

Re: disklabel differences FreeBSD, DragonFly

Source: <http://unix.derkeiler.com/Mailing-Lists/FreeBSD/hackers/2006-07/msg00409.html>

- *From:* Mike Meyer <mwm-keyword-freebsdhackers2.e313df@xxxxxxxxxx>
 - *Date:* Thu, 27 Jul 2006 18:37:25 -0400
-

In <44C93454.5020404@xxxxxxxxxxxxxxxxxx>, Alex Zbyslaw <xfb52@xxxxxxxxxxxxxxxxxx> typed:

Mike Meyer wrote:

You assume that "running out of space" happens over time, but with some runaway process logging to a file, for example, the partition filling up will still happen without you expecting it. It might take a bit longer with a big disk, but 20 minutes instead of 5 minutes isn't much different in terms of warning.

Yes, I'm assuming that "running out of space" happens over time. Sustained I/O speeds on modern hardware was around 100MB/sec last time I looked. So a good, large disk – say a terabyte raid (you need raid to get those performance numbers, so call it 2 500GB disks to keep it simple) – will take about three hours to fill *if you do nothing but write to the disk*. A runaway process – especially one generating log data – is normally doing other things that it's trying to log information about.

I don't have terabyte raids and for me a "big" disk is 250Gb.

In that case, you don't get 100MB/sec of throughput, either. Even if you've got a relatively fast single disk, you're going to be getting maybe 50MB/sec of throughput. So it's *still* going to take hours to fill the disk even if you do nothing but write to disk. And to complete the reprise of the paragraph you elided, if you've got a system that gets a lot more than 100MB/sec to disk, you almost certainly have a lot more disk than a terabyte.

A runaway system demon writing to disk might well do other things. A badly written user program might do nothing but write to disk. If you run servers that just run a bunch of standard ports and system demons, then this is unlikely to happen to you. If you work in an environment

Re: disklabel differences FreeBSD, DragonFly

where one or more fallible programmers churn through gigabytes of data it's depressingly easy to accidentally do *nothing* but write to disk.

You know, that's exactly the kind of environment I work in. We churn through gigaROWS of data. We have processes whose sole job is to pull data and write it to disk. Even major failures (like losing the network connection to half the consumer machines) don't cause the disk to fill in minutes. More like days on a properly configured machine.

That's because, even if your system is spending *all* of it's time doing nothing but writing to the disk, it'll take hours to fill the disk given most modern disk configurations. Disks have gotten bigger faster than they've gotten faster. So while you used to be able to fill a disk in minutes (or could you?), it takes a really strange configuration to do that now.

A further reason to separate partitions is that dump works at the level of a partition. Different partitions may have very different backup requirements, and for those of us without huge tape drives, partitioning to a size that can be dumped on one tape makes life easier.

That's one of the technical reasons I mentioned in the part you didn't quote.

To my mind, it only takes *one* technical reason. If I need multiple partitions to make backups easy, then arguments about log files are irrelevant :-)

If you're going to count 1, 2, many, then we already have "many" partitions, and don't need more. Once you get into finer distinctions of "many", you need to figure out which reasons are actually valid, and which are spurious, so you can pick from among those manys.

Well, there are always special cases. But hardware is so cheap these days, I'm used to fine-tuning the *system*, not just the partition. If something is so critical that it needs it's own partition, it's probably so critical that it needs it's own system as well. In fact, most of the thing I work on these days are so critical that they need several systems, half of them at a second site with automatic failover between them.

Not everyone is in a position to throw money at everything and what's cheap to you isn't cheap to everyone.

Re: disklabel differences FreeBSD, DragonFly

Boxes are cheaper than disk space – my last two low-end boxes cost less than my last small disk drive, even though I ordered them all about the same time. If you can afford the disk for some process, then chances are good you can afford a system instead, or as well.

I can't possibly justify one system for everything that needs a partition, nor do I even feel the need to do that. If anything, it would be a major inconvenience.

My claim is that your "everything that needs a partition" probably includes things that don't. But to prove that, we need to examine the reasons you think those things need a partition. I believe the only one you've given so far – as a space firewall – is specious.

Your arguments remind me of the environments I worked on in the 70s and 80s. Minis and mainframes that did all the computing for an organization. All the daemons that talked to the outside world ran on the same box as the developers ran compiles and tests on, etc. While that worked really well when it came to generating a robust OS, I haven't seen an environment like that in decades. Hell, most of my clients would shit bricks at the thought of putting their source or data on a machine that could be reached from the internet at large at all. Every developer has a box – or three – on their desks. The ETL boxes are distinct from the database boxes are distinct from the internal mail server is distinct from the external mail server, etc. If I want to have a process send email notices about something, I usually have to beat on them if I want a mail server on the box. And so on....

Frankly, if you're really worried about bootable slices, you should be advocating giving FreeBSD the ability to boot from a logical volume.

Who said I didn't? I have no objection to such a facility and would welcome it. It just imagined that extending the number of partitions from 8 to 16 would have been easier than booting from logical slices. If booting from logical slices is easier then I'm all for it.

You're not asking the right question just yet. The question shouldn't be "which is easier to add", but "which provides the most benefit for the least pain" (which subsumes the pain involved in adding it). I believe that the benefits of adding more partitions per slice are minimal – there are at least three ways to add more file systems that aren't bootable, and there's a better fix to the problem of wanting more bootable partitions – and the pain involved in upgrading a system across a change in the bsdlable would be rather high.

Re: disklabel differences FreeBSD, DragonFly

<mike

--

Mike Meyer <mwm@xxxxxxxx> <http://www.mired.org/consulting.html>
Independent Network/Unix/Perforce consultant, email for more information.

freebsd-hackers@xxxxxxxx mailing list

<http://lists.freebsd.org/mailman/listinfo/freebsd-hackers>

To unsubscribe, send any mail to "freebsd-hackers-unsubscribe@xxxxxxxx"