



Running Oracle Databases on Amazon Web Services with InfoScale Enterprise

Best Practices

Last updated: 12/9/2019

VERITAS[™]

The truth in information.

Abstract

Veritas InfoScale™ Enterprise provides a complete suite of storage and clustering capabilities for organizations that want to deploy and run their Oracle database applications in Amazon Web Services (AWS). InfoScale Enterprise builds on the strengths of the AWS platform and extends it to provide scale-out performance, resiliency, and high availability (HA) that protects your databases against the failure of Availability Zones (AZs). The Flexible Storage Sharing (FSS) feature of InfoScale lets you overcome the limitations of the shared-nothing storage in AWS, by providing shared storage and a single namespace across servers and even across AZs. The SmartIO feature of InfoScale, which is a read caching solution, leverages Instance Store volumes to enhance application performance beyond the IOPS limits of EC2 or EBS volumes.

This document describes configuration considerations for customers who choose to run self-managed Oracle databases on Amazon Elastic Compute Cloud (Amazon EC2) instances with InfoScale Enterprise.

Architecting for performance

Optimizing every component of the infrastructure enhances performance and reliability. The following sections provide some recommendations for choosing the instance types and the database storage in an Oracle database implementation on AWS with InfoScale Enterprise.

AWS provides many Amazon EC2 instance types, and you can choose the instance type that best fits your workload. The general recommendation is to choose instance types that have a good balance of memory and CPU. Veritas highly recommends that you use only those instance types that are optimized for Amazon Elastic Block Store (Amazon EBS) volumes. Instances of these type are commonly referred to as EBS-optimized EC2 instances. For better database performance, you should choose instances that provide higher network performance. For more information, refer to the following article on EBS-optimized instances:

<https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/EBSOptimized.html>

For high and consistent IOPS and database performance, Veritas recommends using the General Purpose (GP2) or the Provisioned IOPS (IO1) volume types. For the latest limits of IOPS per volume for both these volume types, refer to the AWS documentation:

<https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/EBSVolumeTypes.html>

The EC2 instances and the EBS volumes have IOPS and throughput limits individually, based on their respective types. So, while choosing an instance type and EBS volume type and size, make sure to estimate the amount of IOPS that are consistently needed for your database. Accordingly, choose the appropriate EC2 instance and allocate enough EBS storage to obtain the estimated IOPS. For information on estimation methods that you can use to determine the IOPS needs of your Oracle database, refer to the *Determining the IOPS Needs for Oracle Database on AWS* whitepaper:

<https://d1.awsstatic.com/whitepapers/determining-iops-needs-for-oracle-database-on-aws.pdf>

If you choose an instance type with local SSD storage, you can also boost the database performance by using the SmartIO feature of InfoScale Enterprise.

