



AUP Remote Lab User Manual (Admin)

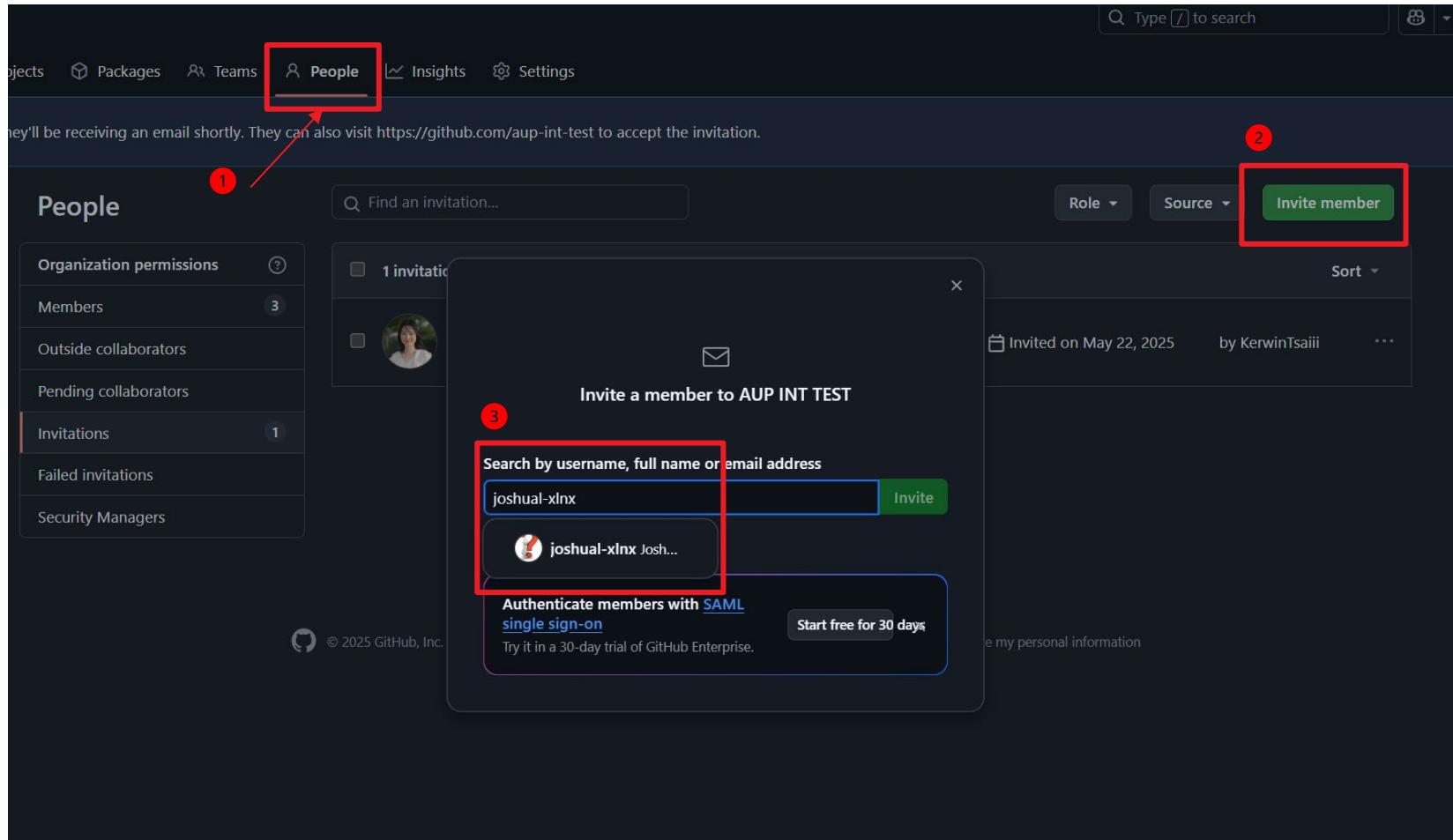
**AMD University Program
AMD Research and Advanced Development**



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...

