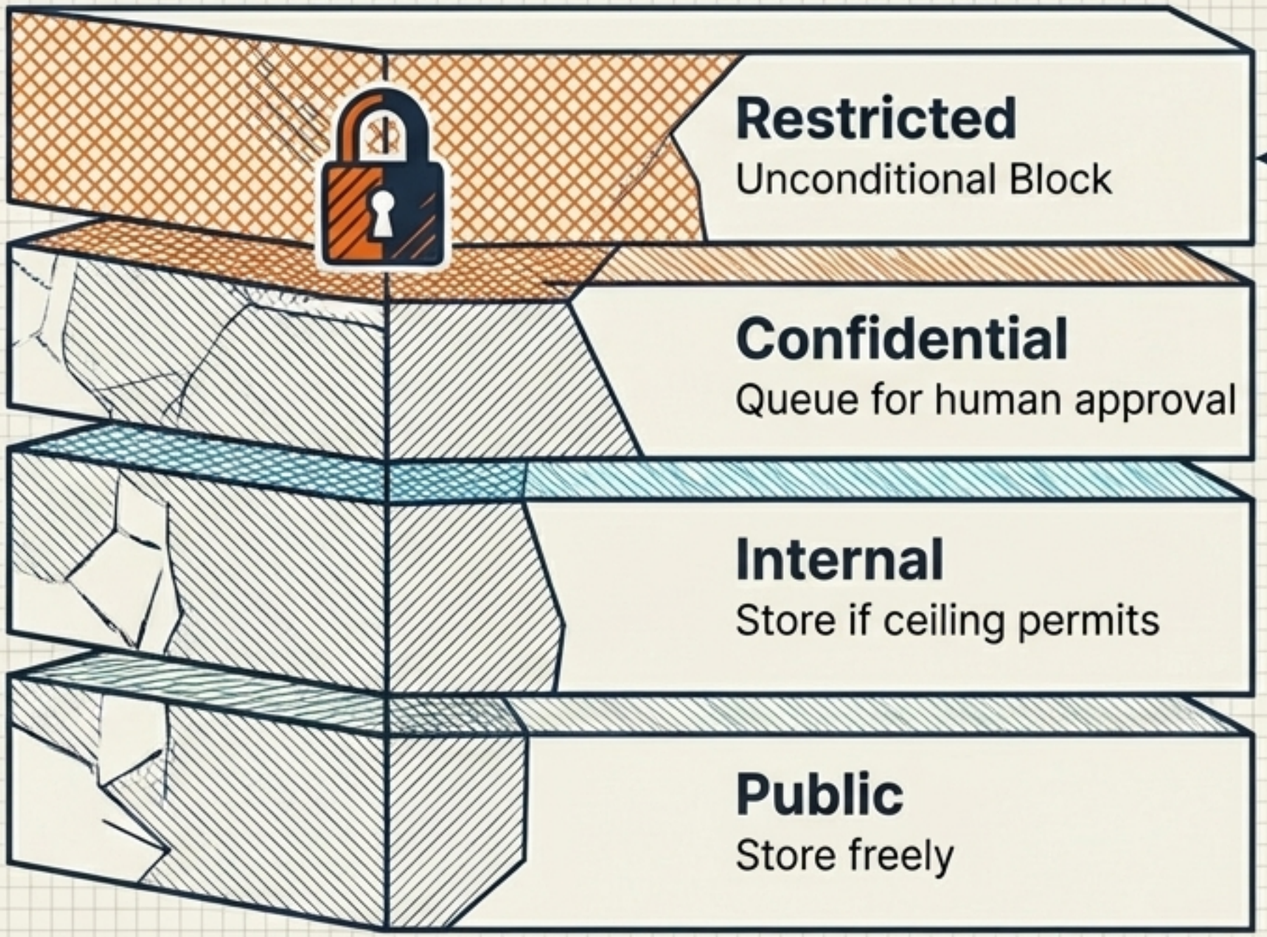
Delegation can only **narrow** permissions, never widen them. A parent cannot grant authority it does not possess.



The child's effective capabilities are the mathematical intersection of the parent's and the child's manifests.

