

星纵 AIoT 推理平台

安装指南



关于本指南

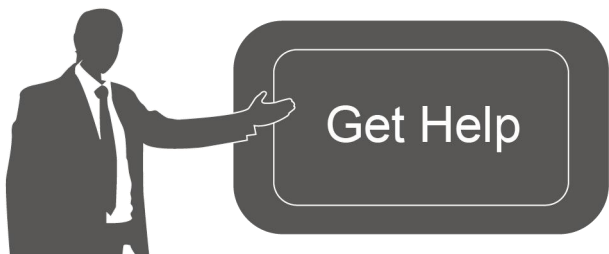
本指南将帮助您在 Ubuntu 上使用 Docker 安装和启动星纵 AIoT 推理平台。

本手册适用于以下用户：

- 分销商
- 网络规划师
- 负责网络配置和维护的网络管理员

版权所有© 2011-2024 星纵物联

保留所有权利。



如需帮助，请联系

星纵物联技术支持：

邮箱：contact@milesight.com

电话：0592-5023060

传真：0592-5023065

地址：厦门市集美区软件园三期 C09 栋

文档修订记录

日期	版本	描述
2024.2.19	V1.0	第一版

目录

星纵 AIoT 推理平台介绍	4
服务器要求	4
硬件要求	4
操作系统要求	4
安装推理平台	4
安装准备	4
安装 Docker	4
安装推理平台	5

星纵 AIoT 推理平台介绍

星纵 AIoT 推理平台是基于 Triton 推理服务器开发的智能 AI 解决方案，旨在识别和处理来自星纵物联感知相机的图片数据并返回给星纵 AIoT 感知平台。

服务器要求

硬件要求

服务器硬件推荐如下：

1~300 台设备接入

- RAM: 4 GB

300~500 台设备接入

- RAM: 8 GB

操作系统要求

- Ubuntu Kinetic 22.10
- Ubuntu Jammy 22.04 (LTS)
- Ubuntu Focal 20.04 (LTS)
- Ubuntu Bionic 18.04 (LTS)

安装推理平台

安装准备

- Ubuntu 服务器
- AIoT 推理平台安装包
- SSH 工具如 Putty、Xshell、SecureCRT 等

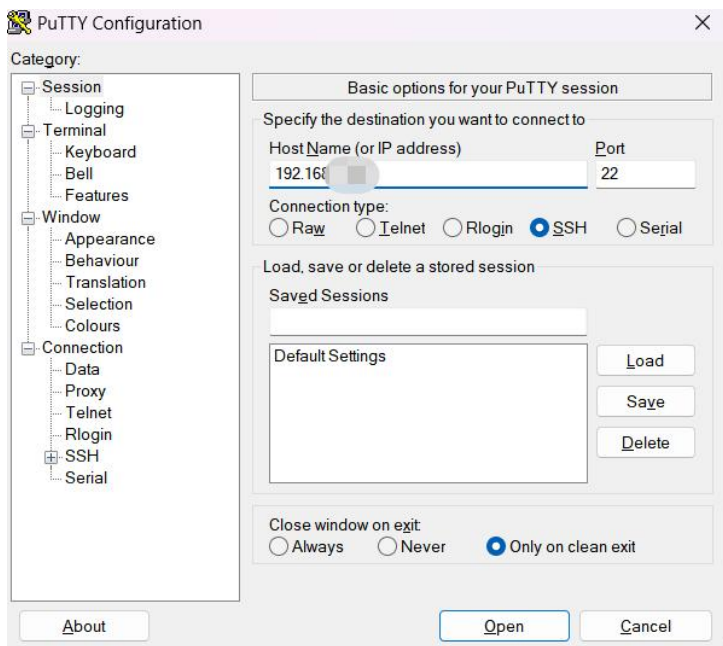
安装 Docker

- for [Ubuntu](#)

安装推理平台

安装步骤

1. 使用 Putty 等远程工具登录服务器



2. 进入 root 权限

```
su root
```

3. 下载推理平台镜像至本地路径下

```
wget https://milesight-cn-resource.oss-cn-hangzhou.aliyuncs.com/share/msaiotinferenceplatform-min-1.0.tar
```

4. 推送本地镜像到 Docker

```
docker load < msaiotinferenceplatform-min-1.0.tar
```

5. 为推理平台创建配置文件

```
nano docker-compose.yml
```

并将下列文本内容加入到 yml 文件中：

```
version: '3.0'
services:
  msinfer:
    image: ms-inference-server-min:1.0
    shm_size: '1gb'
    ports:
      - "5221:8080"
    volumes:
      - /ms-inference-server:/ms-inference-server
    command: infer_daemon
```

命令参数介绍:

- **image:** 镜像名称
 - ✧ **ms-inference-server**: 标准版本镜像, 支持 GPU 推理, 大小 16GB 左右
 - ✧ **ms-inference-server-min**: 最小版本镜像, 纯 CPU 推理, 大小 1.5GB 左右
- **ports:** 端口映射, 冒号右边的端口是镜像内部端口, 不可以修改, 左边的端口可以根据需要修改
 - ✧ **5221:8080**: web 端口
 - ✧ **volumes**: 路径映射, 用来保存数据和日志, 冒号右边的不可以修改, 左边可以根据需要设置

注意: 如果您已经有创建过 compose 文件, 只需将如上灰底部分内容直接复制到 compose 文件。

```
version: '3.0'
services:
  mysp:
    restart: always
    image: "msaiotsensingplatform:1.0.1.0-a4"
    ports:
      - "5220:9090"
      - "1883:1883"
      - "7070:7070"
      - "5683-5688:5683-5688/udp"
    environment:
      TB_QUEUE_TYPE: in-memory
      CASSANDRA_URL: localhost:9042
      CASSANDRA_KEYSPACE_NAME: msaiotsensingplatform
    volumes:
      - /var/mysp-data:/data
      - /var/mysp-logs:/var/log/msaiotsensingplatform

  msinfer:
    image: ms-inference-server-min:1.0
    shm_size: '1gb'
    ports:
      - "5221:8080"
    volumes:
      - /var/ms-inference-server:/ms-inference-server
    command: infer_daemon
```

6. 启动镜像

```
docker compose up -d
```

注意:

从 2.3.4 版本开始, ThingsBoard 支持 Docker Compose V2 (Docker Desktop 或 Compose 插件), 因为 Docker 不再支持 docker-compose 作为独立设置。

我们强烈建议更新到 Docker Compose V2 并使用它。

如果您仍然依赖使用 Docker Compose 作为 docker-compose (带连字符), 那么请执行以下命令来启动 ThingsBoard:

```
# docker-compose up -d
```

7. 执行完上述命令后, 就可以在浏览器中打开推理平台。使用 Ubuntu 系统 ip: 端口号形式登录页面。

默认使用: <http://{your-host-ip}:5221>



8. 出现任何问题, 可以检查服务日志中的错误。

查看推理平台容器日志, 运行以下指令:

```
$ docker compose logs -f msinfer
```

停止推理平台, 运行如下指令:

```
$ docker compose stop msinfer
```

启动推理平台, 运行如下指令:

```
$ docker compose start msinfer
```

注意:

Docker Compose 作为 docker-compose (带连字符) 已被弃用。建议改用 Docker Compose V2。

如果您仍然依赖 docker compose 作为独立版本, 以下是上述命令的列表:

```
# docker-compose logs -f msinfer
```

```
# docker-compose stop msinfer
```

```
# docker-compose start msinfer
```