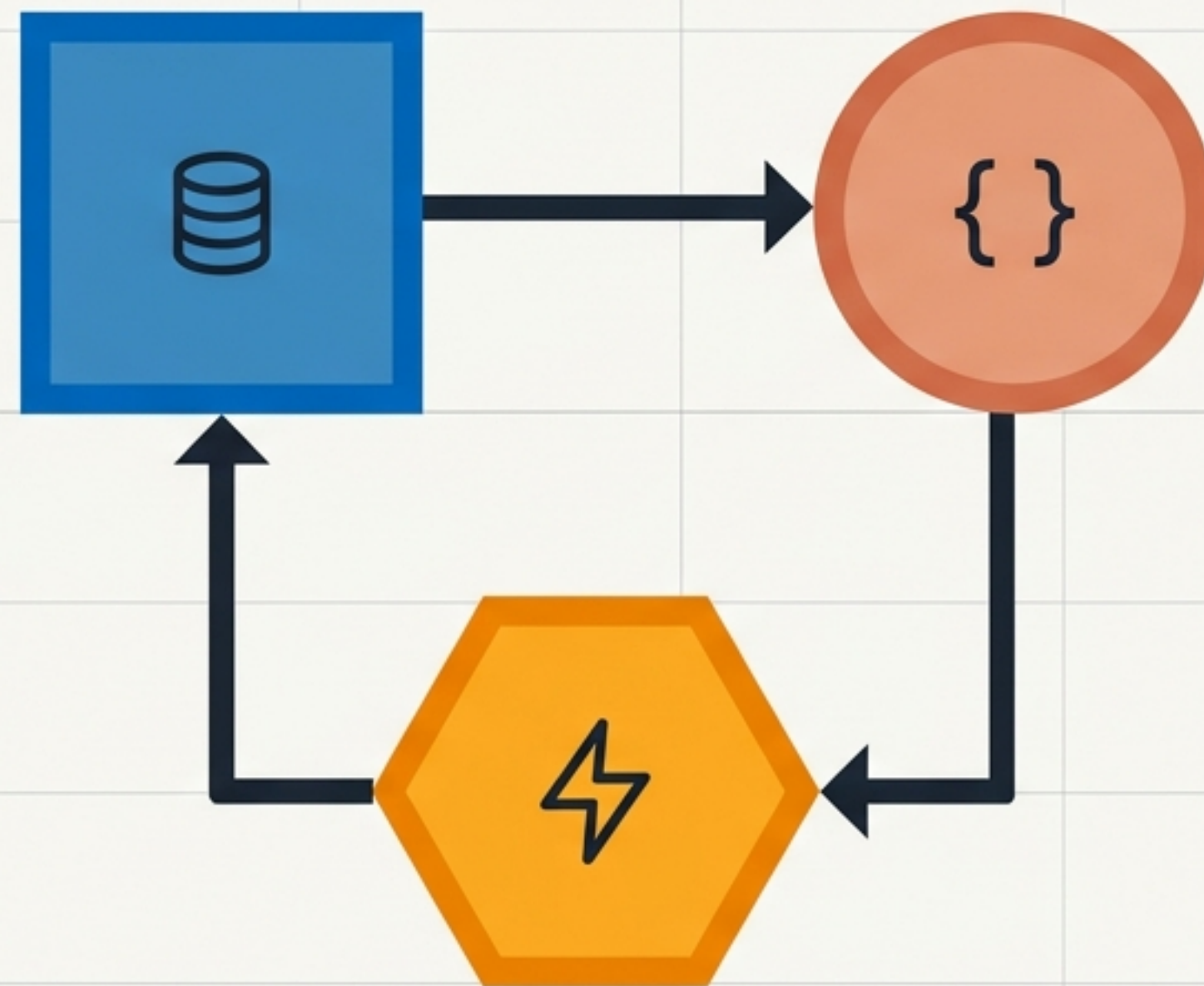


# Evaluating Infrastructure Platforms for Next-Generation AI Agents



# All three platforms share the core primitives of agentic execution



GEAP, Claude Agent SDK, and Cloudflare Agents all reached GA between 2024 and April 2026. Because they share these execution primitives, the real architectural divergence lies entirely in how they handle infrastructure: state management, deployment topology, and vendor lock-in.

