

MaxKB使用指南

1 商品说明

MaxKB = Max Knowledge Brain，是一款强大易用的企业级 AI 助手，支持 RAG 检索增强生成、工作流编排、MCP 工具调用能力。MaxKB 支持对接各种主流大语言模型，广泛应用于智能客服、企业内部知识库问答、员工助手、学术研究与教育等场景。

本商品通过鲲鹏服务器+EulerOS2.0进行安装部署

2 商品购买

您可以在云商店搜索“MaxKB”。

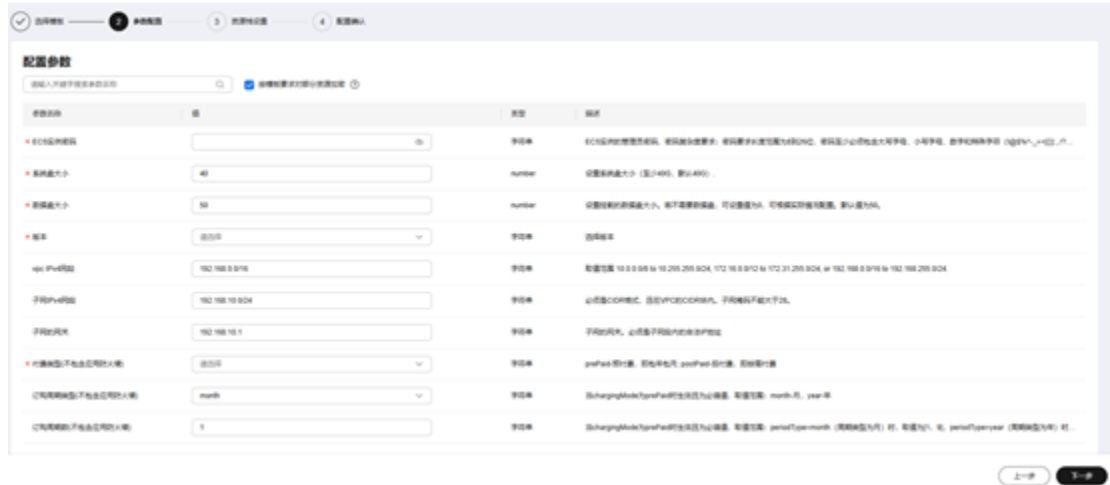
其中，地域、规格、推荐配置使用默认，购买方式根据您的需求选择按需/按月/按年，短期使用推荐按需，长期使用推荐按月/按年，确认配置后点击“立即购买”。

