

Number of Appliances Dedicated to Deduplicating Backup Data Shrinks even as Data Universe Expands

One would think that with the continuing explosion in the amount of data being created every year, the number of appliances that can reduce the amount of data stored by deduplicating it would be increasing. That statement is both true and flawed. On one hand, the number of backup and storage appliances that can deduplicate data has never been higher and continues to increase. On the other hand, the number of vendors that create physical target-based appliances dedicated to the deduplication of backup data continues to shrink.

Data Universe Expands

In November 2018 IDC released a [report](#) where it estimated the amount of data that will be created, captured, and replicated will increase five-fold from the current 33 zettabytes (ZBs) to about 175 ZBs in 2025. Whether one agrees with that estimate, there is little doubt that there are more ways than ever in which data gets created. These include:

- Endpoint devices such as PCs, tablets, and smart phones
- Edge devices such as sensors that collect data
- Video and audio recording devices
- Traditional data centers
- The creation of data through the backup, replication and copying of this created data
- The creation of metadata that describes, categorizes, and analyzes this data

All these sources and means of creating data means there is more data than ever under management. But as this occurs, the

number of the products originally developed to control this data growth – hardware appliances that specialize in the deduplication of backup data after it is backed up such as those from [Dell EMC](#), [ExaGrid](#), and [HPE](#) – has shrunk in recent years.



Here are the top five reasons for this trend.

1. Deduplication has Moved onto Storage Arrays.

Many storage arrays, both primary and secondary, give companies the option to deduplicate data. While these arrays may not achieve the same deduplication ratios as appliances purpose-built for the deduplication of backup data, their combination of lower costs and high levels of storage capacity offset the inabilities of their deduplication software to optimize backup data.

2. Backup software offers deduplication capabilities.

Rather than waiting to deduplicate backup data on a hardware appliance, almost all enterprise backup software products can deduplicate on either the client or the backup server before storing it. This eliminates the need to use a storage device dedicated to deduplicating data.

3. Virtual appliances that perform deduplication

on the rise.

Some providers, such as [Quest Software](#), have exited the physical deduplication backup target appliance market and re-emerged with virtual appliances that deduplicate data. These give companies new flexibility to use hardware from any provider they want and implement their software-defined data center strategy more aggressively.

4. Newly created data may not deduplicate well or at all.

A lot of the new data that companies may not deduplicate well or at all. Audio or video files may not change and will only deduplicate if full backups are done – which may be rare. Encrypted data will not deduplicate at all. In these circumstances, deduplication appliances are rarely if ever needed.

5. Multiple backup copies of the data may not be needed.

Much of the data collected from edge and endpoint devices may only need a couple of copies of data, if that. Audio and video files may also fall into this same category of not needing to retain more than a couple copies of data. To get the full benefits of a target-based deduplication appliance, one needs to backup the same data multiple times – usually at least six times if not more. This reduced need to backup and retain multiple copies of data diminishes the need for these appliances.

Remaining Deduplication Appliances More Finely Tuned for Enterprise Requirements

The reduction in the number of vendors shipping physical target-based deduplication backup appliances seems almost counter-intuitive in the light of the ongoing explosion in data growth that we are witnessing. But when one considers

must of data being created and its corresponding data protection and retention requirements, the decrease in the number of target-based deduplication appliances available is understandable.

The upside is that the vendors who do remain and the physical target-based deduplication appliances that they ship are more finely tuned for the needs of today's enterprises. They are larger, better suited for recovery, have more cloud capabilities, and account for some of these other broader trends mentioned above. These factors and others will be covered in the forthcoming DCIG Buyer's Guide on Enterprise Deduplication Backup Appliances.