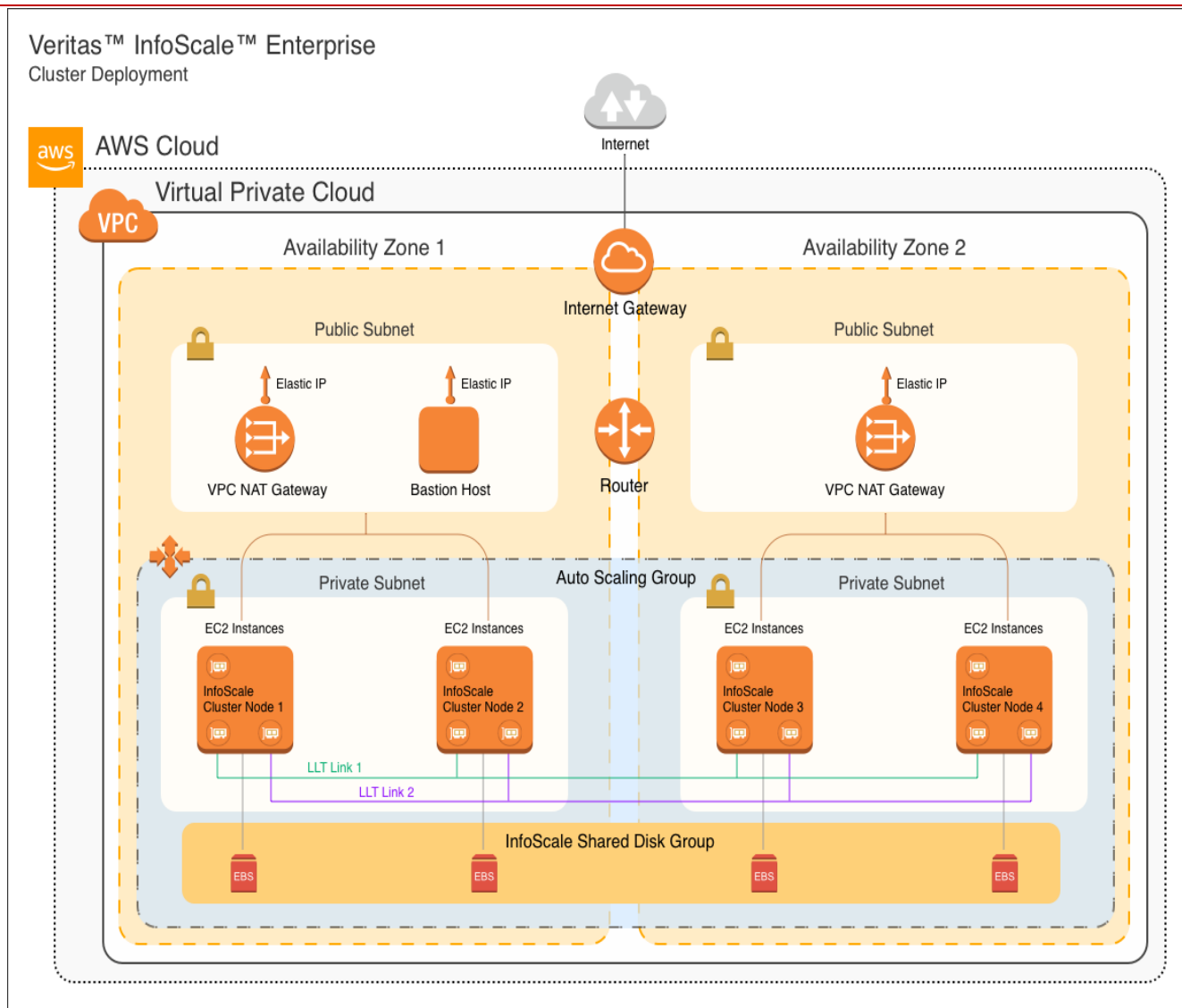
Using Infoscale Enterprise

As described earlier in this document, understand the IOPS requirement for your workload and then choose the appropriate EC2 instance type and the EBS volume type:

- Choose an EC2 instance with better network performance.
- Enable enhanced networking for your EC2 instances to further enhance the network performance.
- Choose EBS-optimized EC2 instances for better storage performance.
- Based on the required IOPS value, provision multiple EBS volumes (GP2 and/or IO1) to get the benefits of striping.

Sample InfoScale Enterprise configuration in AWS

The following graphic depicts a sample four-node InfoScale cluster in AWS:



InfoScale configuration—additional recommendations

Veritas additionally recommends the following settings for your Oracle database configuration in AWS:

- Configure an InfoScale cluster by configuring LLT over UDP. Set 9000 MTU in `/etc/llttab`.
- Configure InfoScale volumes with the mirrored-stripe or the striped-mirror layout for storage HA and better performance by striping across multiple EBS volumes.
- Use SmartIO for read caching, by leveraging Instance Store volumes (NVMe SSDs), for better application performance.
- Set the VxFS tunables `max_diskq` and `initial_extent_size` to 104857600 and 1M respectively.

Reference architecture—Infoscale Enterprise

This section describes the reference architectures of the configurations tested at Veritas.

Note: If you are using InfoScale Enterprise 7.4, ensure that the following packages are installed on the system:

- VRTSglim-7.4.0.0001
- VRTSllt-7.4.0.1402
- VRTSvxvm-7.4.0.1505
- VRTSvxfs-7.4.0.1200

Two-node InfoScale cluster across AZs without SmartIO

A two-node InfoScale cluster is configured across AZs, but the SmartIO feature is not used.

AWS configuration

- EC2 instance type: m4.4xlarge
 - 16 vCPUs, 64GB
 - EBS-optimized
 - Network performance—High
 - 16000 maxIOPS and 250 MB/s maxThroughput
- EBS volume type: 4 GP2 volumes
 - 1333G per node
 - 16000 IOPS per node

InfoScale configuration

- A two-node InfoScale cluster is configured, where the nodes exist in different AZs.
- A data volume is mirrored across the two nodes for resiliency and striped across multiple disks for better performance.
- The tunable values are set as follows:
 - In case of a specific volume or file system, the following parameters are set in `/etc/vx/tunefstab`:
`/dev/vx/dsk/<diskgroup>/<volume_name> max_diskq=104857600,initial_extent_size=1M`

- o In case of all volumes or file systems, the following parameters are set in `/etc/vx/tunefstab`:
`system_default max_diskq=104857600,initial_extent_size=1M`

For details, refer to the `vxtunefs` man page.