# Gemini Enterprise Agent Platform (GEAP) is fully managed and GCP-native

## Architectural Player Card



### Philosophy

An opinionated platform assuming GCP integration is a feature, not a constraint.

### What you gain

- Out-of-the-box compliance (Agent Identity, Security Command Center integration)
- Managed persistence (Memory Bank + Agent Sessions)

### What you give up

- Portability. Migrating to AWS or on-prem means rebuilding the state, registry, and gateway layers from scratch.

# Claude Agent SDK is a code-first framework maximizing portability



## Philosophy

The least opinionated platform. Gives you primitives and leaves infrastructure entirely to you.

## What you gain

- Maximum infrastructure portability (run on AWS, GCP, or bare metal)
- Direct integration with Claude Opus 4.7 for highest-quality reasoning tasks

## What you give up

- Managed services. You must engineer your own long-term memory layer (vector DB + retrieval), session management, and routing.

# Cloudflare Agents execute at the global edge with co-located state



## Architectural Player Card

Philosophy Agents as Durable Objects—microservers running a TypeScript SDK at the edge.

### What you gain

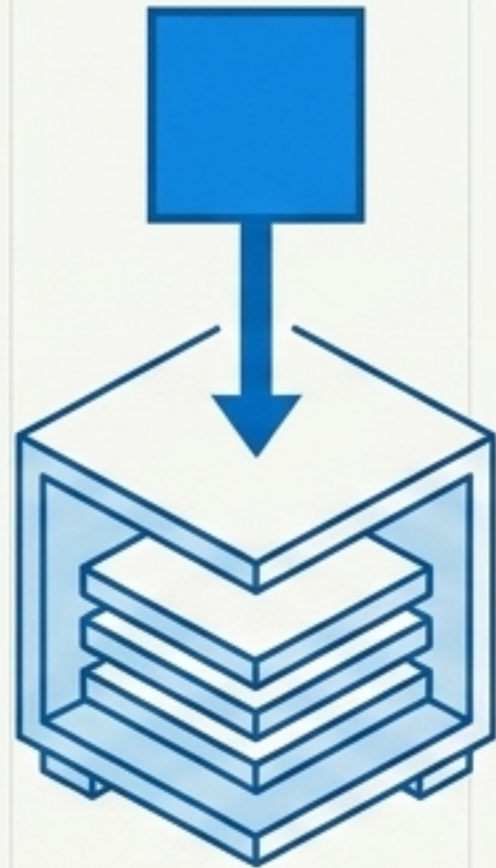
- Sub-millisecond cold starts and built-in WebSocket support for real-time latency
- Zero-round-trip state reads via built-in SQLite database per agent

### What you give up

- Ecosystem integration. No Python support, no built-in compliance pillars, and you must wire your own model provider via API.

# State management dictates data models, failure modes, and migration paths

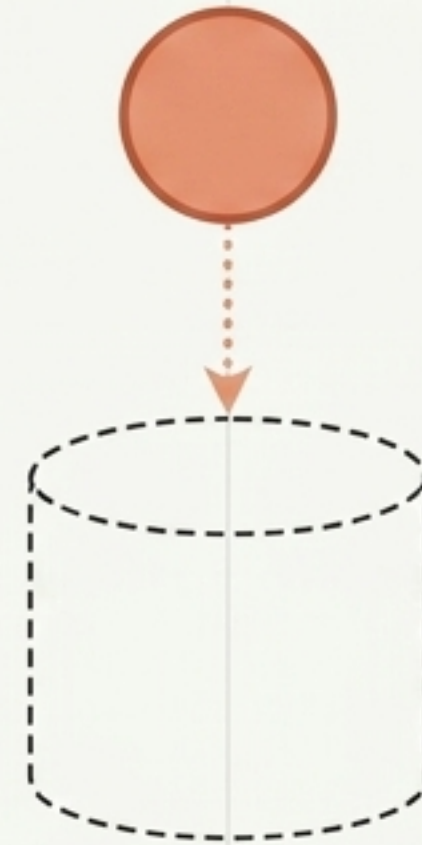
GEAP



Memory Bank

Platform distills cross-session memory. Implicit trust required.

Claude SDK



BYO Database

Complete control, complete responsibility.

