



AUP Remote Lab User Manual (Admin)

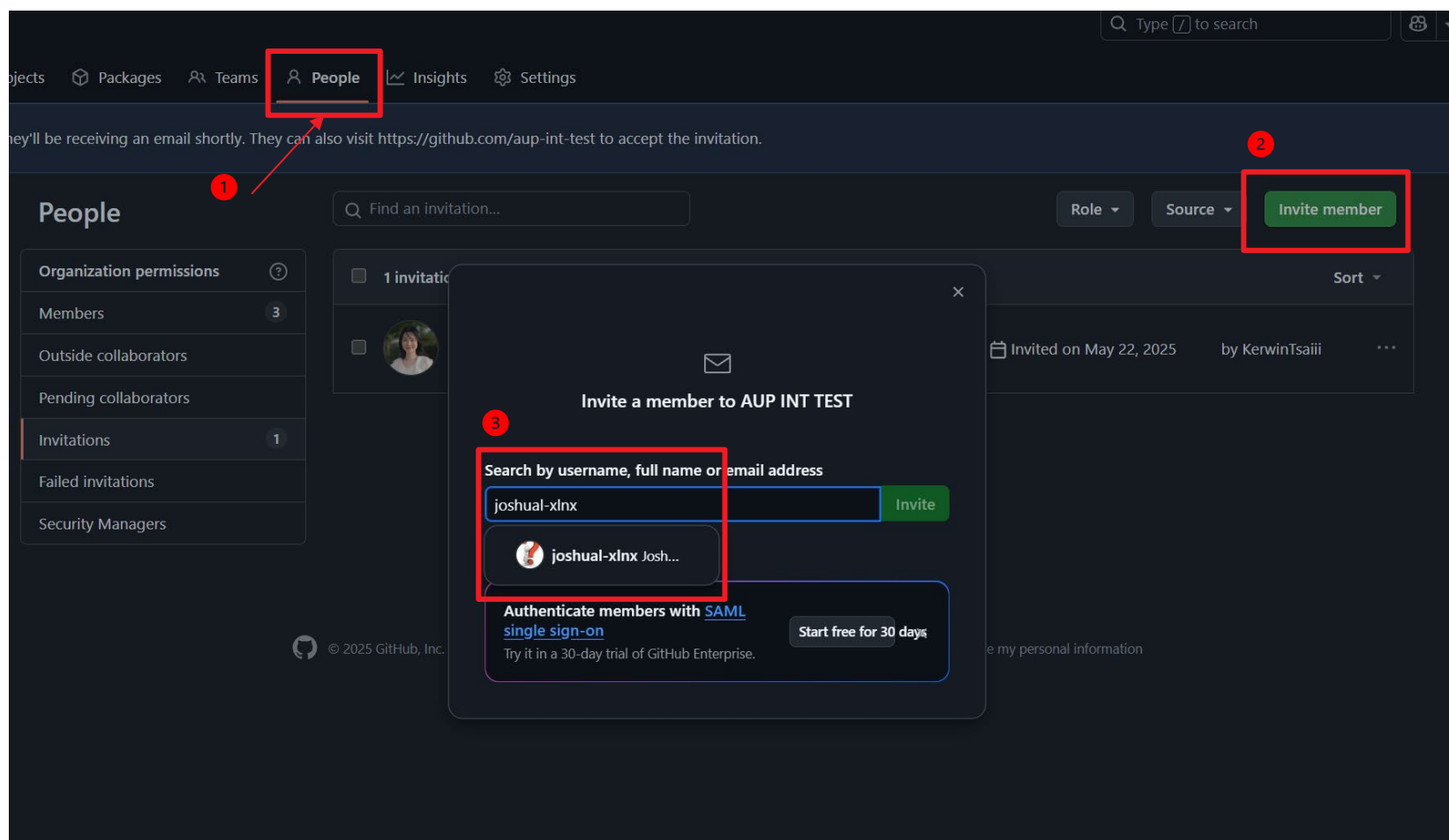
**AMD University Program
AMD Research and Advanced Development**



User Management

Add new user

- We managing user via a GitHub organization AUP INT TEST



Setup user teams

We use user teams to differentiate the computing resources available for different users.

