Unbreakable Enterprise Kernel

Release Notes for Unbreakable Enterprise Kernel Release 4 Update 7

E94694-08
September 2022

Unbreakable Enterprise Kernel Release Notes for Unbreakable Enterprise Kernel Release 4 Update 7,
E94694-08

Contents

Preface

Unbreakable Enterprise Kernel: Release Notes for Unbreakable Enterprise Kernel Release 4 Update 7 provides a summary of the new features, changes, and known issues in the Unbreakable Enterprise Kernel Release 4 Update 7.

Conventions

The following text conventions are used in this document:

Convention	Meaning
boldface	Boldface type indicates graphical user interface elements associated with an action, or terms defined in text or the glossary.
italic	Italic type indicates book titles, emphasis, or placeholder variables for which you supply particular values.
`monospace`	Monospace type indicates commands within a paragraph, URLs, code in examples, text that appears on the screen, or text that you enter.

Documentation Accessibility

For information about Oracle's commitment to accessibility, visit the Oracle Accessibility Program website at https://www.oracle.com/corporate/accessibility/.

For information about the accessibility of the Oracle Help Center, see the Oracle Accessibility Conformance Report at https://www.oracle.com/corporate/accessibility/templates/t2-11535.html.

Access to Oracle Support for Accessibility

Oracle customers that have purchased support have access to electronic support through My Oracle Support. For information, visit https://www.oracle.com/corporate/accessibility/learning-support.html#support-tab.

Diversity and Inclusion

Oracle is fully committed to diversity and inclusion. Oracle respects and values having a diverse workforce that increases thought leadership and innovation. As part of our initiative to build a more inclusive culture that positively impacts our employees, customers, and partners, we are working to remove insensitive terms from our products and documentation. We are also mindful of the necessity to maintain compatibility with our customers' existing technologies and the need to ensure continuity of service as Oracle's offerings and industry

standards evolve. Because of these technical constraints, our effort to remove insensitive terms is ongoing and will take time and external cooperation.

1

New Features, Bug Fixes and Notable Changes

The Unbreakable Enterprise Kernel Release 4 (UEK R4) is Oracle's fourth major release of its heavily tested and optimized operating system kernel for Oracle Linux 6 Update 7 or later, and Oracle Linux 7 Update 1 or later, on the x86-64 architecture. It is based on the mainline Linux kernel version 4.1.12.

UEK R4U7 uses the 4.1.12-124.14.1 version and build of the UEK R4 kernel, which includes security and bug fixes, as well as driver updates. This kernel has been tested within environments running the latest available Oracle Linux releases: Oracle Linux 6 Update 8, Oracle Linux 6 Update 9, Oracle Linux 7 Update 4 and Oracle Linux 7 Update 5.

Oracle actively monitors upstream check-ins and applies critical bug and security fixes to UEK R4.

> **Important:**
>
> Run the `yum update` command regularly to ensure that the latest bug fixes and security errata are installed on your system.

UEK R4 uses the same versioning model as the mainline Linux kernel version. It is possible that some applications might not understand the 4.1 versioning scheme. However, regular Linux applications are usually neither aware of nor affected by Linux kernel version numbers.

Notable Changes

The following notable changes and features are included in this update:

- **RDS support for IPv6 and other enhancements**

  See RDS Improvements for more information.

- **DTrace updates**

  See DTrace Improvements for more information.

RDS Improvements

Reliable Datagram Sockets (RDS) is a high-performance transport protocol that offers low overhead and low latency to deliver datagrams over a variety of transports, such as InfiniBand, loopback or TCP sockets. The following notable changes and features are included in this update:

- **RDS IPv6 support**

  Support for the use of IPv6 addresses has been added to the kernel RDS and related modules. Existing RDS applications using IPv4 addresses are able to continue to run

normally, but applications that require IPv6 addresses can do so by passing the address in struct `sockaddr_in6 to bind()`, `connect()` or `sendmsg()`.

Additional updates have been made to user space packages, such as `rds-tools` and `libibacl` to ensure that support is enabled in tools that can make use of this feature.

- **Improvements to RDS large fragment size implementation**

  A patch was applied to improve the RDS large fragment size implementation by taking advantage of multiple scatter-gather entries and to match the allocation of the scatter-gather entries to the PAGE_SIZE. This improves the performance and resolves issues in congestion handling.

- **Fix for NULL pointer dereference when using RDS with a debug kernel**

  A debug statement in the RDS code could cause a NULL pointer dereference resulting in a stack trace when running RDS on a debug version of the kernel. The fix moves the debug statement to avoid the NULL pointer dereference.

DTrace Improvements

A number of bug fixes and enhancements including module and utility updates are included for DTrace on UEK R4U7, bringing the current version to 1.0.0.

The following are other notable changes that are included:

- **Bug fixes**

  Numerous bug fixes have been applied to provide greater stability and better performance.

- **FBT improvements**

  Several patches and bug fixes were applied to improve performance and expand the capability of the Function Boundary Tracing (FBT) module.

- **Added lockstat probes**

  This update includes support for lockstat DTrace probes. These probes can be viewed using `dtrace -l -P lockstat`. DTrace lockstat support allows for dynamic tracing of kernel locking events. For example, these probes can provide information on which locks are most frequently used, which locks exhibit the most contention and which locks are held longest.

- **SDT probe improvements**

  Improvements were made to the SDT probes (`regular` and `isenabled`) to encode these as a call to a stub function, that could rewrite these calls as NOP sequences during the boot process to minimize the impact of the presence of these probes and to prevent the likelihood of a crash during system boot.

- **Increased precision of timestamps**

  Improvements were made to reintroduce a high resolution timer into DTrace so that time measurements returned in a timestamp variable are more accurate, reducing the possibility of a negative delta in a calculation.

- **Library interface changes**

  An interface problem that can cause DTrace consumers to dereference freed memory when victim processes grabbed via `ustack()`, `umod()`, `usym()` or `dtrace`

`-c` or `-p exec()` has been fixed. This requires changes to certain users of libdtrace, and relinking. The library soname has been bumped to `libdtrace.so.1` correspondingly. All consumers must relink, but consumers not using the `dtrace_proc_*()` APIs need no code changes. All places where code changes are needed elicit a compile-time error. The `dtrace_proc_*()` functions have changed the type they take to an opaque handle, `struct dtrace_proc`. There is a new function `dtrace_proc_getpid()` to get the PID from this opaque handle. `dtrace_proc_grab()` has been renamed to `dtrace_proc_grab_pid()`.

- **Compile-time array bounds checking**

  User space packages were updated to add checks of the bounds of non-associative arrays, both in CTF and in declared arrays. Lvalue arrays used for assignment are also bounds-checked.

KVM Improvements

The following are the notable fixes and improvements that have been made in this update:

- **Fix to remove inappropriate warning messages**

  A minor fix was applied for an issue that generated an innocuous warning message on a host running Oracle Linux 7 Update 4 and using `libvirt` or `virtmanager` to create a QEMU guest. The code that generated the issue has been updated to only print when debugging is enabled.

- **`pvclock-page` value handling issue resolved**

  An issue that triggered when a guest passes KVM its pvclock-page GPA for the first time is resolved to follow standard logic applied to other pvclock shared pages, preventing the page from initializing with an incorrect random value that could cause a system hang.

- **Upstream patches applied for better checks on VM Exit pending events**

  Upstream patches were applied to resolve a blocking error that could trigger when an event was re-injected to L2 and that could cause an L2 guest to exit to L1 even when there was no pending L1 event. The fix adds addition checks for pending events and returns `-EBUSY` if there is one.

- **Security fixes for retpoline support**

  Security fixes have been implemented to mitigate against kernel or cross-process memory disclosure such as the attack vector used by Spectre V2. A backport was introduced to fix an issue that resulted in the use of a stale MSR value generated by a previous VM exit where retpoline support is enabled in the host kernel. This issue directly affected KVM.

