

Re: [patch 5/6] mmu_notifier: Support for drivers with revers maps (f.e. for XPmem)

Re: [patch 5/6] mmu_notifier: Support for drivers with revers maps (f.e. for XPmem)

Source: <http://linux.derkeiler.com/Mailing-Lists/Kernel/2008-02/msg10851.html>

- *From:* Nick Piggin <nickpiggin@xxxxxxxxxxxxx>
 - *Date:* Thu, 21 Feb 2008 15:20:02 +1100
-

On Wednesday 20 February 2008 20:00, Robin Holt wrote:

On Wed, Feb 20, 2008 at 02:51:45PM +1100, Nick Piggin wrote:

On Wednesday 20 February 2008 14:12, Robin Holt wrote:

For XPMEM, we do not currently allow file backed mapping pages from being exported so we should never reach this condition. It has been an issue since day 1. We have operated with that assumption for 6 years and have not had issues with that assumption. The user of xpmem is MPT and it controls the communication buffers so it is reasonable to expect this type of behavior.

OK, that makes things simpler.

So why can't you export a device from your xpmem driver, which can be mmap(ed) to give out "anonymous" memory pages to be used for these communication buffers?

Because we need to have heap and stack available as well. MPT does not control all the communication buffer areas. I haven't checked, but this is the same problem that IB will have. I believe they are actually allowing any memory region be accessible, but I am not sure of that.

Then you should create a driver that the user program can register and unregister regions of their memory with. The driver can do a `get_user_pages` to get the pages, and then you'd just need to set up some kind of mapping so that userspace can unmap pages / won't leak memory (and an `exit_mm` notifier I guess).

Because you don't need to swap, you don't need coherency, and you are in control of the areas, then this seems like the best choice.

Re: [patch 5/6] mmu_notifier: Support for drivers with revers maps (f.e. for XPmem)

Re: [patch 5/6] mmu_notifier: Support for drivers with revers maps (f.e. for XPmem)

It would allow you to use heap, stack, file-backed, anything.

—

To unsubscribe from this list: send the line "unsubscribe linux-kernel" in the body of a message to majordomo@xxxxxxxxxxxxxxxx

More majordomo info at <http://vger.kernel.org/majordomo-info.html>

Please read the FAQ at <http://www.tux.org/lkml/>