Key Backup Challenges: The Dilemma of Backup at Scale

Scaling backup operations is one of the challenges defined by Gartner in their insightful paper, “5 Key Challenges You Must Solve with Your Next Backup Platform.” We’ve made it available here at no cost (registration required). I’ve worked with Gartner for many years, and I’m always impressed at how well their analysts understand the data protection space. This paper is ample evidence of that, and I highly recommend it. It’s a solid 18 pages and provides a wealth of information. For anyone pondering “what’s next?” with their backup environment, it’s a must-read.

Back to scale. Nothing stresses out your backup environment more than growth, and this applies to everything: hardware, software, and operations most of all.

Let’s look at a simple illustration.

Let’s say you had to back up 500 servers a day, a mix of physical and VMs. It’s a pretty difficult task keeping all the job management straight. You also need to ensure you have the capacity and performance to handle the workload within your backup window. 500 servers equal a pretty big environment. But some organizations have *5,000* servers, or even more. Think about how that 10x uptick in scale would impact you.

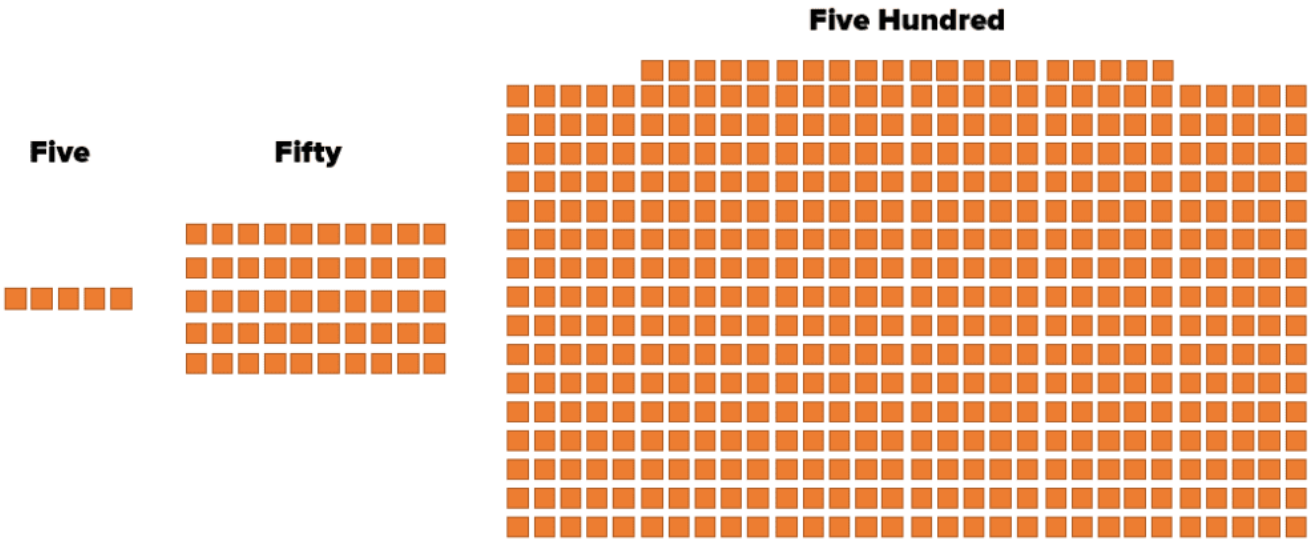
Job management alone becomes crazy. Consider, if you started one backup job a minute, it would take over 83 hours just to *start* 5,000 backups. That’s not meeting anyone’s SLAs. You have to start and run many jobs simultaneously. This can lead to massive performance challenges from network saturation and storage saturation. And think of the troubleshooting. Let’s say you have six backup failures a night to troubleshoot out of 500 backups. Multiply that by 10 and now you’ve got *sixty* backup failures to troubleshoot. Maybe more because the scale itself can lead to higher failure rates. And guess what: you won’t have ten times the staff to handle it.

This is why the traditional backup products are still around and widely used. You know the names: Veritas NetBackup, Commvault, IBM Spectrum Protect, and a few others. They’ve figured out how to scale and protect many thousands of nodes. The new-kids-on-the-block products likely won’t be able to manage the same scale for years. This is not to say they can’t be useful as point products even in a massive backup environment.

Hardware is also a challenge. Whether you’re using a home-grown solution (media server plus disk) or a purpose-built backup appliance, you’ve likely experienced the challenge of scaling up. Just saying the words “fork-lift upgrade” to a backup admin can get you some dirty looks, but many users have never had another option. And then swapping out backup hardware typically means reconfiguring backup jobs to use the new target(s).

Scale is something that Pure can help you with. Consider our FlashBlade™ product as a backup target (and yes, you can use all-flash storage for backup). It scales both capacity and compute simply by adding blades. All upgrades and expansions are non-disruptive, even major product generation updates, making capacity

increases seamless when backup data growth requires more capacity. FlashBlade also has great performance and backup bandwidth letting you run more backup jobs together. And FlashBlade can increase your data recovery time in a big way. FlashBlade uses Pure1® storage management which makes everything easy. You also get Pure's proactive support services. One of the keys to dealing with increased scale is to reduce complexity as much as possible: simple is smart.



Consider Storage Alternatives for Backup

Large organizations have been managing massive backup environments for a long time. But there are smaller, fast-growing companies learning about scale the hard way. Their backup platforms are breaking down under the stress of constant growth. How to relieve the pressure without changing backup products?

Consider alternative data protection methods. Not everything has to be protected with backups. For some workloads, you can switch to a storage-based approach and use a combination of snapshots and replication. This is effective for your busiest or largest workloads, because snap-and-replicate is far more efficient than backup, at least if you're using Pure Storage! If you've experienced snaps and replication with other storage vendors, you may have a bad taste in your mouth. It's common to suffer performance degradation as snap counts rise. Struggles with complex management and annoying space reservations are common. And

then there's the extra licensing fees. At Pure we've solved all those technical challenges. Plus, the array includes snapshot and replication software at no additional cost. That's how you do it right.

Storage-only protection still requires a solution for long-term data retention. This is where you can use your backup software to make periodic (weekly, monthly) backups to tape. Software like Veeam and Commvault can leverage Pure Storage snapshots as backup sources, eliminating host impact (that's a big benefit and a topic for another day). And if you don't want tape, Pure has native cloud integration and the ability to copy snapshots into the cloud. Since Pure snapshots are portable, you can recover them back to any Pure FlashArray™, not only the source. These aren't your grandfather's snapshots!

By moving your most challenging workloads from backup to snap-and-replicate you take a disproportionately large load off your overall backup environment. Not all backups are equal, and it makes sense to use the technology that's most appropriate to a workload.

Whoever Hesitates is Lost

My final advice on this topic is don't wait until you have a problem. If your backup environment is growing, take steps now to head off a potential problem. Because when you reach the tipping point it's going to be a much bigger mess than you might think. And do take a look at the Gartner paper. Scale isn't the only challenge facing backup environments, and the more you know now the better you can plan ahead.