2.1 商品支持自定义 ECS 购买，具体见章节 3.1.1

2.2 使用 RFS 模板直接部署



必填项填写后，点击 下一步



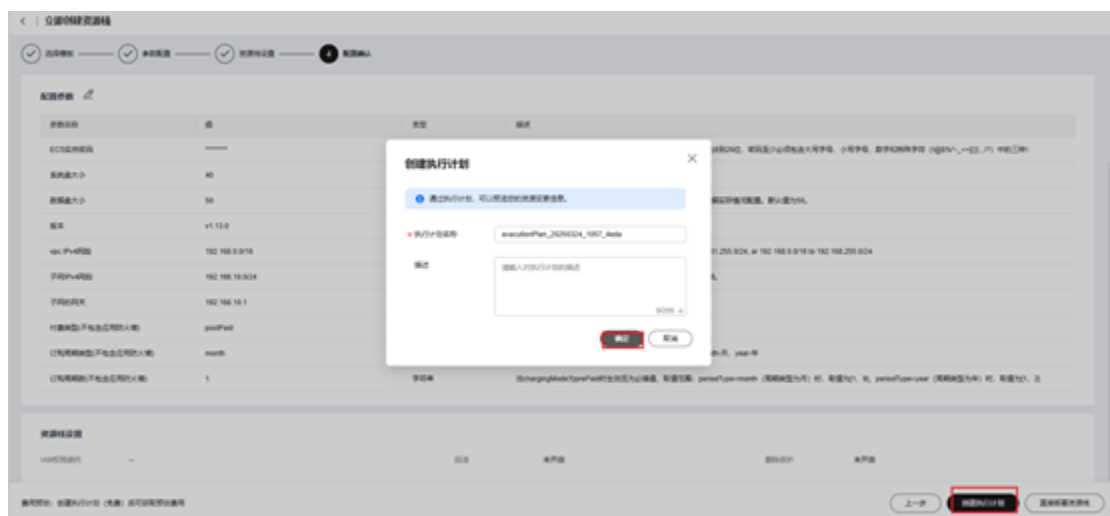
参数名称	值	类型	描述
EC2实例规格		字符串	EC2实例的规格名称。规格名称要求：规格名称长度范围为8到12位，规格至少由1个大写字母、1个小写字母、数字和特殊字符（如\$%^&*+=-~_.,:;@!~）组成。
实例大小	40	number	设置实例大小（范围400，默认400）。
实例大小	50	number	设置实例的实例大小，如果不填则默认，可以设置为0，可设置实例与实例，默认为0。
版本		字符串	指定版本
vpc ID	102-100-0-0-10	字符串	指定VPC ID。VPC ID格式为：102-100-0-0-10，102-100-0-0-10，102-100-0-0-10，102-100-0-0-10。
子网ID	102-100-10-0-10	字符串	指定子网ID。子网ID格式为：102-100-10-0-10，子网ID至少由1个大写字母、1个小写字母、数字和特殊字符（如\$%^&*+=-~_.,:;@!~）组成。
子网掩码	102-100-10-1	字符串	子网掩码，至少由1个大写字母、1个小写字母、数字和特殊字符（如\$%^&*+=-~_.,:;@!~）组成。
计费模式（不指定计费模式）		字符串	指定计费模式，默认为按量计费，可设置为包年包月。
计费周期（不指定计费周期）	month	字符串	指定计费周期，默认为按月计费，可设置为按年计费。
计费周期（不指定计费周期）	1	字符串	指定计费周期，默认为按月计费，可设置为按年计费。



高级设置

安全设置

创建直接计划后，点击 确定



创建直接计划

通过API创建，可以创建直接计划。

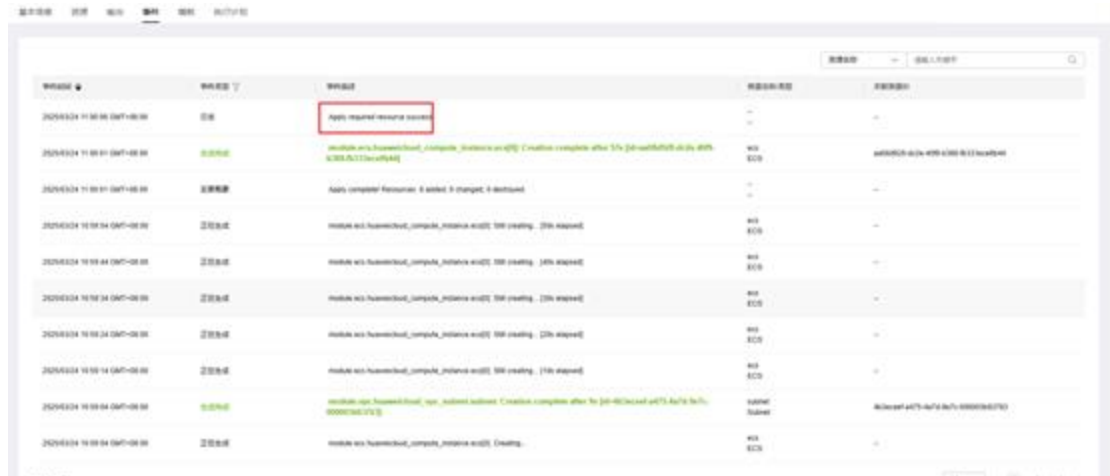
执行计划名称

描述

点击 部署



如下图“Apply required resource success.”即为资源创建完成



3 商品资源配置

商品支持ECS控制台配置，下面对资源配置的方式进行介绍。

3.1 ECS 控制台配置

3.1.1 准备工作

在使用ECS控制台配置前，需要您提前配置好安全组规则。

安全组规则的配置如下：

- 入方向规则放通端口8080，源地址内必须包含您的客户端ip，否则无法访问
- 入方向规则放通CloudShell连接实例使用的端口22，以便在控制台登录调试。
- 出方向规则一键放通

