Thin Provisioning

ESG Storage Innovations Series
Focus on 3PAR

By
Tony Asaro
Senior Analyst
April 2006

Thin Provisioning

ESG believes that one of the more important but often misunderstood advancements within data storage is thin provisioning. This report defines and discusses thin provisioning, its economic impact and how it works.

Thin provisioning is a way to minimize the cumbersome task of storage provisioning while improving capacity utilization. When provisioning storage using traditional methods, system administrators typically dedicate a fixed amount of storage to each particular application.

For example, if a 500 GB volume is allocated to an application with only 100 GB of actual data, the other 400 GB has no data stored on it. That unused capacity is still dedicated to that application and no other application can use it. This means that the unused capacity of that 500 GB is wasted storage, which means that it is also wasted money. And even though all of the storage capacity may eventually be used, it could take years to do so. This is a major problem when managing storage capacity and is often referred to as stranded storage, or allocated-but-unused storage.   The problem is exacerbated by local and remote copy facilities that consume additional storage by duplicating allocated-but-unused capacity (copying "nothing" to "nothing").

With thin provisioning, the storage provisioning process starts out the same as traditional provisioning. Using that same example, the system administrator provisions 500 GB to the application with only 100 GB of actual data. With thin provisioning, the unused 400 GB can still be made available for other applications or need not be purchased at all. This approach allows the application to grow transparently and at the same time ensure that capacity is not wasted. Thin provisioning is really Just-in-Time storage. The application thinks it has 500 GB of storage but the storage system only provides the capacity when needed. The rest of it stays in the pool. System administrators set thresholds for them to be alerted when more disks need to be added into the storage pool.

Thin provisioning offers a simple solution to the problem of stranded capacity. Thin provisioning is a storage system technology that allows users to safely allocate to an application as much logical capacity as needed over its lifetime. Meanwhile, physical capacity is drawn from a common pool of storage on an as-needed basis. That is, only when an application performs writes is physical capacity drawn from the storage pool. Additionally, physical capacity can be added to the storage pool non-disruptively at any time.

Compared to Traditional Volume Expansion

Various storage arrays support a feature referred to as dynamic volume expansion, which is sometimes confused with thin provisioning. Dynamic volume expansion is a form of traditional provisioning that allows customers to increase the size of their volumes as part of an online process. For example, if the size of the original volume is 100 GB but needs to be increased to 200 GB, the system administrator can use an online process to do this, using tools provided by the storage array. This capability addresses the problem of data growth by enabling system administrators to manually add more capacity without having to re-boot servers. However, dynamic volume expansion does not solve the problem of stranded storage.

The inefficiencies of traditional storage provisioning can negatively impact capital costs and storage administration resources. The most obvious inefficiency is the amount of storage that goes unused and therefore increases the total cost of ownership. Additionally, since this allocated but unused storage capacity cannot be reclaimed for other applications, customers have to buy more storage capacity as their environments grow, increasing cost even further. At some point customers may actually be required to buy a completely new storage system in addition to the one they have in place. ESG conducted a survey in June 2005 of 20 customers who focused on the limitations of traditional storage provisioning methods. The highlights of that survey included:

- 54 percent of the customers were aware that they had stranded and unused storage capacity due to inefficient provisioning methods.
- 55 percent of these customers had between 31 and 50 percent of stranded and unused storage. For example, if they provisioned 10 TB of storage capacity, then 3.1 TB to 5 TB of that was stranded.

- 45 percent of the total users surveyed had to buy an additional storage system because they could not utilize their stranded storage. This means that, although these customers had unused storage capacity they've already paid for, they needed to buy a new storage system to meet the needs of their business.
- 30 percent of users were planning to buy an additional storage system in the next twelve months because they could not access their stranded storage. Combining the 45 percent of customers who purchased an additional storage system with the 30 percent who were planning to buy, then the number of users who needed to purchase additional storage to keep their business running increases to 75 percent.
- 80 percent of users felt that storage provisioning was a time and resource drain on their IT organizations.

