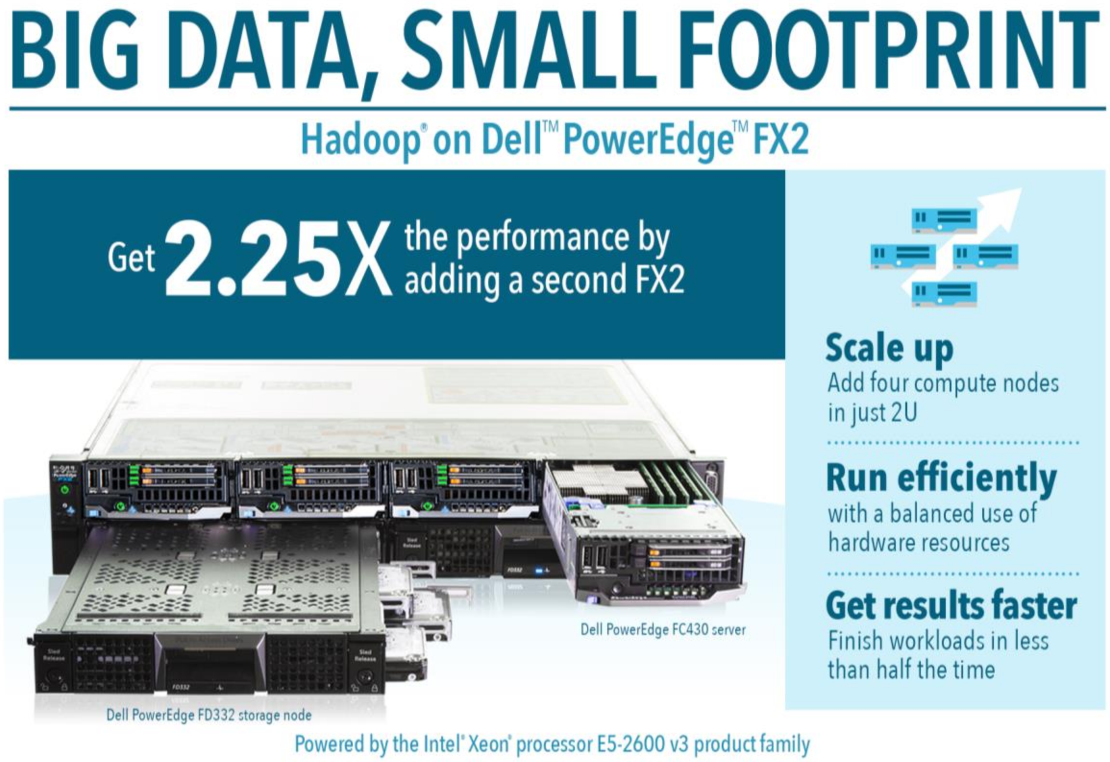
HADOOP INFRASTRUCTURE SCALING WITH THE DELL POWEREDGE FX2



When wading into the Hadoop big data pool, it's important to select a solution that can handle the jobs you run, and one that is flexible enough to scale well as the size of your big data needs increase over time. The Dell PowerEdge FX2 is a datacenter solution that combines all the essential IT elements—servers, storage, and networking blocks—into a very compact 2U chassis. You can tailor the Dell PowerEdge FX2 solution to meet your unique workload needs, such as Hadoop workloads that process big data. In particular, Hadoop thrives with uniform compute scale-out and a high disk-to-compute ratio for Hadoop File System (HDFS) storage capacity, both of which the Dell PowerEdge FX2 provides.

In the Principled Technologies labs, we tested a single Dell PowerEdge FX2 with four PowerEdge FC430 nodes, and found that it completed our Hadoop workload in 25 minutes and 58 seconds. When we added a second Dell PowerEdge FX2, Hadoop performance scaled well: by just adding a second FX2 cluster, it cut the job time by more than half. All the way down to 11 minutes and 31 seconds.

While many Hadoop infrastructures have dozens of nodes, you want to be sure when starting out to choose a flexible and scalable solution. By choosing the Dell PowerEdge FX2 to start your Hadoop infrastructure, you can get all the benefits of its unique converged infrastructure design, which can include fast performance, simplified management, and space savings thanks to its dense nature. And when you decide it's time to scale out your solution, adding a cluster and cutting job times in half is simple thanks to the Dell PowerEdge FX2 all-in-one chassis.

SEPTEMBER 2015
A PRINCIPLED TECHNOLOGIES REPORT
Commissioned by Dell Inc.

BIG DATA IN SMALL SPACES

Sorting and reorganizing the data you collect can help your organization get a handle on how your business runs. Hadoop is an application that breaks big data into smaller sets and spreads them out over multiple server nodes, making big data analysis fast and scalable.

The Dell PowerEdge FX2 solution configured with four server nodes and two storage blocks can run Hadoop workloads, and does it all in just 2U of space. With servers, storage, and networking sharing a common chassis, the Dell PowerEdge FX2 brings all the elements of a traditional datacenter into a single chassis, which can simplify your infrastructure. Because the PowerEdge FX2 can support a number of different configurations of those elements, you can build your organization's PowerEdge FX2 to fit your exact workload needs. These are just some of the kinds of benefits that the Dell PowerEdge FX2 can bring to organizations that traditional server and storage setups can't; it helps you make the most efficient use of each element in your infrastructure.

WHAT WE FOUND

About the results

Our test workload used 300GB of data and performed several common Hadoop operations on large datasets, including data generation, sorting the data, and data validation. Our workload executed a short data integrity check after the data generation and sorting portions. These operations are simple but highly representative of real-world Hadoop workloads that stress the Map-Reduce framework and the Hadoop Filesystem API.

We used Cloudera Distributed Hadoop (CDH) 5.4.2 as our Hadoop cluster software. We set up the first Dell PowerEdge FX2 to house the Edge, Name, and Data Node roles across four nodes. The second Dell PowerEdge FX2 unit had four Data Nodes. See Appendix C for specific Hadoop tuning parameters.

