# FastGPT社区版服务实例部署文档

## 前提条件

部署FastGPT社区版服务实例,需要对部分阿里云资源进行访问和创建操作。因此您的账号需要包含如下资源的 权限。 **说明**:当您的账号是RAM账号时,才需要添加此权限。

权限策略名称	备注
AliyunECSFullAccess	管理云服务器服务(ECS)的权限
AliyunVPCFullAccess	管理专有网络(VPC)的权限
AliyunROSFullAccess	管理资源编排服务(ROS)的权限
AliyunComputeNestUserFullAccess	管理计算巢服务(ComputeNest)的用户侧权限

# 计费说明

FastGPT社区版在计算巢部署的费用主要涉及:

- 所选vCPU与内存规格
- 系统盘类型及容量
- 公网带宽

### 部署流程

### 1. 访问计算巢FastGPT社区版部署链接,按提示填写部署参数:

计算巢	计算巢 / 我的服务 / @	刘建服务实例		100 80	1 钉钉交流群 帮	帮助文档
欢迎页	← 创建服务	实例				< 分享
服务目录						
我的服务	FastG	<b>ГР社区版 三方服务</b>	服务商名称	计算巢社区服务		
服务实例	ー 版本beta		服务商主页	https://market.aliyun.c	om/store/49259	€4
服务审核	视化进行	定一个面引 LLm 为语言被重加为加入中与自然的。 是此为中国政的的政策是之些、恢重项仍受害的。 网络马马尔德之下的马马	服劳阿加加	1 异来社区服务		
服务分销						
服务部署物			$\overline{}$ $\underline{}$	⇒ Ξ	必填 5/5 全部	8/8
运维公告						
学习中心				服务实例名称		
设置	▲ 服务实例名称	1astgptzsol 	刘祐次酒	* 地域		助文档 く分享 4 8/8 × ×
		长度小姐是041 子付,必须以来又子母开关,可已含数子、英文子母、知知线(一)和下划线(_),且需付百服务中定的名称规范。	又时页标	<ul> <li>付费类型配置</li> </ul>		
				付费类型		
应用管理		10 day = (14) 113		<ul> <li>资源配置</li> </ul>		
AppFlow	* 地域	华东1(杭州) •		* 实例类型		~
				* 实例密码		~
				<ul> <li>可用区配置</li> </ul>		
	◇ \\ 预失望能且			* 可用区ID		~
				* 专有网络VPC买例ID		~
	付费类型	按量付费		<ul> <li></li></ul>		
				101-222 (114) (24) (89) (22)		
	保存配置参数	下一步:确认订单 正在计算价格				

- 2. 参数填写完成后可以看到对应询价明细,确认参数后点击下一步:确认订单。
- 3. 确认订单完成后同意服务协议并点击**立即创建**进入部署阶段。
- 4. 等待部署完成后就可以开始使用服务,进入服务实例详情点击服务地址。

计算巢	计算巢 / 服务实	例 / 服务实例详情							體 钉钉交流群 帮助文档
欢迎页	← fastg	ptlpbv (si-8	fef5d9b024	404d1d8	6ce)				删除服务实例
服务目录									
我的服务	概览	资源 事	件 监控	运维管理	运维项	日志管理	升级历史	费用统计	
服务实例	<b>—</b>						40.50		
服务审核	事件					服务的	资源		
服务分销	计划内运维事件	牛 非预期运维事件	+ 实例费用事件	非 实例多	全事件	ECS	6		
服务部署物	0	0	0	0		1			
运维公告									
学习中心									
设置	< 立即使用								
	初始用户名 ⑦	) root				初始密码(	) 3ا	34	
	FastGPT服务	地址 ⑦ http://47.:	16:3000						
应用管理									
AppFlow									
	实例信息								
	状态	❷ 已部署				服务实例名	3称 fa	stgptlpbv	
	状态描述	-				创建时间	2	025年3月18日 18:49:47	
	更改时间	2025年3月18	日 18:51:36			标签	0	•	
	费用分析	立即查看 [2]				咨酒钥		-acfmzmbzoaad5og / default resource group	

5. 加载完成后即可登录使用FastGPT服务。

	■ 简体中文 >
FastGPT	4
用户名	
密码	
使用即代表你同意我们的 服务协议 & 隐私协议	
	í.
无法登录、点击联系	
annai (A) III MAAA	

### 6. 选择模型:

•	▲ 个人信息 △ 第三方账号	可用模型 調整設置 模型渠道 调用日本 模型提供表 6 阿里干向 、 模型类型 语言模型 、			取い機製 配置文件 飲暖機製 Q 機械機製品機能
	模型提供商	<u>病型ID や</u>	模型类型	启用(0)	
		校 gwen-turbo 〔128k〕 工具期用〕	8882		⊳®
() 1.4.11	⊖ 288	珍 gwen-plus G4k ) IIR現現	6882		⊳ @
89		vi≩ qwen-vi-plus (32k) R(38	8882		⊳ @
		珍 qwen-max (32k) 工具現用	语古德型		⊳ ø
		珍 qwen-vi-max (32k) 微能	语意频型		⊳ ©
		が qwq-plus 128k 工成36用	场高限型		⊳ ¢
		1芬 qwq-32b 〔128k〕 ①正闻明用	國家報告		⊳ ø
		128k	语言模型		⊳ @
		1梦 qwen2.5-7b-instruct (128k) 正具現用	通信规范		⊳ ¢
		∲ qwen2.5-14b-instruct [128k] 正用即用	85K2		⊳ ¢
		☆ qwen2.5-32b-instruct (128k) (工具現用)	5285		⊳ ¢
		12 qwen2.5-72b-hstruct 1228	6562		⊳ ø
\$7	• V4.9.1				

### 7. 配置模型的请求地址和APIKey:

☑ 模型参数编辑				×
參数名			参数名	
模型ID ⑦		qwen-turb	支持工具调用 ◎	
模型提供商		珍 阿里千问 ~	支持函数调用 ⑦	
别名 ③		Qwen-turbo	支持图片识别 ◎	
****			支持输出思考 ⊙	
<b>淑</b> 天上下又		128000	用于知识库文件处理	
*知识库最大引用		100000	用于问题分类	
最大响应 tokens ③		8000	用于文本提取	
具+ ↓ □ ∅			用于工具调用节点	
RK/Amile U			默认提示词 ①	
展示 Top-p 参数		•		4
展示停止序列参数		•	自定义问题分类提示词 ①	
响应格式	[ "text", "ison object"			6
	1	4	自定义内容提取提示词 ①	
自定义请求地址 ③		https://dashscope.aliyuncs.com/v1/chat/co	٦	
自定义请求 Key ⑦		sk-7b384456	₽adu 26/1/2/10 ①	
			BUUY BR/P-F-R	4
				恢复默认配置 取消 确认

请求地址为: https://dashscope.aliyuncs.com/compatible-mode/v1/chat/completions

8. 保存后点击启用并测试:

2 个人信息	可用模型 模型配置 模型渠道 调用日志			就认模型 配置文件 题提供
△ 第三方账号	模型提供商 珍 阿里千间 🛛 🗸 模型类型 语言模型 🗸			Q 根据模型名提索
《 模型提供商	展開口 や	模型类型	启用(1)	
② API 密钥 ◎ 设置	0 gwen-turbo (1286) 江风明明	(图)的 (W.Z.)		
⊕ ±±	5 quen-plus 64 IARH	语言规型		
	校 qwen-vi-plus 32x)(我觉)	语言规范		
	☆ qwen-max うZK 【工具現用	语言模型		> *

9. 测试成功后即可正常使用模型。