SC18 UCX BOF: OSSS OpenSHMEM-on-UCX

Tony Curtis <anthony.curtis@stonybrook.edu>
Abdullah Shahneous Bari Sayket
Wenbin Lü
Dr. Barbara Chapman
UCF 2018 Workshop: OSSS-UCX OpenSHMEM

- OpenSHMEM tl;dr
  - 1-sided PGAS library
  - C/C++
  - Fortran (currently deprecated, but you never know)
  - API
    - Init/finalize
    - Rank query
    - RDMA calls: put, get, put_nbi … direct to variables
    - AMOs: fetch/set, swaps, fetch-add, add, bitwise, …
    - Reductions
    - Scatter/gather
    - Broadcast
UCF 2018 Workshop: OSSS-UCX OpenSHMEM

- This is not my beautiful SHMEM: how did I get here?
  - SHMEM developed about the same time as MPI
  - Implementations diverged: no standard
    - E.g. SGI, Cray, Quadrics, GPSHMEM, IBM
    - shmem_init(void) vs. start_pes(int npes)
  - #ifdef mess in programs
  - ~ 2010 began standardizing OpenSHMEM
    - Vendors begin transitioning to OpenSHMEM conformance
  - Now at specification 1.4, working on 1.5
  - OpenSHMEM has 22 “signed up” contributing entities
  - We started using UCX in the 1.3 to 1.4 transition
    - GASNet before that
SC18 UCX BOF: OSSS OpenSHMEM-on-UCX

● Reference Implementation
  ○ Communications: UCX
  ○ Wireup: PMIx
  ○ Collectives: SHCOLL (Rice)

● OpenSHMEM Spec. 1.4
● Will be 1.5, … as specifications released
● Optionally configured extensions
OSSS OpenSHMEM-on-UCX

Tony Curtis (SBU), Howard Pritchard (LANL)

- Reference OpenSHMEM 1.4 ++ Implementation
  - Open Source Software Solutions
  - LANL
  - Stony Brook U
  - Rice U / Georgia Tech
- UCX for communications
  - User and contributor

UCX

- PMIx for startup, resilience
- Program launch via
  - mpiexec: Open-MPI
  - PMIx Reference RunTime Environment
  - PRRTE

OpenSHMEM

- PMIx server
- PMIx client

PE #0

$ oshrun -np N program

UCP

UCX

cma knem xpmem

IBV

GNI

shmem_long_put shmem_quiet shmem_long_atomic_add ...
shmemc_put shmemc_quiet shmemc_add64 ...
ucp_put_nb ucp_worker_flush ucp_atomic_post ...

http://www.openucx.org/
https://github.com/openshmem-org/osss-ucx
https://pmix.github.io/pmix/
https://github.com/open-mpi
https://github.com/pmix/prrte
http://www.openshmem.org/

OpenSHMEM API

Abstr. Comms API

Collectives API

OpenSHMEM API

PE #N-1
UCF 2018 Workshop: OSSS-UCX OpenSHMEM

- OSSS is **Open Source Solutions**
- UCX is … well, I guess we all know that
- PMIx is **Process Management Interface for Exascale**
  - E.g. as supplied by Open-MPI’s mpiexec or the PMIx reference server
  - Provides clients (ranks) startup, position query, locality info, fault tolerance and more
- OSSS-UCX tries to be as performant as possible but is also
  - A learning aid for budding OpenSHMEMers
SC18 UCX BOF: OSSS OpenSHMEM-on-UCX

- OpenSHMEM + threads/tasks
  - How OpenSHMEM interacts with OpenMP, tasking models
    - E.g. Contexts with Charm++, Legion
  - Resource partitioning
  - Thread/task allocation conflicts
  - Explore math methods that could work well
    - E.g. AMR
SC18 UCX BOF: OSSS OpenSHMEM-on-UCX

- OpenSHMEM + graphs
  - Use of Rice’s HOOVER to model mosquito-borne disease
    - More about HOOVER later...
  - Patch agent hybrid model
  - (Wenbin internship summer 2018 @ LANL)
SC18 UCX BOF: OSSS OpenSHMEM-on-UCX

**OSSS-UCX Reference Implementation**

- [https://github.com/openshmem-org/osss-ucx/releases](https://github.com/openshmem-org/osss-ucx/releases)
Thanks for not falling asleep!