**部署手册**

# 部署准备

本文件旨在帮助您在部署 Agent构建平台及智能体应用 之前，确认贵公司的软硬件环境是否满足要求。我们希望此文档能够帮助您更好地理解此次合作，并高效地使用 Agent构建平台及智能体应用 提供的高可用性、多租户、权限认证与管理等一系列企业级功能。

Agent构建平台及智能体应用 是一款可以运行在 Kubernetes（K8S）环境上的软件，本文仅讨论部署单套 Agent构建平台及智能体应用 的情况。请尽可能完全按照我们提供的软硬件需求清单进行准备，以免在部署过程中遇到问题或困难，甚至超出翼海云峰技术团队所能够提供服务的范畴。

翼海云峰技术团队提供以下服务：

**深度技术支持服务：**提供定制化部署、定制化开发、定制化信息咨询等服务。

**标准技术支持服务：**提供常规的软件部署协助服务。

如果您对部署有额外需求，请联系您的销售代表。

## 1. 硬件需求

### 1.1 测试阶段

在测试阶段，您至少需要具备以下环境，以启动最基础的 Agent构建平台及智能体应用 系统：

1、Kubernetes 1.24+ 环境

* 4 核 CPU，16 GB 内存

2、Helm 3.14+

3、Kubectl

4、S3 存储或兼容 S3 协议的持久化存储

* 容量至少为 50 GB（推荐 100 GB）

该系统具有以下特点：

* 具备 Agent构建平台及智能体应用 的全部软件功能
* 低可用性
* 数据可靠性较低
* 不适用于任何生产或类生产环境

您也可以使用 **Minikube** 进行部署。Minikube 是一种轻量级的本地 Kubernetes 集群，适合早期试用功能和学习 Agent构建平台及智能体应用 的软件能力。**请勿在生产环境中使用此方案；**一旦服务器发生问题，您的数据可能无法保存。

在此阶段，请确保引入贵司的运维工程师或相关服务器维护人员，以获取足够的资源并减少信息沟通成本。如果您不熟悉 Kubernetes、Helm Chart、Kubectl 等概念及其具体应用，建议尽快联系具备相关能力的技术人员获得支持，或联系翼海云峰销售团队获取深度技术支持的报价。如果您在合同阶段已购买相应的深度技术支持服务，我们将指派专门技术团队协助您进行部署。

### 1.2 生产环境

在生产环境下，您需要具备以下硬件需求，以测试集群部署环境下的 Agent构建平台及智能体应用 能力：

1、Kubernetes 集群

* 1 个工作节点（Worker Node）：4 核 CPU，16 GB 内存

2、持久化存储

* S3、Azure Blob、阿里云 OSS、Google Cloud Storage 等
* 100 GB 存储空间

3、PostgreSQL 数据库

* 版本 14+
* 2 核 CPU，4 GB 内存
* 256 GB 存储空间

4、Redis 数据库

* 版本 6+
* 1 GB 内存

5、向量数据库虚拟机 - Qdrant 单节点

* 1 台虚拟机
* 4 核 CPU，8 GB 内存
* 100 GB 存储空间

您可以自行选择供应商，但我们建议选择具有良好品牌信誉的企业，以避免在使用过程中遇到硬件问题。

在向量数据库方面，我们首选推荐使用 Qdrant 进行部署。如您希望使用其他向量数据库，请提前与您的销售代表联系，获取支持的向量数据库清单。目前，Agent构建平台及智能体应用 仅对极少数向量数据库提供官方支持。对于不在支持名单中的向量数据库，无法保证其可用性或稳定性，且无法在常规服务中提供任何部署支持。对特定向量数据库的支持属于 深度技术支持服务 的范畴，但我们无法确保在所有情况下都能支持您的数据库选择，因此请联系销售代表进行评估，以免造成误解。

如果您在上述测试中认为软件功能良好，并已准备好将软件投入生产环境，建议准备以下硬件条件，以确保其高可扩展性。该硬件配置可支持每日活跃用户数达 3000 人：

1、Kubernetes 集群

* 3 个主节点（Master Nodes）：每个节点 4 核 CPU，16 GB 内存
* 6 个工作节点（Worker Nodes）：每个节点 8 核 CPU，32 GB 内存