File System Improvements

The following file systems improvements have been made:

btrfs Updates

- Several patches were introduced to resolve a race condition when merging the internal extent map, which caused errors to get returned when performing multiple concurrent DirectIO reads/writes.

ext4 Updates

- A corner case that could cause a kernel panic when an incorrect memory address was calculated in one of the methods used by `kfree` to obtain a link, resulting in an -EIO error, has been fixed.
- Several upstream patches were applied to fix various issues in the SEEK_HOLE implementation.

NFSv4 Updates

- Patches were applied to fix an issue that caused system crashes on the server if a migration failed and resulted in a low reference count.

OCFS2 Updates

- A fix is applied for an issue that caused file system corruption that could not be resolved using `fsck`. This issue was caused by the fact that the first cluster group descriptor is not stored at the start of the group but at an offset from the start. The code has been updated to take this into account when doing an fstrim on the first cluster group.
- A fix was applied for an issue that caused a deadlock when there is a blocked remote lock request waiting for the lock to be down-converted. The problem was related to a change in the code to use generic POSIX ACL infrastructure which is unsuitable for use with ocfs2 inode creation with ACLs, as this code is unaware of cluster wide inode locks.

XFS Updates

- Several upstream patches were applied to fix various issues in the SEEK_HOLE implementation.
- Patches were applied to fix a hang that could occur while unmounting after the file system has gone offline due to storage problems or file system corruption.
- Patches were applied to prevent any attempt to create new files if the file system is already out of space. This fix can prevent file system corruption or a resulting internal error.

Driver Updates

The Unbreakable Enterprise Kernel supports a wide range of hardware and devices. In close cooperation with hardware and storage vendors, several device drivers have been updated or added by Oracle, per the information in the following table.

Driver	Version	Description
`be2net`	11.4.0.0 + Patches	Broadcom/Emulex OneConnect 10Gbps NIC Driver
`bnx2x`	1.713.10 + Patches	QLogic BCM57710/57711/57711E/ 57712/57712_MF/57800/57800_MF/ 57810/57810_MF/57840/57840_MF Driver
`bnxt_en`	1.9.0 + Patches	Broadcom/Emulex BCM573xx NIC Driver
`e1000`	7.3.21-k8-NAPI + Patches	Intel(R) PRO/1000 Network Driver
`e1000e`	3.2.6-k + Patches	Intel(R) PRO/1000 Network Driver

Driver	Version	Description
enic	2.3.0.45	Cisco VIC Ethernet NIC Driver
fnic	1.6.0.34 + Patches	Cisco FCoE HBA Driver
i40e	2.1.14-k + Patches	Intel® Ethernet Connection XL710 Network Driver
i40evf	3.0.0-k + Patches	Intel® XL710 X710 Virtual Function Network Driver
ixgbe	5.1.0-k + Patches	Intel® 10 Gigabit PCI Express Network Driver
ixgbevf	4.1.0-k + Patches	Intel® 82599 Virtual Function Driver
lpfc	11.4.0.7	Broadcom/Emulex LightPulse Fibre Channel SCSI Driver
megaraid_sas	07.704.04.00-rc1	Avago MegaRAID SAS Driver
mlx4_core	2.2-1 + Patches	Mellanox ConnectX HCA low-level driver
mpt3sas	16.100.00.00	LSI MPT Fusion SAS 3.0 Device Driver
nvme nvme-core	1.0 + Patches	NVMe Block Device Driver and Core Support
qla2xxx	9.00.00.00.40.0-k + Patches	QLogic Fibre Channel HBA Driver
qlcnic	5.3.65 + Patches	QLogic 1/10 GbE Converged/ Intelligent Ethernet Driver
qmi_wwan	Patches	Qualcomm MSM Interface (QMI) WWAN driver
smartpqi	1.1.2-126	Microsemi Smart Family Controller Driver
xen-blkback xen-blkfront xen-netback xen-netfront	Patches	Xen Virtual Block Device and Xen Virtual Network Device drivers
xscore	6.0.r8044 + Patches	Oracle Virtual Network Driver Modules for Core Support.

Additional Notes For Driver Updates

The following notes are included at the request of a vendor for the listed driver:

- lpfc: Locked optics support is enabled for LPE32000 HBAs and all variant HBAs of this architecture. With this capability, these HBAs will detect and enable both Avago or Emulex certified SFP and QSFP optics.

  For driver rev 11.0.0.13 and higher unqualified optics will be disabled, the link will not come up, an error message is written to the log file and the lpfc driver will display this message:

  ```
  3176 Port Name [wwpn] Unqualified optics - Replace with Avago optics for Warranty and Technical support
  ```

For driver rev 11.4.0.7 and higher, Target Queue Depth can be tuned dynamically to control performance or to help manage potential queuing problems. The default Target Queue Depth is 65535, which typically offers the greatest performance. This can be changed to a value in the range 10 to 65535. When changing the Target Queue Depth the new target queue depth remains in effect for a minimum of 40 seconds. Any changes made within the 40 second window are deferred and do not take effect until the 40 second window has passed. Note that the 40 second window may be removed in a future version of this driver module. To change the Target Queue Depth on SCSI Host X run the command:

```
# echo new_value > /sys/class/scsi_host/hostX/lpfc_tgt_queue_depth
```

Replace *new_value* with an integer between 10 and 65535.

- `mpt3sas`: As of UEK R4U4, the `mpt2sas` driver has been merged with the `mpt3sas` driver to provide a single driver module that supports both SAS 2.0 and SAS 3.0 HBAs. Changes have been applied to `dracut` to correctly handle the module aliases for the migration to a single driver module.

Technology Preview

The following features that are included in the Unbreakable Enterprise Kernel Release 4 are still under development, but are made available for testing and evaluation purposes:

- **DCTCP (Data Center TCP)**

  DCTCP enhances congestion control by making use of the Explicit Congestion Notification (ECN) feature of state-of-the-art network switches. DCTCP reduces buffer occupancy and improves throughput by allowing a system to react more intelligently to congestion than is possible using TCP.

- **DRBD (Distributed Replicated Block Device)**

  A shared-nothing, synchronously replicated block device (*RAID1 over network*), designed to serve as a building block for high availability (HA) clusters. It requires a cluster manager (for example, pacemaker) for automatic failover.

- **Kernel module signing facility**

  Applies cryptographic signature checking to modules on module load, checking the signature against a ring of public keys compiled into the kernel. GPG is used to do the cryptographic work and determines the format of the signature and key data.

- **Server-side parallel NFS**

  Server-side parallel NFS (pNFS) improves the scalability and performance of an NFS server by making file metadata and data available on separate paths.

Compatibility

Oracle Linux maintains user-space compatibility with Red Hat Enterprise Linux (RHEL), which is independent of the kernel version running underneath the operating system. Existing applications in user space will continue to run unmodified on the Unbreakable Enterprise Kernel Release 4 and no re-certifications are needed for RHEL certified applications.

To minimize impact on interoperability during releases, the Oracle Linux team works closely with third-party vendors whose hardware and software have dependencies on kernel modules. The kernel ABI for UEK R4 will remain unchanged in all subsequent updates to the initial release. In this release, there are changes to the kernel ABI relative to UEK R3 that require recompilation of third-party kernel modules on the system. Before installing UEK R4, verify its support status with your application vendor.

Header Packages for Development

As of UEK-3.8-QU2, the `kernel-uek-headers` package is no longer built and distributed. There are three kernel packages that might be useful for development purposes. The `kernel-headers` package forms part of the API for user space programs. The `kernel-devel` package is used for standard RHCK development and module compilation. The `kernel-uek-devel` package is used for UEK development and module compilation. Neither the `kernel-uek-headers`, nor the `kernel-headers` packages, are needed for kernel development.

The `kernel-headers` package provides the C header files that specify the interface between user-space binaries or libraries and UEK or RHCK. These header files define the structures and constants that you need to build most standard programs or to rebuild the `glibc` package.

The `kernel-devel` and `kernel-uek-devel` packages provide the kernel headers and makefiles that you need to build modules against UEK and RHCK.

