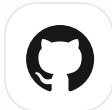




ONE ENDPOINT. NEVER STOP BUILDING.

Claude Code, free & non-stop.

OmniRoute is a free, open-source (MIT) AI gateway that runs on your own computer. You point Claude Code at it once, and it routes every request to free AI models from many providers. When one provider's free limit runs out, it automatically switches to the next one, so your coding session keeps going.

**diegosouzapw / OmniRoute**

github.com/diegosouzapw/OmniRoute · MIT license

[Open repo →](#)**150+**

providers with a free tier

~1.6Bfree tokens / month
(estimate across providers)**15–95%**

fewer tokens with built-in compression

\$0

to start. Runs on your own machine



Token numbers are the project's own estimates of what providers' free tiers give. You connect your own free accounts; each provider sets its own limits.

One address in the middle



Claude Code
your coding agent

OmniRoute

localhost:20128



1 • You ask

Claude Code sends the request to OmniRoute instead of a paid API.

2 • It picks a model

With model **auto**, OmniRoute chooses a healthy provider you connected.

3 • Limit hit? Next one

If a free quota runs out, it falls back to the next provider automatically.

BEFORE YOU START

What you need



Node.js (LTS)

Download the LTS version from nodejs.org.
Check it with `node -v`



A terminal

Terminal on Mac/Linux, PowerShell on Windows.



Claude Code installed

The `cclaude` command should already work in your terminal.



1+ free provider account

e.g. Google AI Studio (Gemini), NVIDIA, OpenRouter. More accounts = more free tokens.

If something doesn't work

- **command not found** → close and reopen the terminal after installing, then try again.
- **No reply from the model** → a fresh install has no providers. Connect at least one in Dashboard → Providers.
- **Key / auth error** → use the key you created in Dashboard → Endpoints, not a provider key.
- **Claude Code can't connect** → make sure the omniroute server is running first.

Set it up in 8 easy steps

1 Install OmniRoute

One command installs it globally:

```
$ npm install -g omniroute
```

2 Run the guided setup

Sets your dashboard password and helps you add the first provider:

```
$ omniroute setup
```

3 Start the server

Keep this terminal window open while you code:

```
$ omniroute
```

4 Open the dashboard

Go to <http://localhost:20128> in your browser and log in.

5 Connect free providers

Dashboard → **Providers** → connect at least one (login with OAuth or paste its free API key). See </dashboard/free-tiers> to pick good free ones.

6 Create your OmniRoute key

Dashboard → **Endpoints** → create an API key. This is the key your tools use.

7 (Optional) Set a fallback chain

Dashboard → **Combos** → order your providers. If one runs out, the next takes over.

8 Connect Claude Code

This writes a ready Claude Code profile for you. Follow what it prints, then start Claude Code as usual:

```
$ omniroute setup-claude
```

Quick test (optional)

Paste your key from step 6. If you get a reply, everything works.

```
curl http://localhost:20128/v1/chat/completions \
-H "Authorization: Bearer YOUR_OMNIROUTE_KEY" \
-H "Content-Type: application/json" \
-d '{"model": "auto", "messages": [{"role": "user", "content": "Hello!"}]}'
```

Works with your favourite tools



Claude Code

omniroute setup-claude



GitHub Copilot

Base URL + key



Codex

omniroute setup-codex



Cursor

omniroute setup-cursor



Cline (VS Code)

omniroute setup-cline



OpenCode · Antigravity

Base URL + key

Any other app? Use Base URL `http://localhost:20128/v1`, your OmniRoute key, and model `auto`.

Best uses

- **Code without stopping.** Long Claude Code sessions keep running even when one free quota ends.
- **Learning & side projects.** Students and beginners can build real projects with \$0 spend.
- **Build your own AI apps.** Any app that speaks the OpenAI format can use one endpoint for 100+ models.
- **Compare models.** Try Gemini, Qwen, DeepSeek and more from one place, without changing code.

Good to know

- **Bring your own accounts.** Free tokens come from providers' free tiers, not from OmniRoute.
- **Limits vary.** Each provider has its own limits and terms; read them.
- **Keep keys private.** Never share your OmniRoute or provider keys.
- **Server must run.** Start omniroute before you open Claude Code.



raul_agentic

AI agents · SaaS · Apps

[instagram.com/raul_agentic](https://www.instagram.com/raul_agentic)

Follow

Tap the card to open Instagram · Tap the repo card on page 1 to open GitHub