

{FTWS}

AutoPilot Scaling Analysis

How many clients can we realistically support? What are the costs, bottlenecks, and infrastructure requirements at each stage of growth?

Free The World Software

Internal Research — March 2026

Classification: Internal / Strategic

1. Current Infrastructure

Before planning scale, here's what we're working with today.

| MACHINE | CHIP | RAM | CURRENT ROLE | AVAILABLE? |
|--------------|------|------|---------------------|------------|
| Mac Studio | M4 | 32GB | Axon (FTWS ops) | Shared |
| Bobby's Mini | M4 | 16GB | Soul'd Out Foods | Dedicated |
| Forge Mini | M2 | 8GB | Dev sandbox | Available |
| Axon's Mini | M4 | 16GB | Marketing (pending) | Available |
| M5 MacBook | M5 | 64GB | Arriving Mar 20-24 | Available |

Cost Structure Per Client Agent

Each AutoPilot client gets a dedicated OpenClaw agent. The cost breaks down into two components:

API Token Costs

The primary cost driver. Each agent consumes API tokens from Anthropic (Claude) for every interaction.

| MODEL | INPUT/ 1M | OUTPUT/ 1M | USE CASE |
|--------|--------------|---------------|---------------------------|
| Haiku | \$1 | \$5 | Simple tasks, routing |
| Sonnet | \$3 | \$15 | Client agents (daily ops) |
| Opus | \$5 | \$25 | Strategy, complex work |

Infrastructure Costs

Hardware, electricity, and network. With our own Macs, this is nearly zero marginal cost per client.

| ITEM | MONTHLY COST |
|----------------------|-----------------|
| Mac Mini electricity | ~\$3-5 |
| Cloudflare hosting | \$0 (free tier) |
| Domain (annual/12) | ~\$1 |
| | ~\$0 |

Estimated per client: A food truck agent using Sonnet with ~50 interactions/day = **\$30-80/month** in API costs.

Internet
(shared)

Infrastructure per client:
~**\$5/month** if running on dedicated Mini.

KEY INSIGHT

At the **\$97/mo Starter** tier, margin is thin (~\$10-60/month profit per client). At the **\$497/mo Professional** tier, margin is healthy (\$350-400/month profit). At **\$1,497/mo Enterprise**, it's very profitable (\$1,300+/month). The business only works at scale with Professional+ clients, or with enough Starter clients to amortize the infrastructure.

2. Capacity by Growth

Stage

How many clients at each stage, and what it takes.

Stage 1: Bootstrap (Now — 5 Clients)

3-5

Clients supported

Using existing hardware. Bobby's Mini serves Soul'd Out Foods (client #1). Forge and Axon's Mini can each host 1-2 more client agents.

Hardware needed: None
(already owned)

API cost: \$150-400/month total

Revenue at \$97/mo avg:
\$291-485/month

Revenue at \$497/mo avg:
\$1,491-2,485/month

Margin: Breakeven at Starter,
profitable at Professional

Stage 2: Growth (6-15 Clients)

6-15

Clients supported

Need 2-3 additional Mac Minis (~\$600 each, M2 refurb). Each Mini runs 2-4 lightweight agents (Sonnet). Shared infrastructure model.

Hardware needed: 2-3 Mac
Minis (\$1,200-1,800)

API cost: \$600-1,200/month

Revenue at \$97/mo avg:
\$582-1,455/month

Revenue at \$497/mo avg:
\$2,982-7,455/month

Key unlock: Multi-tenant agent
hosting (multiple agents per
Mac)

Stage 3: Scale (16-50 Clients)

16-50

Clients supported

Requires rack-mount infrastructure or cloud hybrid. Consider Mac Mini rack, colocation, or cloud VPS for non-Apple-specific agents. Need monitoring dashboard, automated provisioning.

Hardware needed: Mac Mini cluster (8-12 units) or cloud hybrid

Capital: \$5,000-10,000

API cost: \$2,000-4,000/month

Revenue potential:
\$8,000-25,000/month

Key unlock: Automated provisioning, monitoring, billing system

Stage 4: Enterprise (50+ Clients)

50+

Clients supported

Requires dedicated operations person, SLA commitments, enterprise billing, and redundancy. Consider local inference (Exo cluster) to cut API costs by 60-80%.

