

decolua / 9router Public

<> Code Issues 130 Pull requests 51 Discussions Actions Projects Security and quality

master 1 Branch 18 Tags Go to file Go to file <> Code

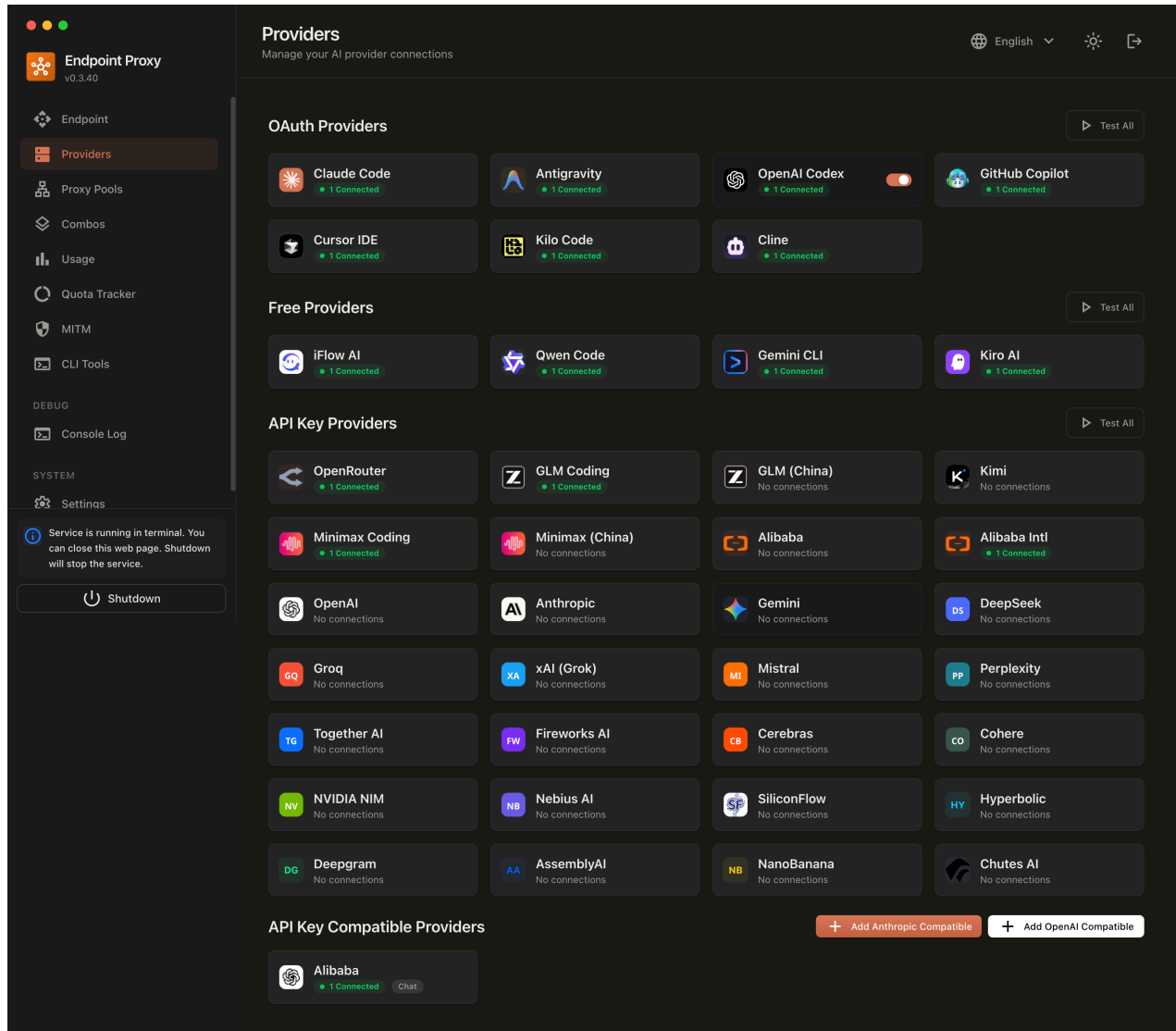
tuanaiseo fix(security)(app): unauthenticated server shutdown endpoint enables d (...)

