

# UXL Foundation: Drive an Open Standard Accelerator Software Ecosystem

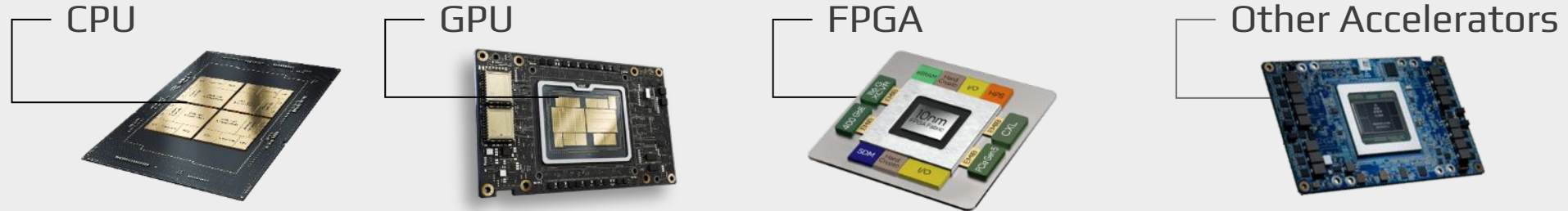
UXL Foundation Steering Committee Members



# Agenda

<b>UXL Foundation introduction</b>	Rod Burns, Steering Committee Chair
<b>UXL Foundation Projects</b> <b>oneAPI Specification</b> <b>Working Groups and SIGs</b>	Andy Wafaa, Arm, Open Source Working Group Chair Robert Cohn, Intel, Specification Working Group Chair Penporn Koanantakool, Google Cloud, AI SIG Chair
<b>What's happening in the UXL Foundation?</b>	Our members
<b>How to contribute</b>	Rod Burns, Steering Committee Chair

# Solving the Challenge of Diverse Hardware Acceleration



70% of developers target heterogeneous systems that use more than one kind of processor or core<sup>1</sup>

Developer Challenges: Multiple Architectures, Vendors, and Programming Models



Open, Standards-based, Multiarchitecture Programming

[Video: What is oneAPI? Overview & Benefits](#)

# Commitment to Open, Scalable Acceleration



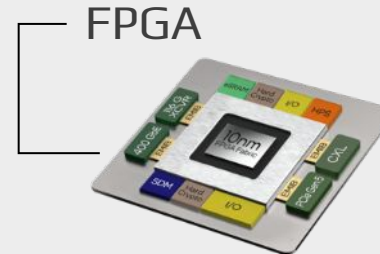
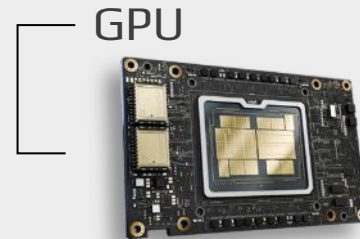
Linux Foundation governed open industry foundation driving a vendor-neutral software ecosystem for multi-architecture accelerated computing

- oneAPI Specification
- Libraries deliver cross-vendor building blocks for developers



Open, Standards-based, Multi-architecture Programming

Performance | Productivity | Freedom from Vendor Lock-In



# Open Standards Programming for All Accelerators

## Mission

- Build a **multi-architecture multi-vendor software ecosystem** for all accelerators
- **Unify** the heterogeneous compute ecosystem **around open standards**
- Build on and expand **open source projects for accelerated computing**



