

**Oracle® Linux**

**Release Notes for Unbreakable Enterprise Kernel Release 4**

**ORACLE®**

E69348-08  
January 2019

---

## Oracle Legal Notices

Copyright © 2019, Oracle and/or its affiliates. All rights reserved.

This software and related documentation are provided under a license agreement containing restrictions on use and disclosure and are protected by intellectual property laws. Except as expressly permitted in your license agreement or allowed by law, you may not use, copy, reproduce, translate, broadcast, modify, license, transmit, distribute, exhibit, perform, publish, or display any part, in any form, or by any means. Reverse engineering, disassembly, or decompilation of this software, unless required by law for interoperability, is prohibited.

The information contained herein is subject to change without notice and is not warranted to be error-free. If you find any errors, please report them to us in writing.

If this is software or related documentation that is delivered to the U.S. Government or anyone licensing it on behalf of the U.S. Government, then the following notice is applicable:

U.S. GOVERNMENT END USERS: Oracle programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, delivered to U.S. Government end users are "commercial computer software" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, use, duplication, disclosure, modification, and adaptation of the programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, shall be subject to license terms and license restrictions applicable to the programs. No other rights are granted to the U.S. Government.

This software or hardware is developed for general use in a variety of information management applications. It is not developed or intended for use in any inherently dangerous applications, including applications that may create a risk of personal injury. If you use this software or hardware in dangerous applications, then you shall be responsible to take all appropriate fail-safe, backup, redundancy, and other measures to ensure its safe use. Oracle Corporation and its affiliates disclaim any liability for any damages caused by use of this software or hardware in dangerous applications.

Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

Intel and Intel Xeon are trademarks or registered trademarks of Intel Corporation. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. AMD, Opteron, the AMD logo, and the AMD Opteron logo are trademarks or registered trademarks of Advanced Micro Devices. UNIX is a registered trademark of The Open Group.

This software or hardware and documentation may provide access to or information about content, products, and services from third parties. Oracle Corporation and its affiliates are not responsible for and expressly disclaim all warranties of any kind with respect to third-party content, products, and services unless otherwise set forth in an applicable agreement between you and Oracle. Oracle Corporation and its affiliates will not be responsible for any loss, costs, or damages incurred due to your access to or use of third-party content, products, or services, except as set forth in an applicable agreement between you and Oracle.

### Abstract

This document contains information on the Unbreakable Enterprise Kernel Release 4. This document may be updated after it is released. To check for updates to this document, and to view other Oracle documentation, refer to the Documentation section on the Oracle Technology Network (OTN) Web site:

<https://www.oracle.com/technology/documentation/>

This document is intended for users and administrators of Oracle Linux. It describes potential issues and the corresponding workarounds you may encounter while using the Unbreakable Enterprise Kernel Release 4 with Oracle Linux 7. Oracle recommends that you read this document before installing or upgrading the Unbreakable Enterprise Kernel Release 4.

Document generated on: 2019-01-02 (revision: 232)

---

---

# Table of Contents

Preface .....	v
1 New Features and Changes .....	1
1.1 Notable Changes .....	1
1.1.1 Containers .....	1
1.1.2 Core Kernel Functionality .....	2
1.1.3 Cryptography .....	4
1.1.4 File Systems .....	4
1.1.5 Memory Management .....	7
1.1.6 Networking .....	7
1.1.7 NUMA .....	9
1.1.8 Real Time .....	10
1.1.9 Security .....	10
1.1.10 Storage .....	11
1.1.11 Virtualization .....	12
1.1.12 Zram .....	13
1.1.13 Zswap .....	14
1.2 Driver Updates .....	14
1.3 iSER and iSCSI LIO Support .....	16
1.4 LSI Logic / Symbios Logic MegaRAID SAS .....	16
1.5 New and Updated Packages .....	16
1.5.1 New and Updated Kernel, Driver, and Firmware Packages .....	16
1.5.2 New and Updated User Space Packages .....	17
1.6 Oracle Linux 7 Issue Fixes Using UEK R4 .....	18
1.6.1 Network Teaming .....	18
1.6.2 systemd Fails to Load the autofs4 and ipv6 Modules .....	18
1.7 Technology Preview .....	18
1.8 Compatibility .....	19
2 Known Issues .....	21
2.1 btrfs .....	21
2.2 DIF/DIX is not supported for ext file systems .....	22
2.3 Console Appears to Hang when Booting .....	22
2.4 DTrace .....	22
2.5 ip_local_port_range Parameter Under LXC .....	22
2.6 Overlayfs Error with Docker .....	22
3 Installation and Availability .....	25
3.1 Installation Overview .....	25
3.2 Subscribing to ULN Channels .....	25
3.3 Enabling Access to Oracle Linux Yum Server Channels .....	26
3.4 Upgrading Your System .....	28
3.5 Installing the Oracle-Supported OFED Packages .....	28



---

# Preface

The *Oracle Linux Unbreakable Enterprise Kernel Release Notes* provides a summary of the new features, changes, and known issues in the Unbreakable Enterprise Kernel Release 4.

## Audience

This document is written for system administrators who want to use the Unbreakable Enterprise Kernel with Oracle Linux. It is assumed that readers have a general understanding of the Linux operating system.

## Related Documents

The latest version of this document and other documentation for this product are available at:

<https://www.oracle.com/technetwork/server-storage/linux/documentation/index.html>.

## Conventions

The following text conventions are used in this document:

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



---

# Chapter 1 New Features and Changes

## Table of Contents

1.1 Notable Changes .....	1
1.1.1 Containers .....	1
1.1.2 Core Kernel Functionality .....	2
1.1.3 Cryptography .....	4
1.1.4 File Systems .....	4
1.1.5 Memory Management .....	7
1.1.6 Networking .....	7
1.1.7 NUMA .....	9
1.1.8 Real Time .....	10
1.1.9 Security .....	10
1.1.10 Storage .....	11
1.1.11 Virtualization .....	12
1.1.12 Zram .....	13
1.1.13 Zswap .....	14
1.2 Driver Updates .....	14
1.3 iSER and iSCSI LIO Support .....	16
1.4 LSI Logic / Symbios Logic MegaRAID SAS .....	16
1.5 New and Updated Packages .....	16
1.5.1 New and Updated Kernel, Driver, and Firmware Packages .....	16
1.5.2 New and Updated User Space Packages .....	17
1.6 Oracle Linux 7 Issue Fixes Using UEK R4 .....	18
1.6.1 Network Teaming .....	18
1.6.2 systemd Fails to Load the autofs4 and ipv6 Modules .....	18
1.7 Technology Preview .....	18
1.8 Compatibility .....	19

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 and later and Oracle Linux 7 Update 1 and later on the x86-64 architecture. It is based on the mainline Linux kernel version 4.1.12. This release also updates drivers and includes bug and security fixes.

Oracle actively monitors upstream checkins and applies critical bug and security fixes to UEK R4.

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.

## 1.1 Notable Changes

The following sections describe the major new features of Unbreakable Enterprise Kernel Release 4 (UEK R4) relative to UEK R3.