1f3d3a8 · yesterday

.github	Fix : MITM	last month
.vscode	fix: resolve SonarQube findings and Ne...	3 months ago
cloud	Refactor config	last month
docs	feat: add pause/resume functionality for...	2 months ago
i18n	docs: add Japanese README (#385)	2 weeks ago
images	Fix : Add custom to model selector	last month
open-sse	fix: merge consecutive userInputMessa...	yesterday
public	feat: add GitLab Duo and CodeBuddy s...	last week
scripts	chore: remove translation script and wo...	last month
src	fix(security)(app): unauthenticated serv...	yesterday
tester/translator	Feat Kiro OAuth, Fix Codex	3 months ago
tests	feat(claude-code): spoof TLS fingerprin...	last week
.dockerignore	feat(docker): add Docker setup, enviro...	2 months ago
.env.example	feat(observability): add toggle for reque...	2 months ago
.gitignore	feat(claude-code): spoof TLS fingerprin...	last week
.gitmodules	Feat : Kiro MITM	3 weeks ago
.npmignore	Initial commit	3 months ago
CHANGELOG.md	feat: add API key visibility toggle in End...	2 months ago
Dockerfile	[fix] fix mitm for docker and enhance do...	2 weeks ago
LICENSE	docs: add MIT LICENSE file	last month
README.md	docs: add Japanese README (#385)	2 weeks ago
eslint.config.mjs	Initial commit	3 months ago
jsconfig.json	Update jsconfig.json and package.json ...	3 months ago
next.config.mjs	Fix AG	last week

package.json	Fix bug	2 days ago
postcss.config.mjs	Initial commit	3 months ago
start.sh	[fix] fix mitm for docker and enhance do...	2 weeks ago

[README](#)
[MIT license](#)
☰



9Router - Free AI Router

Never stop coding. Auto-route to FREE & cheap AI models with smart fallback.

Connect All AI Code Tools (Claude Code, Cursor, Antigravity, Copilot, Codex, Gemini, OpenCode, Cline, OpenClaw...) to 40+ AI Providers & 100+ Models.

npm v0.3.83 downloads 30k/month license MIT

[Quick Start](#) • [Features](#) • [Setup](#) • [Website](#)

[Tiếng Việt](#) • [中文](#) • [日本語](#)

🤔 Why 9Router?

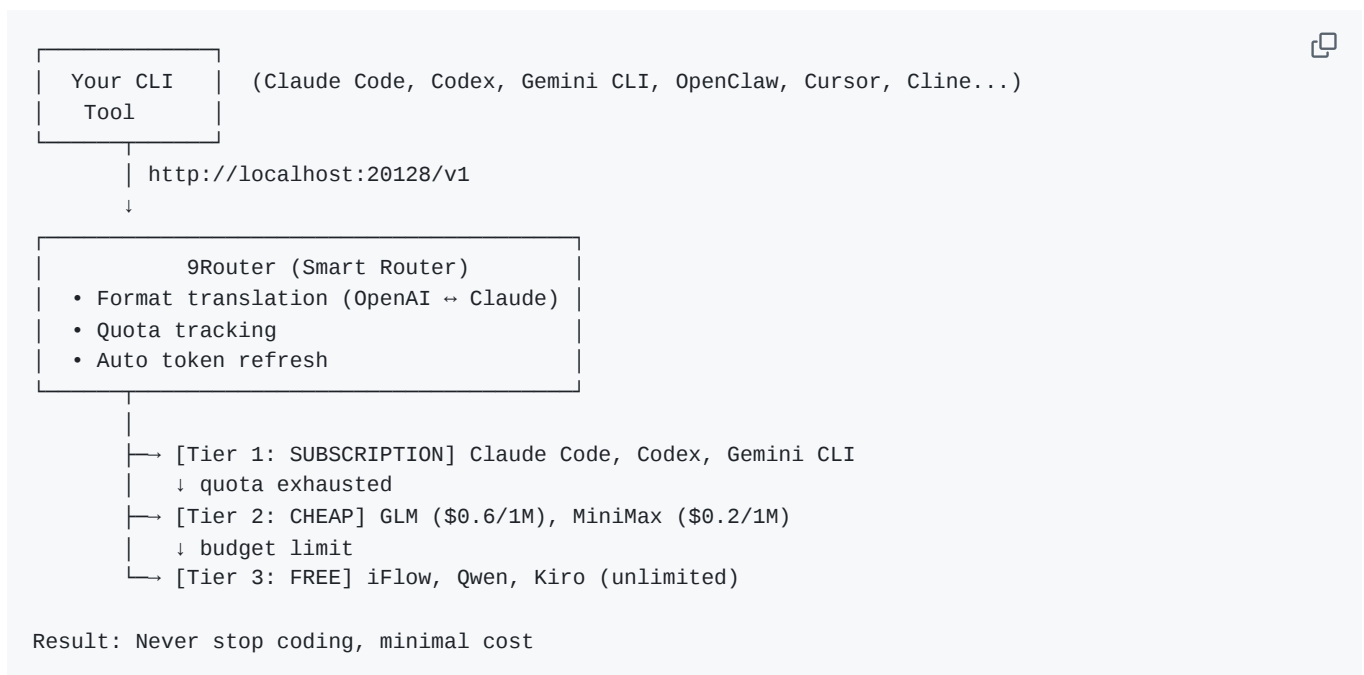
Stop wasting money and hitting limits:

- ❌ Subscription quota expires unused every month
- ❌ Rate limits stop you mid-coding
- ❌ Expensive APIs (\$20-50/month per provider)
- ❌ Manual switching between providers

9Router solves this:

- ✅ **Maximize subscriptions** - Track quota, use every bit before reset
- ✅ **Auto fallback** - Subscription → Cheap → Free, zero downtime
- ✅ **Multi-account** - Round-robin between accounts per provider
- ✅ **Universal** - Works with Claude Code, Codex, Gemini CLI, Cursor, Cline, any CLI tool

🔄 How It Works



⚡ Quick Start

1. Install globally:

```
npm install -g 9router
9router
```

🎨 Dashboard opens at `http://localhost:20128`

2. Connect a FREE provider (no signup needed):

Dashboard → Providers → Connect **Claude Code** or **Antigravity** → OAuth login → Done!

3. Use in your CLI tool:

```
Claude Code/Codex/Gemini CLI/OpenClaw/Cursor/Cline Settings:  
Endpoint: http://localhost:20128/v1  
API Key: [copy from dashboard]  
Model: if/kimi-k2-thinking
```



That's it! Start coding with FREE AI models.

Alternative: run from source (this repository):

This repository package is private (`9router-app`), so source/Docker execution is the expected local development path.

```
cp .env.example .env  
npm install  
PORT=20128 NEXT_PUBLIC_BASE_URL=http://localhost:20128 npm run dev
```



Production mode:

```
npm run build  
PORT=20128 HOSTNAME=0.0.0.0 NEXT_PUBLIC_BASE_URL=http://localhost:20128 npm run start
```



Default URLs:

- Dashboard: `http://localhost:20128/dashboard`
- OpenAI-compatible API: `http://localhost:20128/v1`

Video Tutorial

 **Complete Setup Guide - 9Router + Claude Code FREE**

NEW 2026 🔥 **WHAT • WHY • HOW • COMPLETE GUIDE • STEP-BY-STEP** 🔥 **15 MIN**

9ROUTER + CLAUDE CODE

FREE

COMPLETE TUTORIAL

HAMID

3 SECTIONS

Sonnet 4.5 FREE

Haiku 4.5 FREE

UNLIMITED

BEGINNER OK

✔️ **ZERO COST • UNLIMITED • 15 MIN TUTORIAL**

🎬 Watch the complete step-by-step tutorial:













- ✔️ 9Router installation & setup
- ✔️ FREE Claude Sonnet 4.5 configuration
- ✔️ Claude Code integration
- ✔️ Live coding demonstration

🕒 **Duration:** 20 minutes | 👤 **By:** Developer Community

[▶ Watch on YouTube](#)




🔧 Supported CLI Tools

9Router works seamlessly with all major AI coding tools:





 Claude-Code	 OpenClaw	 Codex	 OpenCode	 Cursor	 Antigravity
 Cline	 Continue	 Droid	 Roo	 Copilot	 Kilo Code

Supported Providers



















OAuth Providers

 Claude-Code	 Antigravity	 Codex	 GitHub	 Cursor
--	--	--	--	---

Free Providers








 iFlow AI 8+ models • Unlimited	 Qwen Code 3+ models • Unlimited	 Gemini CLI 180K/month FREE	 Kiro AI Claude • Unlimited
--	---	--	--

API Key Providers (40+)

 OpenRouter	 GLM	 Kimi	 MiniMax	 OpenAI	 Anthropic
 Gemini	 DeepSeek	 Groq	 xAI	 Mistral	 Perplexity
 Together AI	 Fireworks	 Cerebras	 Cohere	 NVIDIA	 SiliconFlow

...and 20+ more providers including Nebius, Chutes, Hyperbolic, and custom OpenAI/Anthropic compatible endpoints

