

Porting UCX for Tofu-D interconnect of Fugaku (Lightning Talk)

Yutaka WATANABE¹⁾ Mitsuhsa SATO^{3,2)} Miwako TSUJI³⁾ Hitoshi MURAI³⁾ Taisuke BOKU²⁾

1) University of Tsukuba

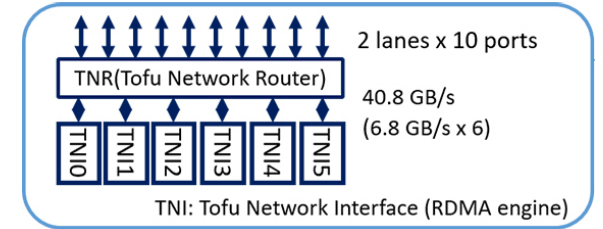
2) Center for Computational Sciences at University of Tsukuba

3) RIKEN Center for Computational Science

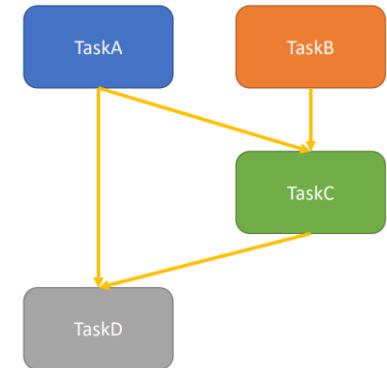
Introduction & Motivation

- Tofu-D interconnect is a 6-dimensional mesh/torus network interface used in Supercomputer Fugaku or Fujitsu FX1000 based systems
 - Tofu-D is integrated in Fujitsu A64FX CPU chip
 - A64FX is an ARM-based many-core processor: 48 compute cores + 2or4 assistant cores
 - “utofu API” is the lowest level of API to use Tofu-D from userland
- UCX does not support Tofu-D natively
 - We cannot fully utilize UCX ecosystem on Fugaku or other systems using Tofu-D
- We are planning to make a task-based distributed programming framework with UCX
 - Task parallel model do not need global synchronization among cores
 - Have a potential to fully utilize the many-core processors
 - UCX will be used as an compatible and lightweight low-level communication layer
 - UCX has less overhead than MPI
 - MPI has problem with multithreaded environment
 - UCX is very useful to make a portable framework with requires node-to-node communication

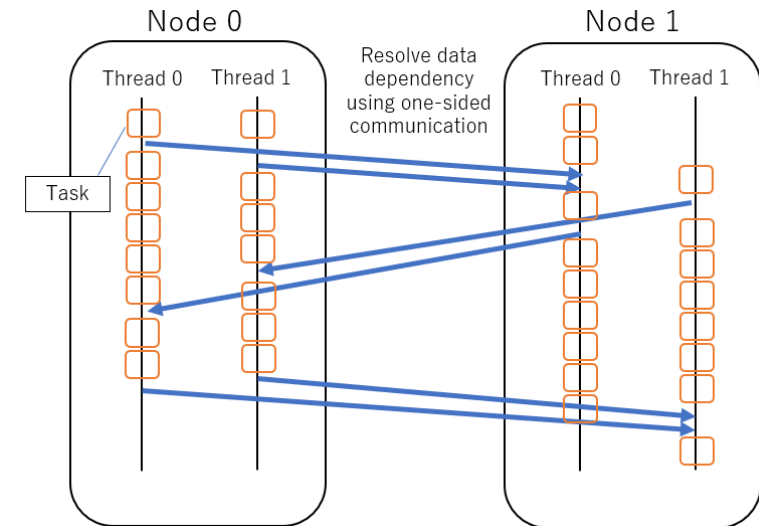
[1] Block diagram of Tofu-D network interface



