RAS Tools Discussion

Suparna Bhattacharya
suparna@in.ibm.com
(with inputs from Eric Biederman, Maneesh Soni)

Quick Update

- In mainline
  - Kexec for i386, x86_64, ppc64
    - user space integration pending on ppc64
  - Kdump for i386
  - Kprobes for i386, x86_64, ppc64, ia64, sparc
  - Debugfs
- Ongoing
  - First failure data capture
    - Kdump for x86_64, ppc64
    - Relocatable kernel image
    - Kernel pages only dump
    - Analysis utilities
  - Tracing
    - Relayfs/tracepipe
    - Probe handler implementations, systemtap
    - Ostra update (acme ?)

Kexec/Kdump

- Kexec on panic and kdump overview (details in OLS talk)
  - New kernel loaded and run from reserved area
    - Capture kernel runs from non-default locations
    - Memory used by old and new kernels mutually exclusive
    - Purgatory unit handles switchover w/ validation
    - Backup sections for areas that may still be overwritten
  - ELF core segments generated by kexec user-space
    - Dying kernel fills in register state of all CPUs
- Reliability
  - Driver re-initialization issues
  - Handling of frame buffer consoles
    - Getting info about the video mode from user space
- Architecture ports of kexec-on-panic and kdump
  - Requires running the kernel at a non-default address
- ELF core dump view improvements
  - Exposing vmalloc area addresses directly to gdb
  - Kernel pages only dump
    - Heuristic (in use or reserved and not on LRU list)
    - Fragmentation increases number of ELF sections

To Ponder and Discuss

- Analysis and filter tools track kernel changes very closely
  - Affects both dump analysis and probe utilities
  - Debug information alone is not enough

- How can we simplify maintenance, reduce fragility ?
  - Allow user-space to share kernel logic ?
    - e.g. radix tree traversal, ukdb idea
  - Maintaining simple utility routines in tree ?
    - built but not directly linked with the kernel
    - included in usr/
  - Improved stack traces
    - crossing exception boundaries
    - interpreting lock sections

Disclaimers and Trademarks

This work represents the view of the authors and does not necessarily represent the view of IBM.

IBM is a registered trademark of International Business Machines Corporation in the United States and/or other countries.

Linux is a registered trademark of Linus Torvalds.

Other company, product, and service names may be trademarks or service marks of others.