### 1.1.1 Containers

The following notable features of containers are implemented in UEK R4:

- Local device cgroup changes are now propagated down the device cgroup hierarchy.

For more information, see [git commit bd2953ebbb533aeda9b86c82a53d5197a9a38f1b](https://github.com/oracle/uek4/commit/bd2953ebbb533aeda9b86c82a53d5197a9a38f1b).

- The `__DEVEL__sane_behavior` option has been introduced for mounting cgroup controllers.  
For more information, see [git commit 873fe09ea5df6ccf6bb34811d8c9992aacb67598](https://git.kernel.org/?q=commit:873fe09ea5df6ccf6bb34811d8c9992aacb67598).
- `memory.numa_stat` now includes hierarchical statistics for child memory cgroups (`memcgs`) in addition to the parent `memcg`.  
For more information, see [git commit 071aee138410210e3764f3ae8d37ef46dc6d3b42](https://git.kernel.org/?q=commit:071aee138410210e3764f3ae8d37ef46dc6d3b42).
- An optional unified control group hierarchy has been introduced.  
For more information, see <http://lwn.net/Articles/601840/>.
- Hierarchy restrictions for `swappiness` and `oom_control` have been removed from `memcgs`.  
For more information, see [git commit 3dae7fec5e884a4e72e5416db0894de66f586201](https://git.kernel.org/?q=commit:3dae7fec5e884a4e72e5416db0894de66f586201).

## 1.1.2 Core Kernel Functionality

The following notable core kernel features are implemented in UEK R4:

- The performance of SPECjbb is improved for a system with more than 10 CPUs by removing contention for the global `epmutex` lock, which is used in `EPOLL_CTL_ADD` and `EPOLL_CTL_DEL` operations. For example, in a typical 16-socket run the performance increases from 35k jOPS to 125k jOPS. Benchmarks also exhibit good scaling from 10 sockets to over 40 sockets.
- The `sysctl_numa_balancing_settle_count` parameter used by the NUMA scheduler has been removed.
- The following tracepoints are now provided to monitor NUMA scheduler activity:

<code>trace_sched_move_numa</code>	Triggered when a task is moved to a node.
<code>trace_sched_stick_numa</code>	Triggered when a NUMA migration fails.
<code>trace_sched_swap_numa</code>	Triggered when a task is swapped for another task.

- The new `SCHED_STACK_END_CHECK` kernel debugging option can be used to check for a stack overrun on calls to `schedule()` on a NUMA system. If the stack end location is overwritten, the system panics as the content of the corrupted region cannot be trusted.
- Sysbench performance has been improved by preventing spurious active NUMA migration.
- CPU clock frequency scaling for performance management. The possible governor settings as displayed by `/sys/devices/system/cpu/cpu*/cpufreq/scaling_governor` are:

<code>ondemand</code>	Sets the CPU clock frequency between the minimum and maximum possible frequencies, according to the current demand usage. The following <code>sysfs</code> parameters are adjustable:	
	<code>ignore_nice_load</code>	Whether processes with a <code>nice</code> value count (0) or do not count (1) toward CPU usage. The default value is 0.
	<code>powersave_bias</code>	How much to reduce the target CPU frequency by as a fraction of



	1000. A value of 0 disables this feature.
<code>sampling_down_factor</code>	A multiplier that the kernel applies to <code>sampling_rate</code> when the CPU is running at its maximum clock frequency. The default value is 1.
<code>sampling_rate_min</code>	Minimum sampling rate.
<code>sampling_rate</code>	Interval in microseconds between assessments of whether the kernel needs to change the clock frequency.
<code>up_threshold</code>	Threshold of average CPU usage as a percentage for the kernel to increase the clock frequency.

`ondemand` is the default governor setting if `tuned` is not configured.

This setting is equivalent to `powersave` for more recent microarchitecture CPUs (for example, Haswell, Broadwell, and later) with which the `pstate` power scaling driver can interact. For older design architecture CPUs (for example, Ivy Bridge, Sandy Bridge, and earlier), `ondemand` is equivalent to `performance` as the cores must be kept in a higher power state to minimize CPU latency.

`performance`

Sets the CPU clock frequency to the maximum possible frequency.



#### Note

`performance` is the default governor setting for the `tuned throughput-performance` profile.

The `performance` profile is appropriate for some real-time applications but it might not be appropriate for all workloads. Running a CPU at maximum frequency can prevent turbo mode from being enabled because doing so would exceed the thermal envelope.

`powersave`

Sets the CPU clock frequency to the minimum possible frequency.

`userspace`

Permits a user-space program running as an effective `root` user to control the CPU clock frequency by creating and using a file named `scaling_setspeed` in the CPU-device directory under `sysfs`.

Oracle recommends that you use `tuned-adm` to select a `tuned` performance profile for your system that is based on its hardware and software configuration, for example:

- If your system has Xeon processors or multiple disks, choose a profile such as `latency-performance` for a cloud server, `throughput-performance` for a database server, or `virtual-host` for a virtual host server.

**Note**

These profiles set the CPU governor setting to `performance`, which might not be appropriate for all workloads.

- For a virtual machine guest, choose the `virtual-guest` profile.
- For a laptop, choose a suitable laptop profile such as `laptop-ac-powersave` or `laptop-battery-powersave`.
- For a desktop machine, choose either the `desktop` or `balanced` profile.

You can use the `tuned-adm list` command to display the available profiles.

If `tuned` is not configured, the default CPU governor setting is `ondemand`, which can cause some bursty, CPU-intensive workloads to run more slowly because of demand hysteresis.

If necessary, you can create your own performance profiles based on the profiles that are provided in the `/etc/tune-profiles` directory hierarchy.

When comparing system performance under different profiles, use benchmarks that simulate your server's typical workload.

For more information, see the `tuned(8)` and `tuned-adm(1)` manual pages, which are available in the `tuned` package.

### 1.1.3 Cryptography

The following notable cryptographic features are implemented in UEK R4:

- Accelerated CRC T10 DIF computation with the `PCLMULQDQ` instruction.
- LZ4 Cryptographic API.
- Support for `sha256_ssse3`, `SHA-224`, `sha512_ssse3`, and `SHA-384`.
- Support for the AMD cryptographic coprocessor, which can be used to accelerate or offload AES, SHA, and other encryption operations.

### 1.1.4 File Systems

The following sections detail the most notable features that have been implemented for file systems in UEK R4:

#### 1.1.4.1 Btrfs

The following btrfs features have been implemented:

- RAID 5 and RAID 6 support.

**Note**

The RAID 5 and RAID 6 features are experimental and potentially not crash-safe.

- Skinny extents introduce a new key type for metadata extent references that uses disk space more efficiently and reduces the space required per extent reference for each tree block from 51 bytes to 33

bytes. As a result, the size of the extent tree decreases by between 30% and 35%, metadata operations are faster as there are fewer copy-on-write operations, and more of the extent tree can be stored in memory. This feature is not supported by previous kernels so it is not enabled by default with UEK R4. To enable the feature, specify the `-x` option to `mkfs.btrfs` when creating a file system or to `btrfsctl` for an existing file system.