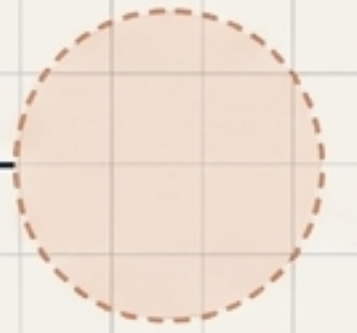
Cloudflare



State co-located with compute. Zero network round-trips.

# Deployment topologies determine baseline latency and geographic flexibility




Claude SDK: Agnostic.  
Deploy anywhere on your own infrastructure.



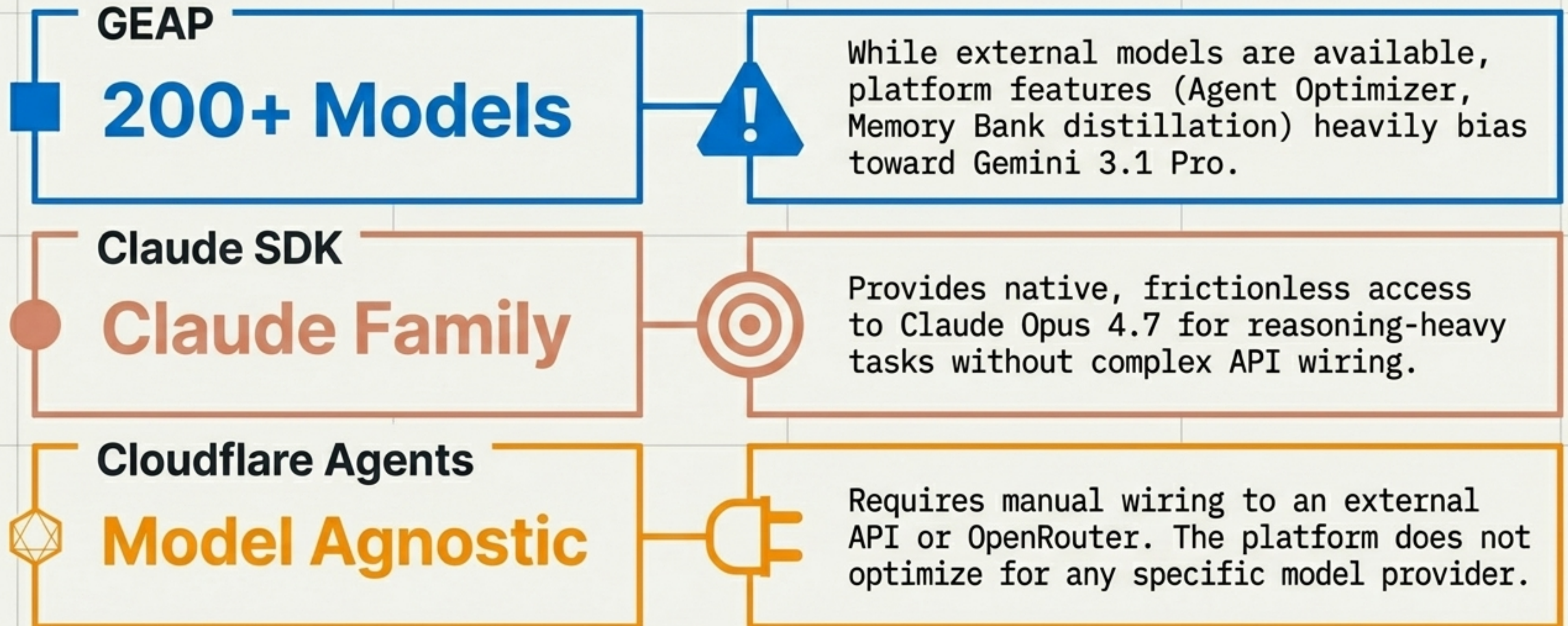
GEAP: Centralized GCP hubs (e.g., us-central1).  
Sub-second pre-warmed instances.

Cloudflare: 300+ global edge locations.  
Sub-millisecond cold starts.

