Contribution ID: **248**

Type: **not specified**

Kernel Livedump

Tuesday, 14 November 2023 12:30 (30 minutes)

Linux Kernel currently has a mechanism to create a dump of a whole memory for further debugging of an observed issue with the help of crashkernel. Unfortunately, we are unable to do this without restarting the host which causes a problem in case of having a high availability service running on the system experiencing some complex issue that cannot be debugged without the complete memory dump and hypervisor-assisted dumps are not an option on bare metal setups. For this purpose, there is a live dump mechanism being developed which was initially introduced by Yoshida Maasanori [1] in 2012. This PoC was already able to create a consistent image of memory with the support of dumping the data into a reserved raw block device.

The PoC remained idle and as the ever-growing Linux community introduces dozens or even hundreds of new features every release, that work obsoleted, especially due to MM changes. I've spent time adapting the patchset to make it work again on Linux v6.4 and I've added a few more features (like vmcore formatting).

In order to put forward the patchset into the upstream again, there is a lot of research and work ahead because of a few tradeoffs that must be better described and understood. Similar to the crashkernel method, which necessitated preallocated space specific to each running instance and was resolved through approximations, there is also reserved preallocated memory. If this memory is not sufficiently large, it may, in certain cases, compromise the consistency of the dumped state. To maintain consistency, one option is to wait within the page fault in kernel memory, but this approach could potentially introduce failures in the original kernel due to synchronization or deadlines in different parts of the kernel.

At LPC I would like to gather as much feedback as possible on my current approach with a discussion about other possible usecases.

[1] https://lore.kernel.org/all/20121011055356.6719.46214.stgit@t3500.sdl.hitachi.co.jp/

Primary author: HRUŠKA, Lukáš

Presenter: HRUŠKA, Lukáš

Session Classification: Linux Kernel Debugging MC

Track Classification: LPC Microconference: Linux Kernel Debugging MC