

SCALE-UP VS SCALE-OUT ARCHITECTURE IN BACKUP STORAGE TECHNOLOGIES

Introduction

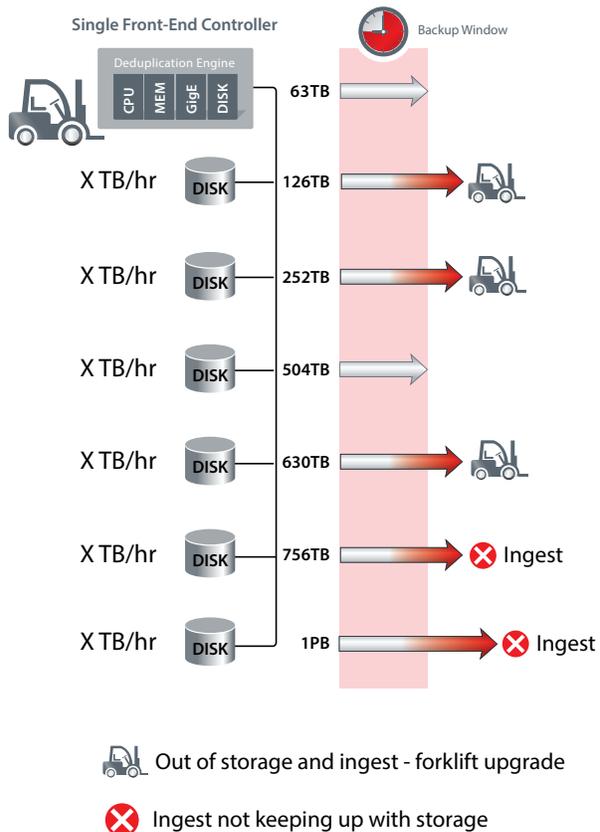
The typical enterprise is seeing data growth of 30% year-over-year, doubling their total data every 2.5 years.

In the past, organizations prepared for backup data growth in one of two ways. They could overbuy a storage server or controller, with the goal of growing into the solution, a wasteful and expensive practice. Or they could use a more traditional standard primary storage – buying a solution aimed to meet their immediate storage needs, and adding physical storage as disk space is needed. This is referred to as Scale-Up approach. With scale-up, the controllers act as the point of access, with a set amount of networking, memory and CPU. Adding physical disks to the existing storage system architecture only increases the storage capacity, it does not increase the networking, memory or CPU of the controller. It is critical to note that in organizations that require the protection of a large amount of data, typically over 30TB up to petabytes, or those that require longer-term retention of their backup data, data deduplication is a critical component of their backup storage solution. Data Deduplication saves only the unique data behind backups to reduce the cost of backup storage. As data grows, the deduplication load also grows requiring an ever-increasing amount of network ports, processor and memory. Unfortunately, traditional data deduplication devices, like Data Domain and HP StoreOnce, use scale-up architecture. With these devices, data deduplication fixes the storage problem however it creates a new compute challenge. They fail to add compute with capacity, and their ability to process backups is degraded and will eventually fail.

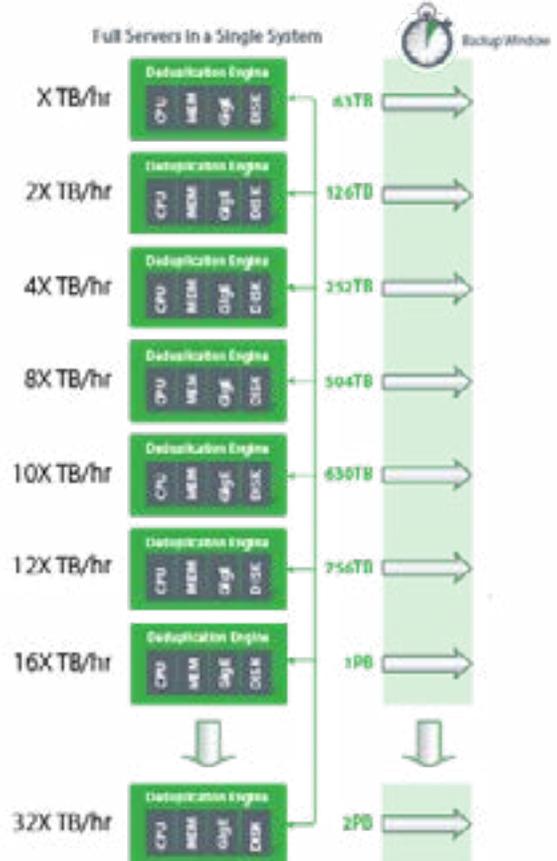
In recent years, a more effective method of scaling storage for backup has emerged – the ability to add in a combination of storage and resources in a cost-effective, scalable appliance solution. This is referred to as Scale-Out. With a scale-out solution, entire new appliances are added into the architecture, not only adding storage, but adding networking, CPU and memory to an interconnected set of appliances. This scale-out capability is intelligent to act a single combined system (including load balancing, deduplication across appliances, and increased processing), without any of the disadvantages of a front-end controller bottleneck. Data deduplication easily keeps up with data growth. This approach can be found in Intelligent Hyperconverged Storage for Backup solutions like ExaGrid.