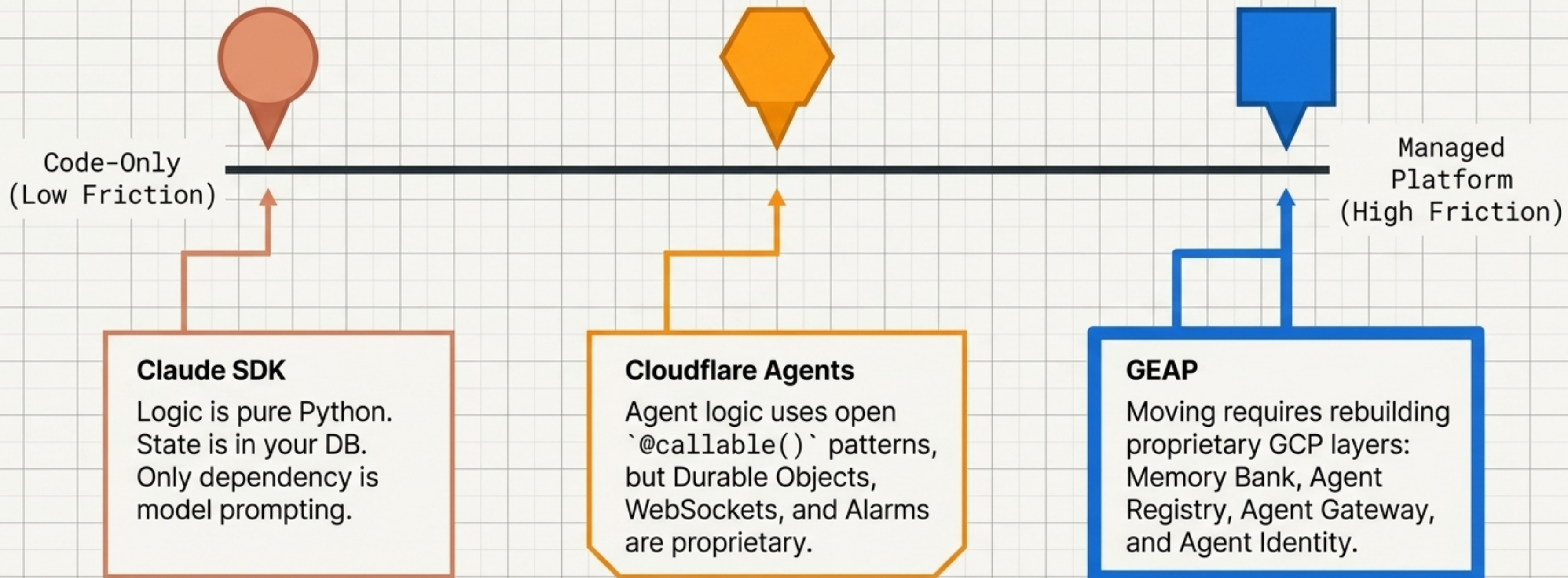
# The Infrastructure Topology Matrix

Capability	 <b>GEAP</b>	 <b>Claude SDK</b>	 <b>Cloudflare Agents</b>
Where it runs	GCP regions	Anywhere	Global Edge
Long-running	Multi-day Agent Runtime	Infra-dependent	Indefinite Durable Objects
WebSocket	Streaming API	Manual setup	Native Built-in
Scheduling	Cloud Scheduler	Manual implementation	Native Alarms
Multi-region	Explicit Config	Manual replication	Automatic Global Distribution