Unified Acceleration Foundation

# Unified Acceleration Foundation

## Steering Members

arm

FUJITSU

Google Cloud

Imagination

intel

Qualcomm

SAMSUNG

vmware<sup>®</sup>  
by Broadcom

## Contributor Members

AKHETONICS

FIXSTARS  
*Speed up your Business*

AXELERA  
ARTIFICIAL INTELLIGENCE

CloudsAI

codasip

CODE RECKONS  
*Science to the CORE*



EMBECOSM<sup>®</sup>



G  
GNAC Intelligence



MARKET  
POTENTIAL

XIANGDIXIAN<sup>®</sup>  
COMPUTING TECHNOLOGY  
象帝先计算技术

UXL  
Unified Acceleration Foundation

# UXL Foundation Governance

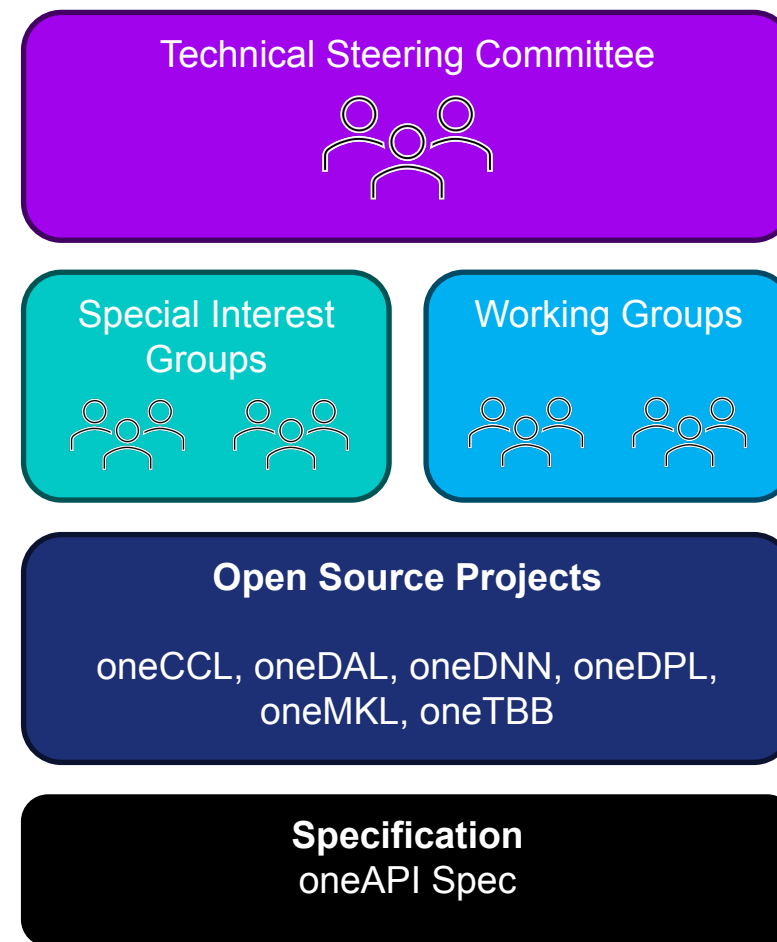
Using Best Practices

Joint Development Foundation governance

**SIGs:** AI, Hardware, Language, Math, Safety Critical

**Working Groups:** Specification, Open Source

Anyone can participate in SIGs and Open Source Projects. Contributors to the Specification must sign the Membership Agreement or Non-Member Feedback Agreement.



Unified Acceleration Foundation

# oneAPI Specification and Projects

KHRONOS<sup>®</sup>  
GROUP



- SYCL powers the UXL Foundation libraries

Heterogeneous, cross-vendor  
programming model



Specification





# Building on Strong Foundations

2019-2020

2021-2022

2023-2024

## SPECIFICATION

- oneAPI provisional specification
- Technical Advisory Boards established

- oneAPI Specification delivered
- Created additional TABs
- Held oneAPI DevSummits

- Members join the UXL Foundation
- Specification migrated to UXL Foundation
- SIGs and Working Group established to coordinate feedback and work

- Open source implementations of oneAPI specification made available
- Initial targets for Intel processors

- Nvidia and AMD targets added to oneMKL and oneDNN
- Arm target added to oneDNN
- Broad adoption of libraries

- Open Source projects migrated to UXL Foundation
- Arm target added to oneDAL

## OPEN SOURCE

Coordinating committee

# Projects Adopting UXL Foundation

Using libraries for cross-vendor portability



## GROMACS

Project uses oneMKL math library to target multiple architectures



## TensorFlow and PyTorch

Projects use oneDNN library for accelerated graph optimization



## Ginkgo

Project uses oneMKL and oneDPL for optimized math and ISO C++ routines



## US National Laboratories

Projects using oneMKL to deploy big science applications across supercomputers and for exascale

