

Re: How to avoid data copies in a driver ?

Source: <http://linux.derkeiler.com/Mailing-Lists/Kernel/2008-05/msg06932.html>

- *From:* "linux-os (Dick Johnson)" <linux-os@xxxxxxxxxxxxx>
 - *Date:* Thu, 15 May 2008 09:16:00 -0400
-

On Thu, 15 May 2008, Francis Moreau wrote:

Hello,

On Wed, May 14, 2008 at 11:23 PM, linux-os (Dick Johnson) <linux-os@xxxxxxxxxxxxx> wrote:

You memory-map the data. Impliment mmap() in your driver.
You can also impliment poll() { select() } so your
application knows when new data are available.

You cannot use a user-mode file-descriptor in the kernel.

Why not ?

I'm suprised because what I need doesn't seem so uncommon, usually
devices send or
receive data to/from files. So a helper (system call ?) to achieve
that other than the basic
read/write seems needed, no ?

--

Francis

The kernel is designed to perform services on behalf of
a caller. The kernel itself doesn't have a process context.
Therefore, a file-descriptor, which requires a process context
to mean anything, is not useful within the kernel unless the
kernel either uses your process context (which happens
efficiently when YOU call the kernel) or it steals one
from somebody else, which means their context gets trashed.
Note that every process is created with at file descriptors
0, 1, and 2. It's only the process context that keeps them

Re: How to avoid data copies in a driver ?

separate.

That said, you can create a kernel-mode task in your driver. That task would have a context. However it wouldn't be YOUR context, so you would need to signal it when data was available, adding overhead and communicate with it from your user-space program context. It would share the same data-space as your driver so it wouldn't need to copy. However, such data would get copied into kernel buffers by the I/O code so it's a waste anyway. You save one copy, which you would do with memory-mapping, plus you have the added communications overhead.

Using memory mapping as previously advised, lets you DMA data directly to a user if your hardware does DMA. It also would allow you to save one copy, even if you don't have DMA capabilities. As far as copies are concerned, there are many copies before your data actually gets to a disk platter. Those copies (usually) occur when the kernel doesn't have anything else to do.

Cheers,

Dick Johnson

Penguin : Linux version 2.6.22.1 on an i686 machine (5588.29 BogoMips).

My book : <http://www.AbominableFirebug.com/>

—

The information transmitted in this message is confidential and may be privileged. Any review, retransmission, dissemination, or other use of this information by persons or entities other than the intended recipient is prohibited. If you are not the intended recipient, please notify Analogic Corporation immediately – by replying to this message or by sending an email to DeliveryErrors@xxxxxxxxxxxxx – and destroy all copies of this information, including any attachments, without reading or disclosing them.

Thank you.

—

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

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

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