To install the packages required to build modules against UEK and the C header files for both UEK and RHCK:

```
# yum install kernel-uek-devel-`uname -r` kernel-headers
```

2

Security Fixes for CVEs

This chapter lists security vulnerabilities and exposures (CVEs) that are specifically addressed in this release. Note that CVEs are continually handled in patch updates that are made available as errata builds for the current release. For this reason, it is absolutely critical that you keep your system up to date with the latest package updates for this kernel release.

You can keep up to date with the latest CVE information at https://linux.oracle.com/cve.

List of CVEs fixed in this release

The following list describes the CVEs that are fixed in this release. The content provided here is automatically generated and includes the CVE identifier and a summary of the issue. The associated internal Oracle bug identifiers are also included to reference work that was carried out to address each issue.

- **CVE-2016-10318**

  A missing authorization check in the fscrypt_process_policy function in fs/crypto/policy.c in the ext4 and f2fs filesystem encryption support in the Linux kernel before 4.7.4 allows a user to assign an encryption policy to a directory owned by a different user, potentially creating a denial of service. (Bug: 25883175)

  See https://linux.oracle.com/cve/CVE-2016-10318.html for more information.

- **CVE-2016-9191**

  The cgroup offline implementation in the Linux kernel through 4.8.11 mishandles certain drain operations, which allows local users to cause a denial of service (system hang) by leveraging access to a container environment for executing a crafted application, as demonstrated by trinity. (Bug: 25062944 27841944)

  See https://linux.oracle.com/cve/CVE-2016-9191.html for more information.

- **CVE-2017-0861**

  Use-after-free vulnerability in the snd_pcm_info function in the ALSA subsystem in the Linux kernel allows attackers to gain privileges via unspecified vectors. (Bug: 27344839)

  See https://linux.oracle.com/cve/CVE-2017-0861.html for more information.

- **CVE-2017-1000112**

  Linux kernel: Exploitable memory corruption due to UFO to non-UFO path switch. When building a UFO packet with MSG_MORE __ip_append_data() calls ip_ufo_append_data() to append. However in between two send() calls, the append path can be switched from UFO to non-UFO one, which leads to a memory corruption. In case UFO packet lengths exceeds MTU, copy = maxfraglen - skb->len becomes negative on the non-UFO path and the branch to allocate new skb is taken. This triggers fragmentation and computation of fraggap = skb_prev->len - maxfraglen. Fraggap can exceed MTU, causing copy = datalen - transhdrlen - fraggap to become negative. Subsequently skb_copy_and_csum_bits() writes out-of-bounds. A similar issue is present in IPv6 code. The bug was introduced in e89e9cf539a2 ("[IPv4/IPv6]: UFO Scatter-gather approach") on Oct 18 2005. (Bug: 26921303)

See https://linux.oracle.com/cve/CVE-2017-1000112.html for more information.

- **CVE-2017-1000405**

  The Linux Kernel versions 2.6.38 through 4.14 have a problematic use of pmd_mkdirty() in the touch_pmd() function inside the THP implementation. touch_pmd() can be reached by get_user_pages(). In such case, the pmd will become dirty. This scenario breaks the new can_follow_write_pmd()'s logic - pmd can become dirty without going through a COW cycle. This bug is not as severe as the original "Dirty cow" because an ext4 file (or any other regular file) cannot be mapped using THP. Nevertheless, it does allow us to overwrite read-only huge pages. For example, the zero huge page and sealed shmem files can be overwritten (since their mapping can be populated using THP). Note that after the first write page-fault to the zero page, it will be replaced with a new fresh (and zeroed) thp. (Bug: 27165913)

  See https://linux.oracle.com/cve/CVE-2017-1000405.html for more information.

- **CVE-2017-1000407**

  The Linux Kernel 2.6.32 and later are affected by a denial of service, by flooding the diagnostic port 0x80 an exception can be triggered leading to a kernel panic. (Bug: 27206805)

  See https://linux.oracle.com/cve/CVE-2017-1000407.html for more information.

- **CVE-2017-10661**

  Race condition in fs/timerfd.c in the Linux kernel before 4.10.15 allows local users to gain privileges or cause a denial of service (list corruption or use-after-free) via simultaneous file-descriptor operations that leverage improper might_cancel queueing. (Bug: 26673877)

  See https://linux.oracle.com/cve/CVE-2017-10661.html for more information.

- **CVE-2017-12154**

  The prepare_vmcs02 function in arch/x86/kvm/vmx.c in the Linux kernel through 4.13.3 does not ensure that the "CR8-load exiting" and "CR8-store exiting" L0 vmcs02 controls exist in cases where L1 omits the "use TPR shadow" vmcs12 control, which allows KVM L2 guest OS users to obtain read and write access to the hardware CR8 register.

  See https://linux.oracle.com/cve/CVE-2017-12154.html for more information.

- **CVE-2017-12190**

  The bio_map_user_iov and bio_unmap_user functions in block/bio.c in the Linux kernel before 4.13.8 do unbalanced refcounting when a SCSI I/O vector has small consecutive buffers belonging to the same page. The bio_add_pc_page function merges them into one, but the page reference is never dropped. This causes a memory leak and possible system lockup (exploitable against the host OS by a guest OS user, if a SCSI disk is passed through to a virtual machine) due to an out-of-memory condition. (Bug: 27062562)

  See https://linux.oracle.com/cve/CVE-2017-12190.html for more information.

- **CVE-2017-12192**

  The keyctl_read_key function in security/keys/keyctl.c in the Key Management subcomponent in the Linux kernel before 4.13.5 does not properly consider that a key may be possessed but negatively instantiated, which allows local users to cause a denial of service (OOPS and system crash) via a crafted KEYCTL_READ operation. (Bug: 27049926)

See https://linux.oracle.com/cve/CVE-2017-12192.html for more information.

- **CVE-2017-12193**

  The assoc_array_insert_into_terminal_node function in lib/assoc_array.c in the Linux kernel before 4.13.11 mishandles node splitting, which allows local users to cause a denial of service (NULL pointer dereference and panic) via a crafted application, as demonstrated by the keyring key type, and key addition and link creation operations. (Bug: 27364588)

  See https://linux.oracle.com/cve/CVE-2017-12193.html for more information.

- **CVE-2017-14106**

  The tcp_disconnect function in net/ipv4/tcp.c in the Linux kernel before 4.12 allows local users to cause a denial of service (__tcp_select_window divide-by-zero error and system crash) by triggering a disconnect within a certain tcp_recvmsg code path. (Bug: 26796038)

  See https://linux.oracle.com/cve/CVE-2017-14106.html for more information.

- **CVE-2017-14140**

  The move_pages system call in mm/migrate.c in the Linux kernel before 4.12.9 doesn't check the effective uid of the target process, enabling a local attacker to learn the memory layout of a setuid executable despite ASLR. (Bug: 27364683)

  See https://linux.oracle.com/cve/CVE-2017-14140.html for more information.

- **CVE-2017-14489**

  The iscsi_if_rx function in drivers/scsi/scsi_transport_iscsi.c in the Linux kernel through 4.13.2 allows local users to cause a denial of service (panic) by leveraging incorrect length validation. (Bug: 26828494)

  See https://linux.oracle.com/cve/CVE-2017-14489.html for more information.

- **CVE-2017-15115**

  The sctp_do_peeloff function in net/sctp/socket.c in the Linux kernel before 4.14 does not check whether the intended netns is used in a peel-off action, which allows local users to cause a denial of service (use-after-free and system crash) or possibly have unspecified other impact via crafted system calls. (Bug: 27386997)

  See https://linux.oracle.com/cve/CVE-2017-15115.html for more information.

