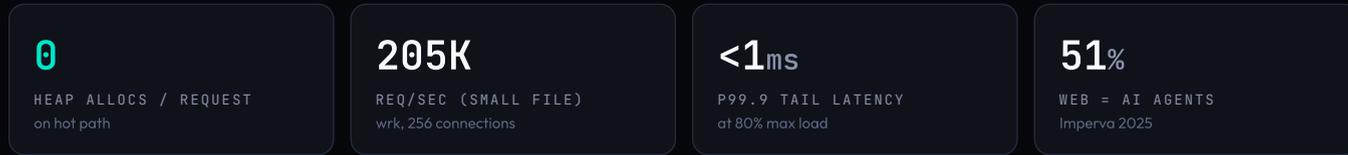# The HTTP Server Built for
the AI Agent Era

Zero-allocation web infrastructure in Rust. When every microsecond costs tokens, SynapServe delivers.

0	**205K**	**<1ms**	**51%**
HEAP ALLOCS / REQUEST	REQ/SEC (SMALL FILE)	P99.9 TAIL LATENCY	WEB = AI AGENTS
on hot path	wrk, 256 connections	at 80% max load	Imperva 2025

01   THE PROBLEM

Today's servers see AI agents as attacks.

When an AI agent receives a task, it sends thousands of HTTP requests per second — hitting APIs, scraping data, coordinating with other agents. To nginx, Apache, and modern Rust frameworks, this looks indistinguishable from a DDoS attack. Agent traffic grew 1,300% in Jan–Aug 2025, and 51% of all web traffic is now non-human.

1,300%	**39,000**	**87%**	**42%**
AI agent traffic growth	req/min per fetcher bot	agents hit product pages	heap from I/O buffers

02   THE SOLUTION

HTTP infrastructure redesigned from byte zero.

Zero-Allocation Hot Path

Every parsed request lives on the stack. Span-based parsing replaces string copies. io_uring provided buffers mean the kernel manages memory. 0 allocations per request, verified by counting allocator.

`0 alloc/req`

Agent-Native Protocols

Built-in IETF Web Bot Auth (Signature-Agent headers), MCP Streamable HTTP, Google A2A discovery, and SSE token streaming. AI agents are the primary design target, not an afterthought.

`MCP + A2A + SSE`

Predictable Tail Latency

Thread-per-core architecture with io_uring. No Arc, no Mutex, no Send bounds. Each core owns its connections, buffers, and allocator. p99.9 stays flat under load — because there's nothing to spike it.

`<1ms p99.9`
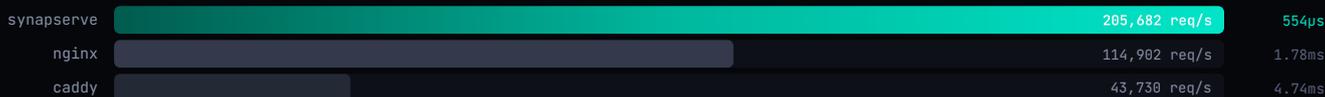
03   TECHNOLOGY

Rust	**io_uring**	**kTLS**	**Thread-per-core**	**SIMD**
Memory safety without GC	Linux 6.1+ zero-copy I/O	Kernel-mode TLS 1.3	Shared-nothing, no locks	SSE2/AVX2 header scan

04   BENCHMARKS

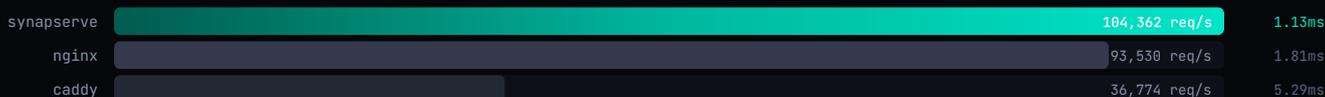STATIC FILE THROUGHPUT — wrk -t4 -c256 -d60s, 8 workers, localhost

small.json 118 bytes

synapserve	205,682 req/s	554µs
nginx	114,902 req/s	1.78ms
caddy	43,730 req/s	4.74ms

medium.json 4,005 bytes

synapserve	178,541 req/s	644µs
nginx	109,958 req/s	1.88ms
caddy	38,398 req/s	5.25ms

large.json 24,203 bytes
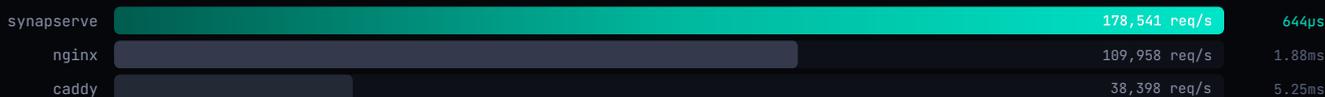
synapserve	104,362 req/s	1.13ms
nginx	93,530 req/s	1.81ms
caddy	36,774 req/s	5.29ms

