
The science behind the report:

# Accelerate I/O with NVMe drives on the new Dell EMC PowerEdge R650 server

This document describes what we tested, how we tested, and what we found. To learn how these facts translate into real-world benefits, read the report Accelerate I/O with NVMe drives on the new Dell EMC PowerEdge R650 server.

We concluded our hands-on testing on March 8, 2021. During testing, we determined the appropriate hardware and software configurations and applied updates as they became available. The results in this report reflect configurations that we finalized on March 3, 2021 or earlier. Unavoidably, these configurations may not represent the latest versions available when this report appears.

# Our results

To learn more about how we have calculated the wins in this report, go to http://facts.pt/calculating-and-highlighting-wins. Unless we state otherwise, we have followed the rules and principles we outline in that document.

Table 1: The input/output operations per second (IOPS) results of our RAID 10 level testing.

| RAID 10 level IOPS | Dell EMC PowerEdge R640 | Dell EMC PowerEdge R650 | Dell EMC PowerEdge R650 times the IOPS |
|---|---|---|---|
| 4KB random 100% read | 505,000 | 3,555,000 | 7.04 |
| 4KB random 100% write | 116,000 | 917,000 | 7.91 |
| 8KB random 70/30 read/write | 294,500 | 1,011,000 | 3.43 |
| 8KB random 50/50 read/write | 214,000 | 1,116,000 | 5.21 |
| 64KB sequential 100% write | 30,700 | 90,100 | 2.93 |

Table 2: The input/output operations per second (IOPS) results of our RAID 6 and RAID 5 level testing.

| | Dell EMC PowerEdge R640 | Dell EMC PowerEdge R650 | Dell EMC PowerEdge R650 times the IOPS |
|---|---|---|---|
| 4KB random read | | | |
| RAID 6 level IOPS | 515,000 | 3,541,000 | 6.88 |
| RAID 5 level IOPS | 514,000 | 3,541,000 | 6.89 |
| 4KB random write | | | |
| RAID 6 level IOPS | 12,600 | 198,000 | 15.71 |
| RAID 5 level IOPS | 27,100 | 186,000 | 6.86 |
| 8KB random 70/30 read/write | | | |
| RAID 6 level IOPS | 47,300 | 407,000 | 8.60 |
| RAID 5 level IOPS | 88,400 | 499,000 | 5.64 |
| 8KB random 100% read | | | |
| RAID 6 level IOPS | 398,000 | 1,793,000 | 4.51 |
| RAID 5 level IOPS | 398,000 | 1,793,000 | 4.51 |
| 64KB sequential 100% read | | | |
| RAID 6 level IOPS | 65,600 | 226,000 | 3.45 |
| RAID 5 level IOPS | 65,500 | 226,000 | 3.45 |

Table 3: The disk bandwidth results of our RAID 10 level testing.

| RAID 10 level disk bandwidth | Dell EMC PowerEdge R640 | Dell EMC PowerEdge R650 | Dell EMC PowerEdge R650 times the MB/s |
|---|---|---|---|
| 4KB random 100% read | 2,072 | 14,600 | 7.05 |
| 4KB random 100% write | 477 | 3,757 | 7.88 |
| 8KB random 70/30 read/write | 2,416 | 8,280 | 3.43 |
| 8KB random 50/50 read/write | 1,756 | 9,144 | 5.21 |
| 64KB sequential 100% write | 2,012 | 5,908 | 2.94 |

Table 4: The disk bandwidth results of our RAID 6 and RAID 5 level testing.

| | Dell EMC PowerEdge R640 | Dell EMC PowerEdge R650 | Dell EMC PowerEdge R650 times the MB/s |
|---|---|---|---|
| 4KB random read | | | |
| RAID 6 level disk bandwidth | 2,108 | 14,500 | 6.88 |
| RAID 5 level disk bandwidth | 2,107 | 14,500 | 6.88 |
| 4KB random write | | | |
| RAID 6 level disk bandwidth | 52 | 811 | 15.75 |
| RAID 5 level disk bandwidth | 111 | 763 | 6.87 |
| 8KB random 70/30 read/write | | | |
| RAID 6 level disk bandwidth | 387 | 3,336 | 8.62 |
| RAID 5 level disk bandwidth | 724 | 4,083 | 5.64 |
| 8KB random 100% read | | | |
| RAID 6 level disk bandwidth | 3,259 | 14,700 | 4.51 |
| RAID 5 level disk bandwidth | 3,259 | 14,700 | 4.51 |
| 64KB sequential 100% read | | | |
| RAID 6 level disk bandwidth | 4,298 | 14,800 | 3.44 |
| RAID 5 level disk bandwidth | 4,295 | 14,800 | 3.45 |

Table 5: The latency results of our RAID 10 level testing.

