**OFI WG Weekly telecom – 04/14/2015**

**Agenda:**

* Issues to discuss
* Look beyond 1.0 release in terms of prioritizing follow-on release
* Complete the Completion discussion

**Release Readiness – ready for release 1.0 RC5**

- Expecting rc5 at the end of this week. Need to touch base with the various providers.

- Hopefully all the changes that would impact the API are now included, prior to Release 1.0

- Last change was to support sockets over OFI a little better.

- Would like this reviewed prior to releasing rc5.

- Sean will create an rc package for fabtests.

**Release schedule**

- should releases be on a three month cycle, or aligned with quarters?

- Release 1.0 will likely be end of April, with quarterly releases thereafter.

- Cisco is relying on dgram support for 1.0, hence happy with release 1.0 on about 4/30 (+/-).

- Next release would likely be end of Q2, possibly as late as two weeks into July.

**Prioritizing Topics for end of June Release**

- Multicast, raw packet support, support in the framework for providers that don’t implement certain features (e.g. verbs doesn’t support RDM, but we could add support for RDM over verbs), provide full capabilities over different providers that don’t support them natively.

- Each provider decides what it wants to provide natively, this is about adding the basics in the framework so at least the feature is there, regardless of the provider. So e.g. there would be a generic RDM service available regardless of whether the provider provides an optimized version of it.

- What applications do we want to prioritize and therefore what features need to be added. e.g. for MPI, RDM is desirable.

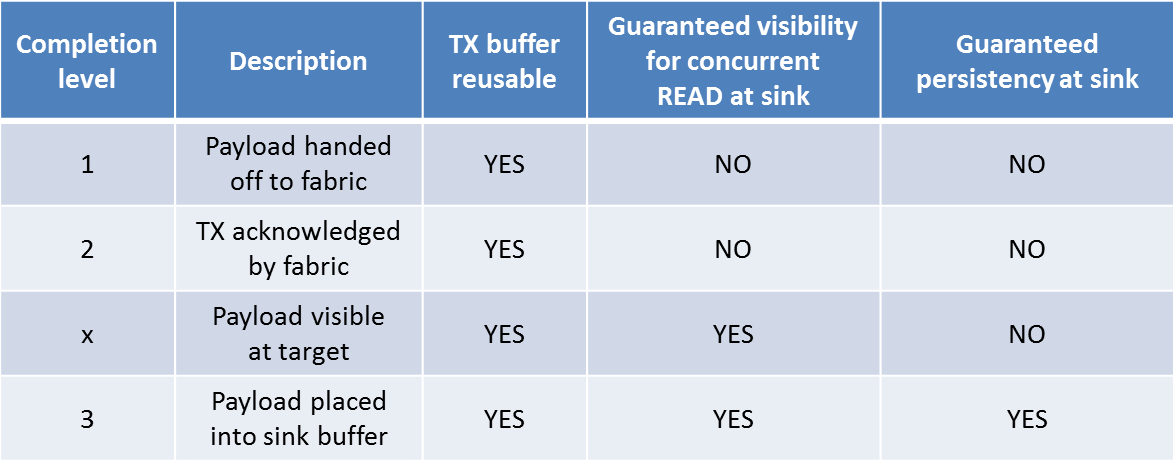
- One approach is to prioritize the API sets / end point type

- RDM over MSG endpoint

- Sean to publish a proposed set of priorities, continue the discussion next week.

**Completions Topic**

Reviewed Bernard’s use case table for remote completion semantics. This table was discussed at length at this week’s DS/DA meeting (previous hour).



An important distinction between level ‘x’ and level ‘3’ is that the major consumer of the completion notification in level ‘3’ is the requester since this allows the requester to release his local resources.

- Storage usage model: a file is updated on the remote side using a series of RDMA write operations. At the conclusion of the writes, the requester wants to atomically update the tree structure reflecting the newly updated file. This can be accomplished using an optimized cache flush operation, and having the acknowledgment of the flush be delayed until the relevant sections of the cache have been flushed to persistent memory. This allows the requester to receive a single notification that all the relevant data has been pushed to persistent memory. “Optimized” in the sense that only specific ranges in cache are flushed, not necessarily the whole cache.

- Byte addressable memory operations, SHMEM: each of these might benefit from receiving an immediate notification that the write operation to persistent memory has completed. This is case #3 in Bernard’s use case table. For byte addressable memory operations, it is considered an optimization to be able to distinguish between writes to volatile memory.

-MPI: today, classical two sided MPI operations use completion levels 1 (local completion) or 2 (remote completion). One-sided operations today exchange an additional handshake notifying the requester that the message has been received. It is felt that one-sided operations would benefit from completion level ‘x’ (i.e. the requester is notified via a completion event that the message has been committed to memory on the responder side and is accessible to the consumer).

Currently, levels 1 & 2 are implemented and documented in the man pages. There is a placeholder in the man pages for level 3, but that will be removed until we can complete the discussion.

Webex link: <https://cisco.webex.com/ciscosales/j.php?MTID=m9389b0513c9ae643d57e2381e254dcf5>  
Webex password: ofi

**Future Agenda Topics:**

* Interfaces and structures for reporting topology data
* Technical issues
  + AV table insert/remove behavior

**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)

**Link to WebEx Recording** - no recording

**Next regular telecon**

Next meeting: Tuesday, 4/7/15

9am-10am Pacific daylight time