Invite Wen Chen to AUP INT TEST

Give them an appropriate role in the organization and add them to some teams to give access to repositories.

Role in the organization

☒ Member

Members can see all other members, and can be granted access to repositories. They can also create new teams and repositories.

☐ Owner

Owners have full administrative rights to the organization and have complete access to all repositories and teams.

Teams — Optional

Find a team...

<div><input type="checkbox"/> cpu</div> <div>this team can only use cpu platform</div>	2 members 0 repositories
<div><input type="checkbox"/> gpu</div> <div>this team can only use gpu resources</div>	1 member 0 repositories
<div><input type="checkbox"/> npu</div>	1 member 0 repositories
<div><input type="checkbox"/> official</div> <div>AUP official staff- can use all resources</div>	2 members 0 repositories

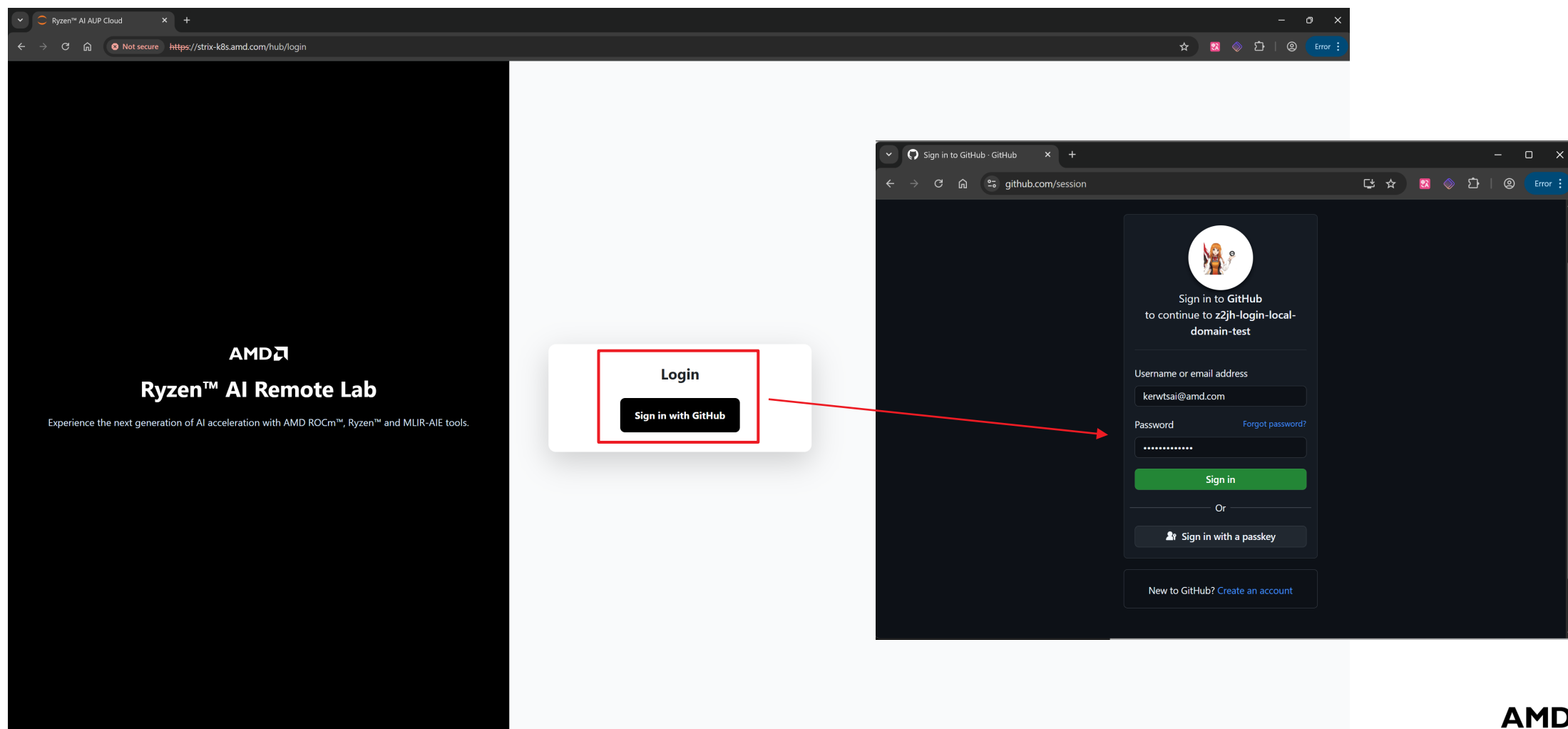
Send invitation



User Interface

GitHub SSO (Single Sign On)

<https://strix-k8s.amd.com>



Select you host

Server Options

Select your desired host

CPU

Run my server for (minutes)

CPU

ROCm Pytorch

Rialto NPU

MLIR AIE NPU

RyzenAI ONNEX NPU

Start

1

The types of hosts you can choose depend on your team's permissions. If you don't see any options, please contact the admin.

Server Options

Select your desired host

CPU

Run my server for (minutes):

10

Start

2

Run my server for e.g. 20 min

3

Click Start

Notebook page

File Edit View Run Kernel Tabs Settings Help

introduction.ipynb

introduction.ipynb

Python 3 (ipykernel)

introduction.ipynb

Markdown

introduction.ipynb Python 3 (ipykernel)

introduction.ipynb

Filter files by name

Examples /

Name	Last Modified
models	15 days ago
resnet50	15 days ago
scenedit	15 days ago
ssd_face_de...	15 days ago
superres	15 days ago
yolov8	15 days ago
0891.png	3 years ago
0892.png	3 years ago
0893.png	3 years ago
0894.png	3 years ago
0895.png	3 years ago
0896.png	3 years ago
0897.png	3 years ago
0898.png	3 years ago
0899.png	3 years ago