3.1.2 创建 ECS

前提工作准备好后，选择ECS控制台配置跳转到购买ECS页面，ECS资源的配置如下图所示



示：

弹性云服务器 / 购买弹性云服务器

<

购买弹性云服务器

1 购买页全面升级，诚邀您参与问卷调查，精选反馈，尊享好礼回馈！立即参与

基础配置

计费模式

包年/包月 按需计费 竞价计费

按需计费实例不支持备案。了解详情

区域

华北-北京四

推荐区域 华北-北京四 (31) 华南-广州 (54) 华东-上海一 (6) 西南-贵阳一 (11) 华东-上海二

云服务器创建后无法更改区域；不同区域之间内网互不相通，请就近选择靠近您业务的区域，减少网络时延。如何选择区域

可用区

随机分配 可用区1 可用区2 可用区3 可用区7 随机至多可用区

<

购买弹性云服务器

规格类型选型 业务场景选型

CPU架构

x86计算 鲲鹏计算

实例筛选

请选择vCPUs 请选择内存 请输入规格名称模糊搜索

隐藏售罄的规格 仅展示待上架类型的规格

鲲鹏通用计算增强型 鲲鹏内存优化型 鲲鹏超高性能型

全部 kc1 kc2

搭载鲲鹏920处理器及25GE智能高速网卡，提供强劲鲲鹏算力和高性能网络，更好满足政府、互联网等各类企业对云上业务高性价比、安全可靠等诉求。

实例类型	规格名称	vCPUs	内存	CPU	基准 / 最大带宽	内网收发包	IPv6	规格参考价
鲲鹏通用计算增强型k...	kc1.large.4	2vCPUs	8GiB	Huawei Kunpeng ...	0.8 / 3 Gbit/s	30万PPS	是	¥0.41/小时
鲲鹏通用计算增强型k...	kc1.xlarge.2	4vCPUs	8GiB	Huawei Kunpeng ...	1.5 / 5 Gbit/s	50万PPS	否	¥0.60/小时
鲲鹏通用计算增强型k...	kc1.xlarge.4	4vCPUs	16GiB	Huawei Kunpeng ...	1.5 / 5 Gbit/s	50万PPS	是	¥0.81/小时
鲲鹏通用计算增强型k...	kc1.2xlarge.2	8vCPUs	16GiB	Huawei Kunpeng ...	3 / 7 Gbit/s	80万PPS	否	¥1.20/小时
鲲鹏通用计算增强型k...	kc1.2xlarge.4	8vCPUs	32GiB	Huawei Kunpeng ...	3 / 7 Gbit/s	80万PPS	是	¥1.63/小时
鲲鹏通用计算增强型k...	kc1.3xlarge.2	12vCPUs	24GiB	Huawei Kunpeng ...	4.5 / 9 Gbit/s	110万PPS	否	¥1.80/小时
鲲鹏通用计算增强型k...	kc1.3xlarge.4	12vCPUs	48GiB	Huawei Kunpeng ...	4.5 / 9 Gbit/s	110万PPS	是	¥2.44/小时

当前规格 鲲鹏通用计算增强型 | kc1.2xlarge.4 | 8vCPUs | 32GiB



操作系统

镜像 ^②

公共镜像

私有镜像

共享镜像

市场镜像

MaxKB1.10.7-arm64-v1.0-HCE2.0(40GiB)

新建私有镜像 [↗](#)使用私有镜像创建弹性云服务器前，请查看[操作系统已知问题](#) [↗](#)☒ 开启安全防护 ^②

主机安全为您提供风险预防、入侵检测、高级防御、安全运营、网页防篡改等安全防护，构建云服务器安全体系。

基础防护

口令检测、漏洞检测等基础防护

免费一个月

企业版

漏洞修复、病毒查杀、等保必备

¥0.18/台/小时

试用1个月，主机安全服务自动停止服务。

网络

虚拟私有云 ^②

ecs-vpc-default(192.168.0.0/16)

新建虚拟私有云 [↗](#)

主网卡

ecs-subnet-default(192.168.10.0/24)



