**OFI WG telecon – 05/07/2019**

**Agenda:**

1. Opens, Agenda bashing
2. Continue Collectives Discussion
3. Feedback on Traffic Classification proposal

**Opens**

**Discussion – Continue Collectives Discussion**

Two parts:

1. Membership – two possibilities.
	1. proposal is that e.g. MPI would fully manage the collective group in a multicast style
	2. push it down below the libfabric level.

Two step process:

1. identify the peer members via new construct ‘AV Set’, very close to an MPI group.
2. join the collective

For some collective operations, the order of members may matter.

Very similar to a multicast group, but with a few differences such as ordering. There are also a few requirements related to data buffers, which are identified at the time of the operation.

Join operation may involve network traffic, depending on whether the provider implements the join or not and whether the join is local or remote.

1. Local vs remote flag – the join operation could go only to the local node, or it could be distributed across the fabric.
2. Expected to be an asynchronous operation (both local and remote, although local is negotiable).
3. Collective operations
	1. New FI\_COLLECTIVE capability bit
	2. Based on existing atomic interface
	3. Includes new static inline wrappers (for user convenience)
	4. Handles most of the popular collective calls
	5. Examples:
		1. Barrier – fi\_barrier static inline 🡪 fi\_atomic call
		2. Broadcast – fi\_broadcast static inline 🡪 fi\_atomic call mapped to FI\_WRITE/FI\_READ depending on which side the broadcast you are on.
		3. All reduce – maps to fi\_fetch\_atomic
		4. All to All – does not map well to any existing call. Can map to atomic fetch but with some restrictions, i.e. type is replaced by number of bytes and all peers are sending the same amount of data (number of bytes)?
		5. A few more to be discussed (left for next meeting)

**Next meeting**

Tuesday, May 21, 2019

9:00 – 10:00AM PST

**Recording:**

|  |
| --- |
| Tuesday, May 7, 2019 |
| 12:04 pm  |  Eastern Daylight Time (New York, GMT-04:00) |
| [Play recording](https://cisco.webex.com/cisco/lsr.php?RCID=e4f731dfd65d439ab1519b36ac01a59d) (59 min 11 sec) |
| Recording password: SrPJt7qb |  |

|  |
| --- |
|   |

**Webex link:** 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)