The Data Classification Hierarchy

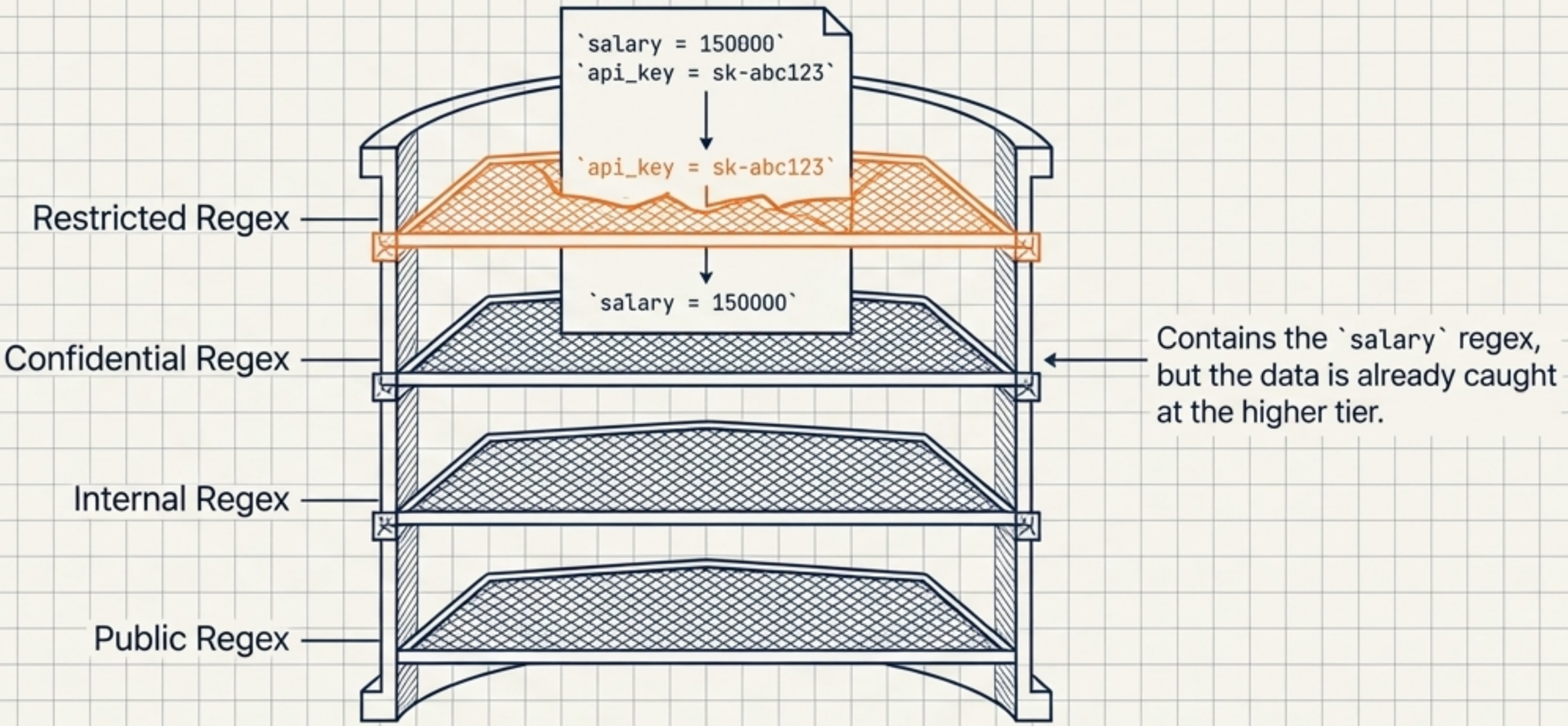
Data ceilings govern memory writes independently of an agent's numeric Trust Level.



Even a Level 5 Orchestrator is unconditionally blocked from writing **Restricted** content to memory.
The Governor denies the write and emits a `memory_blocked` event.