Hardware: Dedicated rack or colo (\$500-1,000/mo)

Staff: 1 ops person (part-time or full-time)

API cost: \$5,000-15,000/month (or \$1,000-3,000 with local inference)

Revenue potential:
\$25,000-75,000/month

Key unlock: Local inference via Exo to crush margins

3. Unit Economics Deep Dive

Per-Client Profitability

| PLAN | MONTHLY REVENUE | API COST | INFRA COST | GROSS MARGIN | MARGIN % |
|--|-------------------|-------------|------------|---------------|----------------|
| Starter (\$97) | \$97 | \$30-80 | \$5 | \$12-62 | 12-64% |
| Professional (\$497) | \$497 | \$50-120 | \$5 | \$372-442 | 75-89% |
| Enterprise (\$1,497) | \$1,497 | \$80-200 | \$10 | \$1,287-1,407 | 86-94% |
| Food Truck Starter (\$499 one-time) | \$499 once | \$0 ongoing | \$2 | \$497 | 99% (one-time) |
| Food Truck AutoPilot (\$999 + \$97/mo) | \$97/mo recurring | \$30-80 | \$5 | \$12-62/mo | + \$999 setup |

CRITICAL FINDING

The \$97/mo tier is a loss leader at low volume. With Sonnet API costs of \$30-80/month per client, a Starter client only generates \$12-62/month in profit. You need 15-20 Starter clients to match one

Professional client. The play is: use Starter/Food Truck to build portfolio and case studies, then upsell to Professional.

Break-Even Analysis

| SCENARIO | MONTHLY FIXED COSTS | CLIENTS NEEDED |
|-------------------------------------|-------------------------------|---------------------------------|
| Solo (KJ only, existing hardware) | ~\$100 (API for Axon + Bobby) | 1 Professional client |
| +2 Mac Minis (\$1,800 amortized/12) | ~\$250 | 1 Professional or 5 Starters |
| +Ops hire (part-time \$2K/mo) | ~\$2,250 | 5 Professional clients |
| Full operation (\$5K/mo costs) | ~\$5,000 | 4 Enterprise or 11 Professional |

API Cost Reduction Strategies

1. **Use Haiku for routing + simple tasks** — 3x cheaper than Sonnet for classification, routing, simple Q&A
2. **Aggressive caching** — Anthropic prompt caching reduces repeat input costs by 90%
3. **Sonnet over Opus** — All client agents on Sonnet (\$3/\$15 vs \$5/\$25). Reserve Opus for FTWS internal only
4. **Local inference via Exo** — When M5 arrives, run 70B models locally for \$0/token. Route simple tasks to local, complex to API
5. **Batch processing** — Anthropic Batch API is 50% cheaper for non-urgent tasks (reports, content generation)
6. **Token budgets per client** — Set monthly caps, overage charged at cost + margin

4. Pros & Cons of Scaling AutoPilot

Pros

- **Near-zero marginal infrastructure cost** — Each Mac Mini hosts multiple agents. No cloud bills. Electricity is \$3-5/machine/month.
- **Software moat** — OpenClaw + custom skills + memory systems aren't easy to replicate. Competitors charge \$497 for a one-time audit. We deliver ongoing autonomous operations.
- **Recurring revenue** — Monthly subscriptions compound. 10 Professional clients = \$4,970/month predictable revenue.
- **Portfolio compounds** — Each client deployment becomes a case study. Soul'd Out Foods sells the next food truck. That sells the next restaurant. Network effects.
- **Local inference is coming** — When Exo cluster is live, API costs drop 60-80%. Every dollar of margin improvement drops straight to profit.
- **No staff required (Stage 1-2)** — Agents are self-running. KJ oversees, Axon manages. Human time per client: ~2-4 hours/month for Professional, near-zero for Starter.
- **Industry-specific packages scale** — Food Truck Package is a template. Clone it for: barber shops, cleaning services, personal trainers, real estate agents. Same tech, different skin.

Cons

- **API dependency** — Anthropic could raise prices, rate-limit, or go down. One provider outage = all clients offline. Need fallback models (OpenAI, Mistral, local).
- **Thin margins at Starter tier** — \$97/mo doesn't leave much room after API costs. A chatty client could eat the entire margin. Need usage caps.
- **Physical hardware management** — Macs need power, internet, occasional reboots. KJ's home is the data center right now. Power outage = all clients down.
- **Client support burden** — "My agent isn't responding" at 2 AM. Need monitoring and alerting before taking on paying clients.
- **No redundancy yet** — If Bobby's Mini dies, Soul'd Out Foods has no agent. Need failover capability (already prototyped with Forge).
- **Security isolation** — Each client agent must be fully isolated. One agent seeing another client's data = catastrophic. Need strict workspace separation.
- **Billing and contracts** — Need Stripe recurring billing, terms of service, SLA definitions. Admin work that doesn't generate revenue.