Key Features

Feature	What It Does	Why It Matters
 Smart 3-Tier Fallback	Auto-route: Subscription → Cheap → Free	Never stop coding, zero downtime
 Real-Time Quota Tracking	Live token count + reset countdown	Maximize subscription value
 Format Translation	OpenAI ↔ Claude ↔ Gemini seamless	Works with any CLI tool
 Multi-Account Support	Multiple accounts per provider	Load balancing + redundancy
 Auto Token Refresh	OAuth tokens refresh automatically	No manual re-login needed
 Custom Combos	Create unlimited model combinations	Tailor fallback to your needs
 Request Logging	Debug mode with full request/response logs	Troubleshoot issues easily

Feature	What It Does	Why It Matters
Cloud Sync	Sync config across devices	Same setup everywhere
Usage Analytics	Track tokens, cost, trends over time	Optimize spending
Deploy Anywhere	Localhost, VPS, Docker, Cloudflare Workers	Flexible deployment options

▶ **Feature Details**

Pricing at a Glance

Tier	Provider	Cost	Quota Reset	Best For
SUBSCRIPTION	Claude Code (Pro)	\$20/mo	5h + weekly	Already subscribed
	Codex (Plus/Pro)	\$20-200/mo	5h + weekly	OpenAI users
	Gemini CLI	FREE	180K/mo + 1K/day	Everyone!
	GitHub Copilot	\$10-19/mo	Monthly	GitHub users
CHEAP	GLM-4.7	\$0.6/1M	Daily 10AM	Budget backup
	MiniMax M2.1	\$0.2/1M	5-hour rolling	Cheapest option
	Kimi K2	\$9/mo flat	10M tokens/mo	Predictable cost
FREE	iFlow	\$0	Unlimited	8 models free
	Qwen	\$0	Unlimited	3 models free
	Kiro	\$0	Unlimited	Claude free

Pro Tip: Start with Gemini CLI (180K free/month) + iFlow (unlimited free) combo = \$0 cost!

Understanding 9Router Costs & Billing

9Router Billing Reality:

- ✓ **9Router software = FREE forever** (open source, never charges)
- ✓ **Dashboard "costs" = Display/tracking only** (not actual bills)
- ✓ **You pay providers directly** (subscriptions or API fees)
- ✓ **FREE providers stay FREE** (iFlow, Kiro, Qwen = \$0 unlimited)
- ✗ **9Router never sends invoices** or charges your card

How Cost Display Works:

The dashboard shows **estimated costs** as if you were using paid APIs directly. This is **not billing** - it's a comparison tool to show your savings.

Example Scenario:

Dashboard Display:

- Total Requests: 1,662
- Total Tokens: 47M
- Display Cost: \$290



Reality Check:

- Provider: iFlow (FREE unlimited)
- Actual Payment: \$0.00
- What \$290 Means: Amount you SAVED by using free models!

Payment Rules:

- **Subscription providers** (Claude Code, Codex): Pay them directly via their websites
- **Cheap providers** (GLM, MiniMax): Pay them directly, 9Router just routes
- **FREE providers** (iFlow, Kiro, Qwen): Genuinely free forever, no hidden charges
- **9Router**: Never charges anything, ever

Use Cases

Case 1: "I have Claude Pro subscription"

Problem: Quota expires unused, rate limits during heavy coding

Solution:

```
Combo: "maximize-claude"
1. cc/claude-opus-4-6      (use subscription fully)
2. glm/glm-4.7            (cheap backup when quota out)
3. if/kimi-k2-thinking     (free emergency fallback)
```

Monthly cost: \$20 (subscription) + ~\$5 (backup) = \$25 total
vs. \$20 + hitting limits = frustration



Case 2: "I want zero cost"

Problem: Can't afford subscriptions, need reliable AI coding

Solution:

```
Combo: "free-forever"
1. gc/gemini-3-flash      (180K free/month)
2. if/kimi-k2-thinking     (unlimited free)
3. qw/qwen3-coder-plus    (unlimited free)
```

Monthly cost: \$0
Quality: Production-ready models



Case 3: "I need 24/7 coding, no interruptions"

Problem: Deadlines, can't afford downtime

Solution:

```
Combo: "always-on"
1. cc/claude-opus-4-6      (best quality)
2. cx/gpt-5.2-codex        (second subscription)
3. glm/glm-4.7            (cheap, resets daily)
```



4. minimax/MiniMax-M2.1 (cheapest, 5h reset)
5. if/kimi-k2-thinking (free unlimited)


Result: 5 layers of fallback = zero downtime

Monthly cost: \$20-200 (subscriptions) + \$10-20 (backup)

Case 4: "I want FREE AI in OpenClaw"

Problem: Need AI assistant in messaging apps (WhatsApp, Telegram, Slack...), completely free

Solution:






Combo: "openclaw-free" 

1. if/glm-4.7 (unlimited free)
2. if/minimax-m2.1 (unlimited free)
3. if/kimi-k2-thinking (unlimited free)







Monthly cost: \$0

Access via: WhatsApp, Telegram, Slack, Discord, iMessage, Signal...