- Snapshot-aware defragmentation does not break the sharing of data by snapshots.
- Offline data deduplication removes copies of repeated data in the file system. This feature is useful for VMs, which often contain similar copies of an operating systems or for backups, where the reduction in storage requirements can be substantial. You can perform deduplication by running a command such as `dedup` or `duperemove` when the file system is mounted and running. Online deduplication, which occurs automatically and transparently as processes write data, is not yet supported.
- New `BTRFS_IOC_SET_LABEL` and `BTRFS_IOC_SET_LABEL ioctl` commands support reading and writing the file system label.
- The `commit` mount option specifies the interval for periodic synchronization to permanent storage, which is 30 seconds by default. Setting too high a value increases the risk of data loss or corruption if the system crashes. A warning is displayed if you set the value greater than 300 seconds.
- Remount now supports the `acl`, `barrier`, `datasum`, `datacow`, `noautodefrag`, `nodiscard`, `noenospc_debug`, `noflushoncommit`, and `treelog` options.
- The extent tree is now more space efficient, which reduces the size of the metadata that it requires.

For more information, see <https://btrfs.wiki.kernel.org/index.php/Changelog>.

#### 1.1.4.2 efivarfs

The Unified Extensible Firmware Interface (UEFI) variable file system (*efivarfs*) is enabled on systems that support UEFI. For Oracle Linux 7, `systemd` automatically mounts *efivarfs*. For Oracle Linux 6, *efivarfs* is not mounted by default. If required, you can mount *efivarfs*, for example:

```
# mount -t efivarfs efivarfs /sys/firmware/efi/efivars
```

#### 1.1.4.3 ext4

The following ext4 features have been implemented:

- Metadata checksumming can be enabled by specifying the `metadata_csum` option when making a file system.
- 64-bit file system support, which allows you to format a file system that is larger than 16 TB, can be enabled by specifying the `64bit` option when making a file system.
- Improved synchronization speed for database workloads.
- Improved write-back performance if delayed allocation is disabled using the `nodealloc` mount option or if ext2 or ext3 compatibility mode is used.
- Improved extent-tree memory caching.
- Improved stabilization of hole punching using `fallocate()`.
- Improved data and hole seeking using `lseek()`.

The following features are considered experimental and are not supported:

- Big allocation (`bigalloc`), which does not currently work with `fallocate()`.
- Inline data, which stores the data for small files in the available space between on-disk inode data structures.
- File-system image creation from a directory using `mke2fs`.
- Specifying an external journal by using the `pathname` mount option.

#### 1.1.4.4 FUSE

The following FUSE features have been implemented:

- Asynchronous I/O support.
- Optimized short direct reads.
- Writepages callback improves memory-mapped writeout by `mmap`.

Cached writeback support is not currently supported by the user-space applications that are provided with Oracle Linux 6 and Oracle Linux 7.

#### 1.1.4.5 NFS

The following NFS features have been implemented:

- Client support for NFSv4.2.

For more information, see <http://tools.ietf.org/html/draft-ietf-nfsv4-minorversion2-20>.

- SELinux Labeled NFS allows many different labels to be used on an NFS share, which is useful for securing virtualization image files and home directories.

#### 1.1.4.6 Overlayfs

The overlayfs file system is an implementation of a union file system that makes several file systems appear as a single file system when mounted. An overlayfs file system consists of a *lower* file system and an *upper* file system which share a single file system namespace. After a file is opened in an overlayfs file system, all operations go directly to the underlying lower or upper file systems, which simplifies the implementation and allows native performance compared to other union file system implementations. A typical use case is to use a read-only OS image as the lower file system and a writeable RAM-backed file system as the upper file system. Modified data is written to the upper file system only and not to the OS image.

Both the upper and lower file systems can be directory trees within the same file system and neither needs to be the root of a file system. The lower file system can be any supported file system, including an overlayfs file system, and does not need to be writable. If the upper file system is writable, as is usually the case, it must support the creation of `trusted.*` extended attributes and it must provide valid `d_type` file type in the `direct` structure returned by `readdir()`. For example, an NFS file system cannot be used for the upper file system.

The overlayfs file system is not available with UEK R3.

For more information, see <https://www.kernel.org/doc/Documentation/filesystems/overlayfs.txt>.

### 1.1.4.7 XFS

The following XFS features have been implemented:

- The new directory entry file type improves the performance of directory recursion by not having to access the inode data from disk.
- Namespace support.
- Defragmentation support for the new CRC file system format.
- The XFS v5 disk format provides metadata CRC, object back references, better crash recovery, and improved `xfs_repair` performance. The metadata CRC feature is experimental and not currently supported.

## 1.1.5 Memory Management

The following notable memory management features are implemented in UEK R4:

- The `MAP_HUGETLB` flag has been implemented in `mmap` to support huge-page memory mapping with `hugetlbfs`.
- Problems have been addressed with `kswapd` and page reclaim behavior during large copy operations or when memory was low.
- Improved page table access scalability in threaded huge-page workloads by reducing lock contention in the page table.

For more information, see <http://lwn.net/Articles/568076/>.

- Improve page-fault scalability in `hugetlb` by handling concurrent page faults. Previously, the kernel could only handle a single `hugetlb` page fault at a time. Typically, the startup time for a 10-gigabyte Oracle database, which generates approximately 5000 page table faults, decreases to 25.7 seconds from 37.5 seconds. Larger workloads should experience even greater improvements in start-up times.
- Support gigantic page allocation in `hugetlb` at runtime in addition to the existing boot-time allocation.
- The unqueued slab allocator (SLUB) is now the default memory allocator for kernel objects. SLUB eliminates the fragmentation that is caused by memory allocation and deallocation by reusing memory that was previously allocated to a data object of the same type.

## 1.1.6 Networking

The following notable networking features are implemented in UEK R4:

- The following VXLAN features have been implemented:
  - Layer 2 redirection with layer 3 switching.
  - Setting destination to a unicast address.
  - UDP tunnel segmentation.
  - IPv6 support.
  - Transmit-side VLAN offload for VXLAN devices.
- Link configuration for transmitting UDPv4 checksums, and transmitting and receiving UDPv6 checksums.

- Switch the network namespace when a packet is encapsulated or unencapsulated.
- Per-socket network polling is supported with the `bnx2x`, `ixgbe`, and `mlx4` network card drivers, which reduces the latency inherent in the NAPI periodic polling method.

For more information, see <http://lwn.net/Articles/551284/> and [2012-lpc-Low-Latency-Sockets-slides-brandenburg.pdf](#).

- The new PIE (Proportional Integral controller Enhanced) network packet scheduler controls the average queuing latency to overcome buffer bloat, ensure low latency and achieve high link utilization under various congestion scenarios with very small overhead.

For more information, see <https://tools.ietf.org/html/draft-pan-tsvwg-pie-00>.

- Support for configuring the SR-IOV virtual function (VF) minimum and maximum transmission rates by using the `ip` command.

For more information, see [git commit ed616689a3d95eb6c9bdbb1ef74b0f50cbdf276a](#).

