UXL Public Test Infrastructure PoC Update

Aaron Dron – Codeplay Software Ltd

August 2024

UXL Public Test Infrastructure PoC Goals

- Enable community contributions
  - Ideally at least a set of smoke tests should be performed on all contributions
- Enable members to enrol and support self-hosted test hardware
  - Having testing across multiple hardware vendors and compute platforms is critical
- Encourage members to develop in UXL repos directly, rather than in private internal repositories
  - This will help avoid large PRs etc
- Where possible prove viability of UXL projects with thorough, public testing

Scope Of Work 2024

- Administrating UXL GitHub organisation
- Help improve security of existing UXL repositories (spec, etc)
- Documenting the process of safely contributing self-hosted test hardware
- Documenting recommendations for working with ancillary test systems
- Migrating 1-2 repositories to the UXL, plus initial CI
- Contributing a small number of Nvidia GPU test systems
- Experimenting with Intel IDC GPU Max systems as GitHub Runners

Progress Update

- Improving OpenSSF scores for relevant repos
- Discussing migration timelines with oneMKL
- Assessing Codeplay's oneAPI Construction Kit as potential PoC migration
- Documenting strategies for self-hosted member provided runners
- Testing our example tooling for working with ancillary test systems
- Initial tests with small cluster of GPU Max systems in Kubernetes

Blockers and Unknowns

- Need a location for private documentation accessible to UXL Members
- Still determining best choices for Nvidia GPUs in Codeplay test systems
- Discussing strategies for driver updates on IDC GPU Max test systems
- Working to determine best practices for allowing IDC GPU Max runners to be used by other projects (e.g. existing containerised environments)

Not Part of the Public Test Infrastructure PoC

- Migrating Intel oneAPI repositories to UXL Organisation
  - The exception potentially being oneMKL's migration
- Providing and maintaining runner systems
  - Aside from the small number directly connected to the PoC
- Building functionality to support Just-In-Time cloud-based runners
  - E.g. efforts to minimise AWS costs
- Creation and maintenance of build environments / containers
  - Aside from our viability assessments on IDC GPU Max systems
- Releases and artifact management

Reach Out

- aaron@codeplay.com
- aaron.dron@intel.com
- UXL Slack