| RAID 10 level latency | Dell EMC PowerEdge R640 | Dell EMC PowerEdge R650 | Percentage lower for the Dell EMC PowerEdge R650 |
|---|---|---|---|
| 4KB random 100% read | 16.2 | 0.6 | 96.30% |
| 4KB random 100% write | 70.3 | 8.9 | 87.34% |
| 8KB random 70/30 read/write | 1.8 | 1.6 | 11.11% |
| 8KB random 50/50 read/write | 3.9 | 0.5 | 87.18% |
| 64KB sequential 100% write | 16.7 | 0.6 | 96.41% |

Table 6: The latency results of our RAID 6 and RAID 5 level testing.

| | Dell EMC PowerEdge R640 | Dell EMC PowerEdge R650 | Percentage lower for the Dell EMC PowerEdge R650 |
|---|---|---|---|
| 4KB random read | | | |
| RAID 6 level latency | 4.0 | 0.6 | 85.00% |
| RAID 5 level latency | 4.0 | 0.6 | 85.00% |
| 4KB random write | | | |
| RAID 6 level latency | 162.8 | 10.3 | 93.67% |
| RAID 5 level latency | 75.7 | 11.0 | 85.47% |
| 8KB random 70/30 read/write | | | |
| RAID 6 level latency | 33.0 | 3.1 | 90.61% |
| RAID 5 level latency | 16.9 | 2.6 | 84.62% |
| 8KB random 100% read | | | |
| RAID 6 level latency | 1.3 | 0.3 | 76.92% |
| RAID 5 level latency | 1.3 | 0.3 | 76.92% |
| 64KB sequential 100% read | | | |
| RAID 6 level latency | 7.8 | 2.3 | 70.51% |
| RAID 5 level latency | 7.8 | 2.3 | 70.51% |

# System configuration information

Table 7: Detailed information on the systems we tested.

| System configuration information | Dell PowerEdge R640 | Dell PowerEdge R650 |
|---|---|---|
| Processor | | |
| Vendor | Intel® | Intel |
| Name | Xeon® Gold | Pre-production 3rd Gen Xeon Scalable CPU |
| Model number | 6142 | n/a |
| Number of cores | 16 | 32 |
| Number of processors | 2 | 2 |
| Memory | | |
| Amount (GB) | 128 | 1,024 |
| Type | PC4-19200 | PC4-23400 |
| Speed (MHz) | 2,400 | 2,933 |
| Storage controller | | |
| Vendor and Model | Dell PERC H730P Mini | Dell PERC H755N Front |
| Firmware | 25.5.8.0001 | 52.14.0-3708 |
| Driver | megaraid_sas 4.18.0-240.10.1.el8_3.x86_64 | megaraid_sas 4.18.0-240.15.1.el8_3.x86_64 |
| Storage | | |
| Vendor and Model | Intel SSDSC2KB019T8R | Dell Ent NVMe AGN MU U.2 |
| Size | 1.9 TB | 3.2 TB |
| Count | 10 | 8 |
| Firmware | XCV1DL63 | 2.0.2 |
| Connectivity/expansion | | |
| Ethernet #1 | Intel Gigabit 4P I350-t rNDC | Intel Ethernet 25G 2P E810-XXV OCP |
| Ethernet #2 | N/A | Broadcom Gigabit Ethernet BCM5720 |
| Operating system | | |
| Vendor and Name | Red Hat Enterprise Linux | Red Hat Enterprise Linux |
| Build number or version | 8.3 | 8.3 |
| BIOS | | |
| BIOS name and version | Dell 2.10.0 | Dell 0.4.3 |

# How we tested

## Changing BIOS System Profile to Performance

1. Power on the System.
2. Enter BIOS, and change the System Profile to Performance.

## Installing Red Hat Enterprise Linux 8.3

1. Boot to the Red Hat Enterprise Linux 8.3 installation media using rhel-8.3-x86_64-dvd.iso.
2. When choosing what to boot, select Install Red Hat Enterprise Linux 8.3.0.
3. To move past the start screen once the installer loads, click Next.
4. In the Installation Summary screen, click Installation Destination.
5. In the Installation Destination screen, accept the default, and click Done.
6. In the Installation Summary screen, click Software Selection.
7. Select Minimal Install, and click Done.
8. In the Installation Summary screen, click Network & Host Name.
9. In Network & Host Name, choose your hostname, turn on your Ethernet connection, and click Configure.
10. In Editing your Ethernet connection, click Connect automatically with priority 0, and click Save.
11. Click Done.
12. Back in the Installation Summary screen, click Begin Installation.
13. In the Configuration screen, click Root Password.
14. In Root Password, enter your root password, and click Done.
15. To reboot into the OS after the installation completes, click Reboot.

