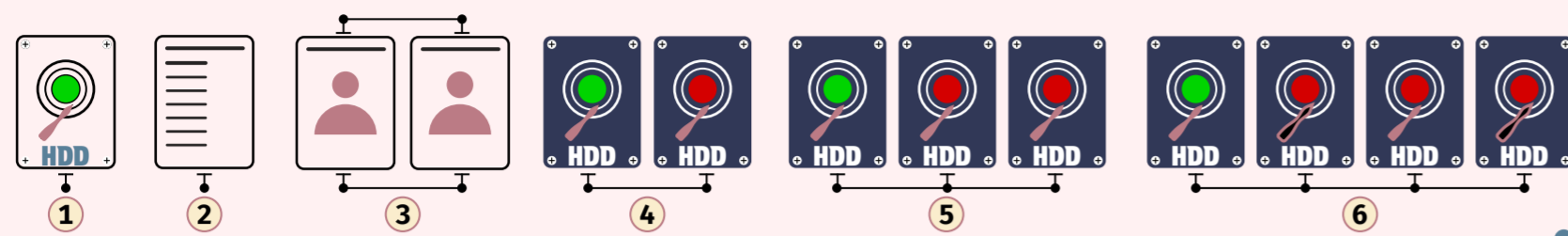


Zpool is a collection of vdevs, a vdev is a collection of devices, you can use whole disk but recommendation is to use partitions, that way you can save yourself from future issues which may arise.

Vdev Types



DISK ①

A block device or disk, if this one disk fails, all data is lost. while creating pool you can specify full path like /dev/da1 or directly da1. also partition can be used same way /dev/da1p1 or da1p1.

MIRROR ③

Mirror is composed of two or more devices, required minimum 2 disks and can tolerate 1 disk failure. data is replicated across the disks.

dRAID ④

Also called as draid1, constructed using multiple internal raidz groups, allows fast resilvering because of integrated distributed hot spares. uses fixed stripe size which affects usable capacity and IOPS. can tolerate 1 disk failure. minimum 2 disk required.

dRAID2 ⑤

Can tolerate 2 disk failure. minimum 3 disks required, rest functionality is as mentioned in dRAID.

dRAID3 ⑥

Can tolerate 3 disk failure. minimum 4 disks required, rest functionality is as mentioned in dRAID.

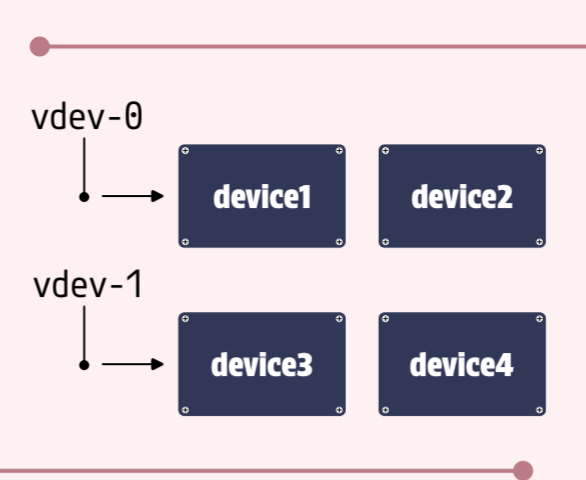
DEDUP

Dedicated device for deduplication tables, if more than one device is used allocations are load balanced.

LOG or Intent Log

This device is used for synchronous transactions, good for databases and NFS, by default these blocks are allocated from the main pool.

ZPOOL



Man Pages
zpoolconcepts(7), zpool(8)

Vdev Types

SPARE LOG DEDUP
SPECIAL CACHE

FILE ②

A regular file, this vdev type is for testing only, while creating pool you have to mention full path of file.

RAIDz ④

Also called as raidz1, minimum 2 disks required, can tolerate 1 disk failure and uses distributed parity. data and parity is striped across all the disk in raidz group, stripe width in raidz is dynamic, more on this later.

RAIDz2 ⑤

Minimum 3 disks required, can tolerate 2 disk failure and uses distributed parity. data and parity is striped across all the disk in raidz group, stripe width in raidz is dynamic.

RAIDz3 ⑥

Minimum 4 disks required, can tolerate 3 disk failure and uses distributed parity. data and parity is striped across all the disk in raidz group, stripe width in raidz is dynamic.