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

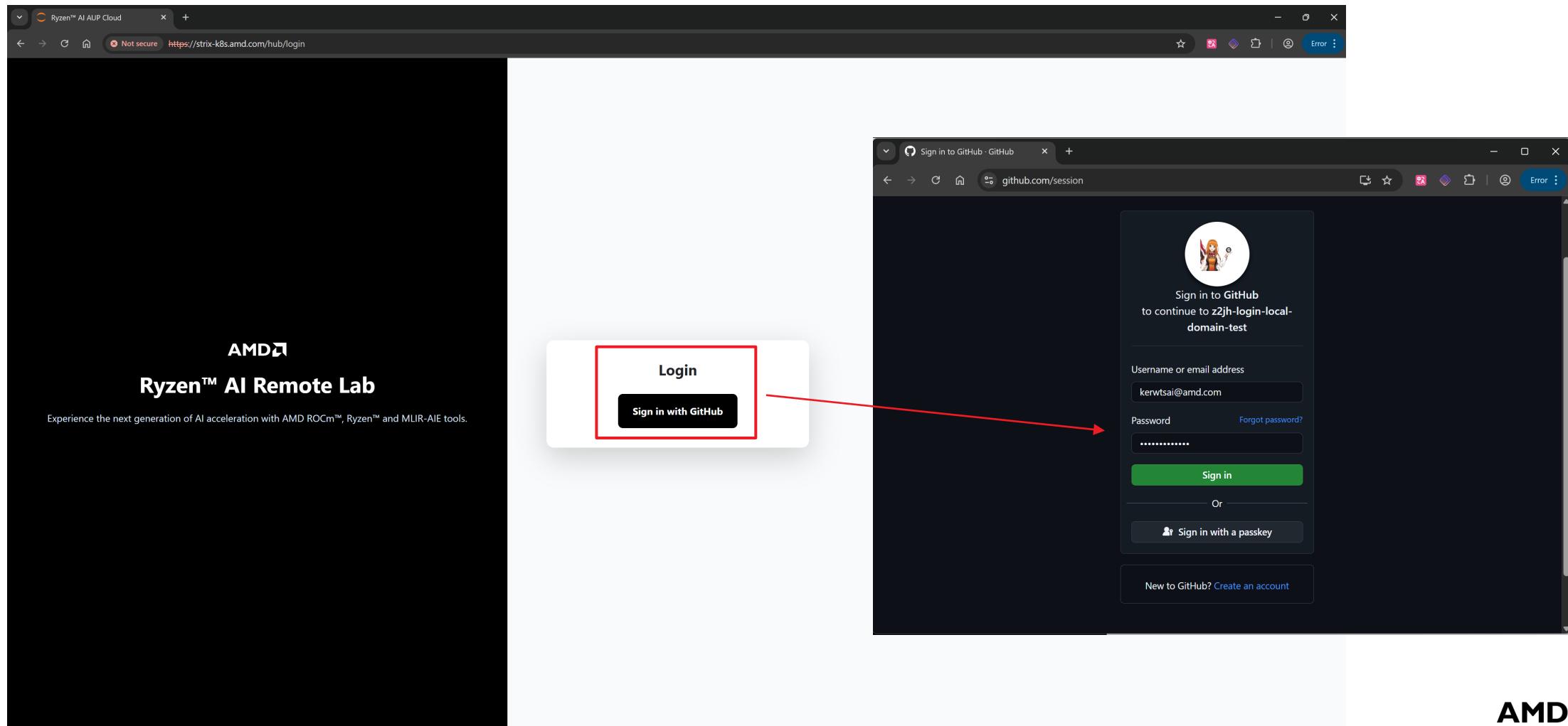
Send invitation



User Interface

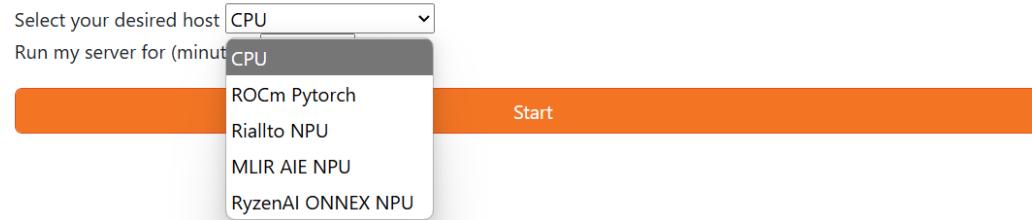
GitHub SSO (Single Sign On)