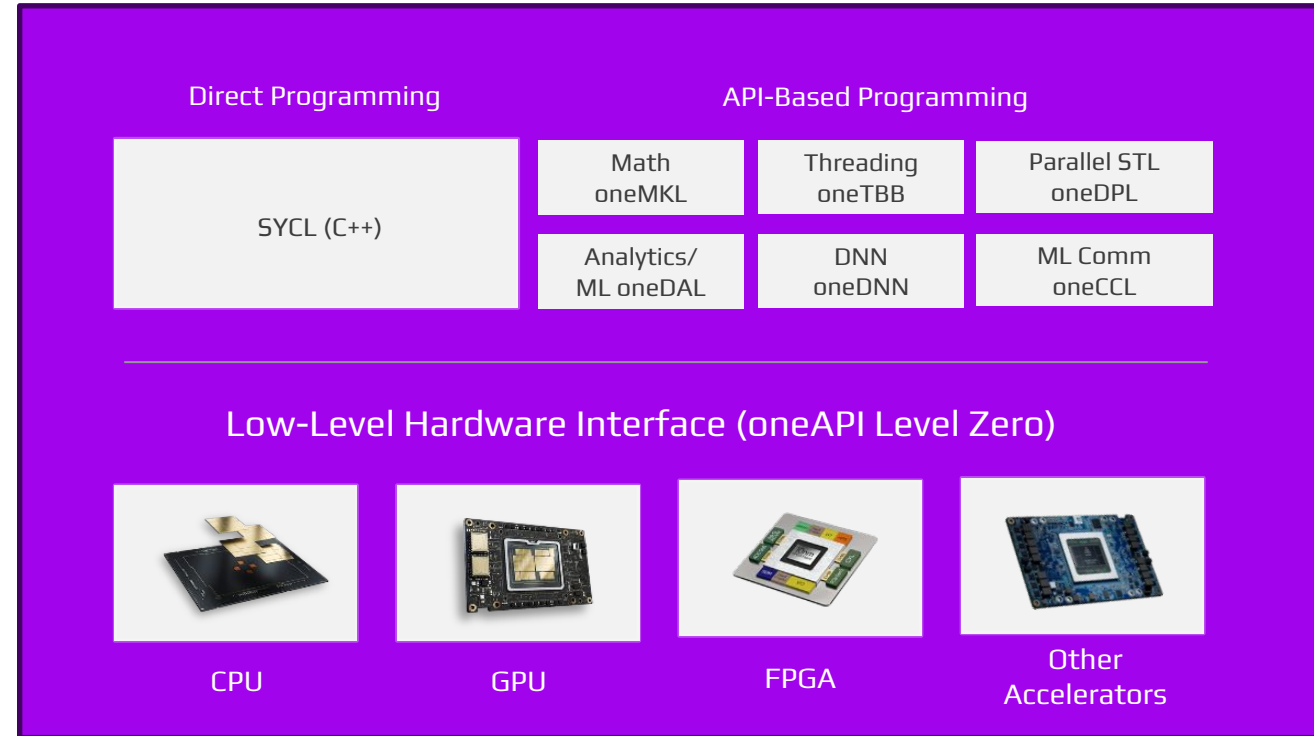
And many more...