The Value of Thin Provisioning

There are eight ways that thin provisioning can save customers money:

1. Less storage is required initially when purchasing a new storage system.
2. Since there will be no stranded storage capacity, less storage is required over the life of the storage system.
3. Additional storage systems will not be required based on having stranded storage.
4. More applications/servers per storage system provide greater levels of consolidation.
5. The time and resource required to perform storage provisioning tasks is reduced.
6. "Thin provisioning-aware" local and remote volume copy services consume less capacity.
7. Customers need to purchase fewer storage system software licenses based on capacity.

Less storage is required initially when purchasing a new storage system. Using traditional storage provisioning methods, customers have no choice but to buy additional capacity up front. However, with thin provisioning, this is not required since capacity does not have to be pre-allocated. Customers can buy what they need with a minimal amount of spare capacity for short-term growth. When a pre-defined threshold is reached, more disks can be added to the storage pool. The amount of money saved can range, depending on the environment. For example, customers can start out with 25 or 50 percent of the capacity they normally would purchase with storage systems that don't support thin provisioning. A customer who previously had to buy 10 TB of capacity because of the limitations of traditional provisioning can instead acquire 2.5 TB or 5 TB with the more efficient thin provisioning method. The point is simple – thin provisioning provides a method for "just in time" provisioning that allows system administrators to be more operationally efficient and lower capital costs.

Since there is no stranded storage capacity, less storage is required over the life of the storage system. It is the experience of ESG that customers have a large amount of capacity that is stranded and unused due to the inefficiencies of traditional provisioning methods. For example, the ESG survey indicated that between 31 and 50 percent of customers' storage was stranded and unused due to poor provisioning. Over the life of a storage system, thin provisioning can save tens of thousands or even hundreds of thousands of dollars, depending on the amount of capacity required. To quantify this, at $20 per GB, 5 TB of stranded storage equates to $100,000 of **wasted cost** to the customer. For 10 TB, the waste becomes $200,000 and so on. This wasted cost of stranded storage is eliminated with thin provisioning. Additionally, companies can delay the acquisition of disk drives and take advantage of the continuous and inevitable reduction of cost and increase in capacity. Of course, since most storage system software licensing is based on installed capacity, additional software cost savings can be achieved with thin provisioning.

Additional storage systems will not be required based on having stranded storage. There is a finite amount of storage capacity that any storage system can support. When the optimal capacity has been reached, customers may need to implement an additional storage system. Referring back to the ESG rapid survey, 45 percent of total users were forced to buy an additional storage system because they could not utilize their stranded storage while another 30 percent were planning to buy a new storage system. Acquiring a completely new storage system because of stranded storage is extremely wasteful. This is especially true since it is more than likely that a substantial amount of unused and inaccessible capacity already resides within the existing storage system. The additional costs associated with buying a new storage system include the up front price of the system, all of its associated software licenses, maintenance fees, and administration costs.

More applications and servers per storage system provide greater levels of consolidation. As described above, since capacity doesn't have to be dedicated and fixed on a per volume basis with thin provisioning, customers can create more volumes. As a result, more applications and servers can be attached to a single storage

system. Storage consolidation is a major initiative and thin provisioning can enable even greater levels of consolidation.

The time and resource required to perform storage provisioning tasks is reduced. ESG regularly hears from system administrators that one of the most mundane tasks they are required to perform is provisioning storage. Thin provisioning can essentially eliminate the need for storage provisioning. Customers can create large logical volumes without any cost to them, allowing the application to keep growing and growing as needed. ESG has spoken to customers who have implemented thin provisioning and they report that they spend literally no time provisioning storage except when creating new volumes.

