Welcome to UCF Virtual Workshop

Pavel Shamis & Gilad Shainer

UCF Workshop, 2020
This is an open, public standards setting discussion and development meeting of UCF. The discussions that take place during this meeting are intended to be open to the general public and all work product derived from this meeting shall be made widely and freely available to the public. All information including exchange of technical information shall take place during open sessions of this meeting and UCF will not sponsor or support any closed or private working group, standards setting or development sessions that may take place during this meeting. Your participation in any non-public interactions or settings during this meeting are outside the scope of UCF’s intended open-public meeting format.
Housekeeping Notes

- Workshop page with up-to-date schedule

- Moderators
  - Pavel Shamis, Cydney Stevens, Brian Sparks

- All presentations and discussions are video recorded
  - Zoom explicitly asks participants for a consent

- Etiquette
  - Please stay on mute unless you are presenting 😊
  - Presenters are encouraged to turn video on
  - If you participate in a discussion or ask a question, it is encouraged to turn the video on
  - If you have questions during a presentation, please wait until end of presentation or use zoom “raise the hand” to signal that you have questions. You also can use Zoom chat to ask questions.
  - You can use Zoom whiteboard and screen sharing for discussion sessions
# Agenda – Day 1

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Topic</th>
<th>Speaker/Moderator</th>
</tr>
</thead>
<tbody>
<tr>
<td>11/30</td>
<td>08:00-09:00</td>
<td>UCF State of the Union</td>
<td>Gilad Shainer, Nvidia&lt;br&gt;Pavel Shamis (Pasha), Arm</td>
</tr>
<tr>
<td></td>
<td>09:00-10:00</td>
<td>GPU memory support</td>
<td>Yossi Itigin, Nvidia</td>
</tr>
<tr>
<td></td>
<td>10:00-10:30</td>
<td>MPICH/UCX Update</td>
<td>Ken Raffenetti, Argonne National Laboratory</td>
</tr>
<tr>
<td></td>
<td>10:30-11:15</td>
<td>RDMA-CORE: DMA-BUF based GPU RDMA Support</td>
<td>Jianxin Xiong, Intel</td>
</tr>
<tr>
<td></td>
<td>11:15-12:15</td>
<td>UCX for Apache Spark</td>
<td>Peter Rudenko, Nvidia</td>
</tr>
<tr>
<td></td>
<td>12:15-13:15</td>
<td>UCX Python – Dask/RAPIDS</td>
<td>Ben Zaitlen, NVIDIA&lt;br&gt;Peter Entschev, NVIDIA&lt;br&gt;Matthew Baker, ORNL</td>
</tr>
</tbody>
</table>
## Agenda – Day 2

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Session</th>
<th>Presenter/Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>12/01</td>
<td>08:00-08:40</td>
<td>UCF - Future directions</td>
<td>Steve Poole, Los Alamos National Laboratory</td>
</tr>
<tr>
<td></td>
<td>08:40-09:40</td>
<td>UCP Protocols v2</td>
<td>Yossi Itigin, Nvidia</td>
</tr>
<tr>
<td></td>
<td>09:40-10:40</td>
<td>UCP Active messages API</td>
<td>Mikhail Brinskii, Nvidia</td>
</tr>
<tr>
<td></td>
<td>10:40-11:40</td>
<td>UCX development in Huawei</td>
<td>Alex Margolin, HPC software architect and team leader, Huawei</td>
</tr>
<tr>
<td></td>
<td>11:40-12:20</td>
<td>Open Smart NIC API - State of the Union</td>
<td>Steve Poole, Los Alamos National Laboratory</td>
</tr>
</tbody>
</table>
## Agenda – Day 3