# Use the UXL Foundation Libraries Today

Download the project binaries via the oneAPI Base Toolkit

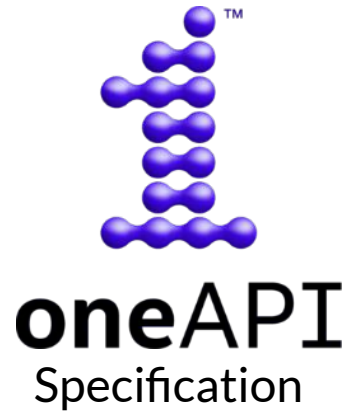
Or build the projects from open source

Target AMD, Arm, Intel and Nvidia processors



# Building alliances

## UXL Foundation



## Project Dependencies



## Future Collaborations



Defines the specification used by  
UXL Foundation projects

Integration of UXL projects and  
demonstrations

# UXL Foundation Working Groups

## Open Source Working Group

### Work Packages

Adopting Open Governance and Best Practices for Open Source, Security

CI/CD for multiple architectures

Demonstration projects

Migration to UXL Foundation Infrastructure

## Specification Working Group

Forum for issues impacting multiple components

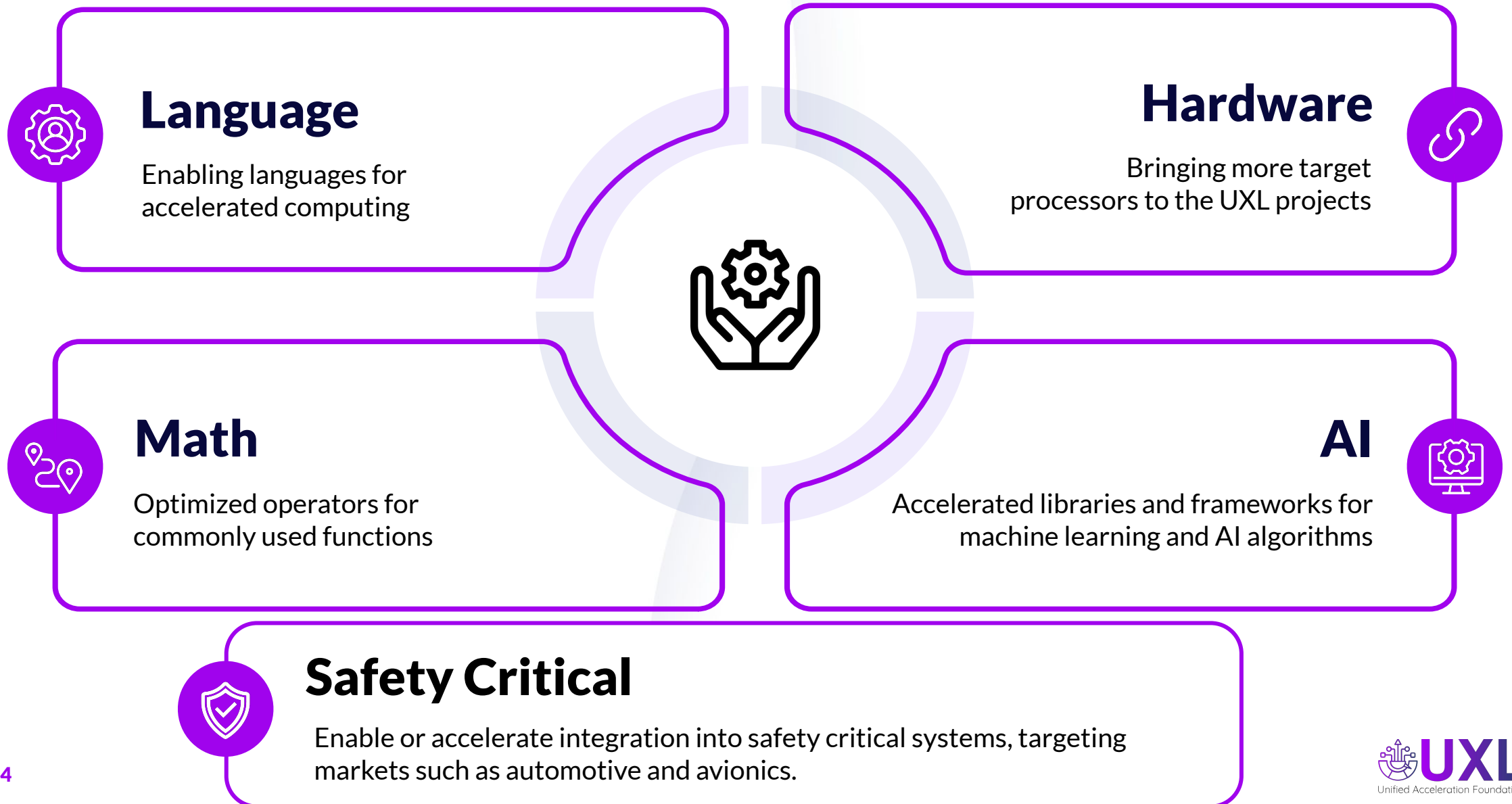
Adopting best practices for community feedback

Coordination of release schedules

Migrating to UXL Foundation Infrastructure

Become a member to join the Working Groups and help shape the projects

# UXL Foundation SIGs



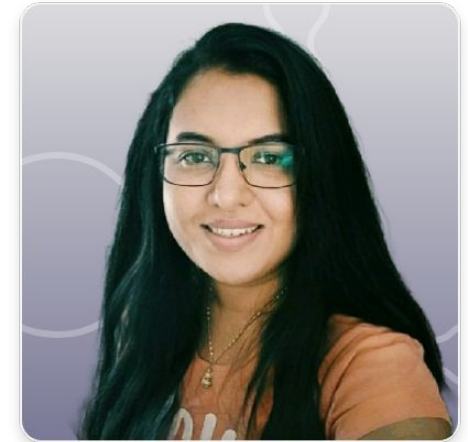
# Fujitsu

## Masahiro Doteguchi & Dr. Priyanka Sharma

Fujitsu's presence in OSS community since 2005, via open-source development in mission-critical systems and in the Supercomputer Fugaku & we are further committed to continuing with this legacy through FUJITSU-MONAKA (2nm Arm CPU)

Optimized and ported the oneDNN DL process library software for the Arm SVE instruction so that it can run at high speed on the Fugaku supercomputer.