"Thin provisioning-aware" local and remote volume copy services consume less capacity. When system administrators create a remote mirror of specific volumes, the data from the volume is not the only thing being copied and transferred. The stranded unused capacity is actually copied and transferred, as well. Processors, cache, and bandwidth are used to transfer an entire volume, consisting of real data and possibly large numbers of empty blocks, over expensive WAN links. When traditionally provisioning a 2 TB volume with only 100 GB of actual data, all 2 TB will still be copied. In some cases, customers will create two or more copies of their primary volumes. With traditional provisioning this would equal six to eight TB, of which only 300 to 400 TB might be actual data, including the primary source.   Of course, this extra capacity also inflates the cost for storage system software which is typically based on installed capacity.

This limitation of traditional storage systems also holds true for other services such as RAID rebuilds and local full volume copies. Thin provisioned volumes only contain real data and the resources needed for remote and local copies as well as RAID rebuilds are always used for real data and not empty blocks.

Customers need to purchase fewer storage system software licenses based on capacity. Some storage software vendors price their software based on the amount of installed capacity.   With thin provisioning, only physical capacity for actual data is installed and therefore software pricing can be greatly reduced.

Conclusion

Thin provisioning is "just in time capacity" that essentially eliminates allocated but unused or stranded storage. It also greatly simplifies storage provisioning tasks, reducing administration costs. Customers should look at storage systems that support thin provisioning differently than those that do not when evaluating the initial capacity configuration, budgeting, capacity growth planning over time, and the total cost of ownership. **The economic impact of thin provisioning can be substantial, saving literally thousands, tens of thousands, and potentially hundreds of thousands of dollars, depending on the scope of the environment.**

Thin versus Traditional

There are a number of benefits provided by thin provisioning compared to traditional methods. With thin provisioning, customers will be able to:

- Create more and larger volumes.
- Buy less storage capacity to start and only add disks as needed.
- Spend less time provisioning storage.
- Potentially support a greater number of servers.
- Experience much better capacity utilization rates.

Thin Provisioning Efficiency
This section compares traditional versus thin provisioning methods. The first example illustrates how thin provisioning is more efficient by providing a larger pool of shared capacity at the outset for multiple applications, as compared to traditional provisioning methods.
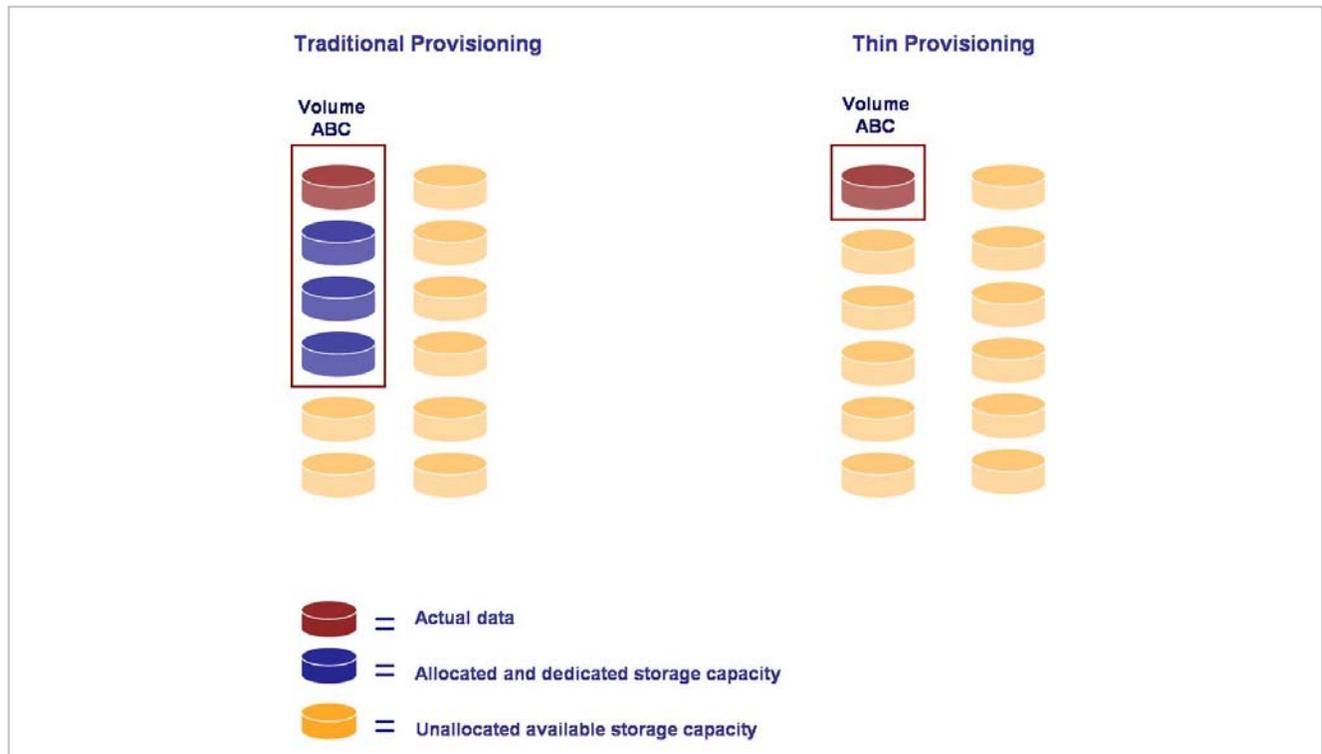
Traditional:
1. The customer's storage system has 3 TB of capacity.
2. The system administrator provisions Volume ABC with 1 TB of capacity.
3. Initially, 200 GB is stored on Volume ABC.
4. The system administrator sees that 2 TB of capacity is left for other applications, although only 200 GB of actual data has been stored.

