



Implementing Kexec with Ironic to Reduce HPC System Downtime

By Kam Killfirst, HPC-DES

7/29/2021

LA-UR-21-27408

What is Ironic?

- Provisioning of bare metal machines
- Manages machine booting through PXE
- Controls and monitors machine using IPMI
- Used independently or as part of the OpenStack Cloud
- Open Source

What is Kexec?

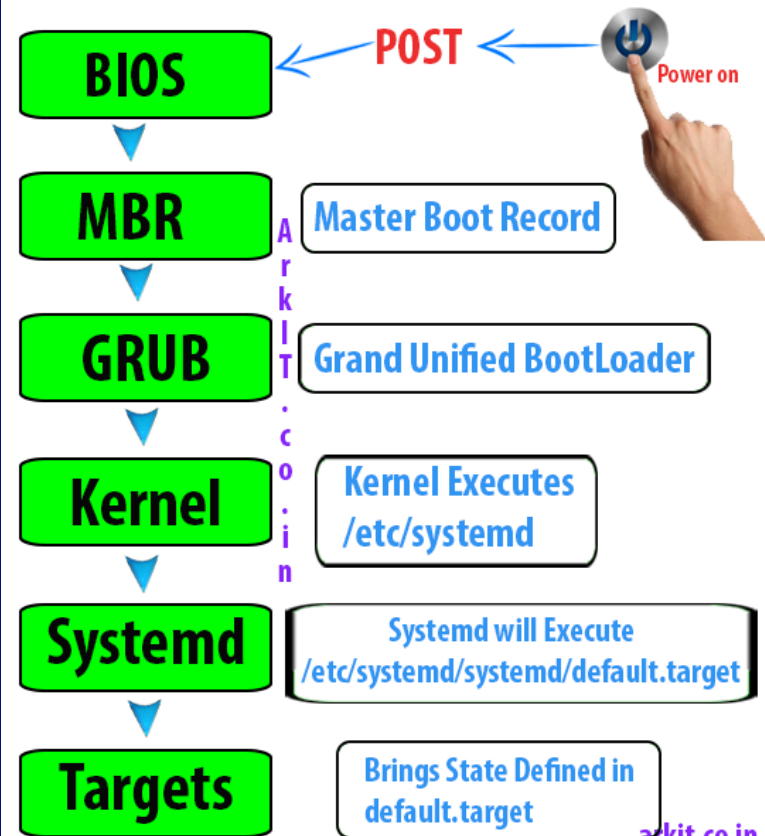
- Tool used to boot into another kernel
- Bypasses power-on self-test (POST), and bootloader
- Loads user-specific kernel and ramdisk images

Why is Kexec important in HPC

- Reboots take a long time
- Downtime means no work is being done
- Larger the machine, the longer the wait

How does Kexec reduce system downtime

- Performs the function of the bootloader
 - Does this from inside current kernel
- Loads kernel and ramdisk image before “rebooting”
- Time is saved in skipping POST and other BIOS functions



