NAME

ibv_advise_mr − Gives advice or directions to the kernel about an address range belongs to a memory region (MR).

SYNOPSIS

```
#include <infiniband/verbs.h>

int ibv_advise_mr(struct ibv_pd *pd,
          enum ibv_advise_mr_advice advice,
          uint32_t flags,
          struct ibv_sge *sg_list,
          uint32_t num_sge)
```

DESCRIPTION

ibv_advise_mr() Give advice or directions to the kernel about an address range belonging to a memory region (MR). Applications that are aware of future access patterns can use this verb in order to leverage this knowledge to improve system or application performance.

Conventional advice values

IBV_ADVISE_MR_ADVICE_PREFETCH

Pre−fetch a range of an on−demand paging MR. Make pages present with read−only permission before the actual IO is conducted. This would provide a way to reduce latency by overlapping paging−in and either compute time or IO to other ranges.

IBV_ADVISE_MR_ADVICE_PREFETCH_WRITE

Like IBV_ADVISE_MR_ADVICE_PREFETCH but with read−access and write−access permission to the fetched memory.

ARGUMENTS

pd The protection domain (PD) associated with the MR.

advice The requested advise value (as listed above).

flags Describes the properties of the advise operation **Conventional advice values** *IBV_ADVISE_MR_FLAG_FLUSH* : Request to be a synchronized operation. Return to the caller after the operation is completed.

sg_list Pointer to the s/g array When using IBV_ADVISE_OP_PREFETCH advise value, all the lkeys of all the scatter gather elements (SGEs) must be associated with ODP MRs (MRs that were registered with IBV_ACCESS_ON_DEMAND).

num_sge

Number of elements in the s/g array

RETURN VALUE

ibv_advise_mr() returns 0 when the call was successful, or the value of errno on failure (which indicates the failure reason).

ENOSYS

libibverbs or provider driver doesn't support the ibv_advise_mr() verb.

ENOTSUP

The advise operation isn't supported.

EFAULT

In one of the following: o When the range requested is out of the MR bounds, or when parts of it are not part of the process address space. o One of the lkeys provided in the scatter gather list is invalid or with wrong write access.

EINVAL

In one of the following: o The PD is invalid. o The flags are invalid.

NOTES

An application may pre−fetch any address range within an ODP MR when using the **IBV_ADVISE_MR_ADVICE_PREFETCH** or **IBV_ADVISE_MR_ADVICE_PREFETCH_WRITE** advice. Semantically, this operation is best−effort. That means the kernel does not guarantee that underlying pages are updated in the HCA or the pre−fetched pages would remain resident.

When using **IBV_ADVISE_MR_ADVICE_PREFETCH** or **IBV_ADVISE_MR_ADVICE_PREFETCH_WRITE** advice, the operation will be done in the following stages: o Page in the user pages to memory (pages aren't pinned). o Get the dma mapping of these user pages. o Post the underlying page translations to the HCA.

If **IBV_ADVISE_MR_FLAG_FLUSH** is specified then the underlying pages are guaranteed to be updated in the HCA before returning SUCCESS. Otherwise the driver can choose to postpone the posting of the new translations to the HCA. When performing a local RDMA access operation it is recommended to use IBV_ADVISE_MR_FLAG_FLUSH flag with one of the pre−fetch advices to increase probability that the pages translations are valid in the HCA and avoid future page faults.

SEE ALSO

ibv_reg_mr(3), **ibv_rereg_mr**(3), **ibv_dereg_mr**(3)

AUTHOR

Aviad Yehezkel <aviadye@mellanox.com>