SPARE

Available hot spares in pool

SPECIAL

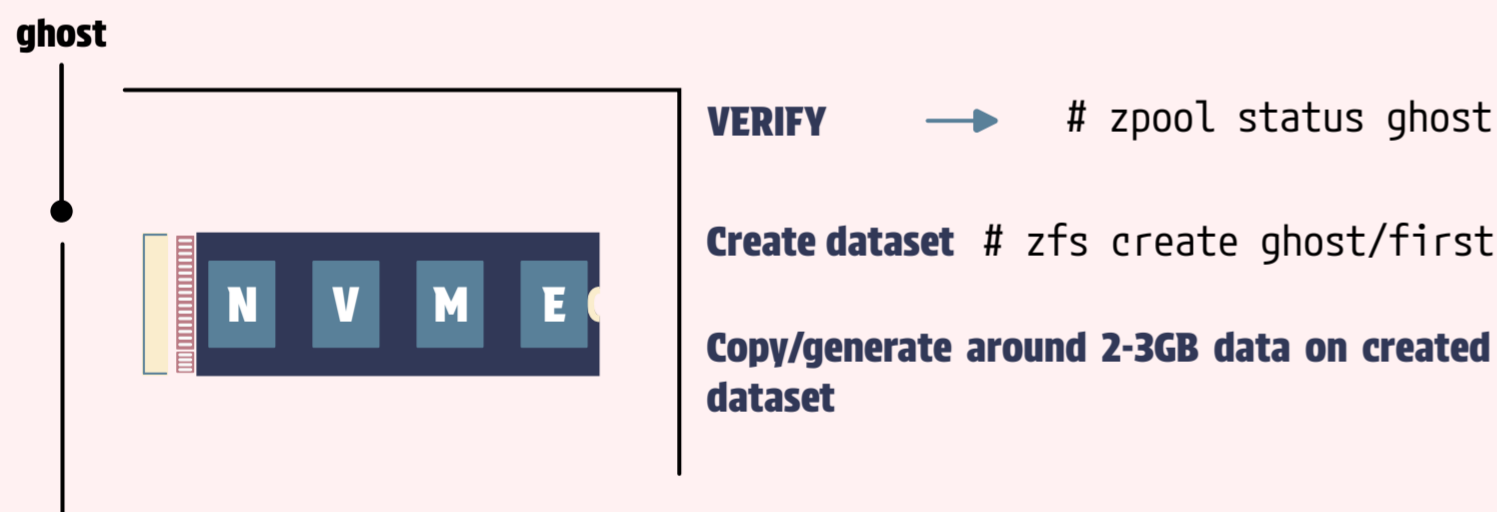
Dedicated device for allocating internal metadata, if more than one device is used allocations are load balanced.

CACHE

Device is used to cache storage pool data.

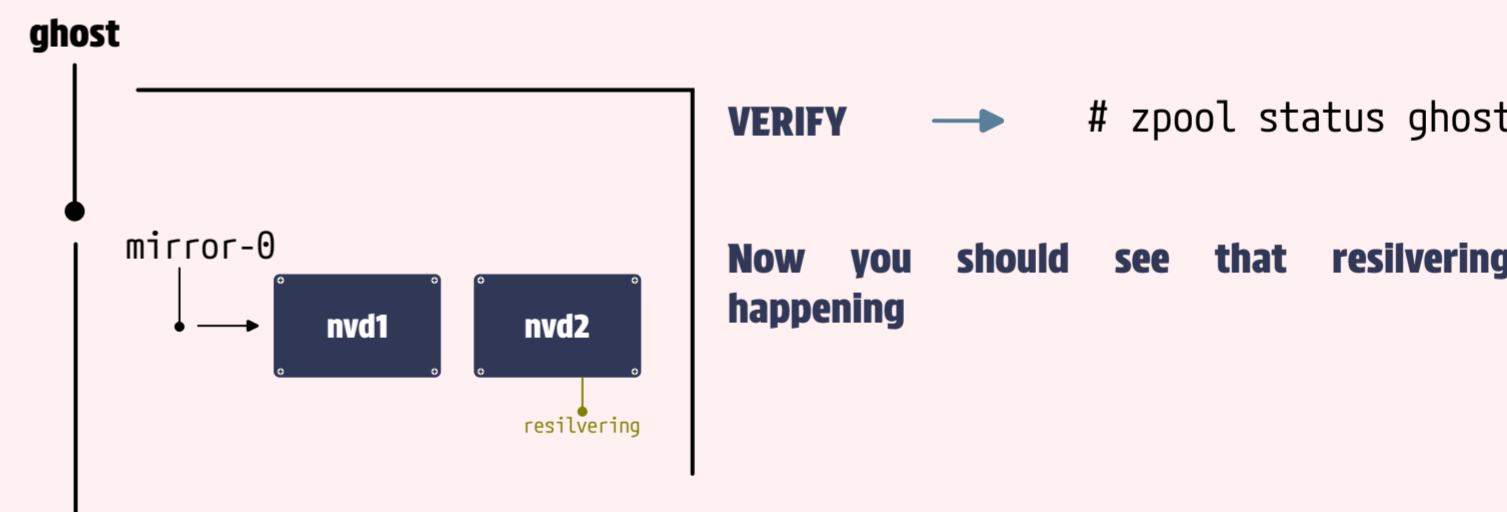
1 Let's create our first pool named ghost with single disk, this pool has no-redundency, if this disks fails, all data is lost.

```
# zpool create ghost nvd1
```