- Support for SR-IOV VF link state control by using the `ip` command. Previously, VF links were always on, regardless of the physical link status, which allows VMs on the same virtual Ethernet bridge to communicate even if the physical function (PF) link state is down. However, if the VFs were bonded in active/standby mode, this configuration prevented failover when the physical link used by a VF went down. You can now use the `ip link set` command to configure the behavior of a VF link:

```
# ip link set device vf number state { auto | enable | disable }
```

The possible settings are:

<code>auto</code>	The VF link state is determined by the PF link state. This setting is suitable for VFs that are bonded in active/standby mode.
<code>disable</code>	The VF link state is permanently down.
<code>enable</code>	The VF link state is permanently up. This is the default setting.

- The following Open vSwitch (OvS) features have been implemented:

- Generic routing encapsulation (GRE) tunnels.
- User-space tunneling interface.
- Stream Control Transmission Protocol (SCTP) support.
- VXLAN tunneling support.
- Wild-carded flow implementation.
- TCP bitwise flag matching.

For more information, see [git commit 5eb26b156e29eadcc21f73fb5d14497f0db24b86](#)

- Allow user space to announce ability to accept unaligned Netlink messages.
- Enable memory-mapped Netlink I/O.
- Enable tunnel generic segmentation offloading (GSO) for Open vSwitch bridge devices so that Open vSwitch can take advantage of hardware offloading to the underlying devices.

- Add `recirc` and `hash` action to support distributing packets between the ports of bond devices.
- Add support for generic network virtualization encapsulation (Geneve) tunneling.
- The `nftables` framework provides packet filtering and packet classification features as a replacement for the `arptables`, `ebtables`, `iptables`, and `ip6tables` frameworks. For example, see <https://lwn.net/Articles/564095/>.

The following `nftables` features have been implemented:

- Replaced `iptables`, while providing backwards compatibility.
- IPv4 and IPv6 masquerading
- Pre-routing and post-routing filtering.
- Extended `NFT_MSG_DELTABLE` call to support flushing the rule set.
- Add filter support for skipping accounting objects.
- Add support for exporting the rule-set generation ID.
- Add CPU attribute support for matching packets against CPU number.
- Add support for matching packet types for the `inet`, `ip`, and `ip6` table families based on link-layer information. For loopback traffic, the packet type is deduced from the network layer header.
- Add support for matching the device group of a packet's incoming or outgoing interface.
- TCP Fast Open optimization is enabled by default in UEK R4 for applications that take advantage of this feature.
- Generic network virtualization encapsulation (Geneve) provides a tunneling framework for establishing layer 2 networks over layer 3 networks.

For more information, see <http://tools.ietf.org/html/draft-gross-geneve-01> and <http://blogs.vmware.com/cto/geneve-vxlan-network-virtualization-encapsulations/>.

- Transmission queue batching defers flushing transmission socket buffers to the network driver to reduce the overall cost of processing the transmission queue and can result in a higher effective packet transmission rate. The `i40e`, `igb`, `ixgbe`, `mlx4`, and `virtio_net` drivers support this feature.

For more information, see <http://lwn.net/Articles/615238/> and <http://netoptimizer.blogspot.com/2014/10/unlocked-10gbps-tx-wirespeed-smallest.html>.

## 1.1.7 NUMA

Many modern multiprocessors have non-uniform memory access (NUMA) memory designs, where the performance of a process can depend on whether the memory range being accessed is attached to the local CPU or to another CPU. As performance is different depending on memory locality, the operating system should ideally schedule a process to run on the CPU whose memory controller is connected to the memory to be accessed.

The following notable NUMA features are implemented in UEK R4:

- Support NUMA affinity for unbound workqueues.

- A new NUMA subsystem provides improved performance for NUMA systems. New NUMA policies attempt to place a process near its memory, can share pages between processes and handle transparent huge pages. (3.8, 3.13)

The following `sysctl` parameters allow you to enable, disable and tune NUMA scheduling:

`numa_balancing_scan_delay` Scan delay in milliseconds used for starting a task when it initially forks.

`numa_balancing_scan_period_maximum` Maximum delay in milliseconds between scanning for tasks.

`numa_balancing_scan_period_minimum` Minimum delay in milliseconds between scanning for tasks.

`numa_balancing_scan_period_reset` Resets the scan delay period.

`numa_balancing_scan_size_mb` Amount of pages in megabytes scanned per scan.

For more information, see <http://lwn.net/Articles/568870/>.

- Add the `numa_balancing` `sysctl` parameter to enable or disable automatic NUMA memory balancing.
- Improved algorithm for NUMA migrations that maximizes the performance of workloads that do not fit on one NUMA node.
- Memory zones are allocated by the page allocator in node order on 64-bit NUMA systems by default.

## 1.1.8 Real Time

The following notable real-time features are implemented in UEK R4:

- Dynamic ticks and full CPU time accounting infrastructure.
- Timerless multitasking support allows the system to run processes without needing to fire up the timer interrupt that is traditionally used to implement multitasking. (3.10, 3.12)

For more information, see <http://lwn.net/Articles/549580/> and <http://lwn.net/Articles/558284/>.

- Deadline scheduling provides `deadline`, `period`, and `runtime` parameters for scheduling processes in the `SCHED_DEADLINE` scheduling class. These process are guaranteed to receive `runtime` microseconds of execution time every `period` microseconds and these `runtime` microseconds are available within `deadline` microseconds from the beginning of the period. The task scheduler runs the process with the lowest `deadline` value.

For more information, see [git documentation 712e5e34aef449ab680b35c0d9016f59b0a4494c](http://git.kernel.org/cgit/linux/kernel/git/rt/linux-rt-wg.git) and <http://lwn.net/Articles/575497/>.

## 1.1.9 Security

The following notable security features are implemented in UEK R4:

- The physical and virtual address at which the kernel image is decompressed is randomized to deter exploit attempts that rely on knowing the location of the kernel internals.
- The Kexec feature, which allows faster rebooting or automatically booting a new kernel after a crash, now incorporates support for allowing only signed Kexec kernels for use with UEFI secure booting.
- The `kexec_load_disabled` `sysctl` parameter can be used to disable Kexec, which allows a system to be better protected against privilege escalation.



- An `exe` field has been added to the auditing log to record the pathname of executables that produce core dumps.
- An `audit_backlog_wait_time` configuration option has been added to the auditing subsystem so that if `auditd` cannot keep up or is blocked, callers are not blocked.
- If the value of the `audit_backlog_limit` parameter is set to zero, the length of the backlog queue is limited only by the amount of system memory.
- By default, errors on `AUDIT_NEVER` rules are now logged.
- The auditing subsystem now logs task information when the state of a feature is changed.
- A netlink multicast socket has been added to read-only user-space clients such as `systemd` to allow read-only access to the audit logs.
- Secure generation of random numbers with the `getrandom` system call. Linux systems traditionally obtained their random numbers from `/dev/[u]random`. This interface is vulnerable to file descriptor exhaustion attacks, where the attacker consumes all available file descriptors, and is also inconvenient for use in containers. The `getrandom` system call, which analogous to the `getentropy` call in OpenBSD overcomes these problems.
- SELinux now reports permissive mode in `avc: denied` messages.