We tested the scalability of the Dell PowerEdge FX2 with four Dell PowerEdge FC430 nodes and two Dell PowerEdge FD332 storage arrays by running the TPCx-HS 300GB workload on one Dell PowerEdge FX2, then adding a second Dell PowerEdge FX2 with the same hardware configuration and measuring the time required to run the same workload. When we added a second Dell PowerEdge FX2 to the cluster, the workload time decreased by 56 percent (see Figure 1).

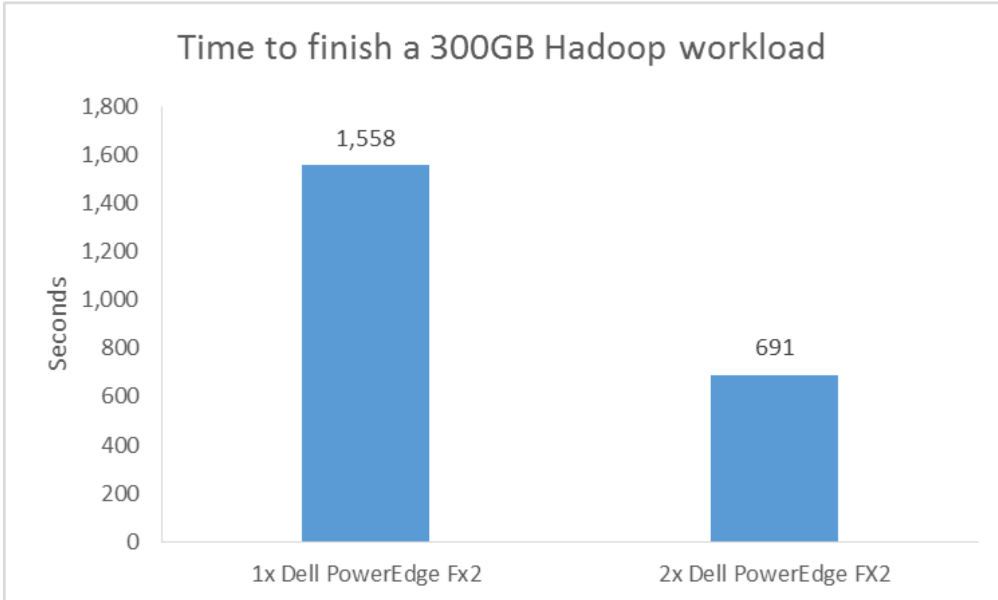


Figure 1: Time to complete our Hadoop workload, in seconds.

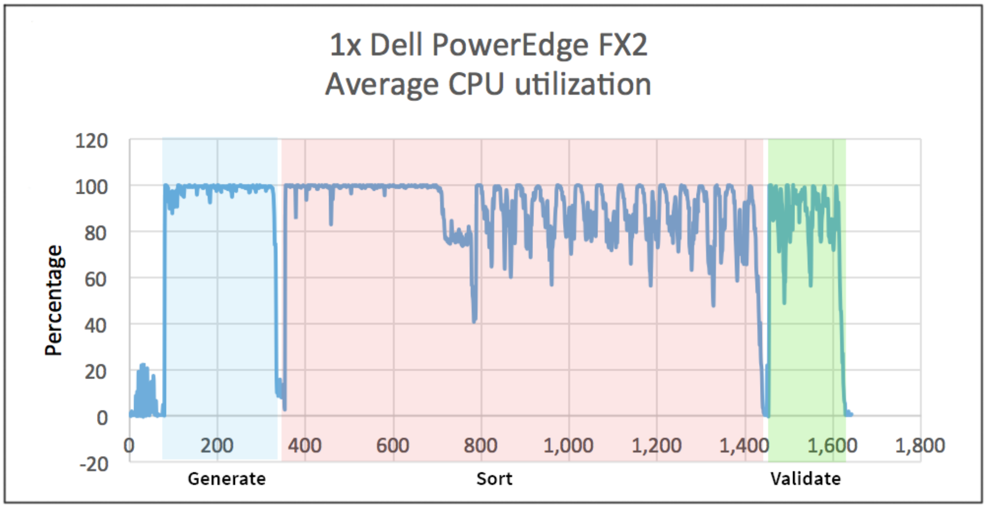
Efficient use of resources

A properly tuned Hadoop cluster can take advantage of all the hardware subsystems (CPU, memory, and storage) you make available to it. Based on Hadoop example workloads TeraGen, TeraSort, and TeraValidate, our workload was dependent on CPU, memory and disk resources, so it was important that all three subsystems were adequately utilized.

Not only did the Dell PowerEdge FX2 unit show excellent scaling, it was also able to provide balanced use of its hardware resources in both phases of testing. Because each of the balanced utilization, an owner of a similarly configured Dell PowerEdge FX2 could run this workload confident that resources are being used efficiently. That same owner could then purchase a second, identical Dell PowerEdge FX2 and be comfortable knowing that their workloads continue to operate without leaving idle hardware on the table.

Figure 2 through 4 show the utilization metrics (averaged across the Data Nodes for each phase) of each hardware subsystem during the first and second phases of our testing.

As Figure 2 shows, CPU utilization remained high for every portion of the workload during the first phase of testing. Adding a second Dell PowerEdge FX2 did not change the CPU utilization performance profile, showing that this workload scales well from a CPU perspective. The slight decrease in CPU activity during sorting is due to the disk-intensive reduce portion of that operation.



