

模型聚合网关 · 新手极速接入指南

从零基础到跑通 Claude Opus 5.5 / Sonnet 5.5 / GPT-6.1 Sol / GPT-6 Astra · 仅需 1 分钟
没有网络要求 · 国内/海外网络均免翻墙直连 · 回国也能直接调用顶级模型

【必看第一步】客户端控制台（自助领 Key / 激活 / 查实时额度）

- 客户端网址（建议收藏）：<https://aghub.lol>
- 管理员激活与技术支持：WeChat: [Huaodnn](#) | Telegram: [+60 1125261626](#) | X: [@ha_hua64139](#)
- 极速启用三步走：
 - 生成密钥：用浏览器打开 <https://aghub.lol>，切换到「申请新密钥」，选择 微信（WX:）或 Telegram（TG:）并填入你的联系账号（邀请码可填好友的 INV-XXXX 或直接留空），生成你的专属 `sk-...` 密钥（自动保存在当前浏览器本地）。
 - 激活密钥：复制 `/client` 网页上显示的 `AG-...` 账户编号发送给管理员（WeChat: [Huaodnn](#) | Telegram: [+60 1125261626](#)）点亮激活（激活前不扣减有效期；超 2 小时未激活将自动释放，可随时重新申领）。
 - 填入软件：按下方表格或配置步骤，将 Base URL 与 `sk-...` 密钥填入任意 AI 软件即可开始使用。

第 0 步：如果你电脑上还没装任何 AI 客户端（按需求 2 选 1 下载）

- 选择 1：图形对话客户端（零基础首选，界面跟网页版 ChatGPT / Claude 一样）
 - 下载免费开源的 Cherry Studio：<https://cherry-ai.com>
 - 或下载 ChatBox：<https://chatboxai.app>（装好后填上 Key 即可直接对话、传文件、写代码）
- 选择 2：终端写代码 Agent + 切换器（CS 编程首选）
 - 终端安装 Claude Code：`npm install -g @anthropic-ai/claude-code`
 - 图形化切换器 CC Switch：<https://github.com/farion1231/cc-switch/releases>

一、全平台统一填写口径（所有客户端通用，收藏这一张表即可）

配置项	统一填写值（直接复制）	填写说明
0. 客户端控制台	https://aghub.lol	自助申领 <code>sk-...</code> 密钥、查看今日美金额度消耗与调用明细
1. 供应商名称	<code>AG-Gateway</code> （或任意名称）	随便起，方便自己识别即可

配置项	统一填写值（直接复制）	填写说明
2. Base URL （接口地址）	<code>https://api.aghub.lol</code>	若软件要求带 <code>/v1</code> （如 Cursor / OpenAI SDK），填 <code>https://api.aghub.lol/v1</code> （网关双向自动兼容；国内免代理直连）
3. API Key （专属密钥）	<code>sk-...</code> （在 <code>/client</code> 网页生成）	复制网页生成的 <code>AG-...</code> 编号联系管理员（微信 <code>Huaodnn</code> / TG <code>+601125261626</code> ）激活后即可调用
4. API 格式 / 协议	优先选 <code>OpenAI (/v1/chat/completions)</code>	选 <code>OpenAI</code> 格式可在一个供应商内同时调用 <code>Claude + GPT + Gemini</code> 全部模型；专门配 <code>Claude Code</code> 时选 <code>Anthropic Messages (/v1/messages)</code>

核心推荐模型 ID 与 Cursor 专用别名（直接复制）

模型定位	标准模型 ID（通用软件填写）	Cursor 编辑器专用别名	计费说明（官方原价 × 通道倍率，缓存读取 0.1x）
Claude 顶级推理旗舰（复杂架构 / 疑难 Bug）	<code>claude-opus-5-5-high</code>	<code>ag-opus</code> （或 <code>gw-claude-opus-5-5</code> ）	<code>\$4.00 / \$20.00</code> 每百万 Token × 编码通道 <code>0.9x</code>
Claude 高速开发主力（日常编码 / 长对话）	<code>claude-sonnet-5-5-high</code>	<code>ag-sonnet</code> （或 <code>gw-claude-sonnet-5-5</code> ）	<code>\$2.00 / \$10.00</code> 每百万 Token × 编码通道 <code>0.9x</code>
OpenAI 6.1 编码主力（高指令遵循 / Agent）	<code>gpt-6.1-sol</code>	<code>ag-gpt6</code> （或 <code>gpt-6.1-sol</code> ）	<code>\$2.00 / \$10.00</code> 每百万 Token × <code>0.2x</code> （最省额度）
OpenAI 6 深度逻辑旗舰（算法推导 / 数学证明）	<code>gpt-6-astra</code>	<code>gpt-6-astra</code>	<code>\$10.00 / \$50.00</code> 每百万 Token × <code>0.2x</code>
Google 3.8 极速轻量（高频问答 / 文档总结）	<code>gemini-3.8-flash-high</code>	<code>ag-flash</code> （或 <code>-medium / -low</code> ）	<code>\$0.75 / \$3.75</code> 每百万 Token × <code>0.9x</code>