## 1.1.10 Storage

The following notable storage features are implemented in UEK R4:

- The device mapper `dm-cache` target allows you to use a fast device such as an SSD as a cache for a slower device such as a rotating disk. You can use various policy plugins to change the selection algorithms for performing actions such as promoting, demoting, cleaning blocks. `dm-cache` supports both writeback and write-through modes. This feature is still flagged as experimental and might not be suitable for production systems.

Updates to `dm-cache` added support for a passthrough mode when the cache contents might not be consistent with the underlying device, cache block invalidation, and cache shrinking.

For more information, see <https://www.kernel.org/doc/Documentation/device-mapper/cache.txt>.

- Bcache is a block layer cache that allows you to use SSDs to cache slower block devices. Bcache can perform both writeback and write-through caching, has no file-system dependencies, is simple to use, and works well on any setup without requiring any configuration.

For more information, see <https://www.kernel.org/doc/Documentation/bcache.txt>, <http://bcache.evilpiepirate.org/>, and <http://lwn.net/Articles/497024/>.

- The new, scalable multiqueue block layer subsystem (`blk-mq`) for supporting high performance SSD storage implements per-CPU submission queues for receiving I/O requests, which are directed to hardware submission queues. The separate per-CPU submission and hardware submission queues balances the I/O workload across multiple CPU cores and reduces latency. The design supports the interface and features of the traditional block layer, but it is also capable of supporting many millions of I/O operations per second by taking advantage of the capabilities of NVM-Express or high-end PCI-E devices and multicore CPUs.

For more information, see <http://lwn.net/Articles/552904/>.

- The device mapper `dm-era` target behaves similarly to the `linear` target with the addition of tracking any blocks that were written within an `era`, which is a time period that you can define. Typical use cases

are tracking the changed blocks in backup software and restoring cache coherency after rolling back a snapshot by partially invalidating the cache contents.

For more information, see <https://www.kernel.org/doc/Documentation/device-mapper/era.txt>.

### 1.1.10.1 OFED Support

The OpenFabrics Enterprise Distribution (OFED) 2.0 stack is integrated with UEK R4, and supports all Oracle branded InfiniBand (IB) hardware, on systems with an x86-64 architecture. This includes:

- Sun InfiniBand Dual Port 4x QDR Host Channel Adapters M2
- Oracle Dual Port QDR Infiniband Adapter M3
- Oracle Dual Port QDR InfiniBand Adapter M4
- Oracle Dual Port EDR InfiniBand Adapter

OFED 2.0 supports the following protocols with UEK R4:

- iSCSI Extensions for remote direct memory access (iSER) provide access to iSCSI storage devices
- Reliable Datagram Sockets (RDS) is a high-performance, low-latency, reliable connectionless protocol for datagram delivery
- Sockets Direct Protocol (SDP) supports stream sockets for RDMA network fabrics
- Ethernet over InfiniBand (EoIB)
- Internet Protocol over InfiniBand (IPoIB)



#### Note

Ethernet tunneling over IPoIB (eIPoIB) is not supported with UEK R4.

OFED 2.0 supports the following RDS features with UEK R4:

- Async Send (AS)
- Quality of Service (QoS)
- Active Bonding (AB)
- Netfilter (NF)
- Shared Request Queue (SRQ)



#### Note

Automatic Path Migration (APM) is not supported with UEK R4.

Support for IB, OFED, and RDS is integrated into the kernel. The OFED user-space RPMs continue to be provided, but the `kernel-ib` and `ofa-kernel` RPMs are not required.

### 1.1.11 Virtualization

The following notable virtualization features are implemented in UEK R4:

- Hyper-V support for `netpoll` allows a network console to be used to debug kernel issues.

- The following Xen features have been implemented:
  - `xen-netback` support for changing the MAC address of an interface
  - ACPI support for CPU and memory hotplug, including a new memory hotplug driver.
  - `xen-netback` support for gathering zerocopy statistics and TX grant mapping.
  - Support for MSI message groups in Dom0.
  - Substantially improved performance of Xen virtual network interfaces by implementing multiple queue support between `xen-netback` and `xen-netfront`.
  - EFI support in Dom0.
  - Xen PVSCSI backend and frontend driver support for high performance passthrough of SCSI devices or LUNs from Dom0 to a Xen PV or HVM guest.
  - Remapping of existing MFNs that were replaced by the identity map to prevent non-contiguous pages occurring in Dom0.
  - Improved PV ticket locks provide more efficient locking of guests for workloads that rely on this mechanism. If a spin lock is not available for more than a brief period, the lock code stops spinning and calls the hypervisor to wait until the lock becomes available again.
  - NUMA topology and I/O exposure to guests.
  - PVH guests now support Paravirtualized Hardware extensions (v3).

## 1.1.12 Zram

Zram compresses everything written to specified block devices in RAM, and is used typically for swap devices to improve the responsiveness of systems that have a limited amount of memory. The following example illustrates how to create and enable a zram swap device:

```
# mkswap /dev/zram0
# swapon /dev/zram0
```

The next example illustrates how to create a file system on a zram device and then mount this file system:

```
# mkfs.ext4 /dev/zram1
# mount /dev/zram1 /tmp
```

The following notable zram features are implemented in UEK R4:

- Zram has been moved out of staging to `drivers/block/zram`.
- Support for LZ4 compression in addition to LZO.
- Performance improvements to concurrent compression of multiple compression streams.
- Support for switching the compression algorithm in `/sys/block/zramN/comp_algorithm`.
- Support for limiting the maximum amount of useable memory for a zram device in `/sys/block/zramN/mem_limit`. You can use memory unit suffixes when setting a value, for example:

```
# echo 1G > /sys/block/zram0/mem_limit
```

To disable the limit, set the value to 0.

- Support for displaying the maximum memory that a zram device has consumed in `/sys/block/zramN/mem_used_max`. Writing 0 to this file resets the counter.

### 1.1.13 Zswap

Zswap is a lightweight, write-behind compressed caching mechanism for swap pages that attempts to compress a page being swapped out to RAM. A successful compression defers and, in many cases, prevents writeback to the swap device, reducing I/O and increasing the performance of a system that is swapping.

For more information, see <http://lwn.net/Articles/537422/>.

## 1.2 Driver Updates

The Unbreakable Enterprise Kernel Release 4 supports a large number of hardware and devices. In close cooperation with hardware and storage vendors, Oracle has updated several device drivers from the versions in mainline Linux 4.1.12.

