Sohum Shah

121056

Abstract

A brief about Linux’s crash dump mechanism after a kernel crash.

KDUMP

Kernel-Dump

# Introduction

A software bug or a hardware failure lead an operating system’s kernel to panic. Due to the kernel panic the kernel crashes down, i.e. it comes across a fatal error from which the kernel cannot recover itself. This would be a very common scenario, if not a frequently occurring scenario, to the kernel developers. Thus there arises a need to restore the previous balanced version of kernel to continue work and also to debug the reason for a kernel crash.

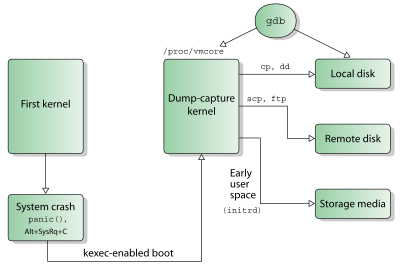
Previous best working image of the kernel can be stored to recover from a kernel crash. Linux particularly, uses such a crash dump mechanism to restore the balance of the kernel: **Kdump.** In the event of a kernel crash, kdump creates a memory image that can be analysed for the purposes of debugging and determining the cause of the crash. The memory image is also called the vmcore, as it is stored as ‘vmcore’ in the /proc/ . It can also be stored in a locally accessible file system or any other attached devices.

# Kdump working

Instead of booting into the same kernel, kdump allows us to boot into a new kernel, also without going through BIOS or UEFI. To preserve system consistency, kdump captures the image of a fresh, properly working kernel and stores it into the RAM. This version of kernel which is stored is also called the capture kernel. There is always a fixed allocation of capture kernel in the RAM, as one may never know when a kernel would crash. This way after the kernel crash the system boots into a clean and reliable environment instead of relying onto a crashed kernel which could cause various issues. To boot into the kernel, the system uses a feature called kexec, explained further:

## Kexec

Kexec is a fast boot mechanism that allows booting a new linux kernel from the context of an already running kernel: a crashed kernel. The main advantage of kexec is that the system does not need to boot into the new kernel through the BIOS firmware. BIOS can be very time consuming especially on big servers with large number of peripherals. This can save a lot of time for kernel developers who end up booting into the system many number of times.



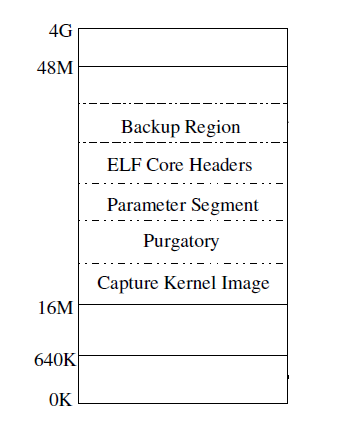
Source: <http://en.wikipedia.org/wiki/Kdump_(Linux)>

The first kernel reserves the following sections of memory:

1) For the capture kernel

2) For loading and booting into the capture kernel

3) Other important parameters and headers and backup regions.



Source: [Kdump - A kexec based dumping mechanism (Presentation slides) - Vievek Goyal et al, OLS 2005](http://lse.sourceforge.net/kdump/documentation/ols2005-kdump-presentation.pdf)

* First 640K of memory is required for booting SMP capture kernel on i386
* Contents of first 640K of memory are backed up in Backup Region
* Purgatory contains the code for backing up first 640K of memory after crash.
* ELF core headers stores the ELF object headers and parameters.

The content of the main memory is preserved while booting into and running the dump-capture kernel by reserving a small amount of the main memory ([RAM](http://en.wikipedia.org/wiki/RAM)) in advance, into which the dump-capture kernel is preloaded so none of the RAM used by the primary kernel is overwritten while a kernel crash is handled. This reserved amount of RAM is used solely by the dump-capture kernel, and it is otherwise unused during normal system operation. Some architectures, including [x86](http://en.wikipedia.org/wiki/X86) and[ppc64](http://en.wikipedia.org/wiki/Ppc64), require a small fixed-position portion of RAM to boot a kernel regardless of where it is loaded; in this case, kexec creates a copy of that portion of RAM so it is also accessible to the dump-capture kernel.

## Current Scenario

Kdump functionality, together with kexec, was merged into the Linux kernel mainline in kernel version 2.6.13, which was released on August 29, 2005.

Various patches have been updated to this functionality for debian machines, and various services have been added in kdump.

## Advantages

* Increased reliability

Dump is captured from a newly booted kernel

* Enhanced flexibility

Dump image can be saved to virtually any storage media supported by kernel.

Filtering mechanism can be plugged in.

* Ease of use

Standard utilities can be used to save the dump image either locally or remotely.

Standard analysis tools like gdb can be directly used for limited debugging.

## Limitations

* Devices are not shutdown/reset after a crash which might result in a driver initialization failure in capture kernel
* Nondisruptive dumping is not possible.

# References

* <http://en.wikipedia.org/wiki/Kdump_(Linux)>
* <http://en.wikipedia.org/wiki/Kernel_panic>
* <http://lse.sourceforge.net/kdump/documentation/ols2005-kdump-presentation.pdf>
* <http://www.dedoimedo.com/computers/kdump.html#mozTocId967714>
* <http://lse.sourceforge.net/kdump/>