Intel i7-8550U, 4C/8T, Linux 6.17.0. ETag & compression disabled. 10s warmup per server per file size. 14.5MB RSS under load.

Three tectonic shifts are converging.

AI agents are the new browsers
- ✓ 51% of web traffic is non-human
- ✓ OpenAI, Anthropic, Google all shipping agents
- ✓ Enterprise deploying autonomous workflows
- ✓ Agent traffic growing 1,300% in 8 months

Protocols are being standardized
- ✓ IETF Web Bot Auth (Signature-Agent)
- ✓ Anthropic MCP (Streamable HTTP)
- ✓ Google A2A (Agent-to-Agent)
- ✓ First-mover advantage is real

io_uring has matured
- ✓ Linux 6.1+ multishot accept
- ✓ Provided buffer rings (zero-copy recv)
- ✓ Zero-copy send (SEND_ZC)
- ✓ The OS is finally ready for this

---

06   MARKET OPPORTUNITY

The AI infrastructure market is *$200B+ by 2030.*

API Gateway for Agent Traffic
Replace nginx/Envoy at edge. Native agent auth, protocol-aware routing, load balancing.

$12B
TAM

LLM Inference Serving
Zero-copy SSE streaming for token delivery. Every byte copy costs real money at scale.

$40B
TAM

Agent-to-Agent Infrastructure
MCP/A2A coordination infrastructure. Machine-to-machine as first-class.

$150B+
TAM

---

07   BUSINESS MODEL

Open Source Core (MIT / Apache 2.0)
- ✓ HTTP/1.1 + HTTP/2 + HTTP/3 server
- ✓ Zero-alloc parser & io_uring I/O
- ✓ TLS 1.3 (kTLS), static files, reverse proxy
- ✓ Community adoption flywheel

SynapServe Cloud (SaaS)
- ✓ Managed agent gateway (edge deploy)
- ✓ Agent auth + signature verification
- ✓ Advanced analytics & enterprise SLAs
- ✓ Per-request billing: recurring revenue

Playbook: Redis → Redis Labs, Kafka → Confluent, Nginx → Nginx Inc (acquired by F5 for $670M)

---

08   ROADMAP & TRACTION

- **DONE** ● **Zero-allocation HTTP/1.1 parser** — Span-based, SIMD scanning, benchmarked at 14.6M req/sec
- **DONE** ● **io_uring I/O layer** — Multishot accept, provided buffer rings, zero-copy send, splice
- **DONE** ● **HTTP/1.1 server framework** — Radix-tree router, virtual hosts, static files, ETag/Range/Brotli
- **DONE** ● **Reverse proxy & upstream** — Keepalive pool, weighted round-robin/least-conn/IP hash, health tracking, retry, backup servers, splice relay, DNS re-resolution
- **DONE** ● **TLS 1.3 with kTLS** — rustls handshake + kernel encryption offload, parallel HTTP/HTTPS listeners, early data support
- **Q2 2026** ● **HTTP/2 + SSE streaming** for LLM token delivery
- **Q3 2026** ● **Agent protocols:** MCP, A2A, IETF Web Bot Auth
- **Q4 2026** ● **HTTP/3 (QUIC)** + SynapServe Cloud managed offering

---

09   THE ASK

Seed round to build the HTTP layer for the agent web.

STAGE
Working Product
HTTP/1.1 server, TLS 1.3 (kTLS), reverse proxy — built & benchmarked

RAISING
Seed
18 months runway: team, development, first deployments

USE OF FUNDS
Ship & Scale
HTTP/2-3, agent protocols, cloud offering, enterprise

- ● 55% Engineering
- ● 20% Cloud infra
- ● 15% Go-to-market
- ● 10% Ops & legal

invest@synapserve.io
synapserve.io | We respond to every serious inquiry within 24 hours