

## ITWG Meeting Minutes – 12/07/2004

### List of those present at the meeting

Company	Name
Broadcomm	Uri E. (UE)
HP	<b>Jim Hamrick (JH)</b>
	<b>Jay Rosser (JR)</b>
	Fred Worley (FW)
IBM	<b>Fredy Neeser (FN)</b>
Network Appliance	Arkady Kanevsky (AK)
Sun	Matthew Pearson (MP)

(People whose names are in **bold** letters were present)

Cascading ascii art attendance diagram. (if you have more than 1 minus visible, you are not eligible to vote.)

```
      hp      ibm      netapp      sun      broadcomm
-----
m-3   +       +       -       -       - <- enforcement starts
m-2   +       +       -       -       -
m-1   +       +       -       -       -
m-0   +       +       -       -       -
```

### Next Meeting

Tuesday 12/14/04

Action items appear in **red** embedded in the document below.

### Discussion

#### Minutes to approve

JR: We have minutes from the 11/23/04 ITWG meeting to approve. Any objections?  
(None noted.) Minutes are approved.

#### AI Review

**AI - JR:** Add discussion of how IRD/ORD negotiation is done into app usage section of it\_ep\_connect man page, and reference to it from app usage section of it\_ep\_accept man page. Status: still pending.

**AI - FN:** Adjust current text describing how to unbind a narrow RMR after QP disconnected.

- Deallocate bound MW
- Deallocate bound MR

Status: Still pending.

**AI - FN:** Generate a proposal for cleaning up access bits. Something like propose a set of bitwise ORable privilege flags that are source code compatible with current IT API and identify invalid combinations.

Status: Closed. (Proposal was sent out.)

**FN:** Current definitions for these bits are listed in email I sent out on 12/07 titled “AI - Revised proposal for cleaning up access privilege bits”. PRIV\_ALL does not represent bitwise OR’ing of all access privilege. Thought intent was that it should. Problems with current privilege bit definitions: 1. IT\_PRIV\_NONE does not have value of 0. 2 Read-only bit definition is incompatible with bitwise OR-ing, because it doesn’t turn on and off read access. 3. IT\_PRIV\_DEFAULT is ambiguous. Meant to be local read and write only.

**FN:** Errata GN 49 resulted from above problems. My solution would make GN 49 go away. My email describes a revised enumeration for privilege bits. There would be four independent bits: local read, local write, remote read, and remote write.

**JH:** How is this going to handle the existing IT-API semantic that if you specify 0 for privileges, you get both local read and local write access? Won’t this break source code compatibility?

**FN:** Yes, it will break source code compatibility, but I think we should clean this up anyway. The existing bit definitions are so confusing that they generated 4 errata in IT-API 1.0.

**JH:** Are we already breaking source code compatibility in this release for other reasons? If so, I’d be inclined to go ahead and clean this up.

**JR:** Requiring the Consumer to post a Send operation to complete connection establishment breaks source code compatibility.

**JH:** Okay, since it appears that we’ll have to break source code compatibility in this release anyway, let’s go ahead and clean this up.

FN: With cleanup we now have the issue that somebody could attempt to request local write access only for an LMR created on an InfiniBand IA. InfiniBand doesn't support local write only permissions.

(More discussion. Somebody points out that we should arrange that the default privileges for an LMR are portable.)

**AI** - FN: Add advice to Consumers that creating write-only local access is a transport-dependent programming practice; providing for both local read and local write is transport-independent. Make sure IT\_PRIV\_DEFAULT is the portable value.

FN: Should we have a value for IT\_PRIV\_NONE. Recommend we get rid of it.

JR: No additional value in doing bind with IT\_PRIV\_NONE.

FN: How about we document IT\_PRIV\_NONE in the Implementer's Guide, but remove it from the spec? That way, we can allow implementers to avoid breaking apps—they can put the value into their header files—but can get rid of the documentation for this. (General consensus that this would be a reasonable way to deprecate this flag.)

JR: Do we need to vote on this proposal?

**AI** – FN: To generate a ballot based on the above email with the date of vote set to 12/3/04.

Status: Still pending.

FN: I will emit a modified proposal for us to vote on.

FN: In process of updating detailed MM requirements. Version 0.99 will have updates.

**AI** – ALL: Think about the consequences of renaming "RMR Context" and suggest terminology.

Status: Closed. (We thought about it during today's ITWG meeting. Discussion below.)

JH: Could typedef old name to new name, drop documentation for old name.

JR: How about remote\_mr\_id\_t for the new name?

FN: How about "exposable MR tag". New name would be emr\_id\_t.

FN: RMR context being used in it\_lmr\_create, it\_lmr\_query, it\_rmr\_bind, it\_rmr\_query. Also used in it\_post\_rdma\_read, it\_post\_rdma\_write.

JR: The new `it_lmr_link` call can cause a change of RMR context, right?

FN: Right.

JH: Why do you want to change this the name, anyway?

FN: Having remote access privileges associated with a Local Memory Region is counterintuitive.