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



Select your host

Server Options



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



2

Run my server for
e.g. 20 min

3

Click Start

Notebook page

The screenshot shows a Jupyter Notebook interface with a red border around the main content area. On the left, there is a sidebar with a file browser and a red box highlighting the 'File' menu. The main area contains a Markdown cell with the following content:

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

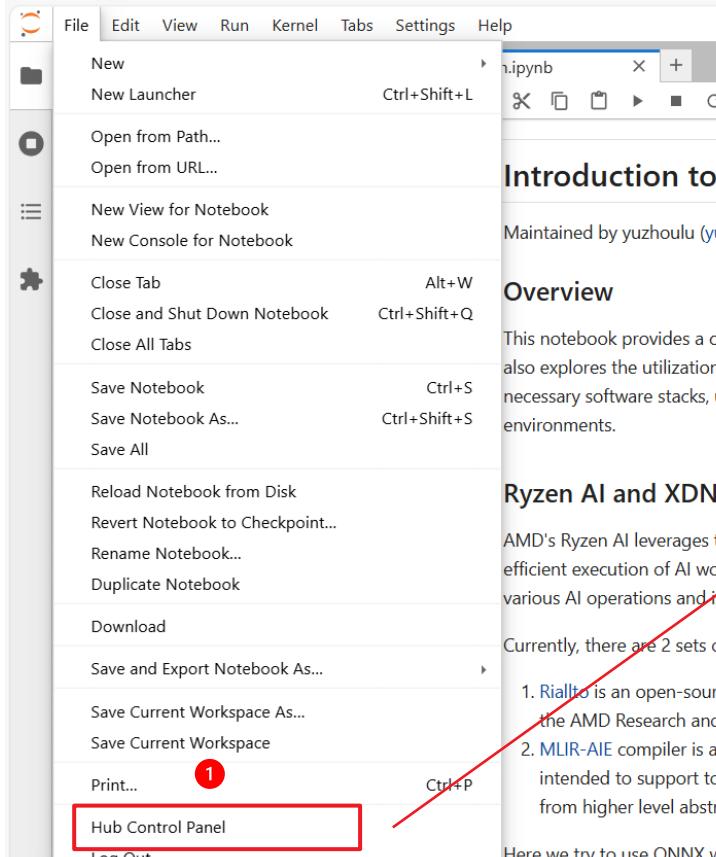
Microsoft Olive Quantizer AMD Vitis™ AI Quantizer

ONNX Runtime Execution Provider

Notebook Workspace with NFS support

Simple 0 2 Python 3 (ipykernel) | Idle Mode: Command Ln 1, Col 1 introduction.ipynb 1 Learning Materials

Choose another host



File Edit View Run Kernel Tabs Settings Help

- New
- New Launcher
- Open from Path...
- Open from URL...
- New View for Notebook
- New Console for Notebook
- Close Tab Alt+W
- Close and Shut Down Notebook Ctrl+Shift+Q
- Close All Tabs
- Save Notebook Ctrl+S
- Save Notebook As... Ctrl+Shift+S
- Save All
- Reload Notebook from Disk
- Revert Notebook to Checkpoint...
- Rename Notebook...
- Duplicate Notebook
- Download
- Save and Export Notebook As...
- Save Current Workspace As...
- Save Current Workspace
- Print... Ctrl+P
- Hub Control Panel**
- Log Out

1

Introduction to R

Maintained by yuzhoulu (yuzh)

Overview

This notebook provides a com also explores the utilization of necessary software stacks, unc environments.

Ryzen AI and XDNA

AMD's Ryzen AI leverages the efficient execution of AI workl various AI operations and is in

Currently, there are 2 sets of si

1. **Rialto** is an open-source i the AMD Research and Ac
2. **MLIR-AIE** compiler is an o intended to support tool | from higher level abstract

Here we try to use ONNX with

2

Stop My Server My Server

3

Home Admin Token

Server Options

Select your desired host CPU

CPU

Run my server for (minutes)

ROCm Pytorch

Rialto NPU

MLIR AIE NPU

RyzenAI ONNEX NPU

Start

Tens

AMD