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Session</th>
<th>Speaker/Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>12/02</td>
<td>08:00-09:00</td>
<td>BlazingSQL with UCX</td>
<td>Rodrigo Aramburu, Felipe Aramburu, BlazingSQL</td>
</tr>
<tr>
<td></td>
<td>09:00-10:00</td>
<td>Charm++ with UCX</td>
<td>Nitin Bhat, Charmworks, Jaemin Choi, University of Illinois Urbana-Champaign</td>
</tr>
<tr>
<td></td>
<td>10:00-10:30</td>
<td>ROCM support in UCX: Status and Roadmap</td>
<td>Sourav Chakraborty, AMD</td>
</tr>
<tr>
<td></td>
<td>10:40-11:40</td>
<td>UCX counters in Score-P and Vampir</td>
<td>Shuki Zanyovka, HPC and Networking architect at Huawei</td>
</tr>
<tr>
<td></td>
<td>11:40-12:40</td>
<td>Unified Communication Datatypes - State of the Union</td>
<td>Pavan Balaji, Argonne National Laboratory</td>
</tr>
<tr>
<td></td>
<td>12:40-13:00</td>
<td>Arm IP building blocks and standards for SmartNIC</td>
<td>Kshitij Sudan, Arm</td>
</tr>
</tbody>
</table>
## Agenda – Day 4

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Topic</th>
<th>Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>12/03</td>
<td>08:00-09:00</td>
<td>UCC: Design and Implementation of Next Generation Collectives Library</td>
<td>Manjunath Gorentla Venkata, Nvidia</td>
</tr>
<tr>
<td></td>
<td>09:00-09:30</td>
<td>One-to-many UCT transports, part I: Shared-memory</td>
<td>Alex Margolin, HPC software architect and team leader, Huawei</td>
</tr>
<tr>
<td></td>
<td>09:30-10:00</td>
<td>One-to-many UCT transports, part II: Multicast</td>
<td>Morad Horany, HPC software developer, Huawei</td>
</tr>
<tr>
<td></td>
<td>10:00-11:00</td>
<td>Until UCC is available - UCG status update</td>
<td>Alex Margolin, HPC software architect and team leader, Huawei</td>
</tr>
<tr>
<td></td>
<td>11:00-11:45</td>
<td>RDMA-CORE Linux kernel and user space updates</td>
<td>Jason Gunthorpe, Nvidia</td>
</tr>
<tr>
<td></td>
<td>11:45-12:45</td>
<td>Scaling Facebook’s Deep Learning Recommender Model (DLRM) with UCC/XCCL</td>
<td>Josh Ladd, Nvidia Srinivas, Facebook</td>
</tr>
<tr>
<td></td>
<td>12:45-13:30</td>
<td>Open Smart NIC API - OpenSHMEM I/O Extensions for Fine-grained Access to Persistent Memory Storage</td>
<td>Megan Grodowitz, Arm</td>
</tr>
</tbody>
</table>

© 2020 UCF Consortium
UCF – State of the Union

Gilad Shainer
**MISSION**: Collaboration between industry, laboratories, and academia to create production grade communication frameworks and open standards for data centric, ML/AI, and high-performance applications

- **Projects & Working Groups**
  - UCX – Unified Communication X – www.openucx.org
  - SparkUCX – www.sparkucx.org
  - OpenSNAPI – Smart NIC Project
  - UCC – Collective Library
  - UCD – Advanced Datatype Engine
  - HPCA Benchmark – Benchmarking Effort

- **Board members**
  - **Jeff Kuehn**, UCF Chairman (Los Alamos National Laboratory)
  - **Gilad Shainer**, UCF President (Nvidia)
  - **Pavel Shamis**, UCF Treasurer (Arm)
  - **Brad Benton**, Board Member (AMD)
  - **Pavan Balaji**, Board Member (Argonne National Laboratory)
  - **Sameh Sharkawi**, Board Member (IBM)
  - **Dhabaleswar K. (DK) Panda**, Board Member (Ohio State University)
  - **Steve Poole**, Board Member (Open Source Software Solutions)

[Join](https://www.ucfconsortium.org) or [info@ucfconsortium.org](mailto:info@ucfconsortium.org)
Unified Communication X (UCX)

A new project based on concept and ideas from multiple generations of HPC networks stacks

- Performance
- Scalability
- Efficiency
- Portability

Decades of community and industry experience in development of HPC & ML network software
Why UCX?