Recent success towards porting of oneDAL on Arm to accelerate ML workloads on Arm. This is also one of the first OSS contributions to UXL foundation



# Google Cloud

Penporn Koanantakool

TensorFlow, JAX, and OpenXLA use oneAPI Deep Neural Network (oneDNN) library to accelerate performance on x86 and aarch64 CPUs, and Intel GPUs.

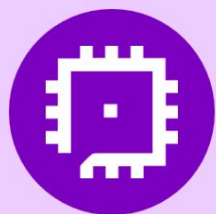
Google Cloud's Cloud HPC Toolkit includes oneAPI Math Kernel Library (oneMKL).





# Imagination

Dave Murray



GPU



CPU



RT



AI

oneMKL

oneDNN

SYCLomatic



Auto



PC



Mobile



IoT



THE  
AUTOWARE  
FOUNDATION



## Robert Cohn

- Unlocking the full potential for accelerated computing requires cross platform software eco-system based on open source & open standards
- Intel contributed projects to UXL for: math, AI, data parallel & distributed computing...
- With UXL and LF, we are fully embracing open governance to build strong developer communities for accelerated computing



Qualcomm

# Qualcomm

Dr. Vinesh Sukumar



# Samsung

Hanwoong Jung

- Parallel programming model for memory-centric computing
  - SYCL extensions for PIM/PNM
- Deep learning compiler/runtime
  - Use oneDNN and unified runtime in oneAPI
- With UXL, we expect the open collaboration will bolster programming models and SW stacks to effectively accelerate AI/HPC applications on NPU, PIM/PNM, and RISC-V CPU/accelerator by Samsung

**SAMSUNG**



# VMware

vmware®  
by Broadcom

Ramesh Radhakrishnan

- Open Source and Standardization is a core tenant of VMware culture
- VMware offers customers choice and flexibility through our growing Private AI Ecosystem
- Streamlining interoperability in both hardware and software is key for adoption of new architectures



# VMware Private AI Open Ecosystem



Falcon



Llama 2



Mistral



MPT



StarCoder



WizardML



anyscale

cnvrg.io



DOMINO



DKube



FedML



Hugging Face



Kubeflow



NVIDIA  
NEMO



oneAPI



PyTorch



run:ai



Weights & Biases

watsonx

HCL



kyndryl

NTT DATA



DELL Technologies

Hewlett Packard  
Enterprise

Lenovo

AMD

intel

NVIDIA



VMware Cloud Foundation™

# Contribute

# How to Contribute?

 **rfc: proposal for block level APIs** RFC

#1852 opened last month by mgouicem

 **rfcs: add proposal on reorganizing GPU abstractions** RFC

#1840 opened on Mar 25 by densamoilov • Approved

 **rfcs: proposal for a verbose logging mechanism** RFC

## Bring your feedback

- Submit issues
- Join the SIGs
- Join the Working Groups
- Contribute to project RFCs

 **Graphs with single StaticReshape or StaticTranspose fail** enhancement

#1858 opened 3 weeks ago by richard-lemurian

 **running destructors before completion of a primitive** enhancement

#1814 opened on Feb 29 by rschoen2

 **GEMM API for efficient LLM inference with W8A16** enhancement platform:arch64

#1788 opened on Jan 20 by oleotiger

## Contribute to projects

- New features
- Expand hardware support
- Documentation

 **cpu: x64: enable groups and multidim along IC for scales in brgemm ma...** main

OpenSSF scorecard #36: Commit c451c68 pushed by perflibs

 **graph: interface: refactor compile partition cache key** main

OpenSSF scorecard #35: Commit 55d48ac pushed by perflibs

 **benchdnn:graph:correct some errors in jsons** main

OpenSSF scorecard #34: Commit 9d737d2 pushed by perflibs

## Contribute resources

- Build infrastructure
- Testing



# Membership

## Steering Member

**\$20,000/year\***

### Steering Member

- Seat on Steering Committee
- Influence direction
- Voting Rights

## General Member

**\$5,000/year\***

### General Member

- Voting rights in Working Groups
- Influence project work packages
- Co-marketing

## Contributor Member

**\$0/year**

### Contributor Member

- Join Working Groups
- Participate in work packages
-

# Join Us!

Check out the UXL Foundation & oneAPI specification



- [UXLFoundation.org](https://uxlfoundation.org)
- [oneAPI.io](https://oneapi.io)



**Join Our  
Mailing Lists**



**Join  
Slack**

# Q&A