Thin Provisioning:
1. The customer's storage system has 3 TB of capacity.
2. The system administrator creates Volume ABC with 1 TB of capacity.
3. Initially, 200 GB of data is stored on Volume ABC, as above.
4. The system administrator sees that 2.8 TB of capacity is available.

### Figure One: Initial Capacity Requirements

In this example, thin provisioning offers an additional 800 GB of capacity to use for other applications over the traditional method. Depending on the number of volumes on the storage system, the delta in available capacity between traditional and thin provisioning can increase dramatically.

In Figure Two, three volumes are created using traditional provisioning methods. The total usable storage capacity of the storage system is 2 TB.

- The total volume size of Volume 1 is 500 GB, of which 100 GB is actual data and 400 GB is allocated unused capacity.
- The total volume size of Volume 2 is 550 GB, of which 50 GB is actual data and 500 GB is allocated unused capacity.
- The total volume size of Volume 3 is 800 GB, of which 200 GB is actual data and 600 GB is allocated unused capacity.

In total, the storage system stores 350 GB of actual data, 1.5 TB of allocated unused capacity, and only 150 GB of available capacity for other applications.

Figure Two: Traditional Versus Thin Provisioning Allocation

Traditional   Thin Provisioning
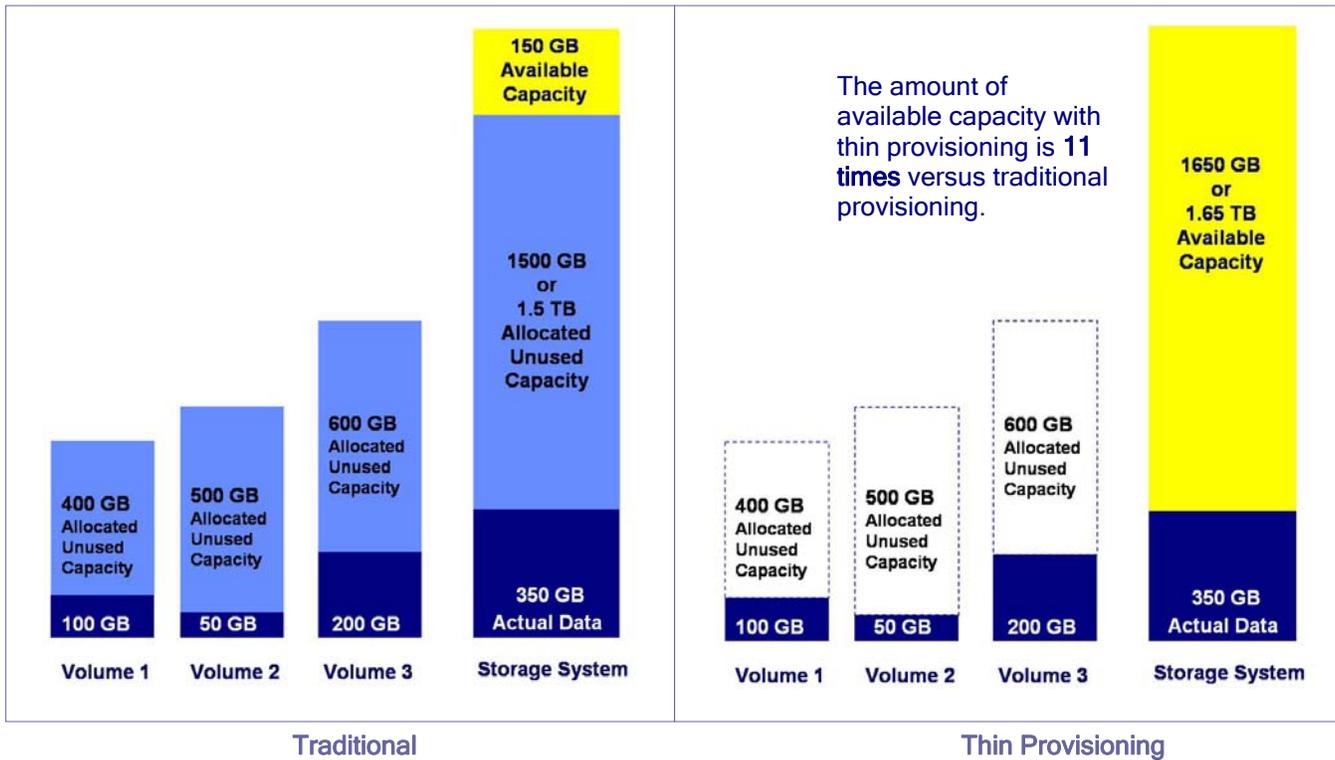
If we look at the same 2 TB storage system with thin provisioning implemented, the differences are quite dramatic. Although the system administrator creates the same size volumes for Volumes 1, 2, and 3, there is no allocated unused capacity. In total, the storage system with thin provisioning stores the same 350 GB of actual data as the other storage system, with 1.65 TB of capacity available for other applications, versus only 150 GB in the traditional storage system. **Thin provisioning effectively increases the amount of available capacity by 11 times.**

Consider the financial impact of having 11 times more available capacity over traditional storage systems, at storage totals of 10 TB, 100 TB, 500 TB, and even 1 PB of storage. As an example, consider a storage system that cost $20 per GB. With a traditional storage system with 1.5 TB of allocated unused capacity, the cost for total storage would be $30,000. By using the thin provisioning feature, there is no cost for the allocated but unused capacity. The cost savings on 10 TB of allocated unused capacity would be $200,000.

Thin provisioning provides customers with a greater pool of storage capacity, allowing system administrators to create more volumes for different applications. As a result, thin provisioning enables customers to allow more servers to share a single storage system (Figure Three).