The following table details commonly used drivers that have been updated since UEK R3 quarterly update 7:

Vendor	Driver	Version	Description
Avago	<a href="#">megaraid_sas</a>	06.807.10.00-rc1	Avago MegaRAID SAS driver
Avago	<a href="#">mpt2sas</a>	20.100.00.00	LSI MPT Fusion SAS 2.0 device driver
Avago	<a href="#">mpt3sas</a>	09.100.00.00	LSI MPT Fusion SAS 3.0 device driver
Chelsio	<a href="#">csiostor</a>	1.0.0	Chelsio FCoE driver (Oracle Linux 7 only)
Chelsio	<a href="#">cxgb4</a>	2.0.0-ko	Chelsio T4 network driver
Chelsio	<a href="#">cxgb4i</a>	0.9.4	Chelsio T4 iSCSI driver
Chelsio	<a href="#">cxgb4vf</a>	2.0.0-ko	Chelsio T4/T5 Virtual Function (VF) network driver
Cisco	<a href="#">enic</a>	2.3.0.12	Cisco VIC Ethernet NIC driver
Cisco	<a href="#">fnic</a>	1.6.0.22	Cisco FCoE HBA driver
Cisco	<a href="#">usnic_verbs</a>	1.0.3	Cisco VIC (usNIC) Verbs driver
Emulex	<a href="#">be2iscsi</a>	10.6.0.1	Emulex OneConnect Open-iSCSI driver
Emulex	<a href="#">be2net</a>	10.6.0.4	Emulex OneConnect 10Gbps NIC driver
Emulex	<a href="#">lpfc</a>	11.0.0.3	Emulex LightPulse Fibre Channel SCSI driver (includes Lancer G6 and multi-queue support)
Emulex	<a href="#">ocrdma</a>	10.6.0.0	Emulex OneConnect RoCE driver
HP	<a href="#">cciss</a>	3.6.26	Driver for HP Smart Array controllers (Oracle Linux 6 only)
Intel	<a href="#">e1000e</a>	3.2.6-k	Intel PRO/1000 network driver
Intel	<a href="#">i40e</a>	1.3.21-k	Driver for Intel Ethernet controller XL710 family
Intel	<a href="#">i40evf</a>	1.3.13	Intel XL710 X710 Virtual Function network driver
Intel	<a href="#">igb</a>	5.3.0-k	Intel Gigabit Ethernet network driver

## Driver Updates

Vendor	Driver	Version	Description
Intel	<a href="#">igbvf</a>	2.0.2-k	Intel Gigabit Virtual Function driver
Intel	<a href="#">iw_nes</a>	1.5.0.1	NetEffect RNIC low-level iWARP driver
Intel	<a href="#">ixgb</a>	1.0.135-k2-NAPI	Intel PRO/10GbE network driver
Intel	<a href="#">ixgbe</a>	4.2.1-k	Intel 10 Gigabit PCI Express network driver
Intel	<a href="#">ixgbevf</a>	2.12.1-k	Intel 82599 Virtual Function driver
Intel	<a href="#">nvme</a>	1.0	Non-volatile memory (Fultondale) driver
Mellanox	<a href="#">mlx4_core</a>	2.2-1	Mellanox ConnectX HCA low-level driver
Mellanox	<a href="#">mlx4_en</a>	2.2-1	Mellanox ConnectX HCA Ethernet driver
Mellanox	<a href="#">mlx4_ib</a>	2.2-1	Mellanox ConnectX HCA InfiniBand driver
Mellanox	<a href="#">mlx5_ib</a>	2.2-1	Mellanox Connect-IB HCA IB driver
Oracle	<a href="#">hxge</a>	1.3.6	10 Gigabit Network Driver
Oracle	<a href="#">sxge</a>	0.09302015	SXGE Ethernet driver
Oracle	<a href="#">sxgevf</a>	0.09302015	SXGEVF Ethernet driver
Oracle	<a href="#">xscore</a>	6.0.r8008	Core services driver
Oracle	<a href="#">xsvhba</a>	6.0.r8008	VHBA Driver
Oracle	<a href="#">xsvnic</a>	6.0.r8008	XSVNIC network driver
Oracle	<a href="#">xve</a>	6.0.r8008	Virtual Ethernet driver
QLogic	<a href="#">bfa</a>	3.2.23.0	Brocade Fibre Channel HBA driver fcpim (informational only, no significant update)
QLogic	<a href="#">bna</a>	3.2.25.1	QLogic BR-series 10G PCIe Ethernet driver (informational only, no significant update)
QLogic	<a href="#">bnx2</a>	2.2.6	QLogic BCM57xx driver
QLogic	<a href="#">bnx2fc</a>	2.9.6	QLogic BCM57712 FCoE driver (informational only, no significant update)
QLogic	<a href="#">bnx2i</a>	2.7.10.1	QLogic BCM57xx/57xxx iSCSI driver
QLogic	<a href="#">bnx2x</a>	1.712.30-0	QLogic 10G/20G Ethernet driver
QLogic	<a href="#">cnic</a>	2.5.22	QLogic CNIC driver
QLogic	<a href="#">ib_qib</a>	1.11	QLogic IB driver
QLogic	<a href="#">netxen_nic</a>	4.0.82	QLogic/NetXen (1/10) GbE Intelligent Ethernet driver (informational only, no significant update)
QLogic	<a href="#">qla2xxx</a>	8.07.00.26.39.0-k	QLogic Fibre Channel HBA driver (informational only, no significant update)
QLogic	<a href="#">qla4xxx</a>	5.04.00-k6	QLogic iSCSI HBA driver
QLogic	<a href="#">qlcnic</a>	5.3.63	QLogic 1/10 GbE Converged/Intelligent Ethernet driver
QLogic	<a href="#">tg3</a>	3.137	Broadcom Tigon3 Ethernet driver
	<a href="#">ib_ipoib</a>	1.0.0	IP-over-InfiniBand net driver

Vendor	Driver	Version	Description
	<code>ib_iser</code>	1.6	iSER (iSCSI Extensions for RDMA) Datamover
	<code>ib_isert</code>	1.0	iSER-Target for mainline target infrastructure
	<code>xen-netback</code>	Upstream patches	Xen virtual network device backend
	<code>xen-netfront</code>	Upstream patches	Xen virtual network device frontend

## 1.3 iSER and iSCSI LIO Support

UEK R4 supports the iSER and iSCSI LIO targets and initiators that are required by the following storage solutions components:

- Emulex Skyhawk OCe14XXX
- FlashGrid

## 1.4 LSI Logic / Symbios Logic MegaRAID SAS

UEK R4 does not support the LSI Logic / Symbios Logic MegaRAID SAS 1078 controller.

## 1.5 New and Updated Packages

To support the newly added functionality that the Unbreakable Enterprise Kernel Release 4 provides, the following sections list kernel and user space binary packages have been added or updated from the ones included in the base distribution. For more information about the ULN channels and Oracle Linux yum server repositories in which these packages are available, see [Chapter 3, \*Installation and Availability\*](#).

### 1.5.1 New and Updated Kernel, Driver, and Firmware Packages

- `crash` (Oracle Linux 6 only)
- `crash-devel` (Oracle Linux 6 only)
- `dtrace-modules`
- `dtrace-modules-provider-headers`
- `dtrace-modules-shared-headers`
- `kernel-uek`
- `kernel-uek-debug`
- `kernel-uek-debug-devel`
- `kernel-uek-devel`
- `kernel-uek-doc`
- `kernel-uek-firmware`
- `libdtrace-ctf`

- `libdtrace-ctf-devel`
- `linux-firmware` (Oracle Linux 6 only)

## 1.5.2 New and Updated User Space Packages

Unless specified otherwise, the following packages are available only in ULN channels.