- **CVE-2017-15537**

  The x86/fpu (Floating Point Unit) subsystem in the Linux kernel before 4.13.5, when a processor supports the xsave feature but not the xsaves feature, does not correctly handle attempts to set reserved bits in the xstate header via the ptrace() or rt_sigreturn() system call, allowing local users to read the FPU registers of other processes on the system, related to arch/x86/kernel/fpu/regset.c and arch/x86/kernel/fpu/signal.c. (Bug: 27050688)

- **CVE-2017-15649**

  net/packet/af_packet.c in the Linux kernel before 4.13.6 allows local users to gain privileges via crafted system calls that trigger mishandling of packet_fanout data structures, because of a race condition (involving fanout_add and packet_do_bind) that leads to a use-after-free, a different vulnerability than CVE-2017-6346. (Bug: 27050772)

  See https://linux.oracle.com/cve/CVE-2017-15649.html for more information.

- **CVE-2017-16525**

The usb_serial_console_disconnect function in drivers/usb/serial/console.c in the Linux kernel before 4.13.8 allows local users to cause a denial of service (use-after-free and system crash) or possibly have unspecified other impact via a crafted USB device, related to disconnection and failed setup. (Bug: 27206824)

See https://linux.oracle.com/cve/CVE-2017-16525.html for more information.

- **CVE-2017-16526**

  drivers/uwb/uwbd.c in the Linux kernel before 4.13.6 allows local users to cause a denial of service (general protection fault and system crash) or possibly have unspecified other impact via a crafted USB device. (Bug: 27206874)

  See https://linux.oracle.com/cve/CVE-2017-16526.html for more information.

- **CVE-2017-16527**

  sound/usb/mixer.c in the Linux kernel before 4.13.8 allows local users to cause a denial of service (snd_usb_mixer_interrupt use-after-free and system crash) or possibly have unspecified other impact via a crafted USB device. (Bug: 27117850)

  See https://linux.oracle.com/cve/CVE-2017-16527.html for more information.

- **CVE-2017-16529**

  The snd_usb_create_streams function in sound/usb/card.c in the Linux kernel before 4.13.6 allows local users to cause a denial of service (out-of-bounds read and system crash) or possibly have unspecified other impact via a crafted USB device. (Bug: 27206916)

  See https://linux.oracle.com/cve/CVE-2017-16529.html for more information.

- **CVE-2017-16530**

  The uas driver in the Linux kernel before 4.13.6 allows local users to cause a denial of service (out-of-bounds read and system crash) or possibly have unspecified other impact via a crafted USB device, related to drivers/usb/storage/uas-detect.h and drivers/usb/storage/uas.c. (Bug: 27206993)

  See https://linux.oracle.com/cve/CVE-2017-16530.html for more information.

- **CVE-2017-16531**

  drivers/usb/core/config.c in the Linux kernel before 4.13.6 allows local users to cause a denial of service (out-of-bounds read and system crash) or possibly have unspecified other impact via a crafted USB device, related to the USB_DT_INTERFACE_ASSOCIATION descriptor. (Bug: 27207211)

  See https://linux.oracle.com/cve/CVE-2017-16531.html for more information.

- **CVE-2017-16532**

  The get_endpoints function in drivers/usb/misc/usbtest.c in the Linux kernel through 4.13.11 allows local users to cause a denial of service (NULL pointer dereference and system crash) or possibly have unspecified other impact via a crafted USB device. (Bug: 27602322)

- **CVE-2017-16533**

  The usbhid_parse function in drivers/hid/usbhid/hid-core.c in the Linux kernel before 4.13.8 allows local users to cause a denial of service (out-of-bounds read and system crash) or possibly have unspecified other impact via a crafted USB device. (Bug: 27207901)

  See https://linux.oracle.com/cve/CVE-2017-16533.html for more information.

- **CVE-2017-16535**

  The usb_get_bos_descriptor function in drivers/usb/core/config.c in the Linux kernel before 4.13.10 allows local users to cause a denial of service (out-of-bounds read and system crash) or possibly have unspecified other impact via a crafted USB device. (Bug: 27207955)

  See https://linux.oracle.com/cve/CVE-2017-16535.html for more information.

- **CVE-2017-16536**

  The cx231xx_usb_probe function in drivers/media/usb/cx231xx/cx231xx-cards.c in the Linux kernel through 4.13.11 allows local users to cause a denial of service (NULL pointer dereference and system crash) or possibly have unspecified other impact via a crafted USB device. (Bug: 27208030)

  See https://linux.oracle.com/cve/CVE-2017-16536.html for more information.

- **CVE-2017-16646**

  drivers/media/usb/dvb-usb/dib0700_devices.c in the Linux kernel through 4.13.11 allows local users to cause a denial of service (BUG and system crash) or possibly have unspecified other impact via a crafted USB device. (Bug: 27215141)

- **CVE-2017-16649**

  The usbnet_generic_cdc_bind function in drivers/net/usb/cdc_ether.c in the Linux kernel through 4.13.11 allows local users to cause a denial of service (divide-by-zero error and system crash) or possibly have unspecified other impact via a crafted USB device. (Bug: 27841392)

- **CVE-2017-16650**

  The qmi_wwan_bind function in drivers/net/usb/qmi_wwan.c in the Linux kernel through 4.13.11 allows local users to cause a denial of service (divide-by-zero error and system crash) or possibly have unspecified other impact via a crafted USB device. (Bug: 27215213)

  See https://linux.oracle.com/cve/CVE-2017-16650.html for more information.

- **CVE-2017-17052**

  The mm_init function in kernel/fork.c in the Linux kernel before 4.12.10 does not clear the ->exe_file member of a new process's mm_struct, allowing a local attacker to achieve a use-after-free or possibly have unspecified other impact by running a specially crafted program. (Bug: 27648200)

  See https://linux.oracle.com/cve/CVE-2017-17052.html for more information.

- **CVE-2017-17712**

  The raw_sendmsg() function in net/ipv4/raw.c in the Linux kernel through 4.14.6 has a race condition in inet->hdrincl that leads to uninitialized stack pointer usage; this allows a local user to execute code and gain privileges. (Bug: 27390679)

  See https://linux.oracle.com/cve/CVE-2017-17712.html for more information.

- **CVE-2017-2618**

  A flaw was found in the Linux kernel's handling of clearing SELinux attributes on /proc/pid/attr files. An empty (null) write to this file can crash the system by causing the system to attempt to access unmapped kernel memory. (Bug: 25660054)

  See https://linux.oracle.com/cve/CVE-2017-2618.html for more information.

- **CVE-2017-5715**

Systems with microprocessors utilizing speculative execution and indirect branch prediction may allow unauthorized disclosure of information to an attacker with local user access via a side-channel analysis. (Bug: 27344012 27365575 27461990 27477743 27542331)

See https://linux.oracle.com/cve/CVE-2017-5715.html for more information.

- **CVE-2017-5753**

  Systems with microprocessors utilizing speculative execution and branch prediction may allow unauthorized disclosure of information to an attacker with local user access via a side-channel analysis. (Bug: 27340445)

  See https://linux.oracle.com/cve/CVE-2017-5753.html for more information.

- **CVE-2017-5754**

  Systems with microprocessors utilizing speculative execution and indirect branch prediction may allow unauthorized disclosure of information to an attacker with local user access via a side-channel analysis of the data cache. (Bug: 27333760 27365431 27378516)

  See https://linux.oracle.com/cve/CVE-2017-5754.html for more information.

- **CVE-2017-7482**

  When a kerberos 5 ticket is being decoded so that it can be loaded into an rxrpc-type key, there are several places in which the length of a variable-length field is checked to make sure that it's not going to overrun the available data - but the data is padded to the nearest four-byte boundary and the code doesn't check for this extra. This could lead to the size-remaining variable wrapping and the data pointer going over the end of the buffer. (Bug: 26376434)

  See https://linux.oracle.com/cve/CVE-2017-7482.html for more information.

- **CVE-2017-7518**

  A flaw was found in the way the Linux KVM module processed the trap flag(TF) bit in EFLAGS during emulation of the syscall instruction, which leads to a debug exception(#DB) being raised in the guest stack. A user/process inside a guest could use this flaw to potentially escalate their privileges inside the guest. Linux guests are not affected by this. (Bug: 27669904)

  See https://linux.oracle.com/cve/CVE-2017-7518.html for more information.