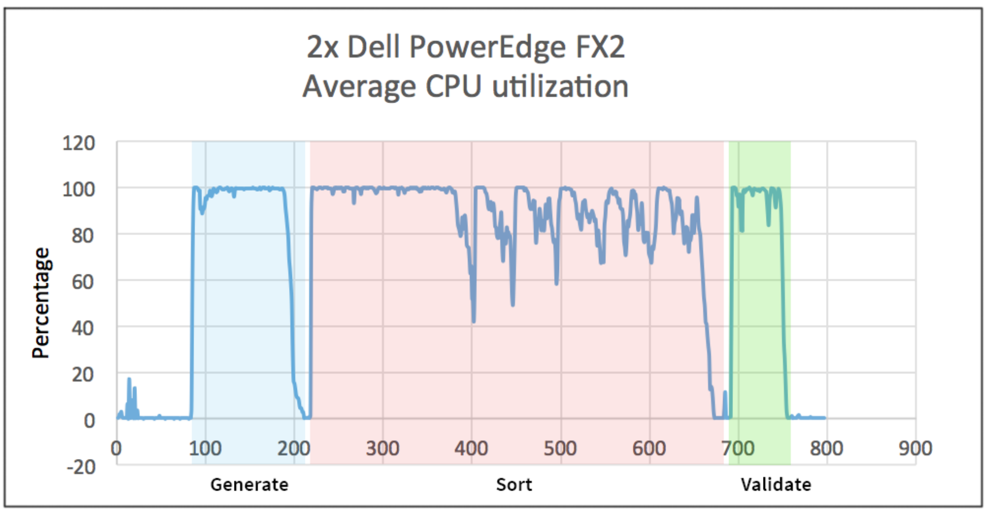
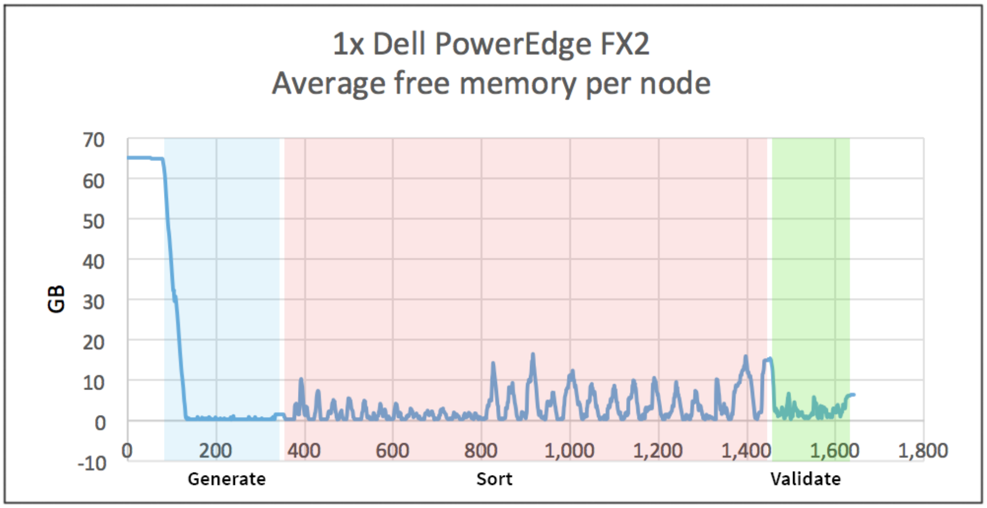


Figure 2: Average Data Node CPU utilization percentages for 1x Dell PowerEdge FX2 and for 2x Dell PowerEdge FX2.

We tuned our Hadoop cluster to take full advantage of the available memory in each node. As Figure 3 shows, the workload was able to make use of all the memory in both phases of testing, indicating that the workload scales well from a memory usage perspective.



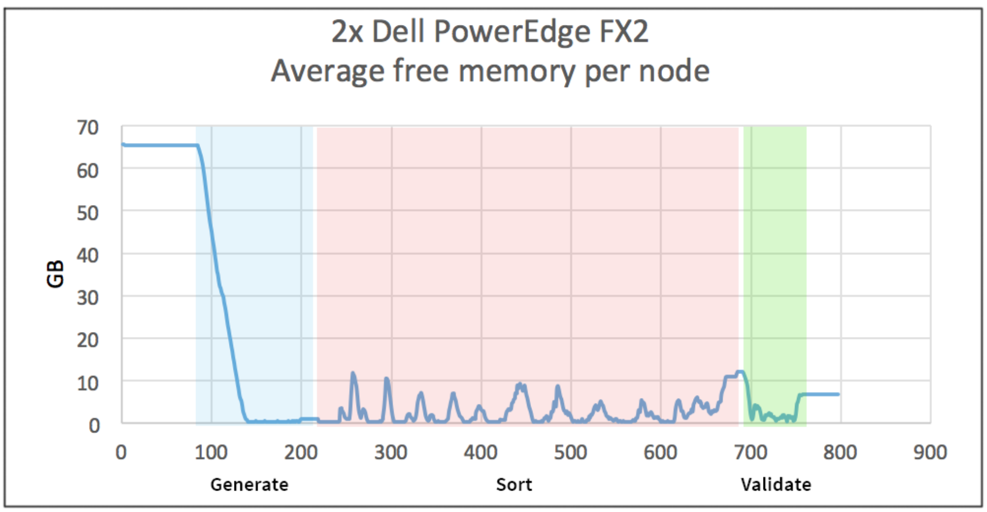


Figure 3: Average free memory per Data Node for 1x Dell PowerEdge FX2 and for 2x Dell PowerEdge FX2.

Disk performance is critical to many Hadoop operations, and the three major operations in our workload are no exception. The Dell PowerEdge FD332 storage blocks and shared RAID controllers allow presentation of the disks in RAID or HBA mode. While a RAID group can add performance and data replication for many common workloads, Hadoop prefers HBA mode as the Hadoop Distributed File System (HDFS) handles replication. Our workload was able to fully utilize the disks during data generation and the reduce portion of the sorting operations. These operations occur in memory whenever possible, which means that disk utilization decreases during data validation and the map portion of sorting As Figure 4 shows, the level of disk utilization was similar in both phases of testing, indicating good scaling of disk resources.