Task-flow graph



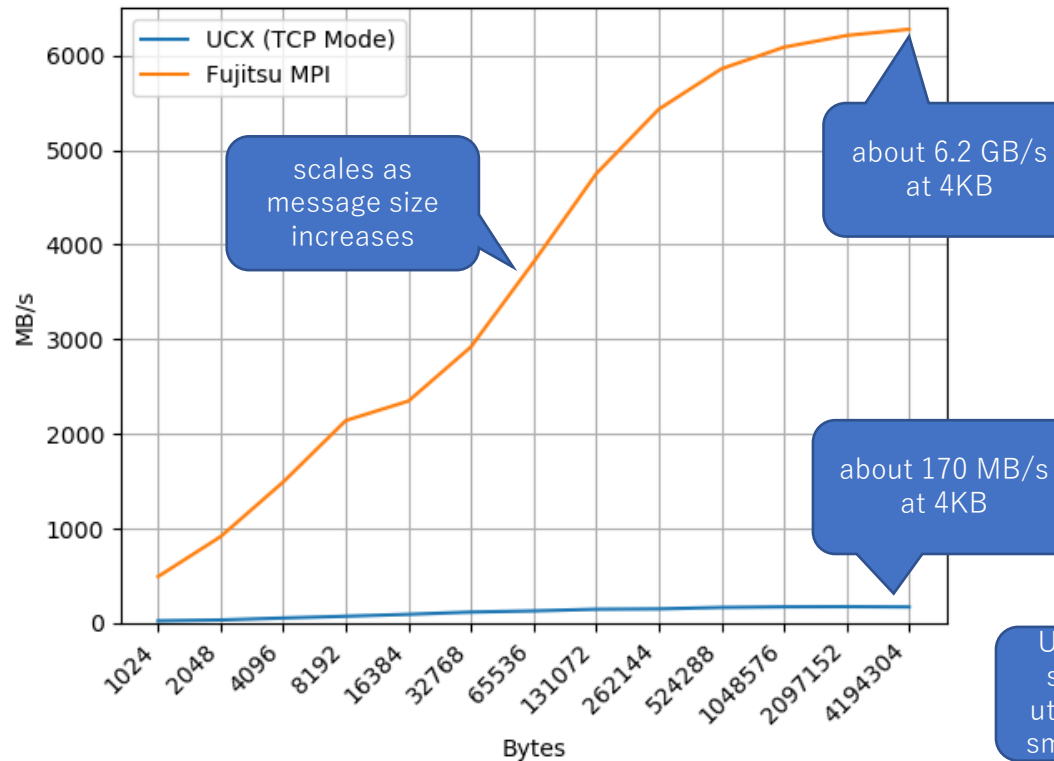
Task-based distributed programming model



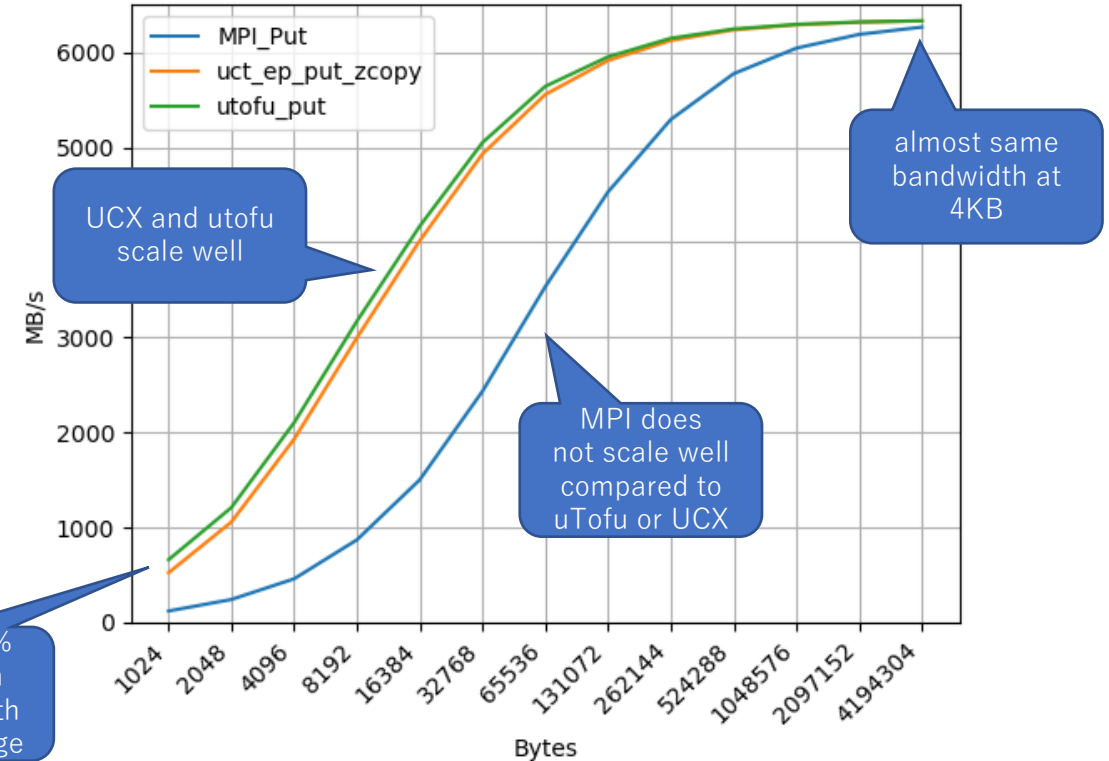
Preliminary performance evaluation using ping-pong benchmark on Supercomputer Fugaku

CPU	Fujitsu A64FX
Interconnect	Tofu-D
Bandwidth per TNI	6.8 GB/s
Injection bandwidth	40.8 GB/s

TCP on Tofu-D: tag message API with original UCX on Fugaku (one-sided API did not work)



Performance of UCT zero copy PUT API with preliminary implementation of UCX for Tofu-D



Current status & Future work

- Current status
 - Implementing Atomic API, Active Message API of UCT for Tofu-D
 - At least continue to implement until OpenSHMEM works correctly with UCX on Tofu-D
- Future plan
 - Scaling test using Supercomputer Fugaku (Fugaku has about 150k nodes)
 - Any scaling test haven't been done yet
 - Optimizing the implementation for Tofu-D based on the performance evaluation
 - Implementing task-based distributed programming framework with UCX for Tofu-D