自动分配IP地址



可用私有IP数量250个

☒ 新增扩展网卡

您还可以增加2块网卡

源/目的检查 ^②

安全组

选择安全组 ^②

default(cd587ac7-a9f8-4840-a56d-561fbc7acce) X



新建安全组

请确保所选安全组已放通22端口（Linux SSH登录），3389端口（Windows远程登录）和ICMP协议（Ping）。[配置安全组规则](#) [↗](#)建议将安全组入方向规则中高危端口的源地址设置为已知IP地址、安全组或IP地址组，避免因网络入侵出现业务中断、数据泄露或数据勒索等严重后果。[查看高危端口](#) [↗](#)展开安全组规则 ^②

公网访问

弹性公网IP ^②

立即购买

使用已有

暂不购买

线路 ^②

全动态BGP

静态BGP

☒ 不低于99.95%可用性保障公网带宽 ^②按带宽计费 ^②

流量较大或较稳定的场景



按流量计费

流量小或流量波动较大场景



加入共享带宽

多业务流量错峰分布场景

指定带宽上限，按实际使用的出公网流量计费，与使用时间无关。

带宽大小 (Mbit/s)

5

10

20

50

100

自定义

带宽范围：1-300 Mbit/s；开启DDoS基础防护 ^② 免费

释放行为

☒ 随实例释放

对于设置了随实例释放的弹性公网IP，将在删除云服务器同时执行删除。

值得注意的是：

- VPC您可以自行创建
- 安全组选择3.1.1章节中配置的安全组
- 弹性公网IP选择现在购买，推荐选择“按流量计费”，带宽大小可设置为10Mbit/s
- 其余默认或按规则填写即可。

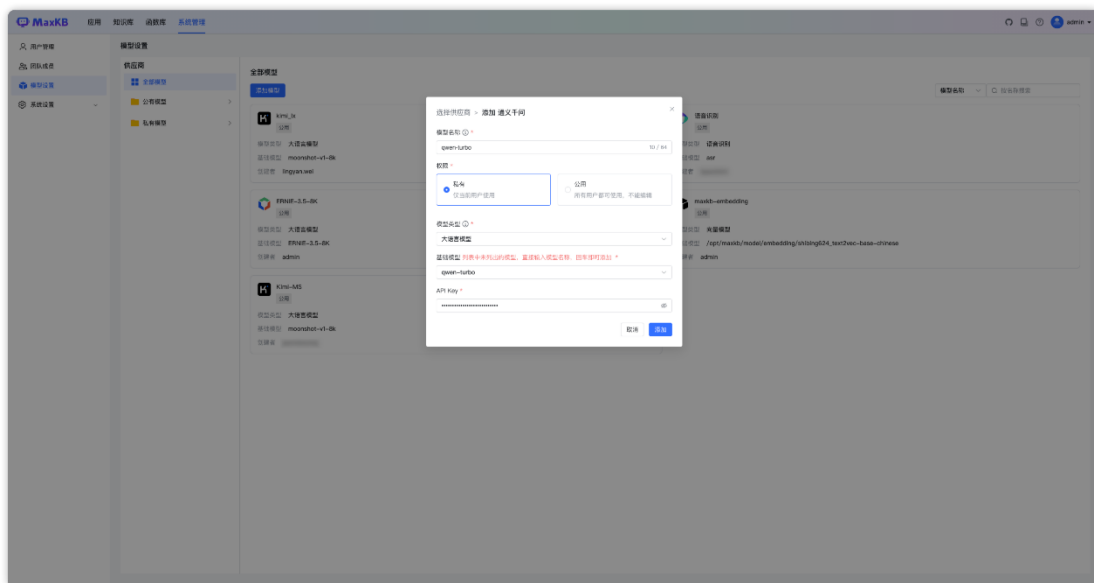
4 商品使用

安装完成后，可通过浏览器访问地址 `http://弹性公网 IP 地址:8080`，并使用默认的管理员用户和密码登录 MaxKB。

用户名：admin

默认密码：MaxKB@123..