- While configuring the InfoScale cluster, LLT is configured over UDP.

The MTU size (9000) is set in the `/etc/llttab` file.

```
link eth0 udp - udp 50000 9000 <IP> -
```

- An InfoScale FSS cluster is configured in AWS:

```
(Linux) (ip-172-31-56-215) ~{1} /opt/VRTS/bin//vxclustadm nidmap
```

Name	CVM Nid	CM Nid	State
ip-172-31-24-198	0	1	Joined: Slave
ip-172-31-56-215	1	0	Joined: Master

For details on configuring an FSS cluster, refer to the Veritas InfoScale documentation.

- An FSS diskgroup is created by using four EBS volumes from each node:

```
(Linux) (ip-172-31-56-215) ~{1} vxdbg -s -o fss init datadg 172-31-24-198_xen-vd0_2 172-31-24-198_xen-vd0_3 172-31-24-198_xen-vd0_4 172-31-24-198_xen-vd0_5 172-31-56-215_xen-vd0_1 172-31-56-215_xen-vd0_2 172-31-56-215_xen-vd0_3 172-31-56-215_xen-vd0_4
```

- A mirror-stripe volume is created with two mirrors and four stripes:

```
(Linux) (ip-172-31-56-215) ~{1} vxassist -g datadg make datavol 1t layout=mirror-stripe ncol=4
```

- On verification, the configuration appears as follows:

```
(Linux) (ip-172-31-56-215) ~{1} vxdisk list
```

DEVICE	TYPE	DISK	GROUP	STATUS
xvda	auto:none	-	-	online invalid
172-31-24-198_xen-vd0_2	auto:cdsdisk	172-31-24-198_xen-vd0_2	datadg	online
shared remote				
172-31-24-198_xen-vd0_3	auto:cdsdisk	172-31-24-198_xen-vd0_3	datadg	online
shared remote				
172-31-24-198_xen-vd0_4	auto:cdsdisk	172-31-24-198_xen-vd0_4	datadg	online
shared remote				
172-31-24-198_xen-vd0_5	auto:cdsdisk	172-31-24-198_xen-vd0_5	datadg	online
shared remote				
172-31-56-215_xen-vd0_1	auto:cdsdisk	172-31-56-215_xen-vd0_1	datadg	online
exported shared				
172-31-56-215_xen-vd0_2	auto:cdsdisk	172-31-56-215_xen-vd0_2	datadg	online
exported shared				
172-31-56-215_xen-vd0_3	auto:cdsdisk	172-31-56-215_xen-vd0_3	datadg	online
exported shared				

```
172-31-56-215_xen-vd0_4 auto:cdsdisk 172-31-56-215_xen-vd0_4 datadg online
exported shared
```

```
(Linux) (ip-172-31-56-215) ~{3} vxprint -uh
```

```
Disk group: datadg
```