统一按美金额度计费：所有模型都从同一份每日美金额度（或充值余额）里扣，不限请求次数和旗舰调用次数；选什么模型就跑什么模型，不会被自动降级。额度用完返回 `429`，每日 `00:00 (UTC+8)` 自动重置。完整规则见「测试说明与配额调度指南」。

Cursor 用户注意：因 Cursor 官方拦截了自定义 Base URL 下 `claude-*` 开头的模型名，在 Cursor 中添加模型时请直接填写上表中的简短别名（`ag-opus`、`ag-sonnet`、`ag-gpt6`、`ag-flash`）；其他软件一律填写标准模型 ID。

思考档位自由调节：支持在模型名后追加 `-low` / `-medium` / `-high` / `-xhigh`（如 `gpt-6.1-sol-xhigh`、`gemini-3.8-flash-medium`）或在软件界面直接调节思考强度。

二、懒人全自动配置：复制下面这段 Prompt 发给你电脑上的 AI

如果你电脑上装有 **Cursor**、**Windsurf**、**Claude Code** 或任何带命令行权限的本地 AI 助手，把下面这段内容复制发给它（替换你的 `sk-...` 密钥），AI 会自动帮你写入配置并完成连通测试：

请帮我全自动配置并验证模型网关与本地 AI 客户端环境。参数如下：

- Base URL: `https://api.aghub.lol` (OpenAI 兼容地址: `https://api.aghub.lol/v1`)
- API Key: <在此粘贴你的 `sk-...` 密钥>
- 模型 ID: `claude-opus-5-5-high`, `claude-sonnet-5-5-high`, `gpt-6.1-sol`, `gpt-6-astra`, `gemini-3.8-flash-high`

请用命令行自动帮我执行以下步骤（不要让我手动改配置文件）：

1. 连通性自检：用 `curl` 请求 `https://api.aghub.lol/v1/chat/completions` (Header: Authorization: Bearer <API Key>, 请求模型 `gpt-6.1-sol` 回复 "pong", `max_tokens: 10`), 确认返回 `200 OK`。
2. 自动安装与配置 Claude Code：检查本机是否有 `claude` 命令；若未安装且有 `npm`, 请执行 ``npm install -g @anthropic-ai/claude-code``；随后自动创建或合并更新 `~/ .claude/settings.json``, 在 ```env``` 字段写入：

```
"ANTHROPIC_BASE_URL": "https://api.aghub.lol",
"ANTHROPIC_AUTH_TOKEN": "<API Key>",
"ANTHROPIC_MODEL": "claude-opus-5-5-high",
"ANTHROPIC_SMALL_FAST_MODEL": "claude-sonnet-5-5-high"
```
3. 若我是 macOS 且装有 Homebrew, 可按需帮我检查或安装 `cherry-studio`` / `cc-switch``, 完成后向我汇报测试结果。

三、四大常用软件 1 分钟手动配置步骤

场景 A：Cherry Studio / ChatBox（零基础图形界面）

1. 打开软件左下角「设置 (齿轮) -> 模型服务商 -> + 添加」, 名称填 `AG-Gateway`, 类型选 `OpenAI`。
2. API 地址填 `https://api.aghub.lol`, API 密钥粘贴你的 `sk-...`。
3. 点击「+ 添加模型」, 填入上方标准模型 ID, 回到聊天页顶部即可切换模型对话。

场景 B：CC Switch / Codex CLI / OpenClaw

1. 名称填 `AG-Gateway`, Base URL 填 `https://api.aghub.lol` (或 `https://api.aghub.lol/v1`), API Key 粘贴 `sk-...`。
2. API 格式：通用选 `OpenAI (/v1/chat/completions)` 或 `OpenAI Responses (/v1/responses)`；专门配 Claude Code 选 `Anthropic Messages (/v1/messages)`。
3. 网关会自动为多轮会话开启粘滞缓存路由（缓存命中率 96%+, 缓存命中的 Token 按 0.1x 计费）。

场景 C：Cursor 编辑器

1. 打开 `Cursor Settings -> Models`, 点击底部 `+ Add Model`, 依次添加: `ag-opus`、`ag-sonnet`、`ag-gpt6`、`ag-flash`。
2. 在 `OpenAI API Key` 填入你的 `sk-...`, 开启 `Override OpenAI Base URL` 填入 `https://api.aghub.lol/v1`, 点击 `Verify` 即可。

场景 D：Claude Code 终端直配（完整支持 2.1.x+ 原生工具链）

1. 打开终端，新建或修改 `~/.claude/settings.json`，写入：

```
{"env": {"ANTHROPIC_BASE_URL": "https://api.aghub.lol", "ANTHROPIC_AUTH_TOKEN": "你的sk-密钥", "ANTHROPIC_MODEL": "claude-opus-5-5-high"}}
```
2. 进入任意代码目录输入 `claude` 启动，输入 `/model` 即可切换模型；网关原生支持 `/v1/messages/count_tokens`（0 额度扣减）与 `ToolSearch`。（若配置遇到问题，直接截图联系管理员微信 `Huaodnn` 或 TG `+60 1125261626` 协助处理）