0900.png	3 years ago
resnet50_tu...	2 days ago
test_sample...	last month

Notebook Workspace with NFS support

Introduction to Ryzen AI NPU, Docker Integration, and AMD XDNA Examples

Maintained by yuzhoulu (yuzhoulu@amd.com)

Overview

This notebook provides a comprehensive introduction to deploying AI models on AMD's Ryzen AI Neural Processing Units (NPUs) using k8s containers. It also explores the utilization of AMD's XDNA architecture through practical examples. The content is designed to guide users through setting up the necessary software stacks, understanding the tools involved, and recognizing the considerations when integrating these technologies into different environments.

Ryzen AI and XDNA Architecture

AMD's Ryzen AI leverages the XDNA architecture, a neural processing unit microarchitecture based on technology acquired from Xilinx. XDNA enables efficient execution of AI workloads directly on the processor, providing significant performance and power efficiency benefits. The architecture supports various AI operations and is integrated into AMD's consumer PC processors, branded as Ryzen AI.

Currently, there are 2 sets of software stack for AMD NPUs.

1. [Rialto](#) is an open-source exploration framework for first time users of the AMD Ryzen AI Neural Processing Unit (NPU) and is developed by teams from the AMD Research and Advanced Development group and the AMD University Program.
2. [MLIR-AIE](#) compiler is an open-source research project from the Research and Advanced Development group (RAD) at AMD. This project is primarily intended to support tool builders with convenient low-level access to devices and enable the development of a wide variety of programming models from higher level abstractions.

Here we try to use ONNX with [Ryzen AI Software Platform](#) to run ML inference examples on NPU.

Trained Models

TensorFlow PyTorch ONNX

Quantization and ONNX Conversion

Microsoft Olive Quantizer AMD Vitis™ AI Quantizer

ONNX Runtime Execution Provider

Learning Materials

Choose another host

