

Re: [perfmon] perfmon2 code review: 32-bit ABI on 64-bit OS

Source: <http://linux.derkeiler.com/Mailing-Lists/Kernel/2006-02/msg04660.html>

- *From:* Stephane Eranian <eranian@xxxxxxxxxx>
 - *Date:* Mon, 13 Feb 2006 12:31:20 -0800
-

David,

On Sun, Feb 12, 2006 at 04:03:44PM -0800, Eric Gouriou wrote:

David Gibson wrote:

On Sat, Feb 11, 2006 at 02:33:54PM -0800, Stephane Eranian wrote:

[...]

The most challenging piece is the IP (program pointer) that is in every sample. Today it is defined as unsigned long because this is fairly natural for a code address. The 64bit OS captures addresses as 64-bit, the 32-bit monitoring tool running on top has to consume them as 64-bit addresses, so u64 would be fine.

But not on a 32-bit kernel with a 32-bit tool, addresses exported as u64 would certainly work but consume double to buffer space, and that is a more serious issue in my mind.

Hmm.. does the sampling buffer collect on userspace PC values, or kernel ones as well?

Either, or both, depending on the measurement settings.

I live in a 64-bit world, so my take on this issue would be to expose the PC as a `uint64_t`, always. There is already so much overhead in the default per-sample header that I wouldn't worry about it.

Re: [perfmon] perfmon2 code review: 32-bit ABI on 64-bit OS

Eric is right, on many architectures, incl. PPC64 I am sure, you can easily configure a counter to measure at any priv levels including at the kernel level. As such a 32-bit monitoring tool could see 64-bit generated samples. Similarly, I don't think it would be unreasonable to have a 32-bit tool monitor 64-bit applications.

The question is whether hardcoding the IP to always be u64 is a valid choice. Eric's comment about overhead is based on the current default sampling format which systematically adds a fixed size header to each sample. That header contains the IP. So adding 4 bytes to this header is not a big deal.

However, because we can define virtual PMDs that map to software resources, it is likely that the default format will evolve to allow an application to specify everything it needs for each sample. For instance, you can have a PMD that maps to the current PID, another one that maps to the interrupt IP. Then you can chose to include those into the sample and you would nto need a fixed size header anymore.

--

-Stephane

-

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

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

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