? Frequently Asked Questions

- ▶  Why does my dashboard show high costs?
- ▶  Will I be charged by 9Router?
- ▶  Are FREE providers really unlimited?
- ▶  How do I minimize my actual AI costs?
- ▶  What if my usage suddenly spikes?

Setup Guide

- ▶  Subscription Providers (Maximize Value)
- ▶  Cheap Providers (Backup)
- ▶  FREE Providers (Emergency Backup)
- ▶  Create Combos
- ▶  CLI Integration
- ▶  Deployment

Available Models

- ▶ View all available models
-



Troubleshooting

"Language model did not provide messages"

- Provider quota exhausted → Check dashboard quota tracker
- Solution: Use combo fallback or switch to cheaper tier

Rate limiting

- Subscription quota out → Fallback to GLM/MiniMax
- Add combo: `cc/claude-opus-4-6` → `glm/glm-4.7` → `if/kimi-k2-thinking`

OAuth token expired

- Auto-refreshed by 9Router
- If issues persist: Dashboard → Provider → Reconnect

High costs

- Check usage stats in Dashboard
- Switch primary model to GLM/MiniMax
- Use free tier (Gemini CLI, iFlow) for non-critical tasks

Dashboard opens on wrong port

- Set `PORT=20128` and `NEXT_PUBLIC_BASE_URL=http://localhost:20128`

First login not working

- Check `INITIAL_PASSWORD` in `.env`
- If unset, fallback password is `123456`

No request logs under `logs/`

- Set `ENABLE_REQUEST_LOGS=true`



Tech Stack

- **Runtime:** Node.js 20+
- **Framework:** Next.js 16
- **UI:** React 19 + Tailwind CSS 4
- **Database:** LowDB (JSON file-based)
- **Streaming:** Server-Sent Events (SSE)
- **Auth:** OAuth 2.0 (PKCE) + JWT + API Keys



API Reference

Chat Completions

POST http://localhost:20128/v1/chat/completions



Authorization: Bearer your-api-key

Content-Type: application/json

```
{
  "model": "cc/claude-opus-4-6",
  "messages": [
    {"role": "user", "content": "Write a function to..."}
  ],
  "stream": true
}
```

List Models

GET http://localhost:20128/v1/models



Authorization: Bearer your-api-key

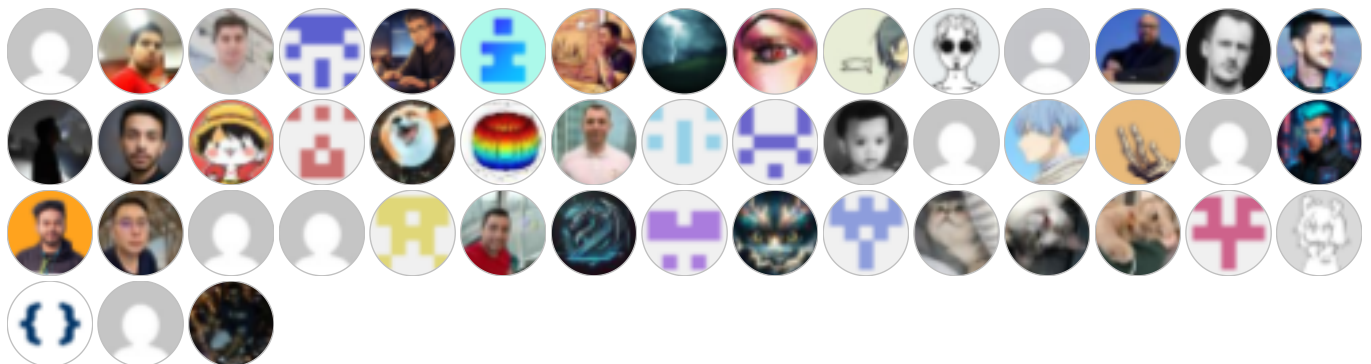
-> Returns all models + combos in OpenAI format

Support

- Website: 9router.com
- GitHub: github.com/decolua/9router
- Issues: github.com/decolua/9router/issues

Contributors

Thanks to all contributors who helped make 9Router better!





Star Chart

Releases 18

fix: Fix OpenRouter custom models not showing after being added version 0.3.83 Latest
20 hours ago

[+ 17 releases](#)

Packages 1

9router

Contributors 49



[+ 35 contributors](#)

Languages