Traditional:
1. The customer's storage system has 3 TB of usable capacity.
2. The system administrator uses the traditional method to provision storage. He or she creates a 500 GB volume for each of the application servers. Each volume contains 200 GB of actual data.
3. The storage system in total has 1 TB of available capacity left, 800 GB of actual data, and 1.2 TB of allocated unused data.
4. The system administrator is limited to supporting only a total of four servers in this example.

Figure Three: Traditional Server Support
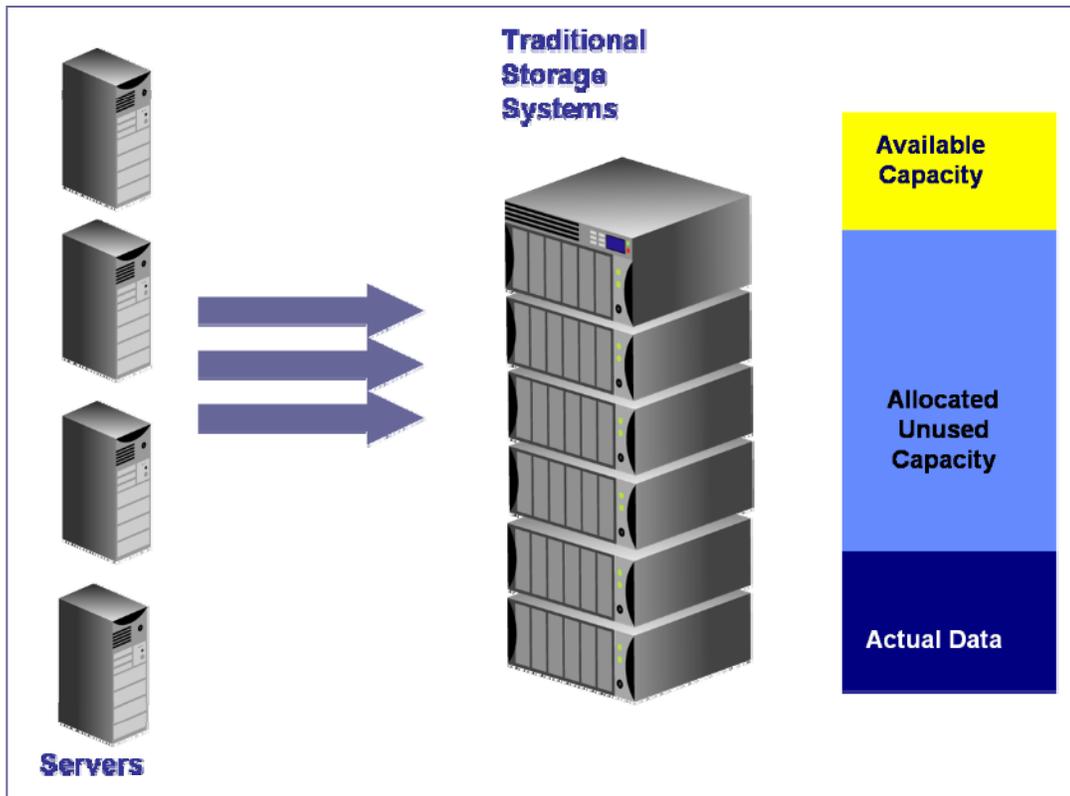
Thin Provisioning
Using thin provisioning in the same example, the system administrator can support a far greater number of servers or applications.

1. The customer's storage system has 3 TB of usable capacity, as above.
2. The system administrator uses the thin provisioning method. He or she creates a 500 GB volume for each of the application servers. Each volume contains 200 GB of actual data.
3. While the storage system has 1 TB of available capacity left, 2 TB of actual data is stored on the storage system.
4. The system administrator is able to connect 10 servers to the storage systems and has plenty of capacity to add more if required.

