**OFI WG telecon – 05/05/2020**

**Agenda:**

1. Opens, agenda bashing
2. RPM – Discussion on current industry status on RPM semantics
3. Enhancements to support additional PM use cases -  [Issue #5874](https://github.com/ofiwg/libfabric/issues/5874)
   1. HPE proposal, reference implementation and associated changes – [PR #5875](https://github.com/ofiwg/libfabric/pull/5875)
   2. Sean’s proposal for a new flag, “FI\_SAVE – [comment on issue #5874](https://github.com/ofiwg/libfabric/issues/5874#issuecomment-622657068)

**Opens**

-none-

**Remote Persistent Memory Discussion – Chet Douglas**

RDMA Memory Placement Extensions new opcodes. Usable for non-volatile memory. Draft should be available maybe tomorrow. IBTA LWG draft is not yet available, but IETF draft is public.

* Write Atomic – guarantees that the NIC will atomically place 8 bytes, used for tail of the log use case.
* Flush – looks like a READ from a completion point of view. Allows Flush to Global Observability (GO) or Flush to Persistence (P). Can be flushed by rkey, or by region within the rkey.

Old semantic: required a write followed by a small read to flush the pipeline. So for a log operation requires a write and small read, followed by pointer write and small read (two round trips).

New semantic: Write Data 🡪 flush, Write pointer 🡪 flush. Now waiting only for the last flush to complete. (Note that ‘Write Atomic’ is really just an 8 byte write operation that follows RDMA Write ordering rules).

IETF added another operation, “Verify” allows verification of data written to persistence. This is not available in the IBTA versions. WRITE 🡪 Flush 🡪 Verify 🡪 Write Ptr 🡪 Flush

Intel Support on Sapphire Rapids Platform:

* Support added to the chipset to allow the NIC to force persistence over PCIe
* Steering bits direct whether data is steered to Globally Observable, or Persistence.
* CPU guarantees write completion within the chipset based on whether it is GO or P.
* Intel chipset now supports two VCs in the chipset: can be used e.g. to prevent a read from being stuck behind a write to persistence.
* Sapphire Rapids now looks at Relaxed Ordering disable. Can be used to eliminate the need for a flush between e.g. an RDMA Write and the following Atomic Write

Comments

* Much of what was presented is related to platform implementation, and
* Seems tightly coupled to verbs

**Enhancements to support added PM use cases – Issue #5874 – Jim Swaro**

We did not get to this today, there will be a special off-cadence meeting next Tuesday to continue the discussion. Zoom logistics to be forthcoming.

**Future Agenda Items**

**Next meeting (special, off-cadence)**

Tuesday, May 12, 2020

9:00 – 10:00AM PST

**Agenda**

1. Enhancements to libfabric APIs to support persistent memory – Jim Swaro, Sean Hefty

**Recording:**

See the OFA central calendar for meeting logistics. <https://openfabrics.org/index.php/ofa-calendar.html>

**OFIWG Download Site:** [www.openfabrics.org/downloads/OFIWG](http://www.openfabrics.org/downloads/OFIWG)

**Github:** <https://github.com/ofiwg/libfabric>

**OFI Software Download Site:** [www.openfabrics.org/downloads/OFI](http://www.openfabrics.org/downloads/OFIWG)