- `acpid` (Oracle Linux 6 only)
- `btrfs-progs` (Available from Oracle Linux Yum Server and ULN)
- `btrfs-progs-devel` (Available from Oracle Linux yum server and ULN)
- `dtrace-utils`
- `dtrace-utils-devel`
- `ibutils`
- `infiniband-diags`
- `libibcm`
- `libibmad`
- `libibmad-devel`
- `libibumad`
- `libibumad-devel`
- `libibverbs`
- `libibverbs-devel`
- `libibverbs-utils`
- `libmlx4`
- `librdmacm`
- `librdmacm-devel`
- `librdmacm-utils`
- `libsdp`
- `mstflint`
- `ofed-docs`
- `opensm-libs`
- `oracle-ofed-release`
- `perftest`
- `qperf`

- [rdma](#)
- [rds-tools](#)

## 1.6 Oracle Linux 7 Issue Fixes Using UEK R4

The following issues are fixed in Oracle Linux 7 by booting with UEK R4:

### 1.6.1 Network Teaming

Teaming is supported with UEK R4. Teaming also works with UEK R3 Quarterly Update 7 or later. (Bug ID 19151770)

### 1.6.2 systemd Fails to Load the autofs4 and ipv6 Modules

When the system is booted with UEK R3, `systemd` fails to load the `autofs4` and `ipv6` modules and errors such as the following are logged:

```
systemd[1]: Failed to insert module 'autofs4'  
systemd[1]: Failed to insert module 'ipv6'
```

The system can load these modules if booted with UEK R4. (Bug ID 18470449)

## 1.7 Technology Preview

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

- **Ceph File System and Object Gateway Federation**

Ceph presents a uniform view of object and block storage from a cluster of multiple physical and logical commodity-hardware storage devices. Ceph can provide fault tolerance and enhance I/O performance by replicating and striping data across the storage devices in a Storage Cluster. Ceph's monitoring and self-repair features minimize administration overhead. You can configure a Storage Cluster on non-identical hardware from different manufacturers.

The Ceph File System (CephFS) and Object Gateway Federation features of Ceph are in technology preview.

- **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.



- **NFS over RDMA interoperation with ZFS and Oracle Solaris**

NFS over RDMA does not yet fully interoperate with ZFS and Oracle Solaris. NFS over RDMA for NFS versions 3 and 4 is supported for Oracle Linux systems using the Oracle InfiniBand stack and is more efficient than using NFS with TCP over IPoIB. Currently, only the Mellanox ConnectX-2 and ConnectX-3 Host Channel Adapters (HCAs) pass the full Connectathon NFS test suite and are supported.

- **NFS server-side copy offload**

NFS server-side copy offload is an NFS v4.2 feature that reduces the overhead on network and client resources by offloading copy operations to one or more NFS servers rather than involving the client in copying file data over the network.

- **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.

## 1.8 Compatibility

Oracle Linux maintains user-space compatibility with Red Hat Enterprise Linux, 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.



---

# Chapter 2 Known Issues

## Table of Contents

2.1 btrfs .....	21
2.2 DIF/DIX is not supported for ext file systems .....	22
2.3 Console Appears to Hang when Booting .....	22
2.4 DTrace .....	22
2.5 ip_local_port_range Parameter Under LXC .....	22
2.6 Overlayfs Error with Docker .....	22

This chapter describes the known issues for the Unbreakable Enterprise Kernel Release 4.

## 2.1 btrfs

- The skinny-metadata feature is not enabled by default as it is incompatible with UEK R3. (Bug ID 22123918)
- 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)
- Commands such as `du` can show inconsistent results for file sizes in a btrfs file system when the number of bytes that is under delayed allocation is changing. (Bug ID 13096268)
- The copy-on-write nature of btrfs means that every operation on the file system initially requires disk space. It is possible that you cannot execute any operation on a disk that has no space left; even removing a file might not be possible. The workaround is to run `sync` before retrying the operation. If this does not help, remount the file system with the `-o nodatacow` option and delete some files to free up space. See <https://btrfs.wiki.kernel.org/index.php/ENOSPC>.
- If you run the `btrfs quota enable` command on a non-empty file system, any existing files do not count toward space usage. Removing these files can cause usage reports to display negative numbers and the file system to be inaccessible. The workaround is to enable quotas immediately after creating the file system. If you have already written data to the file system, it is too late to enable quotas. (Bug ID 16569350)
- The `btrfs quota rescan` command is not currently implemented. The command does not perform a rescan and returns without displaying any message. (Bug ID 16569350)
- 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. (Bug ID 16609467)
- If you run `btrfsck --init-csum-tree` on a file system and then run a simple `btrfsck` on the same file system, the command displays a Backref mismatch error that was not previously present. (Bug ID 16972799)
- Btrfs tracks the devices on which you create btrfs file systems. If you subsequently reuse these devices in a file system other than btrfs, you might see error messages such as the following when performing a device scan or creating a RAID-1 file system, for example:

```
ERROR: device scan failed '/dev/cciss/c0d0p1' - Invalid argument
```

You can safely ignore these errors. (Bug ID 17087097)

- If you use the `-s` option to specify a sector size to `mkfs.btrfs` that is different from the page size, the created file system cannot be mounted. By default, the sector size is set to be the same as the page size. (Bug ID 17087232)
- 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` Oracle Linux yum server repositories. 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` Oracle Linux yum server repositories.

## 2.2 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)

## 2.3 Console Appears to Hang when Booting

When booting Oracle Linux 6 on hardware with an ASPEED graphics controller, the console may 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)

## 2.4 DTrace

- Argument declarations for probe definitions cannot be declared with derived types such as `enum`, `struct`, or `union`.
- The following compiler warning can be ignored for 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 may not have accurate symbol resolution for symbols introduced by the `dlopen()`. (Bug ID 20045149)

## 2.5 ip\_local\_port\_range Parameter Under LXC

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)

## 2.6 Overlayfs Error with Docker

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 breaks Dockerfile builds 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. (Bug ID 21804564)



---

# Chapter 3 Installation and Availability

## Table of Contents

3.1 Installation Overview .....	25
3.2 Subscribing to ULN Channels .....	25
3.3 Enabling Access to Oracle Linux Yum Server Channels .....	26
3.4 Upgrading Your System .....	28
3.5 Installing the Oracle-Supported OFED Packages .....	28

You can install 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.

### 3.1 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 [Section 3.2, “Subscribing to ULN Channels”](#).

If your system is not registered with ULN, you can obtain most of the packages from Oracle Linux yum server. See [Section 3.3, “Enabling Access to Oracle Linux Yum Server Channels”](#).

Having subscribed your system to the appropriate channels on ULN or Oracle Linux yum server, upgrade your system. See [Section 3.4, “Upgrading Your System”](#).

After upgrading to UEK R4, you can replace any existing OFED packages with the Oracle-supported OFED packages, see [Section 3.5, “Installing the Oracle-Supported OFED Packages”](#).