TY	NAME	ASSOC	KSTATE	LENGTH	PLOFFS	STATE	TUTILO	PUTILO
dg	datadg	datadg	-	-	-	-	-	-
dm	172-31-24-198_xen-vd0_2	172-31-24-198_xen-vd0_2	-	1.30t	-	REMOTE	-	-
dm	172-31-24-198_xen-vd0_3	172-31-24-198_xen-vd0_3	-	1.30t	-	REMOTE	-	-
dm	172-31-24-198_xen-vd0_4	172-31-24-198_xen-vd0_4	-	1.30t	-	REMOTE	-	-
dm	172-31-24-198_xen-vd0_5	172-31-24-198_xen-vd0_5	-	1.30t	-	REMOTE	-	-
dm	172-31-56-215_xen-vd0_1	172-31-56-215_xen-vd0_1	-	1.30t	-	-	-	-
dm	172-31-56-215_xen-vd0_2	172-31-56-215_xen-vd0_2	-	1.30t	-	-	-	-
dm	172-31-56-215_xen-vd0_3	172-31-56-215_xen-vd0_3	-	1.30t	-	-	-	-
dm	172-31-56-215_xen-vd0_4	172-31-56-215_xen-vd0_4	-	1.30t	-	-	-	-
v	datavol	fsgen	ENABLED	1.00t	-	ACTIVE	-	-
pl	datavol-01	datavol	ENABLED	1.00t	-	ACTIVE	-	-
sd	172-31-24-198_xen-vd0_2-01	datavol-01	ENABLED	256.00g	0.00	-	-	-
sd	172-31-24-198_xen-vd0_3-01	datavol-01	ENABLED	256.00g	0.00	-	-	-
sd	172-31-24-198_xen-vd0_4-01	datavol-01	ENABLED	256.00g	0.00	-	-	-
sd	172-31-24-198_xen-vd0_5-01	datavol-01	ENABLED	256.00g	0.00	-	-	-
pl	datavol-02	datavol	ENABLED	1.00t	-	ACTIVE	-	-
sd	172-31-56-215_xen-vd0_1-01	datavol-02	ENABLED	256.00g	0.00	-	-	-
sd	172-31-56-215_xen-vd0_2-01	datavol-02	ENABLED	256.00g	0.00	-	-	-
sd	172-31-56-215_xen-vd0_3-01	datavol-02	ENABLED	256.00g	0.00	-	-	-
sd	172-31-56-215_xen-vd0_4-01	datavol-02	ENABLED	256.00g	0.00	-	-	-
dc	datavol_dco	datavol	-	-	-	-	-	-
v	datavol_dcl	gen	ENABLED	70.06m	-	ACTIVE	-	-
pl	datavol_dcl-01	datavol_dcl	ENABLED	70.06m	-	ACTIVE	-	-
sd	172-31-24-198_xen-vd0_2-02	datavol_dcl-01	ENABLED	70.06m	0.00	-	-	-
pl	datavol_dcl-02	datavol_dcl	ENABLED	70.06m	-	ACTIVE	-	-
sd	172-31-56-215_xen-vd0_1-02	datavol_dcl-02	ENABLED	70.06m	0.00	-	-	-

```
(Linux) (ip-172-31-56-215) ~{4} mount | grep datavol /dev/vx/dsk/datadg/datavol on
/mnt/datavol type vxfs (rw,seclabel,delaylog,largefiles,cio,ioerror=mwdisable)
```

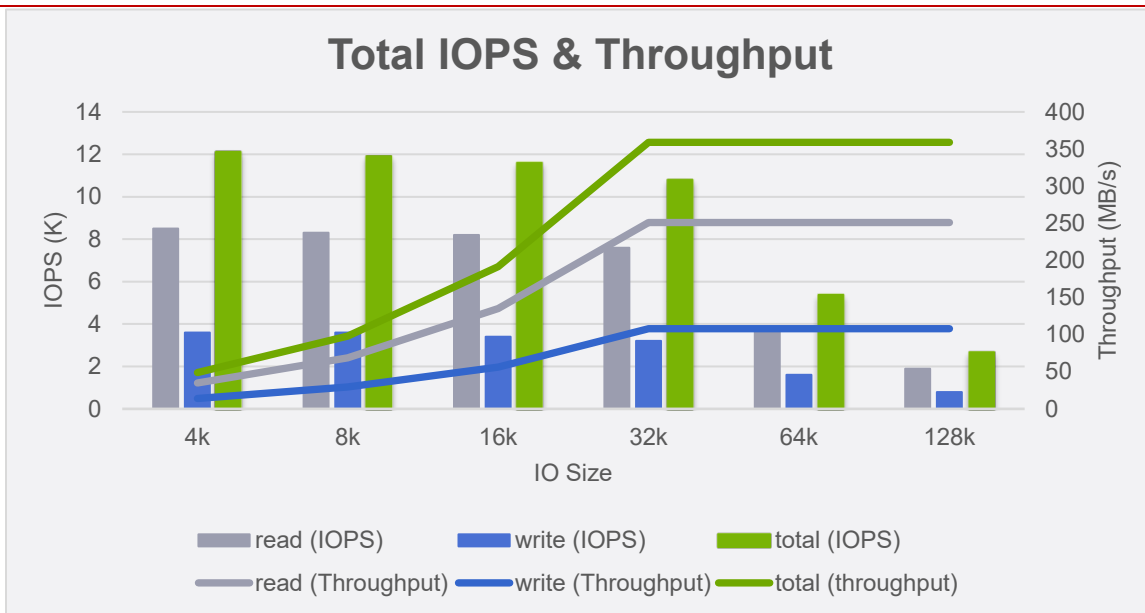
Performance characterization using FIO workload

Workload details:

- Read/write workload: 70/30
- Parameters:


```
[iotest]
ioengine=libaio
rw=randrw
group_reporting=1
rwmixwrite=30
size=1000m
numjobs=16
direct=1
sync=0
blocksize=<bsize>k
time_based
iodepth=8
directory=/mnt/datavol/
```

Results:



With InfoScale Enterprise, a maximum of 12000 IOPS can be achieved by using m4.4xlarge instances and GP2 storage, with a 70/30 read/write workload. IOPS get declined at higher IO sizes, and the total throughput of the application is capped at 359MB/s.

