LINUX PLUMBERS CONFERENCE September 20-24, 2021

PKRAM feature development

Anthony Yznaga, Oracle

What is PKRAM?

- A way to preserve memory pages across kexec
- Provides APIs for preserving collections of byte and in-memory page data,
  - For example, preserve the page data of a tmpfs file and its file attributes.
- Flexible
  - Does not require determining a size and configuring a portion of physical memory as emulated persistent memory.

Use Cases

- Live update of a cloud hypervisor
  - Preserve guest VMs across reboot of the host
  - Guest memory and/or auxiliary guest data (e.g. iommu data)
- Preserve database block caches across reboot
- Others?

Limitations

- Preserving and restoring adds overhead
  - Can mitigate with optimizations
- Does not work for firmware reboot
- Potential failures due to memory fragmentation
- New kernel needs to know about PKRAM
- Cannot preserve HugeTLB pages (yet)

Current State

- RFC v2

  https://lore.kernel.org/linux-mm/1617140178-8773-1-git-send-email-anthony.yznaga@oracle.com/

  - API and supporting functionality
  - Simple use of PKRAM with tmpfs
  - A number of optimizations to improve performance
    - Parallelization of work to preserve and restore tmpfs files
    - Defer initialization of page structs of preserved pages

Questions or Comments?