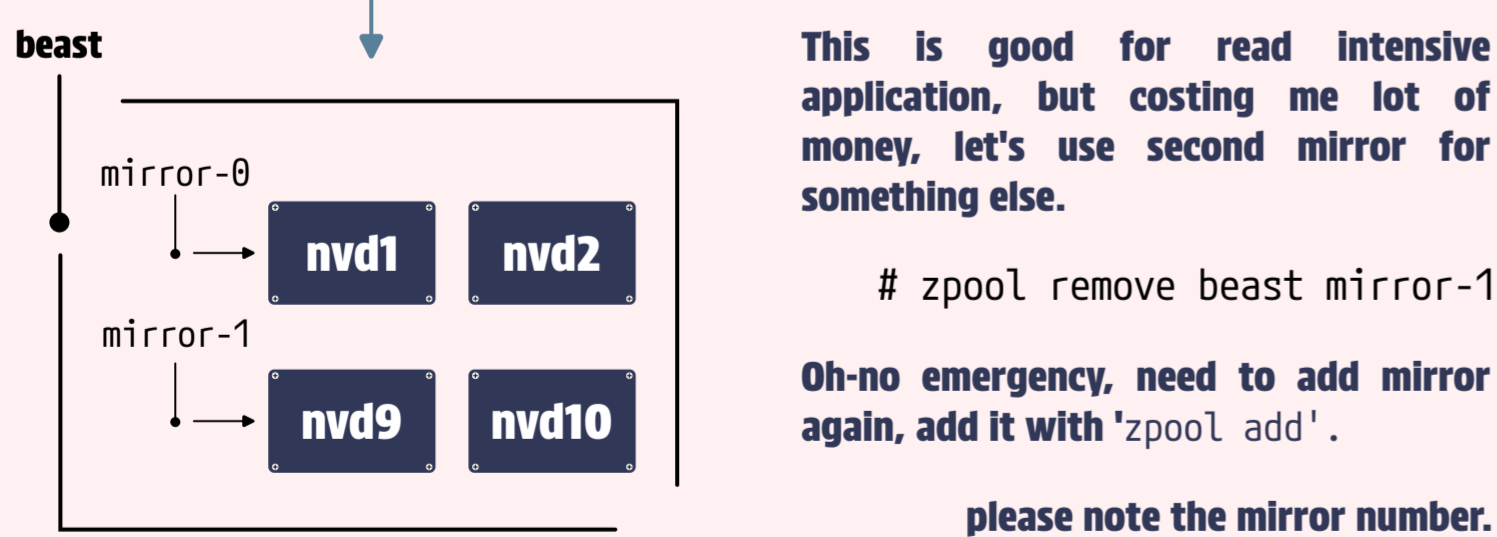
2 We can convert single disk(stripe) pool to mirror vdev by attaching one more disk, now data is mirrored across the disks.

```
# zpool attach ghost nvd1 nvd2
```



4 For whatever reason, I have decided to add one more mirror to existing pool, note that you can add mirror while creating pool also, Let's explore.