Cost of the deployment = \$2458/month (EC2 and EBS volumes for 2 nodes)

\$\$/IOPS (for 16K IO) = \$0.2

Two-node InfoScale cluster across AZs with SmartIO

A two-node InfoScale cluster is configured across AZs, and SmartIO is configured.

AWS configuration

- EC2 instance type: i3.2xlarge
 - 16 vCPUs, 61GB
 - EBS-optimized, 1.9T NVMe SSD
 - Network performance – Up to 10Gbps
 - 12000 maxIOPS
- EBS volume type: 4 GP2 volumes
 - 1000G per node
 - 12000 IOPS per node

InfoScale configuration

- A two-node InfoScale cluster is configured, where the nodes exist in different AZs.
- A data volume is mirrored across the two nodes for resiliency and striped across multiple disks for better performance.
- The tunable values are set as follows:

- In case of a specific volume or file system, the following parameters are set in `/etc/vx/tunefstab`:
`/dev/vx/dsk/<diskgroup>/<volume_name> max_diskq=104857600,initial_extent_size=1M`
- In case of all volumes or file systems, the following parameters are set in `/etc/vx/tunefstab`:
`system_default max_diskq=104857600,initial_extent_size=1M`

For details, refer to the `vxtunefs` man page.

- While configuring the InfoScale cluster, LLT is configured over UDP.

The MTU size (9000) is set in the `/etc/llttab` file.

```
link eth0 udp - udp 50000 9000 <IP> -
```

- The `sfcache` is created using a 1.9T NVMe SSD Instance Store volume and it is used as a read cache for better application performance. As the cache hit ratio increases, SmartIO helps to better utilize the IOPS that are provided by EBS volume for write IOs, by serving the maximum read IOs from the cache.

```
(Linux) (ip-172-31-56-215) ~{1} /opt/VRTS/bin//vxclustadm nidmap
```

Name	CVM Nid	CM Nid	State
ip-172-31-24-198	0	1	Joined: Slave
ip-172-31-56-215	1	0	Joined: Master

For details on configuring an FSS cluster and SmartIO, refer to the Veritas InfoScale documentation.

- An FSS diskgroup is created by using four EBS volumes from each node:

```
(Linux) (ip-172-31-56-215) ~{1} vxdg -s -o fss init datadg 172-31-24-198_xen-vd0_2 172-31-24-198_xen-vd0_3 172-31-24-198_xen-vd0_4 172-31-24-198_xen-vd0_5 172-31-56-215_xen-vd0_1 172-31-56-215_xen-vd0_2 172-31-56-215_xen-vd0_3 172-31-56-215_xen-vd0_4
```

- A mirror-stripe volume is created with two mirrors and four stripes:

```
(Linux) (ip-172-31-56-215) ~{1} vxassist -g datadg make datavol 1t layout=mirror-stripe
ncol=4
```

- Create the sfcache volume using the Instance Store volume.

```
(Linux) (ip-172-31-56-215) ~{20} vxdg init sfcachedg 172-31-56-215_nvme4_0
```

```
(Linux) (ip-172-31-56-215) ~{20} vxassist -g sfcachedg make sfcachevol 1t
```

```
(Linux) (ip-172-31-56-215) ~{20} vxdisk list | grep sfcachedg
```

```
172-31-56-215_nvme4_0 auto:cdsdisk 172-31-56-215_nvme4_0 sfcachedg online
ssdtrim
```

```
(Linux) (ip-172-31-56-215) ~{18} vxprint -g sfcachedg
```

TY	NAME	ASSOC	KSTATE	LENGTH	PLOFFS	STATE	TUTIL0	PUTILO
dg	sfcachedg	sfcachedg	-	-	-	-	-	-

```
dm 172-31-56-215_nvme4_0 172-31-56-215_nvme4_0 - 3710871664 - - -
```

```
v sfcachevol fsgen ENABLED 2147483648 - ACTIVE -
```

```
pl sfcachevol-01 sfcachevol ENABLED 2147483648 - ACTIVE -
```

```
sd 172-31-56-215_nvme4_0-01 sfcachevol-01 ENABLED 2147483648 0 - -
```

```
(Linux) (ip-172-31-56-215) ~{20} sfcache create -t VxFS sfcachedg/sfcachevol
```

```
(Linux) (ip-172-31-56-215) ~{16} sfcache list
```