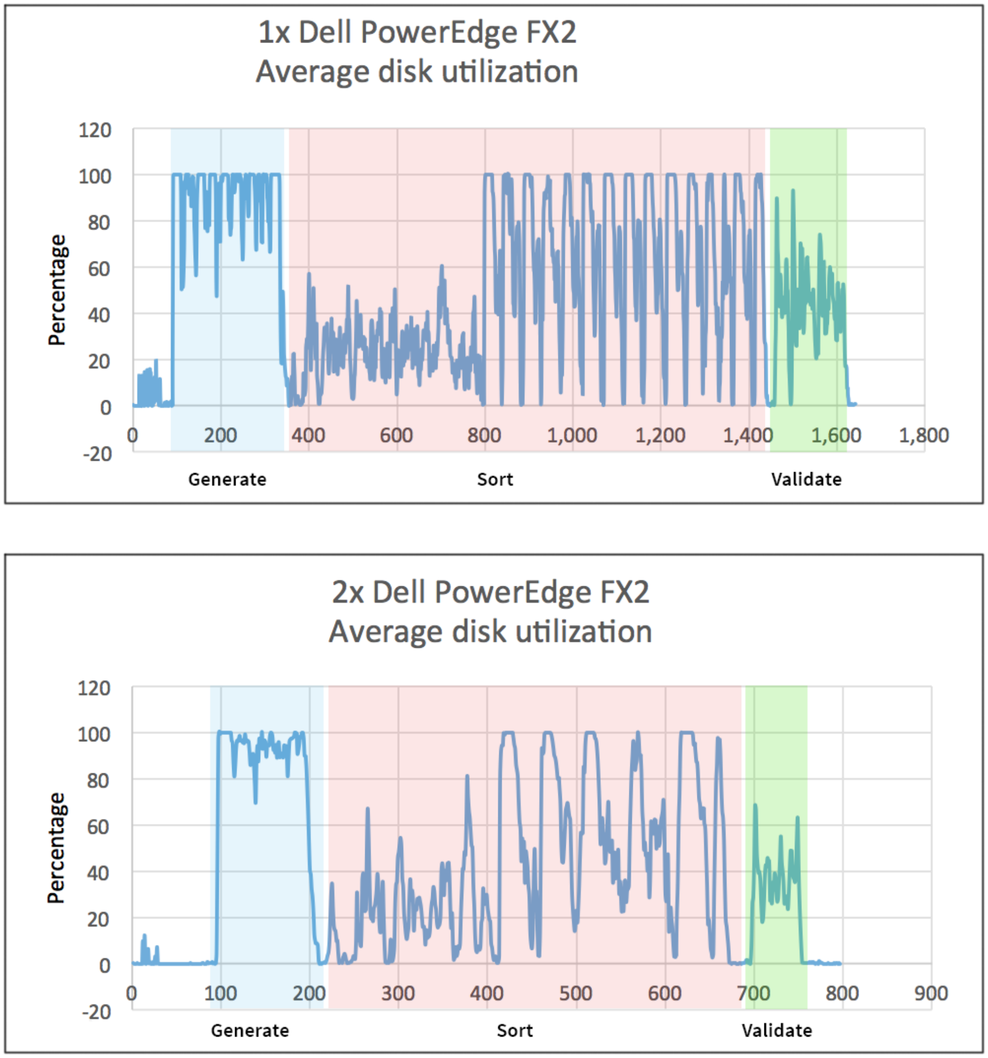


Figure 4: Average disk utilization across all Data Nodes for 1x Dell PowerEdge FX2 and for 2x Dell PowerEdge FX2.

CONCLUSION

The definition of a successful Hadoop solution need not be limited to whether or not the hardware can run the jobs and sort the data. As our tests show, the Dell PowerEdge FX2 was powerful enough to run our Hadoop workload, but more importantly, it scaled well when we added another cluster. Adding a second PowerEdge FX2 chassis complete with four Dell PowerEdge FC430 server nodes and Dell PowerEdge FD332 storage cut the time to run our Hadoop job in half. The all-in-one chassis that brings compute, storage, and networking together can also offer other benefits inherent in its design: the Dell PowerEdge FX2 can sort big data in a small space, which can also deliver space savings and ease the burden of managing the Hadoop solution.

APPENDIX A – ABOUT THE COMPONENTS

About the Dell PowerEdge FX2 enclosure

The shared infrastructure approach of the Dell PowerEdge FX2 enclosure is scalable and can help you make the most of your datacenter space while reducing rack space. The Dell PowerEdge FX2 enclosure has a standard 2U footprint and features a modular design that can hold different combinations of compute and storage nodes to meet your specific goals. The PowerEdge FX2 fits four half-width or eight quarter-width compute nodes to increase the compute density in your rack and optimize the space in your datacenter. You can deploy the FX2 solution like a traditional rack-mounted server while gaining the benefits and features that more expensive dense blade solutions provide. Important features of the FX2 enclosure include the following:

- Up to eight low-profile PCIe® expansion slots
- Two pass-through or optional networking FN I/O Aggregator modules
- Embedded network adapters within the server nodes
- Offers both chassis-based management through the Chassis Management Controller and rack-based management through Integrated Dell Remote Access Controller (iDRAC) with Lifecycle Controller on each compute node

The Dell PowerEdge FX2 enclosure fits a number of server and storage options, including the PowerEdge FM120, FC430, FC630, and FC830 servers, and PowerEdge FD332 storage node—all powered by Intel® Xeon® processors. For more information about the Dell PowerEdge FX2 solution, visit www.dell.com/us/business/p/poweredge-fx/pd.

About the Intel Xeon processor E5-2600 v3 product family

According to Intel, the Intel Xeon processor E5-2600 v3 product family "helps IT address the growing demands placed on infrastructure, from supporting business growth to enabling new services faster, delivering new applications in the enterprise, technical computing, communications, storage, and cloud." It also delivers benefits in performance, power efficiency, virtualization, and security.

