OpenZFS

= ZFS INTENT LOG & ZIL (WRITE/READ CACHE) =

ZFS L2ARC / Read Cache:

In the world of storage, caching can play a big role in improving performance. **OpenZFS** offers some very powerful tools to improve read & write performance.

To improve read performance, ZFS utilizes system memory as an Adaptive Replacement Cache (ARC), which stores your file system's most frequently and recently used data in your system memory. You can then add a Level 2 Adaptive Replacement Cache (L2ARC) to extend the ARC to a dedicated disk (or disks) to dramatically improve read speeds, effectively giving the user **all-flash performance**.

We have to use high-performance SSD/NVME as L2ARC;

Suppose, here **sda15** is our SSD, let's add that SSD as L2ARC,

```
zpool add -f vol1 cache /dev/sda15
```

In case if we want to use two SSD (recommended for reliability/high-availability)

```
zpool add -f vol1 cache mirror /dev/sda15 /dev/sda16
```

```
zpool status
```

```
      cache
        sda15     ONLINE       0     0     0
```

Note: if L2ARC disk fail, pool may freeze for some time but data will be OK!

ZFS ZIL / Write Cache!:

OpenZFS includes something called the ZFS Intent Log (ZIL). The ZIL can be set up on a dedicated disk called a Separate Intent Log (SLOG) similar to the L2ARC, but it is not simply a performance boosting technology.

Many people think of the ZFS Intent Log like they would a write cache. This causes some confusion in understanding how it works and how to best configure it. First of all, the ZIL is more accurately referred to as a "log" whose main purpose is actually for data integrity.

Please read this article http://www.freenas.org/blog/zfs-zil-and-slog-demystified/

We have to use high-performance SSD/NVME as ZIL, Suppose, here **sda16** is our SSD, let's add that SSD as ZIL,

```
zpool add -f vol1 log /dev/sda16
```

In case if we want to use two SSD (recommended for reliability/high-availability)

```
zpool add -f vol1 log mirror /dev/sda16 /dev/sda17
```

```
zpool status
```

```
      logs
        sda16     ONLINE       0     0     0
```

Note: if ZIL disk fail, pool may freeze for some time but data will be OK!

=3=

So our final setup is,
zpool status >

```
root@group1-node3:~# zpool status
  pool: vol1
 state: ONLINE
  scan: resilvered 112K in 0h0m with 0 errors on Tue Jul 31 12:27:36 2018
config:

        NAME        STATE     READ WRITE CKSUM
        vol1        ONLINE       0     0     0
          mirror-0  ONLINE       0     0     0
            sda10   ONLINE       0     0     0
            sda11   ONLINE       0     0     0
          mirror-1  ONLINE       0     0     0
            sda14   ONLINE       0     0     0
            sda13   ONLINE       0     0     0
        logs
          sda16     ONLINE       0     0     0
        cache
          sda15     ONLINE       0     0     0
```

As we have to practice LXC on this, so we will remove cache & log device, as here in lab it is not SSD and likely we will get poorer performance.

```
zpool remove vol1 /dev/sda15
zpool remove vol1 /dev/sda16
```

Task: Add two disks (here sda15, sda16) in mirror mode to increase capacity of our existing zfs pool.

Note: It is best-practice for zfs for Linux is to use disk-id instead of /dev/sdx

ll /dev/disk/by-id/ |grep sd |more

```
lrwxrwxrwx 1 root root  11 Jul 31 10:56 ata-ST9500420AS_5VJCLXJQ-part10 -> ../../sda10
lrwxrwxrwx 1 root root  11 Jul 31 10:56 ata-ST9500420AS_5VJCLXJQ-part11 -> ../../sda11
lrwxrwxrwx 1 root root  11 Jul 31 11:41 ata-ST9500420AS_5VJCLXJQ-part12 -> ../../sda12
lrwxrwxrwx 1 root root  11 Jul 31 11:41 ata-ST9500420AS_5VJCLXJQ-part13 -> ../../sda13
lrwxrwxrwx 1 root root  11 Jul 31 12:27 ata-ST9500420AS_5VJCLXJQ-part14 -> ../../sda14
lrwxrwxrwx 1 root root  11 Jul 31 12:46 ata-ST9500420AS_5VJCLXJQ-part15 -> ../../sda15
lrwxrwxrwx 1 root root  11 Jul 31 13:02 ata-ST9500420AS_5VJCLXJQ-part16 -> ../../sda16
```

In that case pools creation command will be

```
zpool create -o ashift=12 -f vol1 mirror /dev/sda10 /dev/sda11 > its equivalent command as follows,
```

```
zpool create -o ashift=12 -f vol1 mirror \
ata-ST9500420AS_5VJCLXJQ-part10 \
ata-ST9500420AS_5VJCLXJQ-part11
```

To see pool io performance,

```
zpool iostat -v
zpool iostat -v 2
```

Useful reading: https://pthree.org/2012/04/17/install-zfs-on-debian-gnulinux/