

OpenClaw at Scale, Compute Bottlenecks, and the Claude Code ‘HERMES.md’ Trap

Coding Agents Alpha Tracker

2026-04-27

OpenClaw at Scale, Compute Bottlenecks, and the Claude Code HERMES.md Trap

By Coding Agents Alpha Tracker • April 27, 2026

The day’s strongest signal is operational: practitioners are getting real coding-agent throughput, then running into compute, test infrastructure, and billing edge cases. This brief covers OpenClaw’s latest scale numbers, a Codex MVP workflow, DeepSeek V4 inside Claude Cowork, and Simon Willison’s timeless safety rules.

TOP SIGNAL

The clearest signal today: coding-agent leverage is showing up as raw throughput—more backlog cleared and more serious work compressed into a single day [1, 2]. Peter Steinberger says OpenClaw agents `clawsweeper` and `clownfish` closed **10k+ issues** and nearly **5k PRs** this week, **27k issues / 30k PRs** since December, while NandoDF says an OpenAI tool for building causal agents let him do in **1 day** what used to take **~2 weeks**, with **compute** as the only bottleneck [1, 2].

the world is transitioning to a compute-powered economy [3]

TOOLS & MODELS

- **OpenAI causal-agent tool** — NandoDF’s hands-on take: strong task decomposition, *beautiful code*, an OK paper draft with guidance, and a workflow compressed from about **2 weeks** to **1 day**; the bottleneck was compute [2].
- **Codex** — danizeres used it to clean a rough Godot codebase, turn sketches into UI with image generation, and get to a playable MVP in hours; Brockman’s summary was simple: Codex empowers anyone to build [4, 5].

- **Claude Code / Max 20x** — in one reported case, uppercase `HERMES.md` in recent git history corresponded with usage being routed onto API-rate billing despite subscription usage remaining; the affected user says Anthropic support acknowledged an ‘authentication routing issue’ but refused a refund [6].
 - Reported non-triggers: `AGENTS.md`, `README.md`, `HERMES` without `.md`, and lowercase `hermes.md` [6].
- **DeepSeek V4 in Claude Cowork (CC)** — Jason Zhou says the new release already runs inside CC with the full desktop experience, no Claude subscription, and about **90% lower cost** with comparable performance; setup is in the thread [7, 8].

WORKFLOWS & TRICKS

- **Split coding from test execution when local CPU is the bottleneck.**
 1. Keep the edit loop local.
 2. Offload test runs to remote compute.
 3. Steinberger says Codex can spin up **32vCPU** instances on Blacksmith Testbox and rip through the suite [9].
 4. Docs: Blacksmith Testbox overview [9].
- **Codex MVP loop for side projects.**
 1. Start with sketches plus a rough codebase.
 2. Use Codex to clean the codebase.
 3. Generate UI from the sketches.
 4. Stop once it is playable and get people using it. In the reported case, the group played the MVP for **2 hours** the same night [4].
- **Claude Code billing sanity check.** If you’re on Max 20x, search recent commits for uppercase `HERMES.md` before a long session. In the reported case, that string—not `AGENTS.md`, not `README.md`, not lowercase `hermes.md`—was the trigger the user eventually isolated by manually binary-searching repos and commits [6].
- **Two timeless agent-safety rules from Simon Willison.**
 - Don’t run agents anywhere they might access production environment credentials [10].
 - Keep tested backups independent from the production host [10].
- **Pattern to watch:** across both NandoDF’s causal-agent work and Steinberger’s test loop, the next bottleneck was compute, not basic agent competence [2, 9, 3].

PEOPLE TO WATCH

- **Peter Steinberger** — high-signal because he keeps posting real operating data, not demos: today’s useful notes were OpenClaw’s closure counts and a concrete remote-test workaround [1, 9].
- **Simon Willison** — worth tracking whenever agent incidents start getting over-interpreted; his response reduced the story to two durable controls: credential isolation and backup isolation [10].
- **Theo / om_patel5** — operator-grade bug reporting. The `HERMES.md` thread is the kind of edge-case detail that can save real money if you use Claude Code heavily [6].
- **NandoDF** — useful because he gave a workload comparison, not a benchmark take: causal-agent work that shrank from about two weeks to one day [2].

WATCH & LISTEN

- **No qualifying YouTube/podcast clip in today’s provided sources.** Best two-minute detour: Simon Willison’s short post on the agent incident, plus the linked incident article for context [10, 11].

PROJECTS & REPOS

- **OpenClaw stack (clawsweeper, clownfish)** — the clearest adoption signal in today’s notes is output: **10k+ issues** and nearly **5k PRs** closed this week, **27k issues / 30k PRs** since December [1].
- **Blacksmith Testbox** — infra worth tracking for agent-heavy repos: remote **32vCPU** test execution when local machines are CPU-constrained [9].

Editorial take: the edge is shifting from prompt cleverness to throughput engineering—more compute where it matters, tighter control over hidden context, and much safer operating boundaries for agents. [2, 3, 9, 6, 10]

Sources

1. X post by @steipete
2. X post by @NandoDF
3. X post by @gdb
4. X post by @danizeres
5. X post by @gdb
6. X post by @om_patel5
7. X post by @aibuilderclub__
8. X post by @jasonzhou1993
9. X post by @steipete
10. X post by @simonw

11. X post by @lifeof_jer