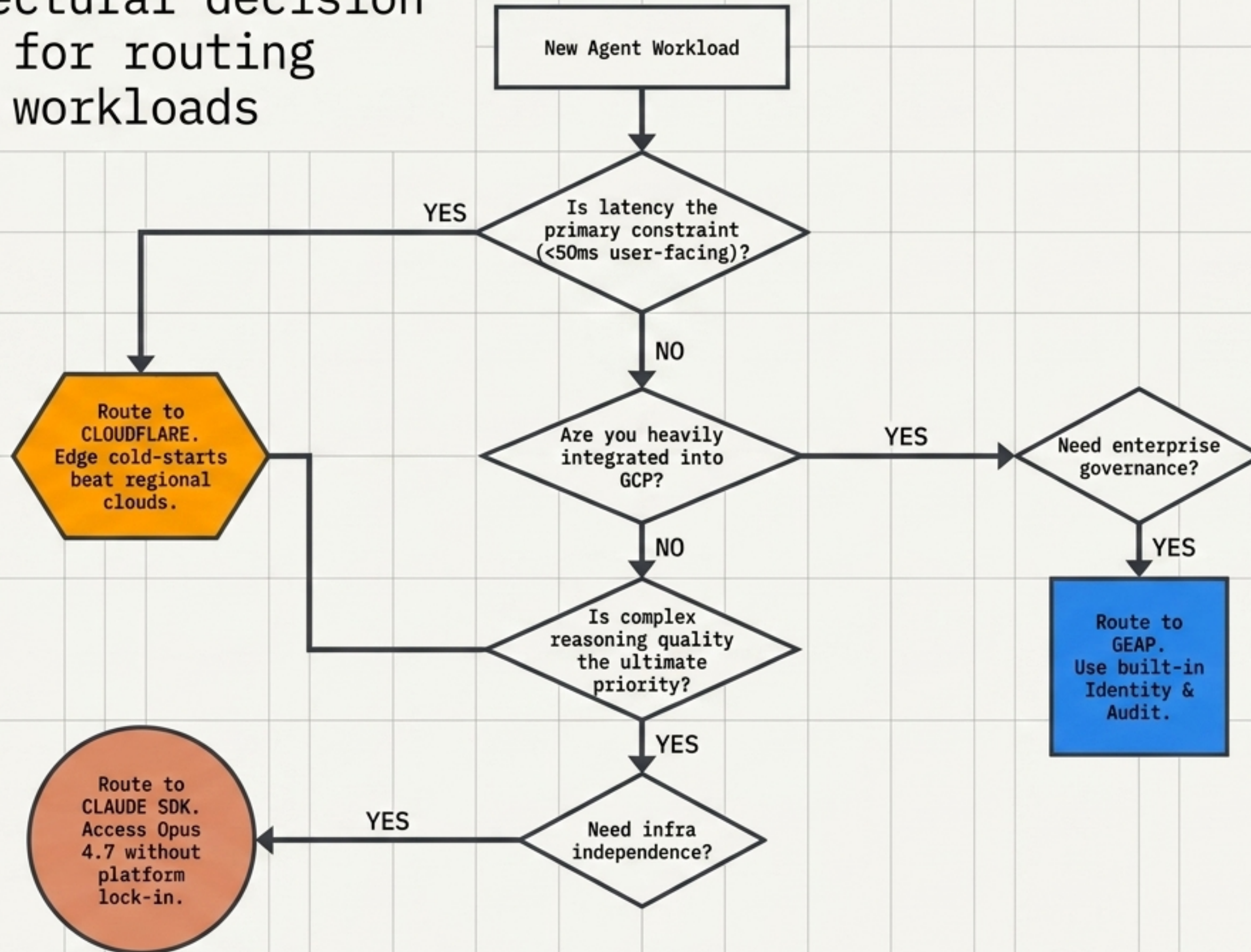
# Model selection involves a trade-off between platform optimization and provider access



# Vendor lock-in is a spectrum measured by the engineering cost of migration



# An architectural decision framework for routing new agent workloads



# Scenario: Fintech compliance monitoring demands built-in enterprise governance

## Workload Constraints

- ✓ Flag unusual transactions
- ✓ Retain 90 days alert history per user
- ✓ Produce regulator audit logs

## The Solution

### Verdict: GEAP Wins

**Agent Identity** provides cryptographic per-invocation audit trails required by regulators out-of-the-box.

**Memory Bank** handles the 90-day retention within GCP's strict regional data residency guarantees, avoiding months of custom DB engineering.

# Scenario: Interactive coding assistants demand sub-millisecond edge latency

## Workload Constraints

- Embed in VS Code extension
- <200ms budget for response time
- Per-user file context isolation
- TypeScript-native development team