2、持久化存储

* 华为云、S3、Azure Blob、阿里云 OSS、Google Cloud Storage 等
* 512 GB 存储空间

3、PostgreSQL 数据库

* 版本 14+
* 4 核 CPU，8 GB 内存
* 512 GB SSD 存储空间/

4、Redis 数据库

* 版本 6+
* 2 GB 内存

5、向量数据库虚拟机 - Qdrant 集群

* 3 台虚拟机
* 每台 8 核 CPU，64 GB 内存
* 100 GB 存储空间

同样，建议选择具有良好品牌信誉的云服务提供商提供相应的硬件。

### 1.3 例外情况

除非另有书面约定，在以下情况下，翼海云峰技术团队将无法为您提供标准部署服务：

1. 您使用的并非标准云服务硬件，而是本地硬件（如物理服务器等）。
2. 您所使用的硬件与其他软件共享。
3. 您希望采用非 Kubernetes 的方案进行部署。
4. 您希望对现有的软件结构、代码等进行修改。
5. 您希望使用翼海云峰技术团队无法支持的某些向量数据库。
6. 您希望使用不被支持的硬件协议。
7. 其他经翼海云峰技术团队明确拒绝支持的情况。

由于上述情况使得原有的标准服务流程与部署程序无法正常运作，必须通过定制化方案进行服务，因此不在我们的标准部署服务范围内。

## 2. 软件需求

您需要知悉，Agent构建平台及智能体应用 中开放了强大的第三方软件接入生态。一旦交易达成，您将得到一份 Agent构建平台及智能体应用 使用手册，该手册中会指出接入第三方软件的方法，敬请参考。但这不意味着翼海云峰技术团队会帮助您部署第三方软件，更不意味着翼海云峰技术团队会为您使用该软件造成的后果负责。请您在使用前自行评估第三方软件的使用方法、风险及人员配置。

​

### 2.1 软件配置

通常情况下，您需要安装并初始化以下运行环境：

* Docker
* Kubernetes（K8S）
* Minikube

同时，需确保您的 Linux 环境 为最新版本，以保证相关依赖库的可用性。除此之外，一般不需要安装其他基础设施。

### 2.2 人员配置

请确保贵司始终有自有的运维工程师，或通过翼海云峰技术团队深度技术支持服务，协助推进Agent构建平台及智能体应用部署的全流程。如果缺乏上述支持，或者您虽有开发团队但缺乏运维经验，可能会导致 Agent构建平台及智能体应用 的部署遇到困难。您需要在商务沟通期间明确提出您的技术力量储备或部署方案计划，以便我们为您匹配最佳的技术解决方案。

Agent构建平台及智能体应用 是一款面向大规模应用的软件，技术力量在部署过程中尤为重要。

### 2.3 更新与维护

合同签订后，您将在企业版部署手册中获取 更新方法。您可以选择不更新，但您的 访问密钥 仅在合同约定的时间内生效。

如果软件出现问题，请联系我们的 技术团队 以获得协助。对于软件问题，我们将提供维护服务。常规的问题响应周期为 24 小时；若出现影响软件生产能力的致命问题，我们将尽快协助您解决。

上述协助指 非介入性 的技术支持，即仅对软件本身的正常运行进行维护，不包括软件使用方面的培训，也不提供定制化开发或部署服务。

## 3. 部署

### 3.1 部署须知

Agent构建平台及智能体应用 是一套软件系统。单独购买软件时，我们将为您提供 一年共计 24 小时 的标准技术支持服务。购买本软件并不意味着您已购买 Agent构建平台及智能体应用 的部署服务。因此，请在阅读本文后，适当评估现有的技术力量，并与销售人员联系，评估技术难度和您的实际需求，以便我们为您提供最适合的部署解决方案。

### 3.2 标准部署服务范围

一般情况下，您将在合同签订并结清款项后收到以下内容：

1. **Agent构建平台及智能体应用 的账号与密码**——用于生成令牌（token），以便访问 Agent构建平台及智能体应用 镜像。
2. **Agent构建平台及智能体应用 部署手册**——请按照手册要求自行部署。

**标准部署服务** 包括：

1. 帮助您确认硬件是否符合标准部署服务需求。
2. 对部署手册的解读和指导。
3. 协助您联系可靠的云服务提供商。
4. 为您提供有限的与部署手册相关的技术咨询。

​

### 3.3 深度技术支持服务

如果您的需求属于 **1.3 例外情况** 中的某一类，或您认为团队无法调集足够的运维力量进行部署，或需要更多信息来培训您的团队，请联系销售人员评估需求。翼海云峰技术团队 为您提供相应的 **深度技术支持服务**，确保您能够经济、高效、轻松地部署并使用所购买的软件。

这些服务通常包括：

1. 深度软件集成定制服务。
2. 深度介入部署流程服务。
3. 深度信息咨询服务。

上述服务均有特定定价，且与您的实际需求密切相关。请联系销售人员获取报价。