The E5-2600 v3 product family has up to 50 percent more cores and cache than processors from the previous generation. Other features include the following:

- Intel Advanced Vector Extensions 2 (AVX2)
- Intel Quick Path Interconnect link
- Up to 18 cores and 36 threads per socket
- Up to 45 MB of last level cache
- Next-generation DDR4 memory support
- Intel Integrated I/O providing up to 80 PCIe lanes per two-socket server
- Intel AES-NI data encryption/decryption

The Intel Xeon processor E5-2600 v3 product family also uses Intel Intelligent Power technology and Per-core P states to maximize energy efficiency. Learn more at www.intel.com/content/www/us/en/processors/xeon/xeon-e5-brief.html.

APPENDIX B – SYSTEM CONFIGURATION INFORMATION

Figure 5 provides detailed configuration information for the test systems, and Figure 6 provides details about the test storage.

Server	Edge Node/Name Node	Data Nodes
Enclosure		
Blade enclosure	Dell PowerEdge FX2	Dell PowerEdge FX2
General dimension information		
Height (inches)	3.5	3.5
Width (inches)	17	17
Depth (inches)	33.5	33.5
Power supplies		
Total number	2	2
Wattage of each (W)	1,600	1,600
Cooling fans		
Total number	8 (2 + 6)	8 (2 + 6)
Dimensions (h × w) of each	3.3 × 3.5 (2), 2.5 × 2.5 (6)	3.3 x 3.5 (2), 2.5 × 2.5 (6)
Voltage (V)	12 (2), 12 (6)	12 (2), 12 (6)
Amps (A)	8 (2), 3.3 (6)	8 (2), 3.3 (6)
General processor setup		
Number of processor packages	2	2
Number of cores per processor package	8	8
Number of hardware threads per core	16	16
System power management policy	Default	Default
CPU		
Vendor	Intel	Intel
Name	Xeon E5-2640 v3	Xeon E5-2640 v3
Stepping	2	2
Socket type	FCLGA2011-3	FCLGA2011-3
Core frequency (GHz)	2.6	2.6
L1 cache	32KB +32KB (per core)	32KB +32KB (per core)
L2 cache	256KB (per core)	256KB (per core)
L3 cache	20MB	20MB
Platform		
Vendor and model number	Dell PowerEdge FC430	Dell PowerEdge FC430
Motherboard model number	03X19KX05	03X19KX05
BIOS name and version	Dell 1.1.5 (05/04/2015)	Dell 1.1.5 (05/04/2015)
BIOS settings	Default w/logical processor disabled	Default w/logical processor disabled
Memory modules		
Total RAM in system (GB)	64	64
Vendor and model number	Hynix HMA42GR7MFR4N-TF	Hynix HMA42GR7MFR4N-TF
Type	PC4-2133	PC4-2133

Server	Edge Node/Name Node	Data Nodes
Speed (MHz)	2,133	2,133
Speed in the system currently running @ (MHz)	1,866	1,866
Timing/latency (tCL-tRCD-iRP-tRASmin)	15-15-15-33	15-15-15-33
Size (GB)	16	16
Number of RAM modules	4	4
Chip organization	Dual	Dual
Hard disks		
Vendor and Model Number	LITE-ON EBT-60N9S	LITE-ON EBT-60N9S
Number of disks in the system	2	2
Size (GB)	60	60
Buffer size (MB)	N/A	N/A
RPM	N/A	N/A
Type	SATA SSD	SATA SSD
Operating system		
Name	Red Hat® Enterprise Linux® 6.5	Red Hat Enterprise Linux 6.5
Build number	2.6.32-573.3.1.el6.x86_64	2.6.32-573.3.1.el6.x86_64
File system	ext4	ext4
Language	English	English
Network adapter 1		
Type	Integrated	Integrated
Vendor and model number	Broadcom® NetXtreme® II 10 Gb Ethernet BCM57810	Broadcom NetXtreme II 10 Gb Ethernet BCM57810
Storage controller 1		
Vendor and model number	Dell PERC S130	Dell PERC S130
Cache size	N/A	N/A
Driver	ahci 3.0	ahci 3.0
Firmware	1.18 (8/5/2015)	1.18 (8/5/2015)
Storage controller 2		
Vendor and model number	N/A	Dell PERC FD33xD
Cache size	N/A	2GB
Driver	N/A	06.902.01.00
Firmware	N/A	25.3.0.0016

Figure 5: System configuration information for the test systems.

Storage array	Dell PowerEdge FD332
Array	Dell PowerEdge FD332
Number of storage controllers	1
Number of drives	16
Disk vendor and model number	Seagate® ST300MM006
Disk size (GB)	300
Disk buffer size (MB)	64
Disk RPM	10K.6
Disk type	SAS HDD

Figure 6: Storage configuration information.

APPENDIX C – HOW WE TESTED

Installing the Dell | Cloudera® Apache® Hadoop Solution

We installed Cloudera Hadoop (CDH) version 5.4 onto our cluster by following the "Dell | Cloudera Apache Hadoop Solution Deployment Guide – Version 5.4" with some modifications. The following is a high-level summary of this process.

Configuring the networking

We used the integrated 10GbE pass-through module on the Dell PowerEdge FX2 to connect to a Dell PowerConnect™ S4810 10GbE switch. We used this switch for management and cluster traffic isolated by VLAN on the switch and the OS. The 10GbE pass-through module did not require any extra configuration.

Configuring the storage

Each of our Dell PowerEdge FX2 units included two Dell PowerEdge FD332 storage arrays. The FD332 can be placed in a single or dual configuration to present its storage to one or both hosts on its side of the array. We placed each of the four FD332 units in split dual mode, so that the storage was presented to all nodes equally (except for the Edge Node, which we did not give any external hard disk storage).