NAME DEVICE	TYPE	SIZE	ASSOC-TYPE	STATE	FSTYPE
sfcachedg/sfcachevol 172-31-56-215_nvme4_0	VxFS	1.00t	AUTO	ONLINE	default

- On verification, the configuration appears as follows:

```
(Linux) (ip-172-31-56-215) ~{19} vxdisk list
```

DEVICE	TYPE	DISK	GROUP	STATUS
xvda	auto:none	-	-	online invalid
172-31-24-198_xen-vd0_2 shared remote	auto:cdsdisk	172-31-24-198_xen-vd0_2	datadg	online
172-31-24-198_xen-vd0_3 shared remote	auto:cdsdisk	172-31-24-198_xen-vd0_3	datadg	online
172-31-24-198_xen-vd0_4 shared remote	auto:cdsdisk	172-31-24-198_xen-vd0_4	datadg	online

```

172-31-24-198_xen-vd0_5 auto:cdsdisk 172-31-24-198_xen-vd0_5 datadg online
shared remote
172-31-56-215_nvme4_0 auto:cdsdisk 172-31-56-215_nvme4_0 sfcachedg online
ssdtrim
172-31-56-215_xen-vd0_1 auto:cdsdisk 172-31-56-215_xen-vd0_1 datadg online
exported shared
172-31-56-215_xen-vd0_2 auto:cdsdisk 172-31-56-215_xen-vd0_2 datadg online
exported shared
172-31-56-215_xen-vd0_3 auto:cdsdisk 172-31-56-215_xen-vd0_3 datadg online
exported shared
172-31-56-215_xen-vd0_4 auto:cdsdisk 172-31-56-215_xen-vd0_4 datadg online
exported shared
(Linux) (ip-172-31-56-215) ~{20} vxdisk list | grep sfcachedg
172-31-56-215_nvme4_0 auto:cdsdisk 172-31-56-215_nvme4_0 sfcachedg online
ssdtrim

```

```

(Linux) (ip-172-31-56-215) ~{3} vxprint -uh
Disk group: datadg

```

TY	NAME	ASSOC	KSTATE	LENGTH	PLOFFS	STATE	TUTIL0	PUTILO
dg	datadg	datadg	-	-	-	-	-	-
dm	172-31-24-198_xen-vd0_2	172-31-24-198_xen-vd0_2	-	1.30t	-	REMOTE	-	-
dm	172-31-24-198_xen-vd0_3	172-31-24-198_xen-vd0_3	-	1.30t	-	REMOTE	-	-
dm	172-31-24-198_xen-vd0_4	172-31-24-198_xen-vd0_4	-	1.30t	-	REMOTE	-	-
dm	172-31-24-198_xen-vd0_5	172-31-24-198_xen-vd0_5	-	1.30t	-	REMOTE	-	-
dm	172-31-56-215_xen-vd0_1	172-31-56-215_xen-vd0_1	-	1.30t	-	-	-	-
dm	172-31-56-215_xen-vd0_2	172-31-56-215_xen-vd0_2	-	1.30t	-	-	-	-
dm	172-31-56-215_xen-vd0_3	172-31-56-215_xen-vd0_3	-	1.30t	-	-	-	-
dm	172-31-56-215_xen-vd0_4	172-31-56-215_xen-vd0_4	-	1.30t	-	-	-	-
v	datavol	fsgen	ENABLED	1.00t	-	ACTIVE	-	-
pl	datavol-01	datavol	ENABLED	1.00t	-	ACTIVE	-	-
sd	172-31-24-198_xen-vd0_2-01	datavol-01	ENABLED	256.00g	0.00	-	-	-
sd	172-31-24-198_xen-vd0_3-01	datavol-01	ENABLED	256.00g	0.00	-	-	-
sd	172-31-24-198_xen-vd0_4-01	datavol-01	ENABLED	256.00g	0.00	-	-	-
sd	172-31-24-198_xen-vd0_5-01	datavol-01	ENABLED	256.00g	0.00	-	-	-
pl	datavol-02	datavol	ENABLED	1.00t	-	ACTIVE	-	-
sd	172-31-56-215_xen-vd0_1-01	datavol-02	ENABLED	256.00g	0.00	-	-	-
sd	172-31-56-215_xen-vd0_2-01	datavol-02	ENABLED	256.00g	0.00	-	-	-

```

sd 172-31-56-215_xen-vd0_3-01 datavol-02 ENABLED 256.00g 0.00 - - -
sd 172-31-56-215_xen-vd0_4-01 datavol-02 ENABLED 256.00g 0.00 - - -
dc datavol_dco datavol - - - - -
v datavol_dcl gen ENABLED 70.06m - ACTIVE - -
pl datavol_dcl-01 datavol_dcl ENABLED 70.06m - ACTIVE - -
sd 172-31-24-198_xen-vd0_2-02 datavol_dcl-01 ENABLED 70.06m 0.00 - - -
pl datavol_dcl-02 datavol_dcl ENABLED 70.06m - ACTIVE - -
sd 172-31-56-215_xen-vd0_1-02 datavol_dcl-02 ENABLED 70.06m 0.00 - - -

