

Reason #5 That You Need File Virtualization



Don MacVittie, 2009-02-04

If you're just joining this series, there is a complete list of the Reasons to date [on my team member page](#).

Are we at reason #5 already? Wow. Okay, this is another one that salesmen will tell you because it is truly compelling, but it is truly a good reason, one of the best.

It is also one of the ones that I eschewed before getting to see real numbers that I could quantify were not marketing material.

The disk savings are real.

Yeah, I said it, and it's true. Sure, you could argue that they're only disk purchase deferrals, I would disagree. If you don't buy disk next year because you don't need it, and you buy the year after... Well, if you would have purchased disk this year and next without it, that's not deferral, it's cost abatement. But I can see the logic behind calling it deferred purchases – you really aren't stopping buying disk, only slowing down the rate of growth.

Either way, depending upon your storage architecture, the savings can be huge. I saw one case study that included a regular old enterprise, and an increase in disk utilization of nearly 70%. That's astounding. And I know the people, so those numbers aren't marketing fluff, they came from IT folks.

It all sounds so simple – you have *BigOldArray1* that has the user directories and the Intranet on it, and seven *LittleArrays* that hold application-specific data. *BigOldArray1* is filling up fast, and you're going to have to expand it. Buy a whole new rack, or upgrade, whatever. But your *LittleArrays* are sitting there at 20% utilization, and between them total more disk space than *BigOldArray1*. That's what a whole lot of data centers look like. More than you'd believe. Because apps often enter the Data Center through Departmental level IT, and only when they need integration with somewhere else, or they grow too big, or IT staffing consolidation occurs do they move to the DC. Or the project was going to grow into this large thing that did more than it does, but that idea was dropped, and it's still running on that hardware purchased from the project budget... Due to mergers, one of the applications that warranted that *LittleArray* is gone... There are a zillion reasons.

But the point is, that 20% utilization number sucks. With File/NAS Virtualization, you can resolve the problem. In the best systems, put them all in a pool and tell the system to manage them. It'll start shifting data to the under-utilized resources, relieving the need to upgrade. In most systems you have to be a bit more participative and tell the system which files on *BigOldArray1* it can move, but you're still shaving space out of nothing. How much? Well, let's just use my example. You can figure out the actual numbers for your data center using the same equations. Let's say that *BigOldArray1* is maxed out at 4TB and is hitting 80% utilization – That would leave 800Gigabytes free space. That's how much room you have in the old system. Let's further say that your *LittleArrays* are each 1 TB, and you have the seven I mentioned, utilized at the 20% I mentioned. That would be 80% of 7TB free – or 5.6 TB that you can recover and use with *BigOldArray1*.

You have more than double the space of your original *BigOldArray1*, you just need to use it more intelligently. Replication (Reason #4) can quickly get stuff off the old machine and onto the new one without deleting it from the old until you're certain all is well.

Of course there are caveats. There always are, anyone who sells you a "more than double your space!" line is lying if they don't admit that.

First, the amount of free space available in your organization won't match my example, nor will it be as clean as my example. Seriously, how many arrays with exactly 20% free space can you have ;-)? But you can counter this by looking at free space yourself.

Second is tiering (Reason #3). If you're using tiering, your options might not be this clean-cut. That's not to say that you can't reap benefits, even benefits as large as my example, but you have to make certain you are moving files to the correct tier, which complicates the equation.

Third is politics. You can do *to* users and just move those departmental apps' data around, or you can run the gauntlet and get buy-in from all the application owners. I would recommend that you move everything when you first install the File/NAS Virtualization device and inform users that the entire architecture is changing (it is), but that requires a lot of forethought – be thinking about it. Otherwise, you know which of the first two options is best for your users and your organization.

Fourth and final is that you know your Data Center better than I or any salesperson ever will. Yeah, I'm a geek, but I have yet to see the "average" data center, how about you? It is possible to have an array appear "under-utilized" from a disk usage perspective but be hammered from an I/O perspective, or on a slower network segment (though that last is generally easy to fix). You need to validate if File/NAS Virtualization will cause you conflicts in that regard. Trust me, *no* File/NAS virtualization vendor wants you to put them in and then have performance issues – the promise is great, they don't want it to backfire on you. ARX can help you evaluate this stuff before making moves (though IIRC the machine needs to be in the virtualization appliance to help, it doesn't have to be part of the virtual directory), presumably other vendors can too.

That's it for this week, I do hope you're finding these posts useful.

Until next time,

Don.

F5 Networks, Inc. | 401 Elliot Avenue West, Seattle, WA 98119 | 888-882-4447 | f5.com

F5 Networks, Inc.
Corporate Headquarters
info@f5.com

F5 Networks
Asia-Pacific
apacinfo@f5.com

F5 Networks Ltd.
Europe/Middle-East/Africa
emeainfo@f5.com

F5 Networks
Japan K.K.
f5j-info@f5.com