- ** Networks **
  - ** MPI **
  - ** PGAS **
  - ** ML/AI Frameworks **
  - ** Data Analytics **
  - ** Storage **
  - ** DPU APIs **

  - ** Host Memory: x86, Arm, Power **
  - ** GPU Memory: AMD RoCM, NVIDIA CUDA **
  - ** DPU Memory **

- ** High-Performance Universal Data Mover **

© 2020 UCF Consortium
UCX High-level Overview

Applications
- HPC (MPI, SHMEM, …)
- Storage, RPC, AI
- Web 2.0 (Spark, Hadoop)

UCP – High Level API (Protocols)
- Transport selection, multi-rail, fragmentation
- HPC API: tag matching, active messages
- I/O API: Stream, RPC, remote memory access, atomics
- Connection establishment: client/server, external

UCT – Low Level API (Transports)
- RDMA
- GPU / Accelerators
- Others
- RC
- DCT
- UD
- iWarp
- CUDA
- AMD/ROCM
- Shared memory
- TCP
- OmniPath
- Cray

Hardware
- OFA Verbs Driver
- Cuda
- ROCM
UCX Performance-portability

- Support for x86_64, Power 8/9, Arm v8
- U-arch tuned code for Xeon, AMD Rome/Naples, Arm v8 (Cortex-A/N1/ThunderX2/Huawei, Fujitsu A64FX)
- First class support for AMD and Nvidia GPUs
- Runs on Servers, Raspberry PI like platforms, SmartNIC, Nvidia Jetson platforms, etc.
We Love Testing

Over 100,000 tests per commit
220,000 CPU hours per release
Annual Release Schedule

- v1.8.0 - July 2020
- v1.9.0 – September 2020
- v1.10.0 – End of December 2020
UCX Latest Releases

- **1.9.0 (September 19, 2020)** - [https://github.com/openucx/ucx/releases/tag/v1.9.0](https://github.com/openucx/ucx/releases/tag/v1.9.0)

**UCX Core**
- Added a **new class of communication APIs** `*_nbx` that enable API extendability while preserving ABI backward compatibility
- Added **asynchronous event** support to UCT/IB/DEVX
- Added support for **latest NVIDIA CUDA** library version
- Added support for **AMD ROCm 3.7 and above**
- Added **new tests** for AMD ROCm
- Added performance optimization for **Fujitsu A64FX** with InfiniBand
- Added support for **relaxed-order PCIe access** in IB RDMA transports
- Added **new TCP connection manager**
- Added **flow control** for RDMA read operations
- Improved **performance in active message** flow

**UCX Java (API Preview)**
- Added support for UCX shared library loading from both classpath and LD_LIBRARY_PATH
- Added configuration map to `ucp_params` to be able to set UCX properties programatically
UCX Users

- MPI implementations: MPICH, Open MPI, Mellanox MPI, Huawei MPI
- PGAS: GasNET
- OpenSHMEM: OSSS SHMEM, Sandia SHMEM, Open MPI SHMEM
- Charm++
- RAPIDS / DASK
- NVIDIA’s NCCL
UCX – Useful links

- Code
  - https://github.com/openucx/

- Website
  - www.openucx.com

- Mailing list
  - https://elist.ornl.gov/mailman/listinfo/ucx-group

- Contributor agreement
  - https://www.openucx.org/license/

- User documentation
  - https://openucx.readthedocs.io/
UCF Project Incubator

Collective Communication API
Web https://www.ucfconsortium.org/projects/ucc/
GIT https://github.com/openucx/ucc
https://github.com/openucx/xucg
https://github.com/openucx/xcl
https://github.com/openucx/ucc_spec
https://github.com/openucx/torch

API for Smart NIC & DPU programmability
Web https://www.ucfconsortium.org/projects/opensnapi/
GIT https://github.com/openucx/shmem-opensnapi

info@ucfconsortium.org