```

```

(Linux) (ip-172-31-56-215) ~{4} mount | grep datavol /dev/vx/dsk/datadg/datavol on
/mnt/datavol type vxfs (rw,seclabel,delaylog,largefiles,cio,ioerror=mwdisable)

```

Performance characterization using FIO workload

Workload details:

- Read/write workload: 70/30

- Parameters:

```

[iotest]
ioengine=libaio
rw=randrw
group_reporting=1
rwmixwrite=30
size=1000m
numjobs=16
direct=1
sync=0
blocksize=<bsize>k
time_based
iodepth=8
directory=/mnt/datavol/

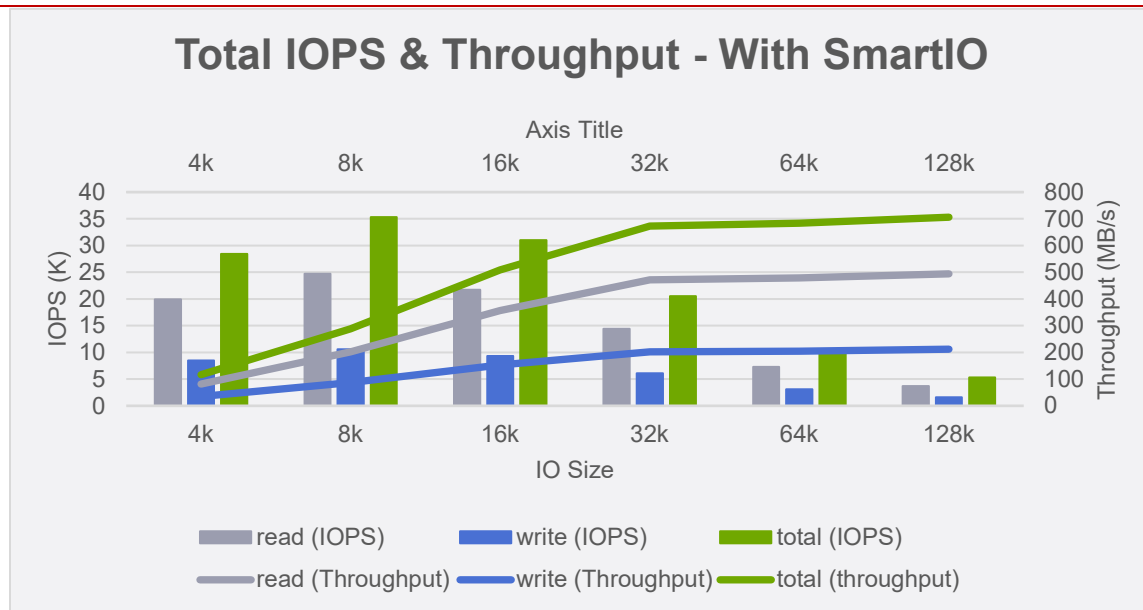
```

Results:

You can use the following command to view the `sfcache` statistics:

```
# sfcache stat -i < interval>
```

The following graphic depicts the statistics with the cache hit ratio beyond 96%:



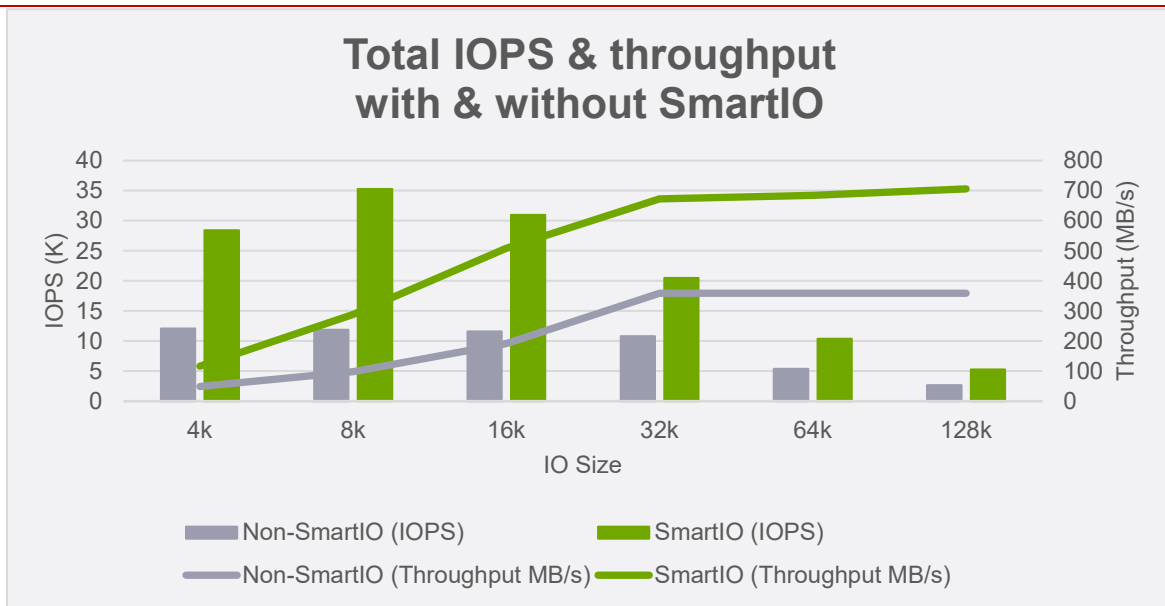
When SmartIO is enabled with InfoScale Enterprise, a maximum of 30000+ IOPS with a 16K IO size can be achieved by using i3.xlarge instances and GP2 storage, with a 70/30 read/write workload.

Cost of the deployment = \$1881.60/month (EC2 and EBS volumes for 2 nodes)

\$\$/IOPS (for 16K IO) = \$0.06

Comparison between the reference architectures with and without SmartIO

The following graphic depicts the comparative IOPS and throughput values of the reference architectures with and without the SmartIO feature.



When SmartIO is enabled as a read caching solution, by leveraging instance store volume, an application performance can be at least doubled in lesser cost (\$0.2/IOPS vs \$0.06/IOPS).

SmartIO brings great value in AWS, where there are IOPS and throughput limit at volume as well as EC level for EBS volumes.

References

Oracle whitepapers for AWS:

<https://aws.amazon.com/oracle/resources/>

Best practices for running an Oracle database in AWS:

<https://docs.aws.amazon.com/aws-technical-content/latest/oracle-database-aws-best-practices/oracle-database-aws-best-practices.pdf>

Veritas InfoScale Enterprise documentation:

[https://www.veritas.com/content/support/en_US/search-results.html?q=*&fq=\(document_type%3A%22document%22\)%20AND%20\(product_name%3A%22InfoScale%20%26%20Storage%20Foundation%22\)%20AND%20\(locale%3A%22en_US%22\)%20AND%20version%3A\(%227.4.1%22\)&docRepo=true&requestedRecords=20](https://www.veritas.com/content/support/en_US/search-results.html?q=*&fq=(document_type%3A%22document%22)%20AND%20(product_name%3A%22InfoScale%20%26%20Storage%20Foundation%22)%20AND%20(locale%3A%22en_US%22)%20AND%20version%3A(%227.4.1%22)&docRepo=true&requestedRecords=20)

ABOUT VERITAS TECHNOLOGIES LLC

Veritas Technologies empowers businesses of all sizes to discover the truth in information—their most important digital asset. Using the Veritas platform, customers can accelerate their digital transformation and solve pressing IT and business challenges including multi-cloud data management, data protection, storage optimization, compliance readiness and workload portability—with no cloud vendor lock-in. Eighty-six percent of Fortune 500 companies rely on Veritas today to reveal data insights that drive competitive advantage. Learn more at veritas.com or follow us on Twitter at [@veritastechllc](https://twitter.com/veritastechllc).

Veritas World Headquarters
2625 Augustine Drive
Santa Clara, CA 95054
+1 (866) 837 4827
www.veritas.com

For specific country offices
and contact numbers,
please visit our website.

