

Use Sparse Files for FreeBSD jails?

Thinking about FreeBSD jails and an [elder post of mine](#) about putting jails within loopback-mounted disk images to enforce disk quota, I asked myself if I should use sparse files or pre-allocated files as virtual disk image for jail-based userland separation.

A sparse file designates a special file of a given size (e.g. 20 gigs), which is neither using nor reserving the whole disk space at once. It won't effectively allocate disk blocks until data is written to the sparse file.

This comes in handy in creating disk images, which can be a very time consuming task.

Let's look at the numbers of creating a 20gig file, which fills up available disk space at once but takes about five minutes to do so. Not to forget about the I/O load the task produces, which affects performance.

```
[root@localhost vz]# time dd if=/dev/zero of=test1.img bs=1024k count=20480
20480+0 records in
20480+0 records out
21474836480 bytes (21 GB) copied, 321.902 seconds, 66.7 MB/s

real 5m21.903s
user 0m0.021s
sys 0m55.282s
```

Creating a 20gig sparse file is done almost immediately, with the difference that disk blocks are not allocated right away.

```
[root@localhost ~]# dd of=test2.img bs=1024k count=0 seek=20480
0+0 records in
0+0 records out
0 bytes (0 B) copied, 1.4744e-05 seconds, 0.0 kB/s

real 0m0.002s
user 0m0.000s
sys 0m0.002s
```

Some possible drawbacks when using sparse files:

- Performance degradation when having multiple sparse files to which data is written randomly. This may end up in heavy fragmentation as the sparse files are likely not to be written contiguously in that case, which in term causes slow read access performance
- Race conditions when creating many sparse files which would exceed the available disk space (e.g. 20 sparse files of 20 gigs each, but the disk is only 200 gigs in size). To handle this, a special monitoring script would be needed to consider logical vs. physical space allocation.

Up until today, I always used pre-allocated disk images, but at the expense of non-usable disk space.

I think that sparse files might be a good choice actually, at least as long as there's not much random data written to it, so the payload

stays mostly identical over longer periods of time.

For sure, a sparse file used as image for a jail-based database server is definitely a bad idea.

Nevertheless, I'm keen to try and see, if this might be a considered options for some real-live scenarios.