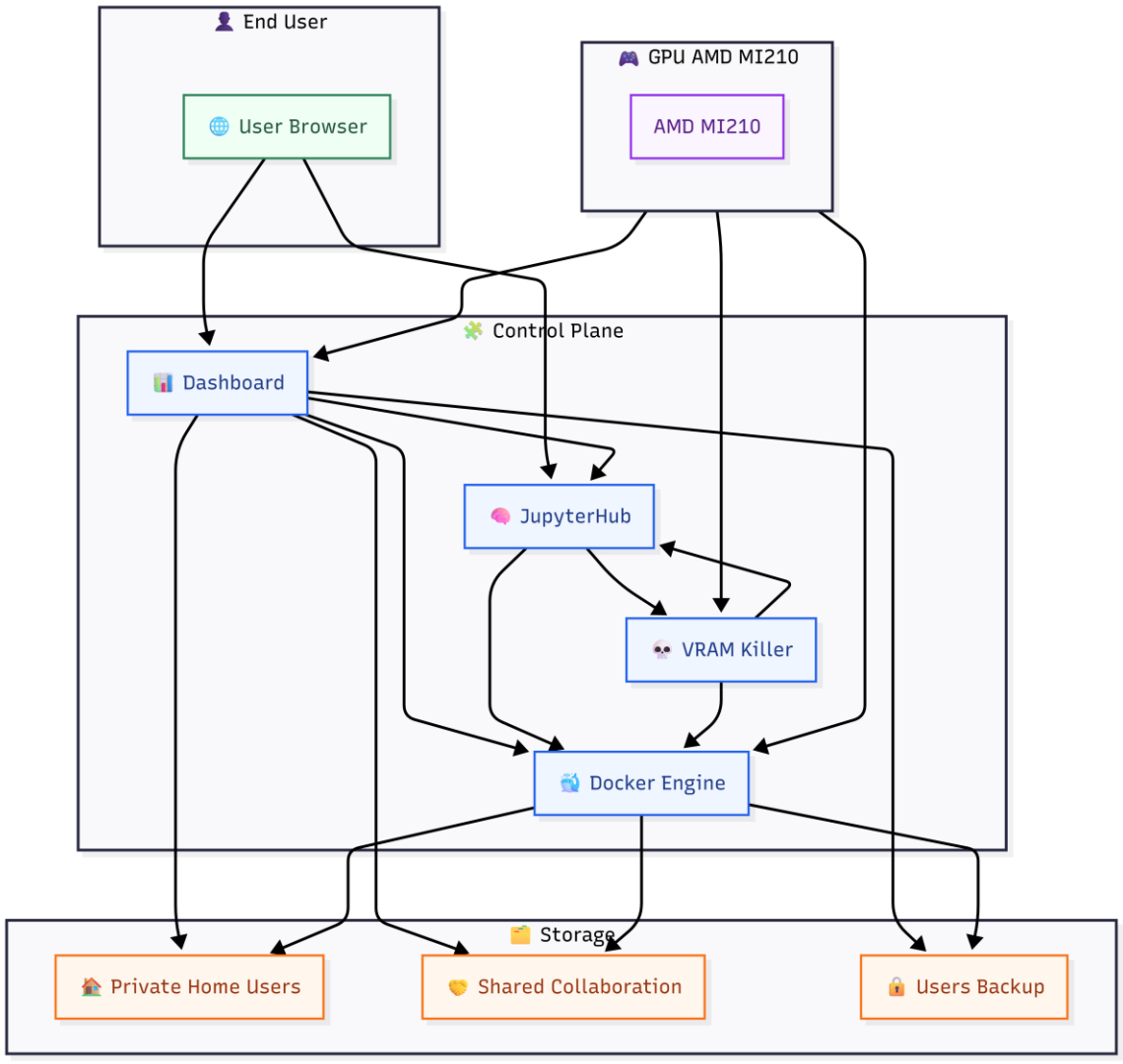
SIT JupyterHub System Architecture

System Overview

This service is a single-VM JupyterHub environment for shared-GPU notebook workloads, as shown in the diagram.



Main Architecture

- VM Operating System: Ubuntu Server 24.04

- GPU: AMD MI210 (shared by all users)
- Orchestration: Docker Compose
- Jupyter frontend: JupyterHub with DockerSpawner
- Authentication: local JupyterHub users with NativeAuthenticator
- User isolation: one Docker container per user
- Storage: a private home directory for each user
- Collaboration storage: separately mounted and read-only by default
- GPU policy: shared GPU with a best-effort target VRAM limit of 5 GB per user
- Operations view: dashboard

Service Specifications

- Supports up to 12 concurrent users
- Each user receives a 100 GB private home directory isolated from other users
- The VRAM limit for running jobs is 5 GB. If exceeded, the process will be terminated within 1 minute.
- Accessible through a web browser
- When a user's access ends, the system removes the user account, backs up the user's data, and retains it for 30 days.

Access URLs

- https://jupyterhub.sit.kmutt.ac.th is the access URL
- https://jupyterhub.sit.kmutt.ac.th/dashboard is the dashboard URL

Storage Layout

- /srv/jupyter/users/<username>: private home directory for each user
- /srv/jupyter/shared: shared collaboration storage
- /srv/jupyter-backup: per-user archive backups
- /srv/container-runtime: runtime data and operation-related mounts
- /opt/jupyterhub: production compose files and hub configuration
- /opt/jupyterhub/user-manage/: user management scripts

Monitoring Dashboard

- The dashboard provides a read-only view of current system information.
- Command-line checks are the source of truth.
- GPU usage is primarily obtained from rocm-smi, with fallback parsing when needed.
- vram-killer monitors GPU usage and logs policy enforcement.

Service Workflow

1. The user opens the URL in a web browser.
2. For first-time access, the user sign up and sets a password.
3. For subsequent access, sign in using your username and the password you set during your first login.
4. Use Notebook.

Notes

- Notebook kernels and running jobs continue to operate even if the browser disconnects, as long as the server is not stopped and the process is not terminated.