- **CVE-2017-7541**

  The brcmf_cfg80211_mgmt_tx function in drivers/net/wireless/broadcom/brcm80211/brcmfmac/cfg80211.c in the Linux kernel before 4.12.3 allows local users to cause a denial of service (buffer overflow and system crash) or possibly gain privileges via a crafted NL80211_CMD_FRAME Netlink packet. (Bug: 26540118)

  See https://linux.oracle.com/cve/CVE-2017-7541.html for more information.

- **CVE-2017-7542**

  The ip6_find_1stfragopt function in net/ipv6/output_core.c in the Linux kernel through 4.12.3 allows local users to cause a denial of service (integer overflow and infinite loop) by leveraging the ability to open a raw socket. (Bug: 26540159)

  See https://linux.oracle.com/cve/CVE-2017-7542.html for more information.

- **CVE-2017-7618**

crypto/ahash.c in the Linux kernel through 4.10.9 allows attackers to cause a denial of service (API operation calling its own callback, and infinite recursion) by triggering EBUSY on a full queue. (Bug: 25882988)

See https://linux.oracle.com/cve/CVE-2017-7618.html for more information.

- **CVE-2017-8824**

  The dccp_disconnect function in net/dccp/proto.c in the Linux kernel through 4.14.3 allows local users to gain privileges or cause a denial of service (use-after-free) via an AF_UNSPEC connect system call during the DCCP_LISTEN state. (Bug: 27290292)

  See https://linux.oracle.com/cve/CVE-2017-8824.html for more information.

- **CVE-2018-1068**

  A flaw was found in the Linux 4.x kernel's implementation of 32-bit syscall interface for bridging. This allowed a privileged user to arbitrarily write to a limited range of kernel memory. (Bug: 27774012)

3
Known Issues

This chapter describes the known issues in this update.

dmi: Firmware registration failure message in dmesg output

A benign error message may appear in dmesg output to display similarly to:

```
[    0.118041] dmi: Firmware registration failed.
```

The issue does not relate to a firmware registration issue, but rather to a minor issue creating a sysfs file needed by DMI. (Bug ID 27687990)

i40e driver can cause a system hang when a high number of VFs are created

When attempting to create the maximum number of SR-IOV Virtual Functions (VFs), on an X7-2 or X7-8 system using the Intel Corporation Ethernet Controller X710/X557-AT 10GBASE-T (rev 01), the command can hang and can eventually cause the system to hang. Specifically, setting the VFs to the maximum value works on the first attempt, but iterative resetting of the value causes the issue on higher values. (Bug ID 27567377)

KVM guests with less than 4 GB of memory might fail to auto reserve crash kernel memory

KVM guests that are running UEK R4 might fail to auto reserve memory for a crash kernel if the `crashkernel=auto` setting is used on a guest with less than 4 GB of physical memory.

To prevent this problem from occurring on KVM guests with less than 4 GB of physical memory, you can reserve memory for a crash kernel by explicitly requesting a reservation size, for example, `crashkernel=128M`. (Bug ID 26933217)

btrfs, ext4 and xfs: Kernel panic when freeze and unfreeze operations are performed in multiple threads

Freeze and unfreeze operations that are performed across multiple threads on any supported file system can cause the system to hang and the kernel to panic. This problem is the result of a race condition that occurs when the unfreeze operation is triggered before it is actually frozen. The resulting unlock operation attempts a write operation on a non-existent lock, resulting in the kernel panic. (Bug ID 25321899)

btrfs Issues

The following are known `btrfs` issues:

- **Send operation causes soft lockup on large deduped file**

  Using `btrfs send` on a large deduped file results in a soft lockup or out-of-memory issue. This problem occurs because the `btrfs send` operation cannot handle a large deduped file containing file extents that are all pointing to one extent, as these types of file structures create tremendous pressure for the `btrfs send` operation.

  To prevent this issue from occurring, do not use `btrfs send` on systems with less than 4 GB of memory. (Bug ID 25306023)

- **Kernel oops when unmounting during a quota rescan or disable**

  Operations that trigger a quota rescan or to disable the quota on a mounted file system cause a kernel oops message when attempting to unmount the file system. This can cause the system to hang. (Bug ID 22377928)

- **Kernel oops when removing shared extents using qgroup accounting**

  The removal of shared extents where quota group (qgroup) accounting is used can result in a kernel oops message. This relates to an issue where inaccurate results are obtained during a back reference walk, due to missing records when adding delayed references. (Bug ID 21554517)

- **No warning when balancing file system on RAID**

  The `btrfs filesystem balance` command does not warn that the RAID level can be changed under certain circumstances, and does not provide the choice of cancelling the operation. (Bug ID 16472824)

- **Double count of overwritten space in qgroup show**

  When you overwrite data in a file, starting somewhere in the middle of the file, the overwritten space is counted twice in the space usage numbers that `btrfs qgroup show` displays. Using the `btrfs quota rescan` does not help fix this issue either. (Bug ID 16609467)

- **Location of `btrfs-progs` and `btrfs-progs-devel` packages**

  The `btrfs-progs` and `btrfs-progs-devel` packages for use with UEK R4 are made available in the `ol6_x86_64_UEKR4` and `ol7_x86_64_UEKR4` ULN channels and the `ol6_UEKR4` and `ol7_UEKR4` channels on the Oracle Linux yum server. In UEK R3, these packages were made available in the `ol6_x86_64_latest` and `ol7_x86_64_latest` ULN channels and the `ol6_latest` and `ol7_latest` channels on the Oracle Linux yum server.

ext4 Issues

The following are known `ext4` issues:

- **System hangs on unmount after an append to a file with negative `i_size`**

  While it is invalid for a file system to load an inode with a negative `i_size`, it is possible to create a file like this and append to it. However, doing so causes an

integer overflow in the routine's underlying writeback, resulting in the kernel locking up. (Bug ID 25565527)

- **Hang occurs during dynamic expansion of inode size**

  A hang occurs with the `ext4` file system during the dynamic expansion of inode size when using the inode's `i_extra_size` field. (Bug ID 25718971)

xfs Issues

The following are known `xfs` issues:

- **Invalid corrupted file system error resulting from a problem with log recovery on v5 superblocks**

  A problem with log recovery on v5 superblocks that causes the metadata LSN not to update for buffers that it writes out, can result in a corruption error similar to the following:

```
[1044224.901444] XFS (sdc1): Metadata corruption detected at
xfs_dir3_block_write_verify+0xfd/0x110 [xfs], block 0x1004e90
[1044224.901446] XFS (sdc1): Unmount and run xfs_repair
...
[1044224.901460] XFS (sdc1): xfs_do_force_shutdown(0x8) called from line 1249
of file fs/xfs/xfs_buf.c.  Return address = 0xffffffffa07a8910
[1044224.901462] XFS (sdc1): Corruption of in-memory data detected.  Shutting
down filesystem
[1044224.901463] XFS (sdc1): Please umount the filesystem and rectify the
problem(s)
[1044224.904207] XFS (sdc1): log mount/recovery failed: error -117
[1044224.904456] XFS (sdc1): log mount failed"
```

  This problem is encountered because the log attempts to replay a buffer update that is no longer valid due to subsequent replayed updates. The result is a corruption error, when in fact, the file system is fine. (Bug ID 25380003)

- **System hangs on unmount after a buffered append to a file with negative `i_size`**

  While it is invalid for a file system to load an inode with a negative `i_size`, it is possible to create a file like this, and in the case where a buffer appends to it, an integer overflow in the routine's underlying writeback results in the kernel locking up. A direct append does not cause this behavior. (Bug ID 25565490)

- **System hangs during `xfs_fsr` on two-extent files with speculative preallocation**

  During an `xfs_fsr` process on extents that are generated by speculative preallocation, the code that determines whether all of the extents fit inline miscalculates because the `di_nextents` call that is used does not account for these extents. This results in corruption of the in-memory inode, and ultimately the code attempts to move memory structures using incorrectly calculated ranges. This causes a kernel panic. (Bug ID 25333211)

- **XFS quotas are disabled after a read-only remount on Oracle Linux 6**

  Quotas are disabled on XFS if the file system is remounted with read-only permissions on Oracle Linux 6. (Bug ID 22908906)

- **Overlay file system is unable to mount on XFS where there is no `d_type` support**

  Overlay file systems rely on a feature known as `d_type` support. This feature is a field within a data structure that provides some metadata about files in a directory entry within the base file system. Overlay file systems use this field to track many file operations such

as file ownership changes and whiteouts. `d_type` support can be enabled in XFS when the file system is created, by using the `-n ftype=1` option. When `d_type` support is not enabled, an overlay file system might become corrupt and behave in unexpected ways. For this reason, this update release of UEK R4 prevents the mounting of an overlay file system on an XFS base, where `d_type` support is not enabled.