1. Log into the Dell PowerEdge FX2 CMC web GUI.
2. In the left-hand navigation pane, click the first storage slot.
3. Click the Setup tab.
4. Select the Split Dual Host radio button, and click Apply.
5. Repeat these steps for the three remaining storage trays.

Configuring the BIOS, firmware, and RAID settings on the hosts

We used the Dell PowerEdge FX2 CMC to update the firmware across the nodes. We also set all BIOS settings to defaults and then disabled logical processors (Intel Hyper-Threading).

1. Log into the Dell PowerEdge FX2 CMC web GUI.
2. Click Server Overview, and then click Update.
3. Check the checkboxes for the desired firmware to be updated, and enter the location of the update file (attainable from Dell Drivers and Downloads).
4. Click Update and allow the Lifecycle Controller to complete the process on each node.
5. Enter the BIOS Setup on each node and set the BIOS settings to defaults. Then, disable logical processors.

Installing the OS on the hosts

We installed Red Hat Enterprise Linux 6.5 using a kickstart file (shown in Appendix C). The kickstart file created our partitions and mount points automatically, as well as disabled SELinux and Iptables and configured our network settings. We performed these steps on each node.

1. Boot into a minimal RHEL Boot ISO and press Tab at the splash screen to enter boot options.
2. Enter the kickstart connection string and required options, and press Enter to install the OS.
3. When the OS is installed, register the system with Red Hat, run yum updates on each node, and reboot to fully update the OS.

Installing Cloudera Manager and distributing CDH to all nodes

We used Installation Path A in the Cloudera support documentation to guide our Hadoop installation. We chose to place Cloudera Manager on the Edge Node so that we could easily access it from our lab network.

1. On the Edge Node, use `wget` to download the latest `cloudera-manager-installer.bin`, located on `archive.cloudera.com`.
2. Run the installer and select all defaults.
3. Navigate to Cloudera Manager by pointing a web browser to `http://<Edge_Node_IP_address>:7180`.
4. Log into Cloudera Manager using the default credentials `admin/admin`.
5. Install the Cloudera Enterprise Data Hub Edition Trial with the following options:
    a. Enter each host's IP address.
    b. Leave the default repository options.
    c. Install the Oracle® Java® SE Development Kit (JDK).
    d. Do not check the single user mode checkbox.
    e. Enter the root password for host connectivity.
6. After the Host Inspector checks the cluster for correctness, choose the following Custom Services:
    a. HDFS
    b. YARN (MR2 Included)
7. Assign roles to the hosts using the information in Figure 7. We used the first node (nn01) in the first Dell PowerEdge FX2 to host the Edge Node and Name Node roles, and the remaining nodes (dn01-dn07) as Data Nodes.

Service	Role	Node(s)
HDFS		
	NameNode	nn01
	Secondary NameNode	dn01
	Balancer	nn01
	HttpFS	nn01
	NFS Gateway	nn01
	DataNode	dn[01-07]
Cloudera Management Service		
	Service Monitor	nn01
	Activity Monitor	nn01
	Host Monitor	nn01
	Reports Manager	nn01
	Event Server	nn01
	Alert Publisher	nn01
YARN (MR2 Included)		
	ResourceManager	nn01
	JobHistory Server	nn01
	NodeManager	dn[01-07]

Figure 7: Role assignments.

8. At the Database Setup screen, copy down the embedded database credentials and test the connection. If the connections are successful, proceed through the wizard to complete the Cloudera installation.

Tuning the Cloudera installation

We used a tuning guide from Cloudera to help choose parameters for optimal Hadoop performance. The configuration parameters that were changed are listed in Figure 8:

Parameter	New value
dfs.block.size	512 MB
mapreduce.map.cpu.vcores	1
mapreduce.reduce.cpu.vcores	1
mapreduce.map.java.opts	820 MB
mapreduce.reduce.java.opts	1,638 MB
mapreduce.map.memory.mb	1,024 MB
mapreduce.reduce.memory.mb	2,048 MB
mapreduce.job.reduces	56
yarn.nodemanager.resource.memory-mb	40 GiB
yarn.nodemanager.resource.cpu-vcores	24
yarn.scheduler.maimum-allocation-mb	40 GiB

Figure 8: YARN resource parameter adjustments.

APPENDIX D – RHEL KICKSTART INSTALLATION FILES

We used kickstart files to automate the Red Hat Enterprise Linux installation. Within the kickstart files, we included options to partition the disks, disable SELinux and the Linux firewall, and configure the networking. The kickstart files for the Edge/Name Node and the Data Nodes differ slightly as there was no external storage presented to the Edge/Name Node.