The Memory Governor Sieve

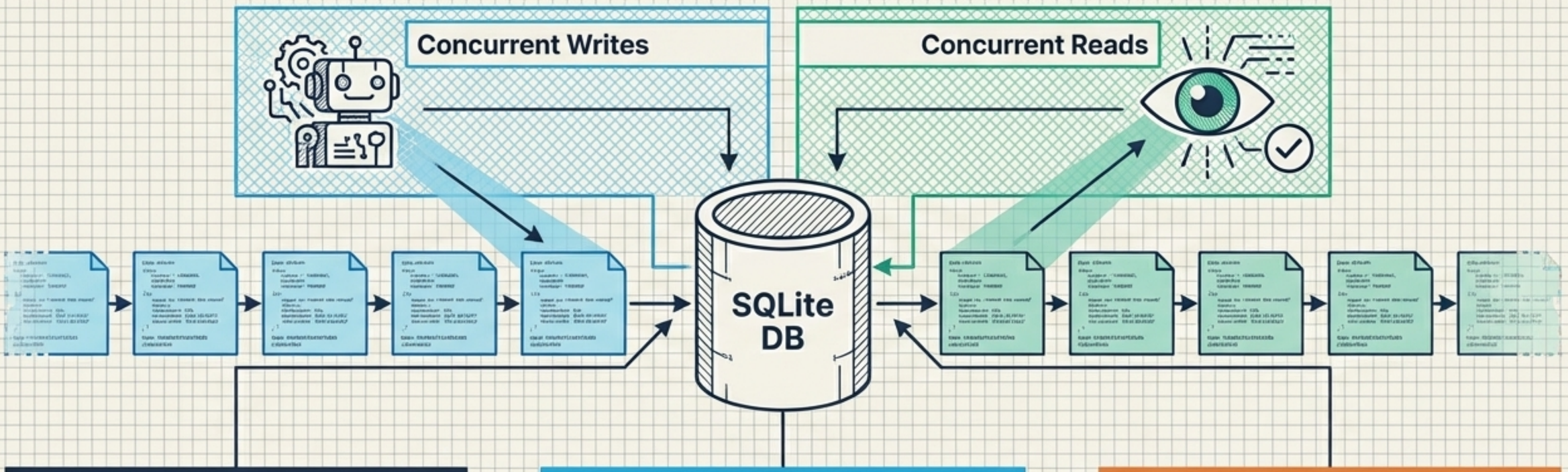
Content is classified using regex patterns. The highest match always wins, ensuring sensitive data is never downgraded by mixed context.



The engine scans from the most restrictive level down to the least restrictive.

Immutable Accountability: The Audit Bus

The governance system's infallible memory. Events are written but never updated or deleted.



WAL Mode

Write-Ahead Logging enables concurrent reads and writes without blocking active agents.

Append-Only

Guarantees tamper-resistance for compliance and post-incident forensics.

Buffer Fallback

Bounded memory queues prevent system crashes if the disk locks.

Anatomy of an Audit Event

The exact record of what happened, who did it, and why.

```
{
  "event_type": "tool_blocked",
  "reason": "tool requires trust level 4,
    agent has level 3"
}
```

Category 1	Category 2	Category 3
Lifecycle	**Execution**	**Delegation**

`agent_start`
`agent_end`
`handoff`
	`tool_use`
	`tool_blocked`
	`error`
		`delegation`
		`delegation_end`

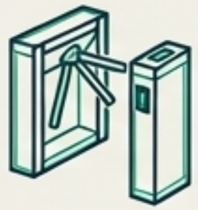
Category 3	Category 4	Category 5
Delegation	**Memory**	**Control**

`delegation`
`delegation_end`
	`memory_write`
	`memory_blocked`
		`gate_pass`
		`gate_fail`
		`escalation`
		`state_change`

System Resilience Decision Matrix

The engineering philosophy dictating when to bypass errors and when to trigger a hard stop.