The `root` partition on Oracle Linux is automatically formatted with `-n ftype=0`, where XFS is selected as the file system. Thus, for backward compatibility reasons, if you have overlay file systems in place already and these are not hosted on alternate storage, you must migrate them to a file system that is formatted with `d_type` support enabled.

To check that the XFS file system is formatted correctly:

```
# xfs_info /dev/sdb1 |grep ftype
```

Replace */dev/sdb1* with the path to the correct storage device. If the information returned by this command includes `ftype=0`, you must migrate the overlay data held in this directory to storage that is formatted correctly.

To correctly format a new block device with the XFS file system with support for overlay file systems, do:

```
# mkfs -t xfs -n ftype=1 /dev/sdb1
```

Replace */dev/sdb1* with the path to the correct storage device. It is essential that you use the `-n ftype=1` option when you create the file system.

If you do not have additional block storage available, it is possible to create an XFS file system image and loopback that can be mounted. For example, to create a 5 GB image file in the `root` directory, you could use the following command:

```
# mkfs.xfs -d file=1,name=/OverlayStorage,size=5g -n ftype=1
```

To temporarily mount this file, you can enter:

```
# mount -o loop -t xfs /OverlayStorage /mnt
```

Adding an entry in `/etc/fstab` to make a permanent mount for this storage, might look similar to the following:

```
                        /OverlayStorage     /mnt        xfs
loop            0 0
```

This configuration can help as a temporary solution to solve upgrade issues. However, using a loopback mounted file system image as a form of permanent storage is not recommended for production environments. (Bug ID 26165630)

DIF/DIX is not supported for ext file systems

The Data Integrity Field (DIF) and Data Integrity Extension (DIX) features that have been added to the SCSI standard are dependent on a file system that is capable of correctly handling attempts by the memory management system to change data in the buffer while it is queued for a write.

The ext2, ext3 and ext4 file system drivers do not prevent pages from being modified during I/O which can cause checksum failures and a "Logical block guard check failed" error. Other file systems such as XFS are supported. (Bug ID 24361968)

Console appears to hang when booting

When booting Oracle Linux 6 on hardware with an ASPEED graphics controller, the console might appear to hang during the boot process after starting `udev`. However, the system does boot properly and is accessible. The workaround is to add `nomodeset` as a kernel boot parameter in `/etc/grub.conf`. (Bug ID 22389972)

Docker Issues

The following are known Docker issues:

- **Running `yum install` within a container on an overlayfs file system can fail with the following error:**

  ```
  Rpmdb checksum is invalid: dCDPT(pkg checksums): package_name
  ```

  This error can break Dockerfile builds but is expected behavior from the kernel and is a known issue upstream (see https://github.com/docker/docker/issues/10180.)

  The workaround is to run `touch /var/lib/rpm/*` before installing the package.

  Note that this issue is fixed in any Oracle Linux images available on the Docker Hub or Oracle Container Registry, but the issue could still be encountered when running any container based on a third-party image. (Bug ID 21804564)

- **Docker can fail where it uses the `overlay2` storage driver on XFS-formatted storage**

  A kernel patch has been applied to prevent overlay mounts on XFS if the `ftype` is not set to 1. This fix resolves an issue where XFS did not properly support the whiteout features of an overlay filesystem if `d_type` support was not enabled. If the Docker Engine is already using XFS-formatted storage with the `overlay2` storage driver, an upgrade of the kernel can cause Docker to fail if the underlying XFS file system is not created with the `-n ftype=1` option enabled. The root partition on Oracle Linux 7 is automatically formatted with `-n ftype=0` where XFS is selected as the file system. Therefore, if you intend to use the `overlay2` storage driver in this environment, you must format a separate device for this purpose. (Bug ID 25995797)

DTrace Issues

The following are known DTrace issues:

- Argument declarations with USDT probe definitions cannot be declared with derived types such as `enum`, `struct`, or `union`.
- The following compiler warning can be ignored for USDT probe definition arguments of type `string` (which is a D type but not a C type):

  ```
  provider_def.h:line#: warning: parameter names (without
  types) in function declaration
  ```

-
- Multi-threaded processes under `ustack()`, `usym()`, `uaddr()` and `umod()`, which perform `dlopen()` in threads other than the first thread might not have accurate symbol resolution for symbols introduced by `dlopen()`. (Bug ID 20045149)

Error, some other host already uses address *xxx.xxx.xxx.xxx*

The following error message might be triggered in certain instances:

```
Error, some other host already uses address  xxx.xxx.xxx.xxx
```

The following are the two instances in which this error message might be triggered:

- When active-bonding is enabled, and you run the `ifup` *ib-interface* command.
- When you run the `service rdma start` command.

You can ignore this message, as in both cases, the InfiniBand interface is brought up successfully. (Bug IDs 21052903, 26639723)

Increased `dom0` memory requirement when using Mellanox HCAs on Oracle VM Server

Oracle VM Servers running UEKR4u2 and upward in `dom0` require at least 400MB more memory to use the Mellanox® drivers. This memory requirement is a result of the default size of the SRQ count being increased from 64K to 256K in later versions of the kernel and the `scale_profile` option is now enabled by default in the `mlx_core` module.

In the case where out-of-memory errors are observed in `dom0`, the maximum `dom0` memory size should be increased. Alternative workarounds might involve manually setting the module parameters for the `mlx4_core` driver. To set these parameters, edit `/etc/modprobe.d/mlx4_core.conf` and set `scale_profile` to `0`. Alternately, set `log_num_srq` to `16`. The preferred resolution to this issue is to increase the memory allocated to `dom0` on an Oracle VM Server. (Bug ID 23581534)

LXC Issues

The following are known LXC issues:

- **The `lxc-net` service does not always start immediately after installation on Oracle Linux 6**

  The `lxc-net` service does not always start immediately after installation on Oracle Linux 6, even though this action is specified as part of the RPM post-installation script. This can prevent the `lxcbr0` interface from coming up. If this interface is not up after installation, you can manually start it by running `service lxc-net start`. (Bug ID 23177405)

- **LXC read-only `ip_local_port_range` parameter**

With `lxc-1.1` or later and UEK R4, `ip_local_port_range` is a read-writable parameter under `/proc/sys/net/ipv4` in an Oracle Linux container rather than being read-only. (Bug ID 21880467)

Kdump fails to produce a `vmcore` file on systems running Oracle Linux 6 and using an Oracle NVMe PCIe 3.0 Switch Card V2

On a system running Oracle Linux 6, after a crash is triggered, `kdump` fails to generate a `vmcore` file if it is configured to dump the file to an NVMe device that is connected to an Oracle NVMe PCIe 3.0 Switch Card V2, and the system is running UEK R4U6 or later.

As a workaround, check that the NVMe solid-state drive (SSD) works by adding the `pci_aspm=off` kernel option to the `KDUMP_COMMANDLINE_APPEND` variable in `/etc/sysconfig/kdump`.