## Configuring Red Hat Enterprise Linux 8.3

1. Log onto the server as root.
2. Update the kernel to use blk_mq scsi_mod:

```
grubby --update-kernel=/boot/vmlinuz-4.18.0-240.10.1.el8_3.x86_64 --args="scsi_mod.use_blk_mq=1 dm
mod.use_blk_mq=y"
```

3. Update the configuration to use exact IRQ balancing:

```
vi /etc/sysconfig/irqbalance
change the line: #IRQBALANCE_ARGS=
to: IRQBALANCE_ARGS="-h exact"
```

4. Tune the system for throughput performance:

```
mkdir /etc/tuned/updated-profile
touch /etc/tuned/tuned.conf
vi /etc/tuned/tuned.conf
```

5. Type the following into the file:

```
[main]
include=throughput-performance

[cpu]
force_latency=1
governor=performance
energy_perf_bias=performance
min_perf_pct=100
```

6. Enable the profile:

```
tuned-adm profile updated-profile
```

7. Update the host:

```
yum update -y
```

8. Reboot the host.

## Installing test software

1. Install FIO:

```
yum install -y fio
```

## Configuring the RAID drives

1. Download PERCCLI from the Dell support website to the server.
2. Install PERCCLI:

```
run rpm -ivh <percli-x.xx-x.noarch.rpm>
```

3. Change directory to /opt/MegaRAID/perccli.
4. To set up each RAID volume, type the following:

```
./perccli64 /c0 add vd type=raid{10,6,5} name=RAID{10,6,5}-1 drives=252:0,1,2,3 pdcache=off
wt nora strip=64
./perccli64 /c0 add vd type=raid{10,6,5} name=RAID{10,6,5}-2 drives=252:4,5,6,7 pdcache=off
wt nora strip=64
./perccli64 /c0/v238 start init full Force
./perccli64 /c0/v239 start init full Force
```

5. To configure each volume once they finish initializing, type the following:

```
echo none > /sys/block/sdb/queue/scheduler none
echo 0 > /sys/block/sdb/queue/add_random 0
echo 2 > /sys/block/sdb/queue/nomerges 2
echo 2048 > /sys/block/sdb/queue/nr_requests 2048
echo 0 > /sys/block/sdb/queue/rotational 0
echo 2 > /sys/block/sdb/queue/rq_affinity 2
echo 1024 > /sys/block/sdb/queue/max_sectors_kb 1024
echo 1024 > /sys/block/sdb/device/queue_depth 1024
echo 1 > /sys/block/sdb/queue/iostats 1
echo none > /sys/block/sdc/queue/scheduler none
echo 0 > /sys/block/sdc/queue/add_random 0
echo 2 > /sys/block/sdc/queue/nomerges 2
echo 2048 > /sys/block/sdc/queue/nr_requests 2048
echo 0 > /sys/block/sdc/queue/rotational 0
echo 2 > /sys/block/sdc/queue/rq_affinity 2
echo 1024 > /sys/block/sdc/queue/max_sectors_kb 1024
echo 1024 > /sys/block/sdc/device/queue_depth 1024
echo 1 > /sys/block/sdc/queue/iostats 1
```

6. Precondition reads and writes on each volume:

```
fio --iodepth=256 --numjobs=1 --runtime=5h /root/fio/jobs/64k_sequential_read.fio
fio --iodepth=256 --numjobs=1 --runtime=9h /root/fio/jobs/64k_random_read.fio
fio --iodepth=256 --numjobs=1 --runtime=5h /root/fio/jobs/64k_sequential_write.fio
fio --iodepth=256 --numjobs=1 --runtime=9h /root/fio/jobs/64k_random_write.fio
```

## Running the benchmark

1. On the server, run the benchmark script:

```
fio --runtime=10m --output=/root/fio/results/output.txt --output-format=normal --terse-version=3 /root/
fio/jobs/{config}.fio
```

## Configuring FIO files

The following sections list the parameters for the I/O profiles we tested.

### 4k_random_read.fio

```
[global]
name=4k_random_read
scramble_buffers=1
buffered=0
sync=0
ioengine=libaio
norandommap
do_verify=0
randrepeat=0
random_generator=tausworthe64
offset=0
thinktime=0
time_based=1
ramp_time=0
direct=1
overwrite=0
group_reporting=1
bs=4k
blockalign=4k
rw=randread
numa_cpu_nodes=0
cpus_allowed_policy=shared
iodepth=32,64
numjobs=32,64

[part1]
filename=/dev/sdb

[part2]
filename=/dev/sdc
```