<http://opensourceadda.blogspot.com/2014/05/linux-boot-process-in-6-stages.html>

How Ironic boots a system

- Ironic serves kernel and ramdisk images
- Reboot happens before node is usable

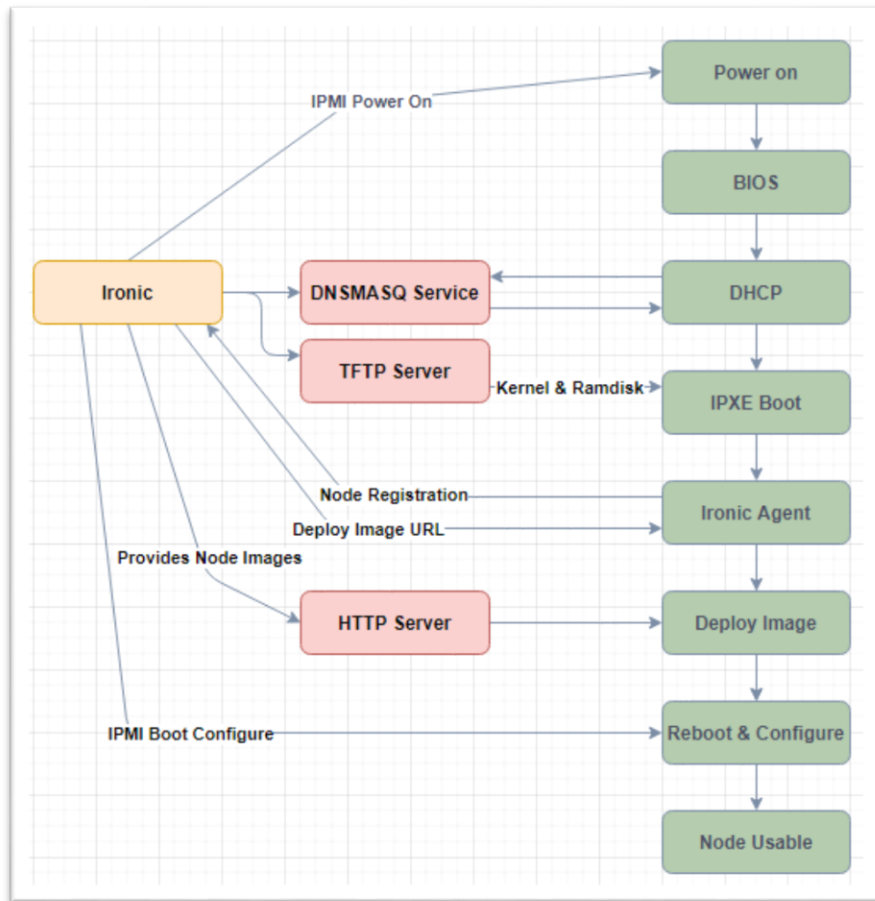


Image created by Kam Killfirst with Draw.io

How Ironic can use Kexec

- Kexec can bypass POST & BIOS

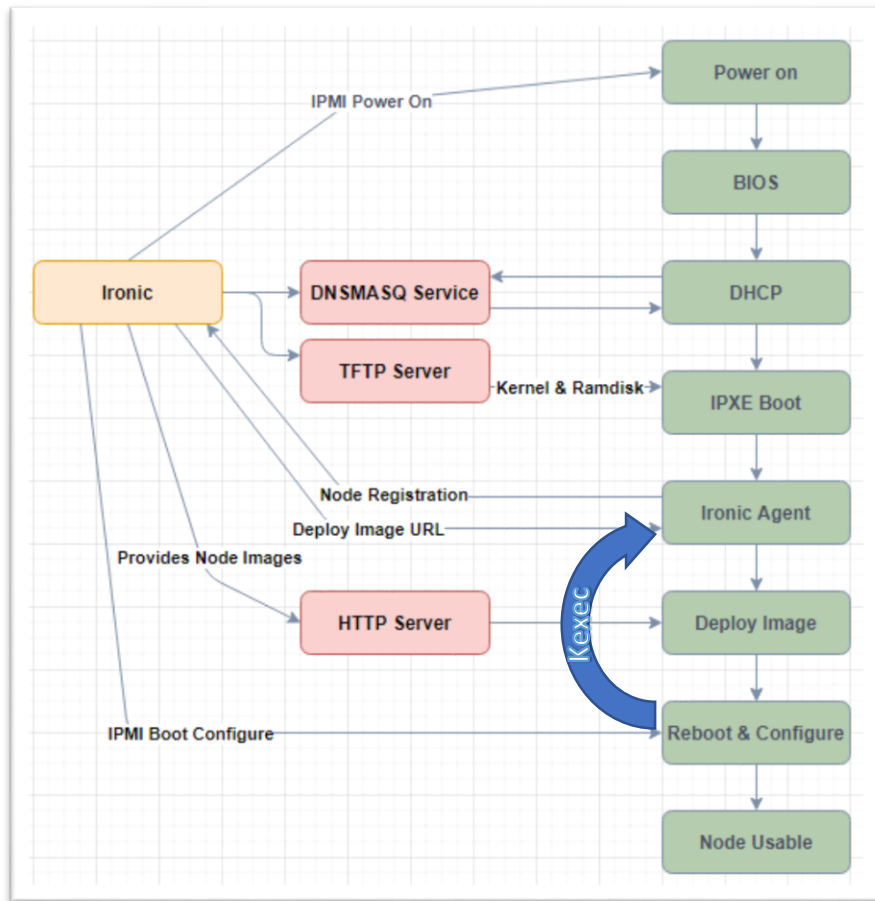


Image created by Kam Killfirst with Draw.io

Testing Methods

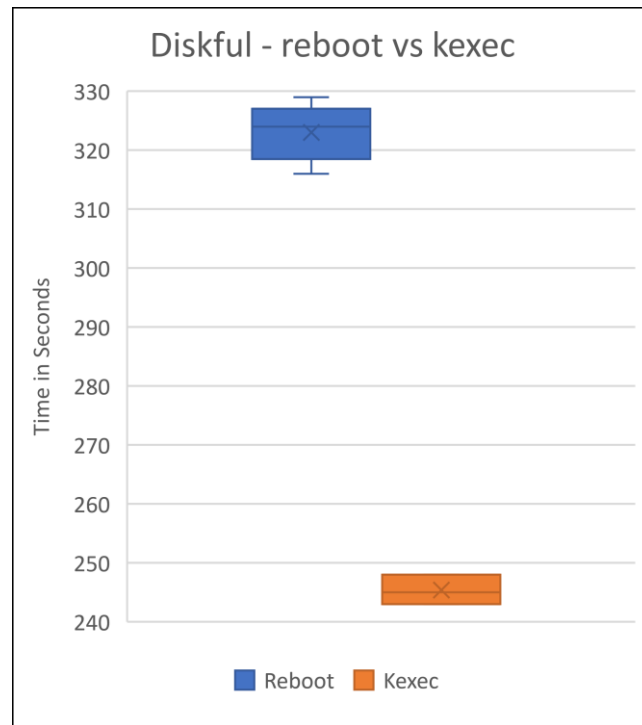
- Tested kexec boot times vs regular reboot
- Time until usable based on ssh connectivity
- Time kept with a bash script
 - Attempts to ssh into machine
 - Stops timer when ssh connects
- Reboot and Kexec initiated with “systemctl”
- Sample size of 5 with each configuration

Testing Setup

- Direct deployment – Diskful
 - Operating system : Ubuntu
 - CPU 32 cores @ 3.2 ghz
 - Memory size : 192 GiB
 - Hard Drive size : 5.5 terabytes
- Ramdisk deployment – Diskless