As an alternate workaround, consider using Oracle Linux 7. (Bug ID 27642801)

NVMe devices not found under the `/dev` directory after PCI rescan

After removing the PCI bus of NVM Express (NVMe) adapter card devices and running a rescan of the PCI bus, no NVMe adapter card devices are found under the `/dev` directory.

The workaround for this issue is to also remove the PCI slot that the NVMe adapter card device is plugged into before running a rescan of the PCI bus. (Bug ID 26610285)

OFED iSER target login fails from an initiator on Oracle Linux 6

An Oracle Linux 6 system with the `oracle-ofed-release` packages installed and an iSER (iSCSI Extensions for RDMA) target configured, fails to login to the iSER target as an initiator. On the Oracle Linux 6 initiator machine, the following behavior is typical:

```
# iscsiadm -m node -T iqn.iser-target.t1 -p 10.196.100.134 --login
Logging in to [iface: default, target: iqn.iser-target.t1, portal:
10.196.100.134,3260] (multiple)
iscsiadm: Could not login to [iface: default, target: iqn.iser-target.t1,
portal: 10.196.100.134,3260].
iscsiadm: initiator reported error (8 - connection timed out)
iscsiadm: Could not log into all portals
```

This is expected behavior resulting from an errata fix for CVE-2016-4564, to protect against a write from an invalid context.

(Bug ID 23615903)

Open File Description (OFD) locks are not supported on NFSv4 mounts

NFS is not designed to handle OFD locking. (Bug ID 22948696).

Oracle VM Server MSI-X interrupt allocation failure for 16 GB QLogic FC HBA

The Intel `ixgbe`/`ixgbevf` and QLogic `qla2xxx` drivers compete for MSI-X resources when using a 16 GB QLogic Fibre Channel HBA on systems that are running Oracle VM Server 3.4. As a result, if both drivers are used in a system, and an attempt is made to create the maximum number of Virtual Function (VF) devices that are allowed for the `ixgbe`/`ixgbevf` driver, an interrupt allocation failure occurs during the creation of the last VF device.

This issue is fully resolved by using the latest Oracle-supported firmware for the QLogic 16GB card (FC Firmware v 8.07.71 or later). (Bug IDs 25952728, 26916827)

Possible kernel crash during manual unloading of QLogic FC HBA driver module

A kernel crash might occur while manually unloading the QLogic Fibre Channel (FC) Host Bus Adapter (HBA) driver module for the Oracle 7101674 16 GB HBA model on an Oracle Sun Server X4-2 that is running UEK R4U6. (Bug ID 27248515)

RDMA service is not set to start at boot time

Because the `ibacm` service starts at boot time, but the `rdma` service does not start, an error indicating the `ibacm` service failed to start is displayed when the system boots. This error is also logged in `/var/log/ibacm.log` immediately after the system boots.

The workaround for this issue is to manually start the `ibacm` service after every boot by running the `service ibacm start` command.

(Bug IDs 26883485 and 27043535)

SDP performance degradation

The Sockets Direct Protocol (SDP), which was designed to provide an RDMA alternative to TCP over InfiniBand networks, is known to suffer from performance degradation on more recent kernels such as UEK R4U2 and later. There is no active development on this protocol.

Although the library for this protocol is still available for this kernel, support is limited. You should consider using TCP on top of IP over InfiniBand as a more stable alternative. (Bug ID 22354885)

Shared Receive Queue (SRQ) is an experimental feature for RDS and is disabled by default

The SRQ function that optimizes resource usage within the `rds_rdma` module is experimental and is disabled by default. A warning message is displayed when you enable this feature by setting the `rds_ib_srq_enabled` flag. (Bug ID 23523586).

Unloading or removing the `rds_rdma` module is unsupported

Once the `rds_rdma` module has been loaded, you cannot remove the module using either `rmmod` or `modprobe -r`. Unloading of the `rds_rdma` module is unsupported and can trigger a kernel panic. Do not set the `module_unload_allowed` flag for this module. (Bug ID 23580850).

4

Installation and Availability

You can install the Unbreakable Enterprise Kernel Release 4 on Oracle Linux 6 Update 7 or later, or Oracle Linux 7 Update 1 or later, running either the Red Hat compatible kernel or a previous version of the Unbreakable Enterprise Kernel. If you are still running an older version of Oracle Linux, first update your system to the latest available update release.

The Unbreakable Enterprise Kernel Release 4 is supported on the x86-64 architecture, but not on x86.

Installation Overview

If you have a subscription to Oracle Unbreakable Linux support, you can obtain the packages for Unbreakable Enterprise Kernel Release 4 by registering your system with the Unbreakable Linux Network (ULN) and subscribing it to additional channels. See Subscribing to ULN Channels.

If your system is not registered with ULN, you can obtain most of the packages from the Oracle Linux yum server. See Enabling Access to Oracle Yum Repositories.

Having subscribed your system to the appropriate channels on ULN or the Oracle Linux yum server, upgrade your system. See Upgrading Your System.

After upgrading to UEK R4, you can replace any existing OFED packages with the Oracle-supported OFED packages, see Installing the Oracle-Supported OFED Packages.

Subscribing to ULN Channels

The kernel image and user-space packages are available on the following ULN channels for Oracle Linux 6:

- `ol6_x86_64_latest` (latest user-space packages for Oracle Linux 6 other than DTrace, OFED, and DRBD packages)
- `ol6_x86_64_UEKR4` (`kernel-uek*`, `dtrace-modules-*`, and `libdtrace-*`)
- `ol6_x86_64_UEKR4_DTrace_userspace` (`dtrace-utils*`)
- `ol6_x86_64_UEKR4_OFED` (latest OFED tools packages)
- `ol6_x86_64_mysql-ha-utils` (`drbd84-utils`)

The kernel image and user-space packages are available on the following ULN channels for Oracle Linux 7:

- `ol7_x86_64_latest` (all of the latest user-space packages for Oracle Linux 7 other than DTrace, OFED, and DRBD packages)
- `ol7_x86_64_latest_optional` (the latest optional user-space packages for Oracle Linux 7 other than DTrace, OFED, and DRBD packages)
- `ol7_x86_64_UEKR4` (`kernel-uek*`, `dtrace-modules-*`, and `libdtrace-*`)

- `ol7_x86_64_UEKR4_DTrace_userspace` (`dtrace-utils*`)
- `ol7_x86_64_UEKR4_OFED` (latest OFED tools packages)
- `ol7_x86_64_mysql-ha-utils` (`drbd84-utils`)

The following procedure assumes that you have already registered your system with ULN.

To subscribe your system to a channel on ULN:

1. Log in to https://linux.oracle.com with your ULN user name and password.
2. On the Systems tab, click the link named for the system in the list of registered machines.
3. On the System Details page, click **Manage Subscriptions**.
4. On the System Summary page, select each required channel from the list of available channels and click the right arrow to move the channel to the list of subscribed channels.

   For Oracle Linux 6, subscribe the system to the `ol6_x86_64_latest` and `ol6_x86_64_UEKR4` channels. If required, you can also add the channels for the DTrace, OFED, and DRBD packages. You do not need to subscribe the system to the `ol6_x86_64_UEK_latest` or `ol6_x86_64_UEKR3_latest` channels.

   For Oracle Linux 7, subscribe the system to the `ol7_x86_64_latest` and `ol7_x86_64_UEKR4` channels. If required, you can also add the channels for the DTrace, OFED, and DRBD packages. You do not need to subscribe the system to the `ol7_x86_64_UEKR3` channel.
5. Click **Save Subscriptions**.

For information about using ULN, see Oracle® Linux: Unbreakable Linux Network User's Guide for Oracle Linux 6 and Oracle Linux 7

Enabling Access to Oracle Yum Repositories

On the Oracle Linux yum server at https://yum.oracle.com/, the kernel image and userspace packages are available on the following repositories.