### 4k_random_write.fio

```
[global]
name=4k_random_write
scramble_buffers=1
buffered=0
sync=0
ioengine=libaio
norandommap
do_verify=0
randrepeat=0
random_generator=tausworthe64
offset=0
thinktime=0
time_based=1
ramp_time=0
direct=1
overwrite=0
group_reporting=1
bs=4k
blockalign=4k
```

```
rw=randwrite
numa_cpu_nodes=0
cpus_allowed_policy=shared
iodepth=32,64
numjobs=32,64

[part1]
filename=/dev/sdb

[part2]
filename=/dev/sdc
```

### 8k_random_50-50.fio

```
[global]
name=8k_random_50-50
scramble_buffers=1
buffered=0
sync=0
ioengine=libaio
norandommap
do_verify=0
randrepeat=0
random_generator=tausworthe64
offset=0
thinktime=0
time_based=1
ramp_time=60
direct=1
overwrite=0
group_reporting=1
bs=8k
blockalign=4k
rw=randrw
numa_cpu_nodes=0
cpus_allowed_policy=shared
iodepth=32
numjobs=8,16,32

[part1]
filename=/dev/sdb

[part2]
filename=/dev/sdc
```

### 8k_random_70-30.fio

```
[global]
name=8k_random_70-30
scramble_buffers=1
buffered=0
sync=0
ioengine=libaio
norandommap
do_verify=0
randrepeat=0
random_generator=tausworthe64
offset=0
```

```
thinktime=0
time_based=1
ramp_time=60
direct=1
overwrite=0
group_reporting=1
bs=8k
blockalign=4k
rw=randrw
rwmixread=70
rwmixwrite=30
numa_cpu_nodes=0
cpus_allowed_policy=shared
iodepth=32
numjobs=8,16,32

[part1]
filename=/dev/sdb

[part2]
filename=/dev/sdc
```

### 8k_random_read.fio

```
[global]
name=8k_random_read
scramble_buffers=1
buffered=0
sync=0
ioengine=libaio
norandommap
do_verify=0
randrepeat=0
random_generator=tausworthe64
offset=0
thinktime=0
time_based=1
ramp_time=60
direct=1
overwrite=0
group_reporting=1
bs=8k
blockalign=4k
rw=randread
numa_cpu_nodes=0
cpus_allowed_policy=shared
iodepth=32
numjobs=8,16,32

[part1]
filename=/dev/sdb

[part2]
filename=/dev/sdc
```

### 64k_sequential_read.fio

```
[global]
```

```
name=64k_sequential_read
scramble_buffers=1
buffered=0
sync=0
ioengine=libaio
norandommap
do_verify=0
randrepeat=0
random_generator=tausworthe64
offset=0
thinktime=0
time_based=1
ramp_time=60
direct=1
overwrite=0
group_reporting=1
bs=64k
blockalign=4k
rw=read
numa_cpu_nodes=0
cpus_allowed_policy=shared
iodepth=256
numjobs=1

[part1]
filename=/dev/sdb

[part2]
filename=/dev/sdc
```

### 64k_sequential_write.fio

```
[global]
name=64k_sequential_write
scramble_buffers=1
buffered=0
sync=0
ioengine=libaio
norandommap
do_verify=0
randrepeat=0
random_generator=tausworthe64
offset=0
thinktime=0
time_based=1
ramp_time=60
direct=1
overwrite=0
group_reporting=1
bs=64k
blockalign=4k
rw=write
numa_cpu_nodes=0
cpus_allowed_policy=shared
iodepth=256
numjobs=1
```

```
[part1]
filename=/dev/sdb

[part2]
filename=/dev/sdc
```

### 64k_random_read.fio

```
[global]
name=64k_sequential_read
scramble_buffers=1
buffered=0
sync=0
ioengine=libaio
norandommap
do_verify=0
randrepeat=0
random_generator=tausworthe64
offset=0
thinktime=0
time_based=1
ramp_time=60
direct=1
overwrite=0
group_reporting=1
bs=64k
blockalign=4k
rw=randread
numa_cpu_nodes=0
cpus_allowed_policy=shared
iodepth=256
numjobs=1

[part1]
filename=/dev/sdb

[part2]
filename=/dev/sdc
```

### 64k_random_write.fio

```
[global]
name=64k_random_write
scramble_buffers=1
buffered=0
sync=0
ioengine=libaio
norandommap
do_verify=0
randrepeat=0
random_generator=tausworthe64
offset=0
thinktime=0
time_based=1
ramp_time=60
direct=1
overwrite=0
group_reporting=1
```

```
bs=64k
blockalign=4k
rw=randread
numa_cpu_nodes=0
cpus_allowed_policy=shared
iodepth=256
numjobs=1

[part1]
filename=/dev/sdb

[part2]
filename=/dev/sdc
```

**Read the report at http://facts.pt/qvST5X5 ▶**

This project was commissioned by Dell EMC.

Principled Technologies®

Facts matter.®