Kickstart file for Edge/Name Node

```
lang en_US
keyboard us
timezone America/New_York --isUtc
#platform x86, AMD64, or Intel EM64T
url --url=http://10.130.200.10/distro/rhel-6.5
#
zerombr
clearpart --initlabel --all
bootloader --location=mbr --driveorder=sdb --append="rhgb quiet crashkernel=auto"
#
part /boot/efi --fstype=ext4 --ondisk=sdb --size=1024
part /boot     --fstype=ext4 --ondisk=sdb --size=1024
part pv.01            --grow --ondisk=sdb --size=1
part pv.02            --grow --ondisk=sdc --size=1

volgroup vg.01 --pesize=4096 pv.01
logvol /     --fstype=ext4 --name=lv_root --vgname=vg.01 --grow --size=48000 --
maxsize=48000
logvol swap               --name=lv_swap --vgname=vg.01 --grow --size=3072   --
maxsize=3072
logvol /home --fstype=ext4 --name=lv_home --vgname=vg.01 --grow --size=1024   --
maxsize=1024
#logvol /var  --fstype=ext4 --name=lv_var  --vgname=vg.01 --grow --size=1      --
percent=100

volgroup vg.02 --pesize=4096 pv.02
logvol /var --fstype=ext4 --name=lv_var --vgname=vg.02 --grow --size=1 --percent=100

#
rootpw --iscrypted
$6$Tj/aOuRg.uWSN9pT$EDmC9Z26ZQylKVP7153tSBn5h96qMLxrKsGEhQ/BHIcWIi7vWg3o39.6Qjv9MhnmtfKT0
M5xcnLtlbUvHGNxT1
authconfig --passalgo=sha512 --useshadow
selinux --disabled
firewall --disabled
#
skipx
firstboot --disable
```

```
#
%post

## misc. configuration

for i in autofs cups ip6tables iptables mdmonitor netfs nfslock postfix rpcbind rpcgssd ;
do
  chkconfig $i off
done
cat >> /etc/rc.local <<EOF_RC
echo never > /sys/kernel/mm/redhat_transparent_hugepage/defrag
sysctl -w vm.swappiness=1
EOF_RC

## time configuration
chkconfig ntpd on
sed -i.orig -e 's|^server|##server|' -e 's|^restrict -6|#restrict -6|' /etc/ntp.conf
cat >> /etc/ntp.conf <<EOF_NTP
server 10.130.200.10    iburst
EOF_NTP

## resource limits for Hadoop uids
cat >> /etc/security/limits.conf <<EOF_LIMITS
hdfs    -       nofile  32768
mapred  -       nofile  32768
hbase   -       nofile  32768
hdfs    -       nproc   32768
mapred  -       nproc   32768
hbase   â       nproc   32768
EOF_LIMITS

# disable IPv6
echo "options ipv6 disable=1" > /etc/modprobe.d/ipv6.conf
echo "NETWORKING_IPV6=no"    >> /etc/sysconfig/network

## disable network manager
chkconfig NetworkManager off
for i in /etc/sysconfig/network-scripts/ifcfg-* ; do
 sed -i 's|NM_CONTROLLED=.*|NM_CONTROLLED=no|' $i
done

# misc network configuration
echo "GATEWAY=10.128.0.1"   >> /etc/sysconfig/network
echo "nameserver 10.41.0.10" > /etc/resolv.conf
cat >> /etc/hosts <<EOF_HOSTS
```

```
## management network

10.128.219.110  ad-nn01
10.128.219.111  ad-dn01
10.128.219.112  ad-dn02
10.128.219.113  ad-dn03

10.128.219.114  ad-dn04
10.128.219.115  ad-dn05
10.128.219.116  ad-dn06
10.128.219.117  ad-dn07

## cluster network

192.168.50.110 ad-nn01
192.168.50.111 ad-dn01
192.168.50.112 ad-dn02
192.168.50.113 ad-dn03

192.168.50.114 ad-dn04
192.168.50.115 ad-dn05
192.168.50.116 ad-dn06
192.168.50.117 ad-dn07
EOF_HOSTS

# create em1
cat > /etc/sysconfig/network-scripts/ifcfg-em1 <<EOF_EM1
DEVICE=em1
ONBOOT=yes
BOOTPROTO=none
USERCTL=no
NM_CONTROLLED=no
EOF_EM1

# create em1.128
cat > /etc/sysconfig/network-scripts/ifcfg-em1.128 <<EOF_EM1128
DEVICE=em1.128
VLAN=yes
ONBOOT=yes
BOOTPROTO=static
IPADDR=10.128.219.110
NETMASK=255.255.0.0
USERCTL=no
NM_CONTROLLED=no
```

```
EOF_EM1128

# create em1.215
cat > /etc/sysconfig/network-scripts/ifcfg-em1.215 <<EOF_EM1215
DEVICE=em1.215
VLAN=yes
ONBOOT=yes
BOOTPROTO=static
IPADDR=192.168.50.110
NETMASK=255.255.0.0
USERCTL=no
NM_CONTROLLED=no
EOF_EM1215

%end
%packages
@performance
@network-file-system-client
@large-systems
@base
%end
```