For Oracle Linux 6:

- `ol6_latest` (latest user-space packages for Oracle Linux 6 other than the OFED tool packages)
- `ol6_UEKR4` (`kernel-uek*`, `dtrace-modules-*`, and `libdtrace-*`)
- `ol6_UEKR4_OFED` (latest OFED tools packages)

For Oracle Linux 7:

- `ol7_latest` (latest user-space packages for Oracle Linux 7 other than the OFED tool packages)
- `ol7_UEKR4` (`kernel-uek*`, `dtrace-modules-*`, and `libdtrace-*`)
- `ol7_UEKR4_OFED` (latest OFED tools packages)

> **Note:**
>
> To be able to install UEK R4, enable the appropriate `ol6_UEKR4` or `ol7_UEKR4` repository and disable the `ol6_UEKR3_latest` or `ol7_UEKR3` repository.
>
> The DRBD (Distributed Replicated Block Device) packages are not available on the Oracle Linux yum server.

To enable access to the Oracle Linux 6 repositories on the Oracle Linux yum server, use `yum-config-manager`. For example, to enable access to the `ol6_latest` and `ol6_UEKR4` repositories, run the following:

```
# yum-config-manager --enable ol6_latest,ol6_UEKR4
```

To enable access to the Oracle Linux 7 repositories on the Oracle Linux yum server, use `yum-config-manager`. For example, to enable access to the `ol7_latest` and `ol7_UEKR5` repositories, run the following:

```
# yum-config-manager --enable ol7_latest,ol7_UEKR5
```

> **Note:**
>
> You can only use `yum-config-manager` to enable or disable repositories where you already have a configuration file for the specified repository. Repository configurations are typically stored in `/etc/yum.repos.d`. The repository configurations required to install UEK on Oracle Linux 7 are included in the `oraclelinux-release-el7` package. The repository configurations required to install UEK on Oracle Linux 6 are included in the `oraclelinux-release-el6` package. If you do not have the `yum-config-manager` command available you may need to install the `yum-utils` package.

See https://docs.oracle.com/en/operating-systems/oracle-linux/6/admin/ol_yum.html in Oracle® Linux 6: Administrator's Guide or https://docs.oracle.com/en/operating-systems/oracle-linux/7/admin/ol7_yum.html in Oracle® Linux 7: Administrator's Guide for more information.

Upgrading Your System

To upgrade your system to UEK R4:

1. After enabling access to the appropriate channels, including `ol6_UEKR4` or `ol7_UEKR4`, on the Oracle Linux yum server or `ol6_x86_64_UEKR4` or `ol7_x86_64_UEKR4` on ULN, run the following command:

   ```
   # yum update
   ```

2. After upgrading the system, reboot it, selecting the UEK R4 kernel (version 4.1.12) if this is not the default boot kernel.

   See https://docs.oracle.com/en/operating-systems/oracle-linux/7/admin/ol7-bootconf.html for more information on updating the default boot kernel on Oracle Linux 7.

See https://docs.oracle.com/en/operating-systems/oracle-linux/6/admin/ol_bootconf.html for more information on updating the default boot kernel on Oracle Linux 6.

For instructions on how to install the Oracle-supported OFED packages after upgrading to UEK R4, see Installing the Oracle-Supported OFED Packages.

If you are upgrading from Oracle Linux 7 Update 3 or Oracle Linux 7 Update 4 to Oracle Linux 7 Update 5 and you already have Oracle-supported OFED packages for UEK R4 installed on your system, follow the upgrade procedures that are described in the Oracle® Linux 7: Release Notes for Oracle Linux 7 Update 5.

See https://docs.oracle.com/en/operating-systems/oracle-linux/6/admin/ol_yum.html in Oracle® Linux 6: Administrator's Guide or https://docs.oracle.com/en/operating-systems/oracle-linux/7/admin/ol7_yum.html in Oracle® Linux 7: Administrator's Guide for more information.

The kernel's source code is available via a public git source code repository at https://oss.oracle.com/git/?p=linux-uek.git;a=summary.

Installing the Oracle-Supported OFED Packages

The following procedure describes how to install the OFED packages that are provided by Oracle, including how to remove any existing OFED packages.

> **Note:**
>
> For any additional preparation that is required prior to installing or upgrading OFED packages, refer to the release notes for the Oracle Linux release that you are running.

To install the OFED packages that are provided by Oracle:

1. If your system is registered with ULN, subscribe the system to the `ol6_x86_64_UEKR4_OFED` or `ol7_x86_64_UEKR4_OFED` channel on ULN as appropriate.

   By default, the `ol7_x86_64_UEKR4` and `ol7_x86_64_latest` channels are enabled when you register an Oracle Linux 7 system with ULN; and the `ol6_x86_64_UEKR4` and `ol6_x86_64_latest` channels are enabled when you register an Oracle Linux 6 system with ULN. Check that these channels are still enabled before you begin installing the OFED packages provided by Oracle. If enabled, disable the `ol6_x86_64_optional latest` or `ol7_x86_64_latest` channel, or you may encounter dependency issues.

   If your system uses the Oracle Linux yum server ensure that your system is up to date and that you have transitioned to use the modular yum repository configuration by installing the `oraclelinux-release-el6` or `oraclelinux-release-el7` package and running the `/usr/bin/ol_yum_configure.sh` script. For example, run the following as root:

   ```
   # yum install oraclelinux-release-el7

   # /usr/bin/ol_yum_configure.sh
   ```

Then use `yum-config-manager` to enable the `ol7_UEKR4_OFED` repository for Oracle Linux 7 or the `ol6_UEKR4_OFED` repository for Oracle Linux 6. By default, `ol7_latest` and `ol7_UEKR4` are already enabled on Oracle Linux 7 and the `ol6_latest` and `ol6_UEKR4` are already enabled on Oracle Linux 6. If these repositories are not enabled on your system, you should enable these as well. For example as root, run:

```
# yum-config-manager --enable ol7_latest ol7_UEKR4 ol7_UEKR4_OFED
```

Use `yum-config-manager` to disable the `ol7_optional_latest` or `ol6_optional_latest` repository or you may encounter dependency issues if this repository is enabled. For exampe:

```
# yum-config-manager --disable ol7_optional_latest
```

2. If you are running Oracle Linux 7, stop and disable the `rdma.service` service.

```
# systemctl stop rdma.service
# systemctl disable rdma.service
```

3. Remove any existing OFED packages:

```
# yum remove 'ibacm*'
# yum remove 'ib-bonding*'
# yum remove 'ibutils*'
# yum remove 'infiniband-diags*'
# yum remove 'libibacl*'
# yum remove 'libibcm*'
# yum remove 'libibmad*'
# yum remove 'libibumad*'
# yum remove 'libibverbs*'
# yum remove 'libmlx4*'
# yum remove 'librdmacm*'
# yum remove 'libsdp*'
# yum remove 'mstflint*'
# yum remove 'ofed-docs*'
# yum remove 'ofed-scripts*'
# yum remove 'opensm*'
# yum remove 'perftest*'
# yum remove 'qperf*'
# yum remove 'sdpnetstat*'
# yum remove 'rdma*'
# yum remove 'rds-tools*'
```

4. Clean all `yum` cached files from all enabled repositories:

```
# yum clean all
```

5. Run one of the following commands, based on server type:

   - For a bare metal server, install the OFED packages for UEK R4 as follows:

     ```
     # yum install oracle-ofed-release
     ```

   - For a server that will function as a guest, install the OFED packages for UEK R4 as follows:

     ```
     # yum install oracle-ofed-release-guest
     ```

6. Enable the RDMA service by entering the following command:

```
# chkconfig rdma on
```

Each UEK release requires a different set of OFED packages. If you change the kernel on your system to a UEK release earlier than UEK R4, remove the existing UEK R4-based OFED packages before installing the correct packages for the new kernel by running the following command:

```
# yum remove --setopt=clean_requirements_on_remove=1 oracle-ofed-release
```

> **Caution:**
>
> Downgrading UEK versions is not advisable, except for testing purposes.

To update OFED packages that are already installed for UEK R4, run this command:

```
# yum update oracle-ofed-release
```

To update the OFED packages that are already installed on the guest, run this command on the guest:

```
# yum update oracle-ofed-release-guest
```