### 3.2 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` (latest 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 the *Oracle Linux Unbreakable Linux Network User's Guide* at [https://docs.oracle.com/cd/E37670\\_01/index.html](https://docs.oracle.com/cd/E37670_01/index.html).

### 3.3 Enabling Access to Oracle Linux Yum Server Channels

At the Oracle Linux yum server repository at <https://yum.oracle.com/>, the kernel image and user-space packages are available on the following channels for Oracle Linux 6:

- `ol6_latest` (latest user-space packages for Oracle Linux 7 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` channel and disable the `ol6_UEKR3_latest` or `ol7_UEKR3` channel.



The DTrace utility and DRBD packages are not available on Oracle Linux yum server.

To enable access to the Oracle Linux 6 channels on Oracle Public Yum, create entries such as the following in `/etc/yum.conf` or in a repository file in the `/etc/yum.repos.d` directory:

```
[ol6_latest]
name=Oracle Linux $releasever Latest ($basearch)
baseurl=https://yum.oracle.com/repo/OracleLinux/OL6/latest/$basearch/
gpgkey=file:///etc/pki/rpm-gpg/RPM-GPG-KEY-oracle
gpgcheck=1
enabled=1

[ol6_UEK_latest]
name=Latest Unbreakable Enterprise Kernel for Oracle Linux $releasever ($basearch)
baseurl=https://yum.oracle.com/repo/OracleLinux/OL6/UEK/latest/$basearch/
gpgkey=file:///etc/pki/rpm-gpg/RPM-GPG-KEY-oracle
gpgcheck=1
enabled=0

[ol6_UEKR4]
name=Latest Unbreakable Enterprise Kernel Release 4 for Oracle Linux $releasever ($basearch)
baseurl=https://yum.oracle.com/repo/OracleLinux/OL6/UEKR4/latest/$basearch/
gpgkey=file:///etc/pki/rpm-gpg/RPM-GPG-KEY-oracle
gpgcheck=1
enabled=1

[ol6_playground_latest]
name=Latest mainline stable kernel for Oracle Linux 7 ($basearch) - Unsupported
baseurl=https://yum.oracle.com/repo/OracleLinux/OL6/playground/latest/$basearch/
gpgkey=file:///etc/pki/rpm-gpg/RPM-GPG-KEY-oracle
gpgcheck=1
enabled=0

[ol6_UEKR4_OFED]
name=OFED supporting tool packages for Unbreakable Enterprise Kernel on Oracle Linux 6 ($basearch)
baseurl=https://yum.oracle.com/repo/OracleLinux/OL6/UEKR4_OFED/$basearch/
gpgkey=file:///etc/pki/rpm-gpg/RPM-GPG-KEY-oracle
gpgcheck=1
enabled=0
```

To enable a channel, set the value of the `enabled` parameter for the channel to 1.

To disable a channel, set the value of the `enabled` parameter for the channel to 0.

In this example, access is enabled to the `ol6_latest` and `ol6_UEKR4` channels but not to the `ol6_UEK_latest`, `ol6_playground_latest` and `ol6_UEKR4_OFED` channels.

To enable access to the Oracle Linux 7 channels, create entries such as the following:

```
[ol7_latest]
name=Oracle Linux $releasever Latest ($basearch)
baseurl=https://yum.oracle.com/repo/OracleLinux/OL7/latest/$basearch/
gpgkey=file:///etc/pki/rpm-gpg/RPM-GPG-KEY-oracle
gpgcheck=1
enabled=1

[ol7_UEK_latest]
name=Latest Unbreakable Enterprise Kernel for Oracle Linux $releasever ($basearch)
baseurl=https://yum.oracle.com/repo/OracleLinux/OL7/UEK/latest/$basearch/
gpgkey=file:///etc/pki/rpm-gpg/RPM-GPG-KEY-oracle
gpgcheck=1
enabled=0
```

```
[ol7_UEKR4]
name=Latest Unbreakable Enterprise Kernel Release 4 for Oracle Linux $releasever ($basearch)
baseurl=https://yum.oracle.com/repo/OracleLinux/OL7/UEKR4/latest/$basearch/
gpgkey=file:///etc/pki/rpm-gpg/RPM-GPG-KEY-oracle
gpgcheck=1
enabled=1

[ol7_UEKR4_OFED]
name=OFED supporting tool packages for Unbreakable Enterprise Kernel on Oracle Linux 7 ($basearch)
baseurl=https://yum.oracle.com/repo/OracleLinux/OL7/UEKR4_OFED/$basearch/
gpgkey=file:///etc/pki/rpm-gpg/RPM-GPG-KEY-oracle
gpgcheck=1
enabled=0
```

In this example, access is enabled to the `ol7_latest` and `ol7_UEKR4` channels but not to the `ol7_UEK_latest` and `ol7_UEKR4_OFED` channels.

You can find more information about installing the software at <https://yum.oracle.com/>, from where you download a copy of a suitable repository file (<https://yum.oracle.com/public-yum-ol6.repo> or <https://yum.oracle.com/public-yum-ol7.repo>).

## 3.4 Upgrading Your System

To upgrade your system to UEK R4:

1. After enabling access to the appropriate channels, including `ol6_UEKR4` or `ol7_UEKR4`, in the Oracle Linux yum server repository or `ol6_x86_64_UEKR4` or `ol7_x86_64_UEKR4` on ULN, run the following command to upgrade the system to UEK R4:

```
# 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.

For instructions on how to install the Oracle-supported OFED packages after upgrading to UEK R4, see [Section 3.5, “Installing the Oracle-Supported OFED Packages”](#).

If you have questions regarding configuring or using `yum` to install updates, refer to the [Oracle Linux Unbreakable Linux Network User's Guide](#).

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>.

## 3.5 Installing the Oracle-Supported OFED Packages

This section describes how to install the OFED packages provided by Oracle, including how to remove any existing OFED packages.

To install the OFED packages 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.

If you want to install the packages from Oracle Linux Yum Server:

- a. Move the existing yum repository file to a backup file, for example:

```
# mv /etc/yum.repos.d/public-yum-olN.repo /etc/yum.repos.d/public-yum-olN.repo.bck
```

- b. Download the latest yum repository file for Oracle Linux 6 or Oracle Linux 7 from Oracle Linux yum server.

```
# wget -O /etc/yum.repos.d/public-yum-olN.repo https://yum.oracle.com/public-yum-olN.repo
```

- c. Edit the downloaded yum repository file and enable the `ol6_UEKR4_OFED` or `ol7_UEKR4_OFED` repository as appropriate.

2. Remove any existing OFED packages.

```
# yum remove 'ibacm*'
# yum remove 'ib-bonding*'
# yum remove 'ibutils*'
# yum remove 'infiniband-diags*'
# 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 'perfctest*'
# yum remove 'qperf*'
# yum remove 'sdpnstat*'
# yum remove 'rdma-*'
# yum remove 'rds-tools*'
```

3. Clean all yum cached files from all enabled repositories.

```
# yum clean all
```

4. Install the OFED packages for UEK R4.

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

5. If the SDP feature is required, install the `libsdp` package.

```
# yum install libsdp
```

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

```
# chkconfig rdma on
```



#### Note

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

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