 - Operating system : Debian
 - CPU 32 cores @ 3.2 ghz
 - Memory size : 192 GiB

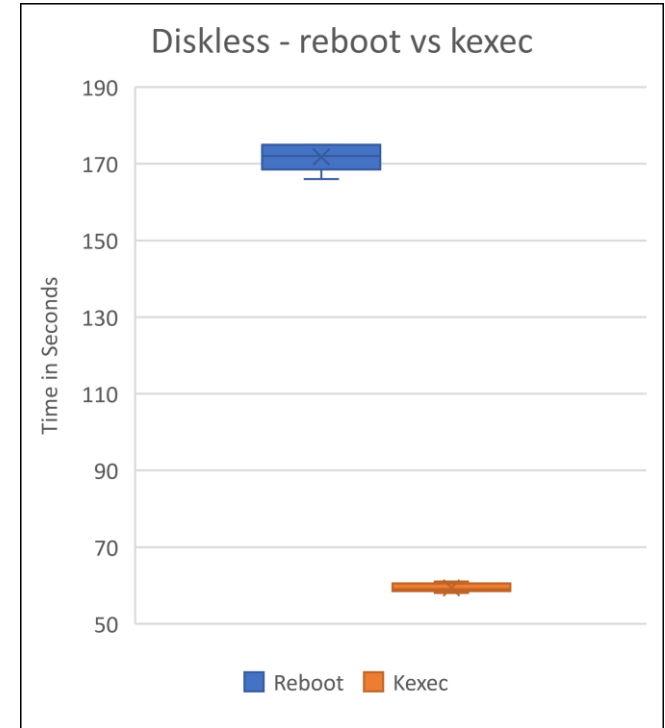
Testing Results : Diskful

- Rebooting using “systemctl reboot”
 - Average time until useable ~ 5 minutes 23 seconds
 - Minimum boot time = 5 minutes 16 seconds
- Rebooting using “systemctl kexec”
 - Average time until useable ~ 4 minutes 5 seconds
 - Minimum boot time = 4 minutes 3 seconds
- Boot time reduction ~ 24%



Testing Results : Diskless

- Rebooting using “systemctl reboot”
 - Average time until useable ~ 2 minutes 52 seconds
 - Minimum boot time = 2 minutes 46 seconds
- Rebooting using “systemctl kexec”
 - Average time until useable ~ 1 minutes
 - Minimum boot time = 58 seconds
- Boot time reduction ~ 65%



Conclusions & Future Work

- Kexec provided a faster boot process
- Kexec provided more consistent boot times
- Adding kexec functionality to Ironic upstream
 - In progress
- Testing kexec on large scale clusters