## The Solution

### Verdict: Cloudflare Agents Wins

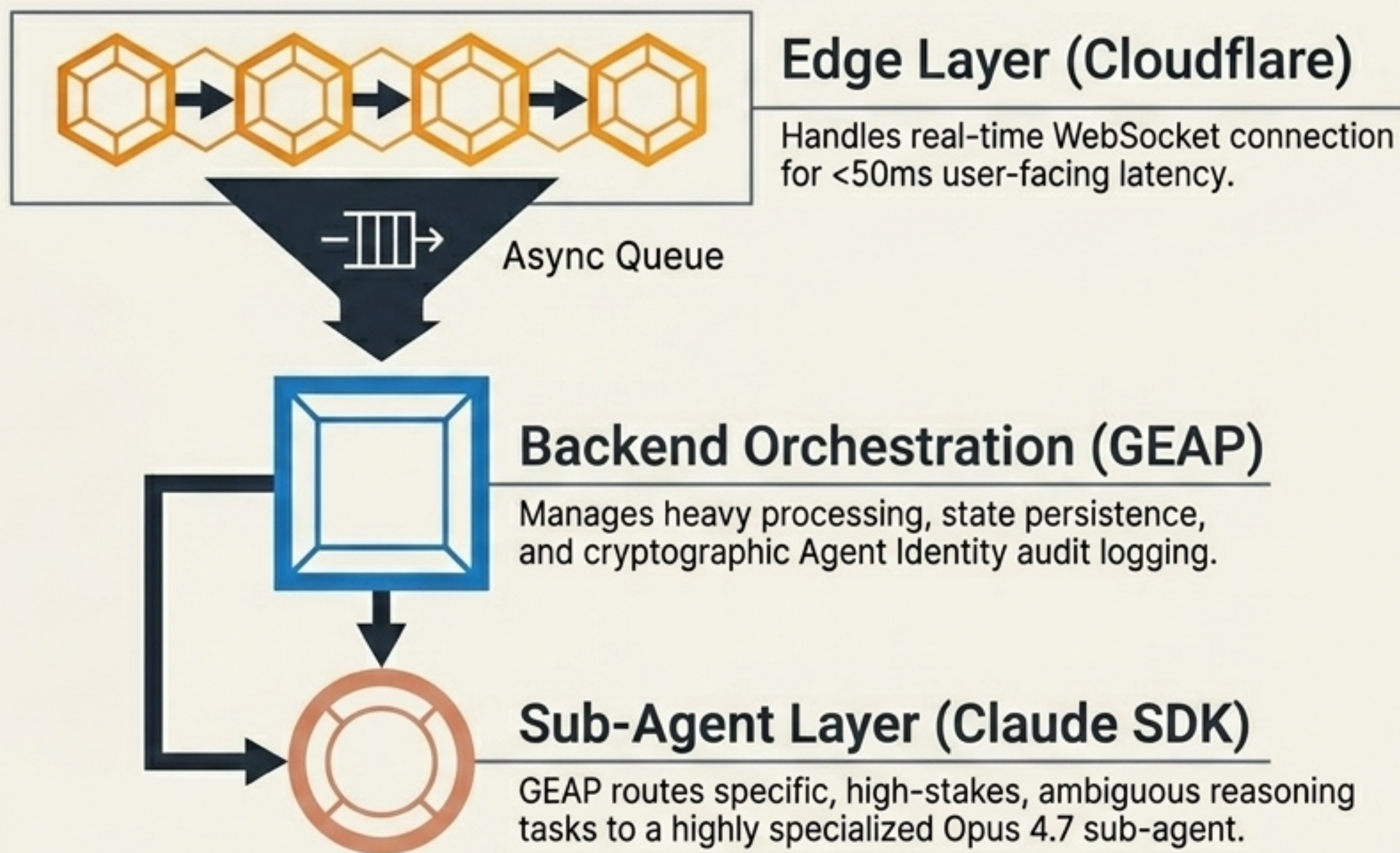
Sub-millisecond edge cold starts easily beat the strict <200ms latency budget.

Per-user file context fits natively into isolated Durable Object state via `this.setState()`. TypeScript SDK matches team skills perfectly.

# The Comparative Evaluation Matrix

	<b>GEAP</b>	<b>Claude SDK</b>	<b>Cloudflare</b>
State Model	Managed Layered SQL	BYO Database	Co-located SQLite
Topology	Regional (Sub-second)	Agnostic (Infra-dependent)	Global Edge (Sub-millisecond)
Model Access	Gemini-optimized, 200+ available	Claude-native reasoning	Model-agnostic wiring
Lock-in Risk	High (Identity, DB, Gateway)	Low (Code-only)	Medium (Durable Objects)
Best Fit	Enterprise multi- agent governance	Complex, ambiguous reasoning	Real-time, low- latency execution

# Enterprise architecture combines platform strengths rather than treating them as mutually exclusive



**ARCHITECTURE\_IS\_COMPOSITION\_NOT\_COMPETITION**