Kickstart file for Data Nodes

```
lang en_US
keyboard us
timezone America/New_York --isUtc
#platform x86, AMD64, or Intel EM64T
url --url=http://10.130.200.10/distro/rhel-6.5
#
zerombr
clearpart --initlabel --all
bootloader --location=mbr --driveorder=sdj --append="rhgb quiet crashkernel=auto"
#
part /boot/efi --fstype=ext4 --ondisk=sdj --size=1024
part /boot     --fstype=ext4 --ondisk=sdj --size=1024
part pv.01            --grow --ondisk=sdj --size=1
part pv.02            --grow --ondisk=sdk --size=1

volgroup vg.01 --pesize=4096 pv.01
logvol /     --fstype=ext4 --name=lv_root --vgname=vg.01 --grow --size=48000 --
maxsize=48000
logvol swap              --name=lv_swap --vgname=vg.01 --grow --size=3072   --
maxsize=3072
logvol /home --fstype=ext4 --name=lv_home --vgname=vg.01 --grow --size=1024   --
maxsize=1024
```

```
#logvol /var  --fstype=ext4 --name=lv_var  --vgname=vg.01 --grow --size=1      --
percent=100

volgroup vg.02 --pesize=4096 pv.02
logvol /var --fstype=ext4 --name=lv_var --vgname=vg.02 --grow --size=1 --percent=100
#
rootpw --iscrypted
$6$Tj/aOuRg.uWSN9pT$EDmC9Z26ZQylKVP7153tSBn5h96qMLxrKsGEhQ/BHIcWIi7vWg3o39.6Qjv9MhnmtfKT0
M5xcnLtlbUvHGNxT1
authconfig --passalgo=sha512 --useshadow
selinux --disabled
firewall --disabled
#
skipx
firstboot --disable
#
%post

## misc. configuration

for i in autofs cups ip6tables iptables mdmonitor netfs nfslock postfix rpcbind rpcgssd ;
do
  chkconfig $i off
done
cat >> /etc/rc.local <<EOF_RC
echo never > /sys/kernel/mm/redhat_transparent_hugepage/defrag
sysctl -w vm.swappiness=1
EOF_RC

## time configuration
chkconfig ntpd on
sed -i.orig -e 's|^server|##server|' -e 's|^restrict -6|#restrict -6|' /etc/ntp.conf
cat >> /etc/ntp.conf <<EOF_NTP
server 10.130.200.10    iburst
EOF_NTP

## resource limits for Hadoop uids
cat >> /etc/security/limits.conf <<EOF_LIMITS
hdfs    -       nofile  32768
mapred  -       nofile  32768
hbase   -       nofile  32768
hdfs    -       nproc   32768
mapred  -       nproc   32768
hbase   â       nproc   32768
EOF_LIMITS
```

```
# disable IPv6
echo "options ipv6 disable=1" > /etc/modprobe.d/ipv6.conf
echo "NETWORKING_IPV6=no"    >> /etc/sysconfig/network

## disable network manager
chkconfig NetworkManager off
for i in /etc/sysconfig/network-scripts/ifcfg-* ; do
 sed -i 's|NM_CONTROLLED=.*|NM_CONTROLLED=no|' $i
done

# misc network configuration
echo "GATEWAY=10.128.0.1"   >> /etc/sysconfig/network
echo "nameserver 10.41.0.10" > /etc/resolv.conf
cat >> /etc/hosts <<EOF_HOSTS

## management network

10.128.219.110  ad-nn01
10.128.219.111  ad-dn01
10.128.219.112  ad-dn02
10.128.219.113  ad-dn03

10.128.219.114  ad-dn04
10.128.219.115  ad-dn05
10.128.219.116  ad-dn06
10.128.219.117  ad-dn07

## cluster network

192.168.50.110 ad-nn01
192.168.50.111 ad-dn01
192.168.50.112 ad-dn02
192.168.50.113 ad-dn03

192.168.50.114 ad-dn04
192.168.50.115 ad-dn05
192.168.50.116 ad-dn06
192.168.50.117 ad-dn07
EOF_HOSTS

# create em1
cat > /etc/sysconfig/network-scripts/ifcfg-em1 <<EOF_EM1
DEVICE=em1
ONBOOT=yes
```

```
BOOTPROTO=none
USERCTL=no
NM_CONTROLLED=no
EOF_EM1

# create em1.128
cat > /etc/sysconfig/network-scripts/ifcfg-em1.128 <<EOF_EM1128
DEVICE=em1.128
VLAN=yes
ONBOOT=yes
BOOTPROTO=static
IPADDR=10.128.219.111
NETMASK=255.255.0.0
USERCTL=no
NM_CONTROLLED=no
EOF_EM1128

# create em1.215
cat > /etc/sysconfig/network-scripts/ifcfg-em1.215 <<EOF_EM1215
DEVICE=em1.215
VLAN=yes
ONBOOT=yes
BOOTPROTO=static
IPADDR=192.168.50.111
NETMASK=255.255.0.0
USERCTL=no
NM_CONTROLLED=no
EOF_EM1215

# HDFS disk configuration on data notes (tries to fail safe):
# create a run-once script in /etc/rc.local ; the contents of this script
# will run only if the file /etc/sysconfig/local-runonce exists

if [ "yes" = "yes" ]; then
 touch /etc/sysconfig/local-runonce
 cat >> /etc/rc.local <<'EOF_RUNONCE'

### code to be run once after the OS install

if [ -f /etc/sysconfig/local-runonce ] ; then
 # create partitions
 for i in {a..h} ; do
  dv=/dev/sd$i
  if [ -b "$dv" ]; then
   parted -s "$dv" mklabel gpt
```

```
   parted -s "$dv" mkpart primary "1 -1"
  fi
  done
  sync; sleep 10; sync
  # create file systems in parallel
  for i in {a..l} ; do
   dv=/dev/sd${i}1
   if [ -b "$dv" ]; then
    mkfs.ext4 "${dv}" &
   fi
  done
  wait
  # update fstab and create mount points
  for i in {a..l} ; do
   dv=/dev/sd${i}1
   if [ -b "$dv" ]; then
    mkdir -p "/data/$i"
    uuidd=$(blkid "$dv" | sed 's/.*\(UUID="[^"]*"\).*/\1/')
    echo "$uuidd /data/$i ext4 defaults,noatime,nodiratime 0 0" >> /etc/fstab
   fi
  done
  rm -f /etc/sysconfig/local-runonce
  mount -a
 fi

EOF_RUNONCE
fi

%end
%packages
@performance
@network-file-system-client
@large-systems
@base
%end
```

ABOUT PRINCIPLED TECHNOLOGIES

Principled Technologies, Inc.
1007 Slater Road, Suite 300
Durham, NC, 27703
www.principledtechnologies.com

We provide industry-leading technology assessment and fact-based marketing services. We bring to every assignment extensive experience with and expertise in all aspects of technology testing and analysis, from researching new technologies, to developing new methodologies, to testing with existing and new tools.

When the assessment is complete, we know how to present the results to a broad range of target audiences. We provide our clients with the materials they need, from market-focused data to use in their own collateral to custom sales aids, such as test reports, performance assessments, and white papers. Every document reflects the results of our trusted independent analysis.

We provide customized services that focus on our clients' individual requirements. Whether the technology involves hardware, software, Web sites, or services, we offer the experience, expertise, and tools to help our clients assess how it will fare against its competition, its performance, its market readiness, and its quality and reliability.

Our founders, Mark L. Van Name and Bill Catchings, have worked together in technology assessment for over 20 years. As journalists, they published over a thousand articles on a wide array of technology subjects. They created and led the Ziff-Davis Benchmark Operation, which developed such industry-standard benchmarks as Ziff Davis Media's Winstone and WebBench. They founded and led eTesting Labs, and after the acquisition of that company by Lionbridge Technologies were the head and CTO of VeriTest.