5. What We Need to Account

For

Before Taking Client #2

| REQUIREMENT | STATUS | EFFORT |
|---|----------------------|---------------------------------|
| Monitoring + alerting (agent health, uptime) | Not built | 2-3 days |
| Automated failover (agent moves to backup machine) | Prototyped on Forge | 1-2 days |
| Client workspace isolation (strict separation) | Proven (Bobby model) | Template ready |
| Onboarding playbook (discovery to launch) | Not documented | 1 day |
| Stripe recurring billing | Not set up | 1 day |
| Terms of Service / SLA | Not drafted | 1 day |
| Client dashboard (usage, status) | Not built | 3-5 days (optional for Stage 1) |
| UPS / power protection | None | \$100-200 purchase |

Before 10+ Clients

| REQUIREMENT | WHY | COST |
|--------------------------------|--|--------------|
| Dedicated internet line | Client agents can't share residential bandwidth at scale | \$50-100/mo |
| Mac Mini rack/shelf | Physical organization, cooling, cable management | \$100-300 |
| Automated provisioning | Spin up new client agents without manual SSH setup | 2-3 days dev |
| | | 1-2 days dev |

| | | |
|----------------------------|---|---------------------|
| Multi-model routing | Route tasks to cheapest capable model (Haiku/Sonnet/Opus) | |
| Backup system | Client memory/workspace backed up to cloud (R2/S3) | 1 day dev + ~\$5/mo |
| Legal entity | Contracts, liability protection, business insurance | \$500-1,000 |

Before 50+ Clients

- **Colocation or dedicated space** — Running 15+ Macs from home is not sustainable. Need a small colo rack (\$300-500/mo)
- **Exo cluster for local inference** — At 50 clients, API costs could be \$5,000-15,000/month. Local inference cuts this to \$1,000-3,000
- **Part-time ops hire** — Monitoring, client onboarding, troubleshooting
- **Client self-service portal** — Dashboard for clients to see agent status, usage, and interact
- **Multi-region redundancy** — Clients expect 99.9% uptime. Need at minimum 2 physical locations

6. Revenue Projections

12-Month Conservative Scenario

Assumption: 1 new client/month starting Month 2, mix of 60% Starter + 30% Professional + 10% Food Truck one-time.

| MONTH | CLIENTS | MRR | API COST | NET | CUMULATIVE NET |
|-------|----------------|---------|----------|---------|----------------|
| 1 | 1 (Soul'd Out) | \$0 | \$50 | -\$50 | -\$50 |
| 2 | 2 | \$97 | \$100 | -\$3 | -\$53 |
| 3 | 3 | \$594 | \$150 | \$444 | \$391 |
| 4 | 4 | \$691 | \$200 | \$491 | \$882 |
| 5 | 5 | \$788 | \$250 | \$538 | \$1,420 |
| 6 | 7 | \$1,382 | \$350 | \$1,032 | \$2,452 |
| 9 | 10 | \$2,270 | \$500 | \$1,770 | \$6,762 |
| 12 | 14 | \$3,458 | \$700 | \$2,758 | \$14,036 |

Note: Includes one-time Food Truck setup fees (\$499-999) not shown in MRR but added to cumulative net.

Aggressive Scenario (2 clients/month, 50% Professional mix)

| MONTH | CLIENTS | MRR | NET/MONTH | CUMULATIVE |
|-------|---------|---------|-----------|------------|
| 6 | 13 | \$4,159 | \$3,509 | \$12,800 |
| 12 | 25 | \$8,225 | \$6,975 | \$52,000 |

BOTTOM LINE

Realistic Year 1: 10-14 clients, \$2,500-3,500 MRR, \$14,000-20,000 net profit. **Aggressive Year 1:** 25 clients, \$8,000+ MRR, \$50,000+ net. Both scenarios assume zero additional hardware purchases after the initial Mac Mini investments already made.

7.

Recommendations

DO NOW

1. **Close the food truck prospect** as client #2 using the proposal we built today
2. **Build monitoring** — simple health check script that pings each agent and alerts via FTWS Command
3. **Document the onboarding flow** — what we did for Soul'd Out Foods, templated
4. **Set up Stripe recurring** for \$97/mo and \$497/mo plans

DO WHEN M5 ARRIVES

1. **Set up Exo cluster** — M5 + Mac Studio over TB5 for local inference
2. **Route client agent tasks to local models** where possible (simple queries, content gen)
3. **Benchmark cost savings** — measure actual API spend reduction

DO AT 5 CLIENTS

1. **Buy 2 refurb Mac Minis** (\$600 each) for dedicated client hosting
2. **Build automated provisioning** — script that sets up a new client agent in minutes
3. **Create industry templates** — Food Truck, Barber Shop, Cleaning Service, Personal Trainer

4. Draft Terms of Service and basic SLA (99% uptime guarantee)

DO NOT

- **Don't scale hardware before revenue** — existing machines support 5 clients
- **Don't over-invest in Starter tier** — push clients to Professional where margins exist
- **Don't skip monitoring** — one client agent going dark for 48 hours kills trust
- **Don't share agent contexts** — strict isolation between every client, every time
- **Don't ignore local inference** — Exo cluster is the margin multiplier that makes this a real business

{FTWS} Free The World Software

AutoPilot Scaling Analysis — Internal Research

March 2026 — Version 1.0