JH: Once upon a time we had clean separate of local and remote, but group felt when creating IT-API 1.0 that we should expose the `R_Key` associated with Memory Regions. This created situation where an LMR can have an associated RMR Context. This “conflict” in the terminology has existed since IT-API 1.0. Unless we’re creating some new, worse conflict in IT-API 2.0, I think we should leave the terminology in this area alone. Our user community has already learned the terminology.

JR: Tend to agree with Jim.

FN: Still think this terminology is confusing.

JH: Is it confusing in a new way for IT-API 2.0? For example, can we remotely invalidate something in iWARP that only has local access privileges? If we could, that would be a problem that might merit changing the terminology.

**AI** – JR: Determine if there are any remote operations in iWARP that can be performed on things that only have local access permission.

**AI** - FN: Add a sentence per RDMA DTO man page describing access privileges for sinks and sources is sufficient.

Status: Still pending.

#### Requirements vote

JR: We voted in detailed MM requirements. Anything more to discuss?

FN: No, don’t think so. Changes implied by the requirements that we voted in will be in next version of man pages.

FN: Had related email exchange on unlinkable attributes or LMRs and RMRs. Attribute is no longer present per RMR, has been moved to being an IA attribute. Same goes for LMRs. Unlinkable attribute intended to show if there was support for unlinking. Whole bunch of conditions needing to be satisfied in order for LMR to be unlinked. Confusing, so got rid of it.

JR: But we can still create LMRs that can’t be unlinked with 1.2 IB hardware, right?

FN: Right. `it_lmr_create` doesn't always give you an LMR that you can unlink even if you are running on IB 1.2. Users who want to be sure their LMR allows for unlinking should use `it_lmr_create_unlinked` to create their LMR

JR: Just remembered an AI I neglected to give myself.

**AI** - JR: Generate a proposal for a mapping service to generate a PBL given an input address.

JR: Fredy, when do you plan to publish updated MM man pages?

FN: By end of this week.

**AI** - FN: Publish updated MM man pages by 12/10/04.

#### CM issues

JR: Any meta issues?

JH: How are we dealing with the issue of turning off the IOH for DAT compatibility?

JR: John Carrier is preparing to generate a new draft that would supercede the IOH. (At least he was before he got laid off from Adaptec.) I think the proposed draft would always incorporate an IOH-equivalent.

JH: If IT-API expects IOH data in MPA REQ/REP but uDAPL doesn't provide it, who's in the wrong here?

FN: Think we should get this issue resolved through the RDDP working group. John said this week or early next week he wants to work on this.

JR: One way to resolve this issue would be if implementation that didn't make use of Implementation Private data was required to send all zeros. We could detect that, and so if uDAPL didn't use this we would know.

FN: First 256 bytes of MPA Private Data are connection parameters containing IOH-type data, second 256 bytes are ULP-private Private Data in John's proposal?

JR: Yes.

JH: So, this would be documented in an informational draft, and we'd just say our TI interfaces only work on networks implementing this informational draft?

JR: Yes.

FN: Would RDDP WG go for variable-length header? Could then specify in MPA Header whether or not connection parameters are present.

JR: MPA Header is near last call. Probably will meet more resistance asking for change in it rather than just asking for defined use model for existing Private Data portion of MPA REQ/REP.

JH: Think it would be good idea to see if we can get alignment with John Carrier on his information draft so that it could subsume the IOH. We could then drop the IOH in a future release, and concurrent with that we would get uDAPL compatibility.

**AI** - JR: Produce requirements for RDDP WG informational draft to support IOH functionality.

FN: Don't think RDDP will like fixed 512 byte mechanism for transferring connection parameters and ULP-private Private Data.

JR: Any other meta issues?

FN: Generic question. Haven't covered interoperability draft in terms of error handling. Have different RNICs we have to deal with: RDMAC, permissive/non-permissive IETF RNIC. When are we going to address this? For example, if preference is to turn off markers and Consumer notices incoming connection request is from an RDMAC RNIC, will request markers on if using an IETF RNIC.

JR: Does consumer-specified marker preference override RNIC permissiveness?

JR: Do we even need marker control to be visible to Consumers if the IETF proposal for interoperability between RDMAC and IETF RNICs goes through?

JH: It might be a good idea to just get rid of marker control in the IT-API. Don't want to force this onto our Consumers unless necessary, and it increasingly looks like it won't be necessary. If it turns out the IETF doesn't solve this problem with its proposal, we could always add support to `it_socket_convert` for marker control in a future release.

FN: Not sure that this is good idea; might still be value in providing marker control to Consumers.

**AI** - FN: Point out cases where marker control still has value even if the IETF proposal for RNIC compatibility is passed.

JR: Anything else?

FN: Question on version numbering. What are we calling the next release of the IT-API?

JR: Suggest our next release be called IT-API 2.0.

FN: But currently it is called "Issue 1.0".

JR: Really? Who did that? (Martin Kirk identified as likely culprit.)

FN: So, should next release be called Issue 2.0 or Version 2.0?

JR: Martin works for us! Version 2.0.

Meeting ends at ~11:37am PST.