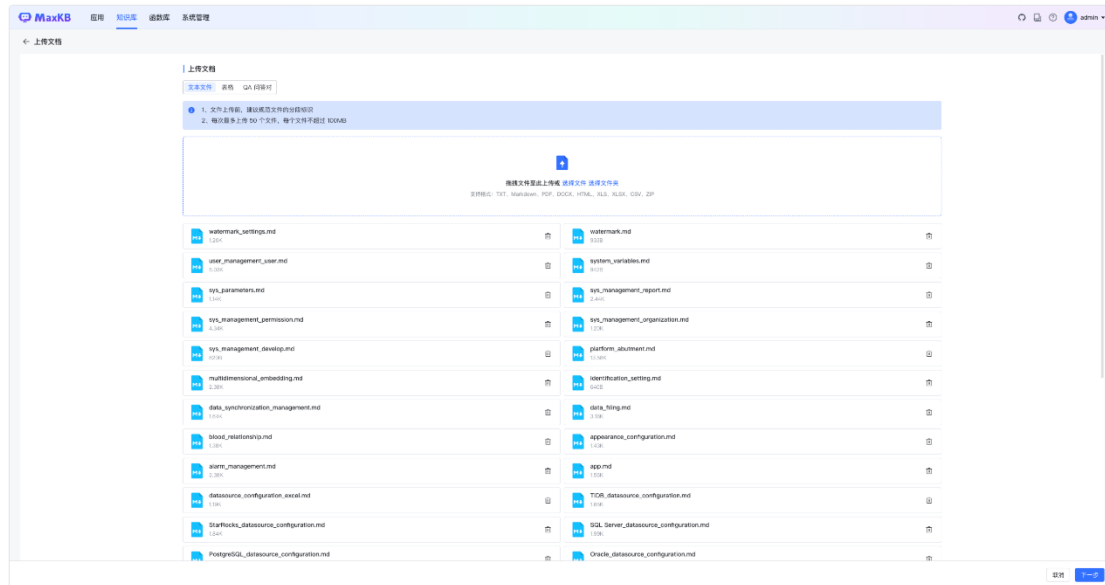
4.1.1 添加模型



4.1.2 创建知识库

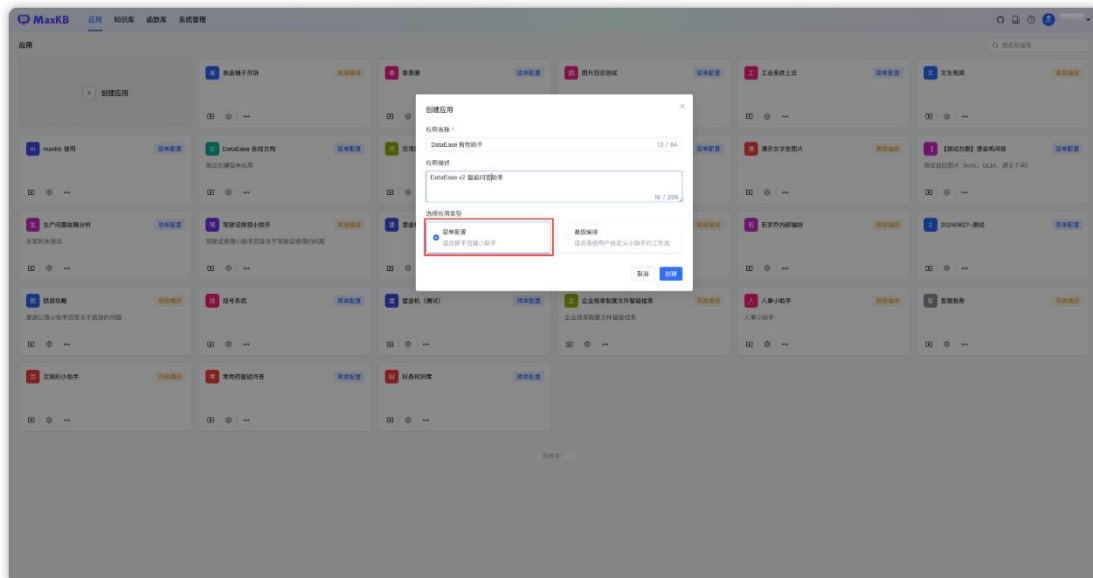
打开【知识库】页面，点击【创建知识库】，输入知识库名称、知识库描述、选择内置的向量模型 MaxKB-Embedding，并设置知识库类型为通用型，然后将离线文档通过拖拽方

式或选择文件上传方式进行上传。



4.1.3 创建应用

点击【创建应用】，输入应用名称，选择【简易配置应用】，点击【创建】



4.2 参考文档

[官方文档](#)