To understand why scale-out is necessary, we must examine the impact data growth has on each implementation.

Scale-Up Architecture



ExaGrid Scale-Out Architecture



How Does Backup Data Growth Impact More than Disk Storage?

Backup data growth is about more than the amount of data stored. To support backup data growth, IT teams must ensure the backup storage infrastructure can store the amount of backup data required to meet their retention policies – and - IT teams must also ensure the backup processing that takes place on storage can scale to meet the growth of backup data intake without negatively impacting their backups. Backups which are impacted can cause a ripple effect – with multiple delays causing backups to fall outside the backup window.

Backup Data Processing as Part of Backup Storage.

The amount of data that can be taken in to a backup storage solution is the amount of data that can be processed by the storage during the time when backups are not disruptive to the enterprise. This caveat is critical.

Backup jobs have to be completed during off hours, called the backup window. If a backup storage solution is able to take in 50TB/hr., its rate of processing is 300TB in a 6 hour backup window. The total data processed is typically restricted to backing up data during off hours – in many cases backup windows are 6 to 8 hours.

In scale-up architectures, this rate of processing number is the high-end number. This does not account for the performance degradation that will occur as data intake sizes grow. In some scale-up cases, depending on the backup software solution utilized, backup intake processing done on the controller / storage can be impacted by deduplication, recoveries, and replication processing. This activity may delay or reduce the amount of data that can be processed. The stated intake on any scale-up solution may be dramatically lower than the actual performance. Users of scale-up architectures that outgrow their backup window find themselves at an impasse – a failure to meet backup windows is not due to the lack of disk storage, but due to overloaded controllers. Adding more disk to scale is not a solution. Updating the controllers creates silos of systems that don't share data, especially for management and global deduplication. So what do they have to do? The users are required to bring in a new larger system, needing a migration of data from the old system. This is referred to in the industry as a rip-and-replace or forklift upgrade.

Why Doesn't This Happen to a Scale-Out Solution?

Unlike the scale-up approach where you need to guess which sized front-end controller is required, the scale-out approach allows you to simply add the appropriate sized appliances as your data grows.

Scale-out architectures are full appliances (processor, memory, bandwidth, and disk) in a single system that can be appended to a system in a network of backup storage. As data grows, all resources are added including additional primary storage, bandwidth, processor, and memory as well as disk capacity. This keeps the backup window fixed in length regardless of data growth, which eliminates expensive forklift upgrades.

Unlike the scale-up approach where you need to guess which sized front-end controller is required, the scale-out approach allows you to simply add the appropriate sized appliances as your data grows.

When choosing a scale-out system, verify that any size or age appliance can be mixed in a single system, which allows IT departments to buy compute and capacity as they need it. This evergreen approach also eliminates both product obsolescence and forklift upgrades.

How to Choose if Scale-Out or Scale-Up is Right for You?

Organizations with a relatively small data set and low expected growth over a 3 to 5 year period are likely able to find a scale-up solution that meets their needs. Enterprises with 30TB to petabytes of backup data to protect and expectations of average to large data growth should look into scale-out solutions like ExaGrid to dramatically reduced backup windows and significantly lower cost of storage up front and over time.

REQUEST A DEMO NOW
INTELLIGENT HYPERCONVERGED STORAGE FOR BACKUP



United States: 350 Campus Drive | Marlborough, MA 01752 | (800) 868-6985

United Kingdom: 200 Brook Drive | Green Park, Reading, Berkshire RG2 6UB | +44 (0) 1189 497 051

Singapore: 1 Raffles Place, #20-61 | One Raffles Place Tower 2 | 048616 | +65 6808 5574

ExaGrid reserves the right to change specifications or other product information without notice. ExaGrid and the ExaGrid logo are trademarks of ExaGrid Systems, Inc. All other trademarks are the property of their respective holders.

©2019 ExaGrid Systems, Inc. All rights reserved.