Fail-Open: Graceful Degradation



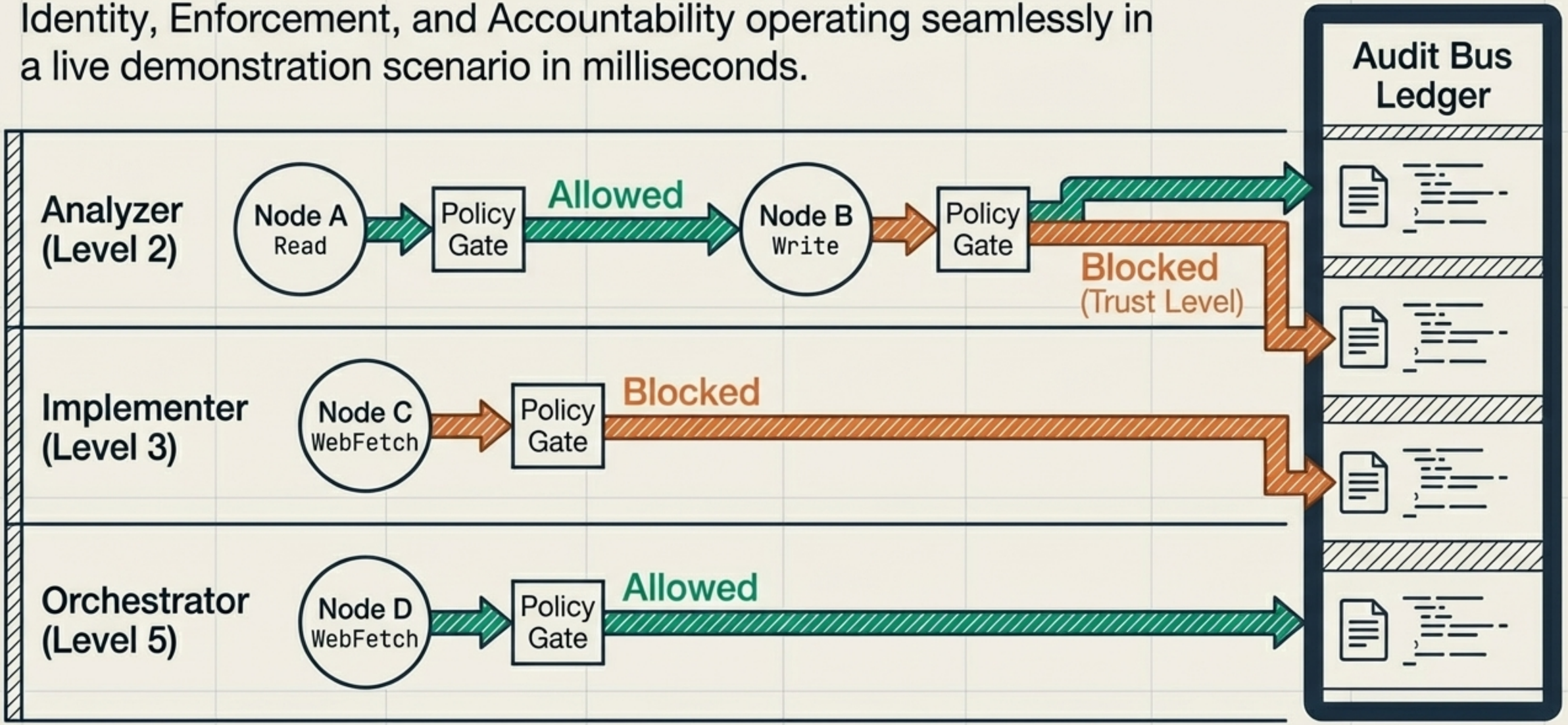
Governance plugin not installed	→	Proceed to avoid breaking existing workflows. (Non-critical value-add)
Memory deduplication unreachable	→	Store anyway to prevent data loss.

Fail-Closed: Hard Stop

Manifest unreadable	→	Block all tools; unidentified agents are a critical threat. (Security-critical gate)
Data classification error	→	Deny storage to prevent accidental restricted leaks.

Synthesis: The Governed Agent Lifecycle

Identity, Enforcement, and Accountability operating seamlessly in a live demonstration scenario in milliseconds.



CC-401 Practical Assessment Checklist

The exact deliverables required to pass the 50-point Part 2 Practical.

✓	**1. Agent Manifests**	3 valid YAMLs (Trust levels 2, 3, 5) with correct bounds.
✓	**2. Tool Tiers**	YAML classifying 10+ tools (Exempt, Standard, Elevated).
✓	**3. Policy Engine**	Implement the 6-step check flow with edge case handling.
✓	**4. Audit Bus**	SQLite WAL mode, append-only, supporting 14 event types.
✓	**5. Demonstration**	Successfully execute the 6-scenario script and query the audit trail.