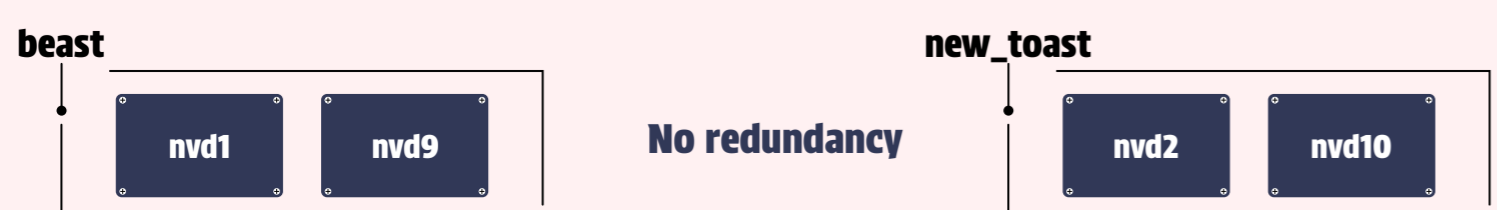
```
# zpool create beast mirror nvd7 nvd8 mirror nvd{9,10}
# zpool add ghost mirror nvd{9,10} ← can add to existing pool
```



For no reason let's 'split' the pool and create new pool out of it, don't split if pool is in the process of resilvering.

```
# zpool split beast new_toast
```

Note that last device from each mirror is split by default. but you can mention devices while splitting to include in new pool. # zpool import new_toast



3 Although I have redundancy in place, it would be great to have spare drive kick in, in case one of the drive fails, let's add "spare vdev" to my ghost pool, but before that let's create one more pool so we can see spare can be shared with multiple pools.

```
# zpool create ghost_replica mirror nvd5 nvd6 spare nvd{3,4}
# zpool add ghost spare nvd{3,4}
```

Two ways to get spare working

1. Manually

Assuming you have copied 2-3 GB data to created dataset

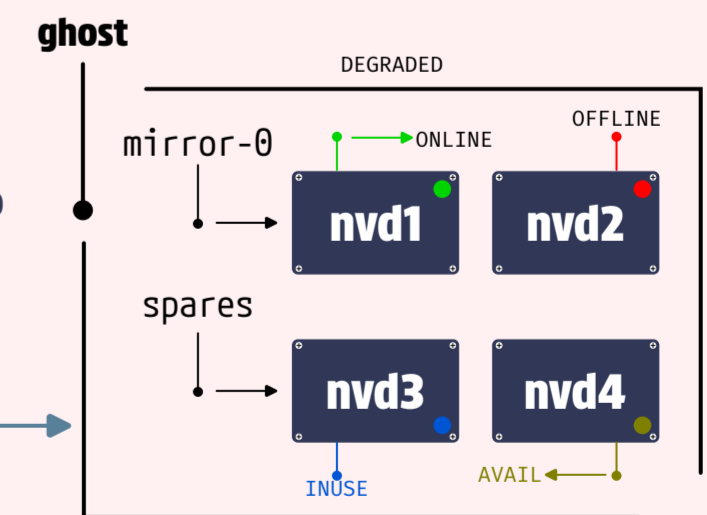
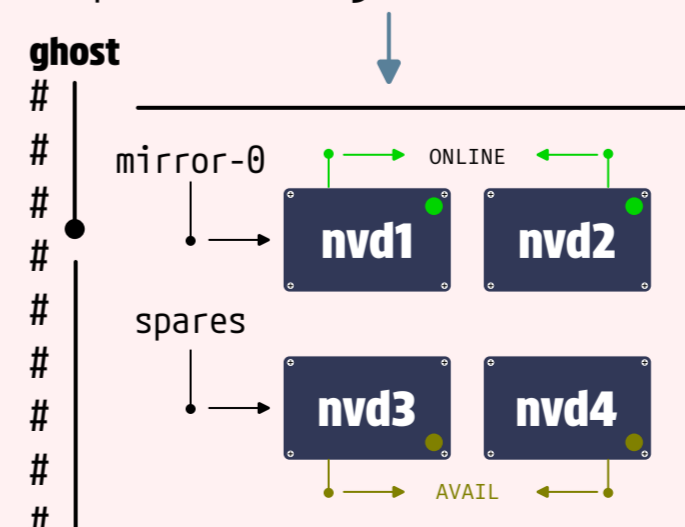
```
# zpool offline ghost nvd2
# zpool replace ghost nvd2 nvd3
# zpool status ghost
```

Resilvering should be happening

```
# zpool status ghost_replica
```

Copy another 1-2 GB data on /ghost/first

```
# zpool online ghost nvd2
# zpool status ghost
# zpool detach ghost nvd3
```



2. Automatic

service zfsd onestart
Offline the nvd2, spare should kick in, once resilvering done, bring nvd2 online all should fall in place as it was.