Figure Four: Traditional Server Support
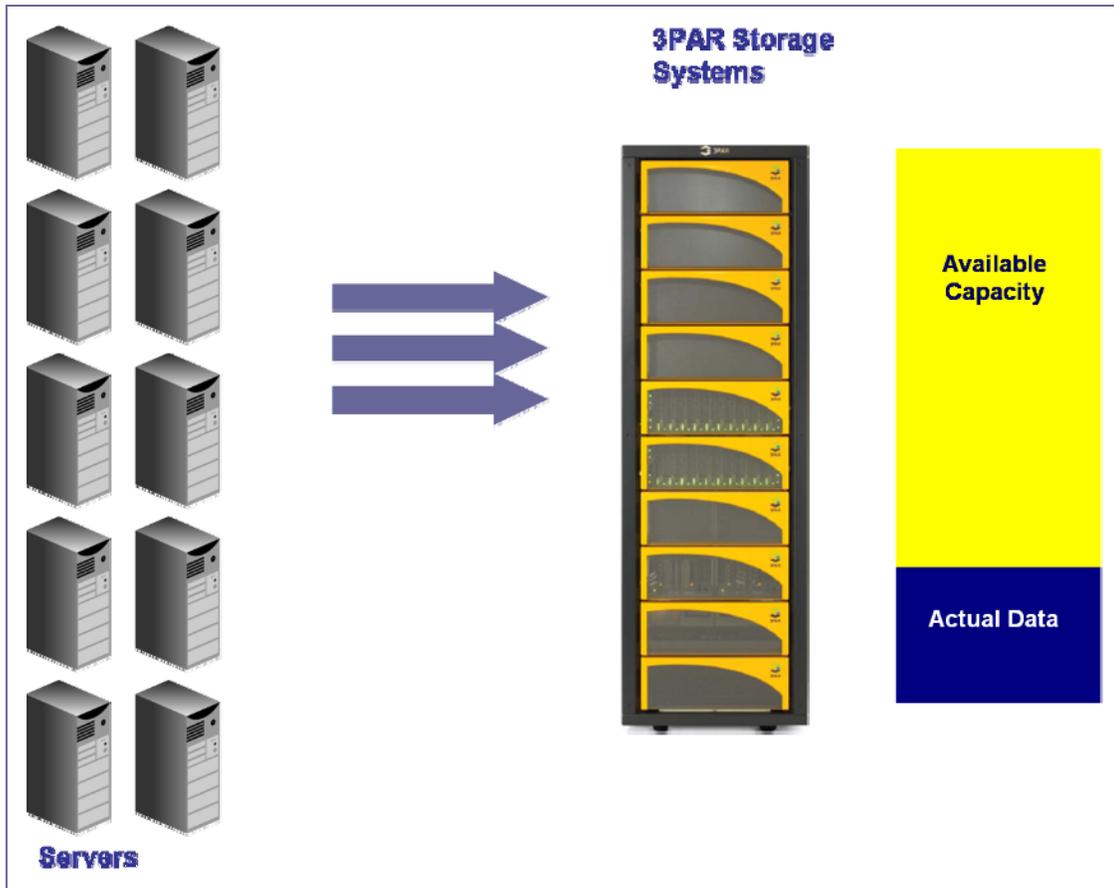
Another major benefit provided by thin provisioning is the ability for customers to buy less storage initially and over the life of the storage system. Traditional provisioning requires system administrators to calculate the volume size of their applications based on the actual data, anticipated growth of each volume over time, plus additional capacity as a buffer just in case the data grows faster than expected.

Thin provisioning requires system administrators to buy a pool of storage large enough to support the actual data and capacity based on short-term requirements – three months, six months, or even a year. As the storage pool is used and a predetermined threshold is crossed, the system administrator is alerted and can add more capacity as needed transparently and online.

In Figure Four, the amount of actual data and available capacity is the same with both methods but with thin provisioning, no additional storage is needed for the allocated unused storage capacity. ESG has seen customers reduce their initial capacity requirements to as low as 25 percent of what they originally would have needed to acquire with traditional provisioning.

Figure Five: Capacity Requirements
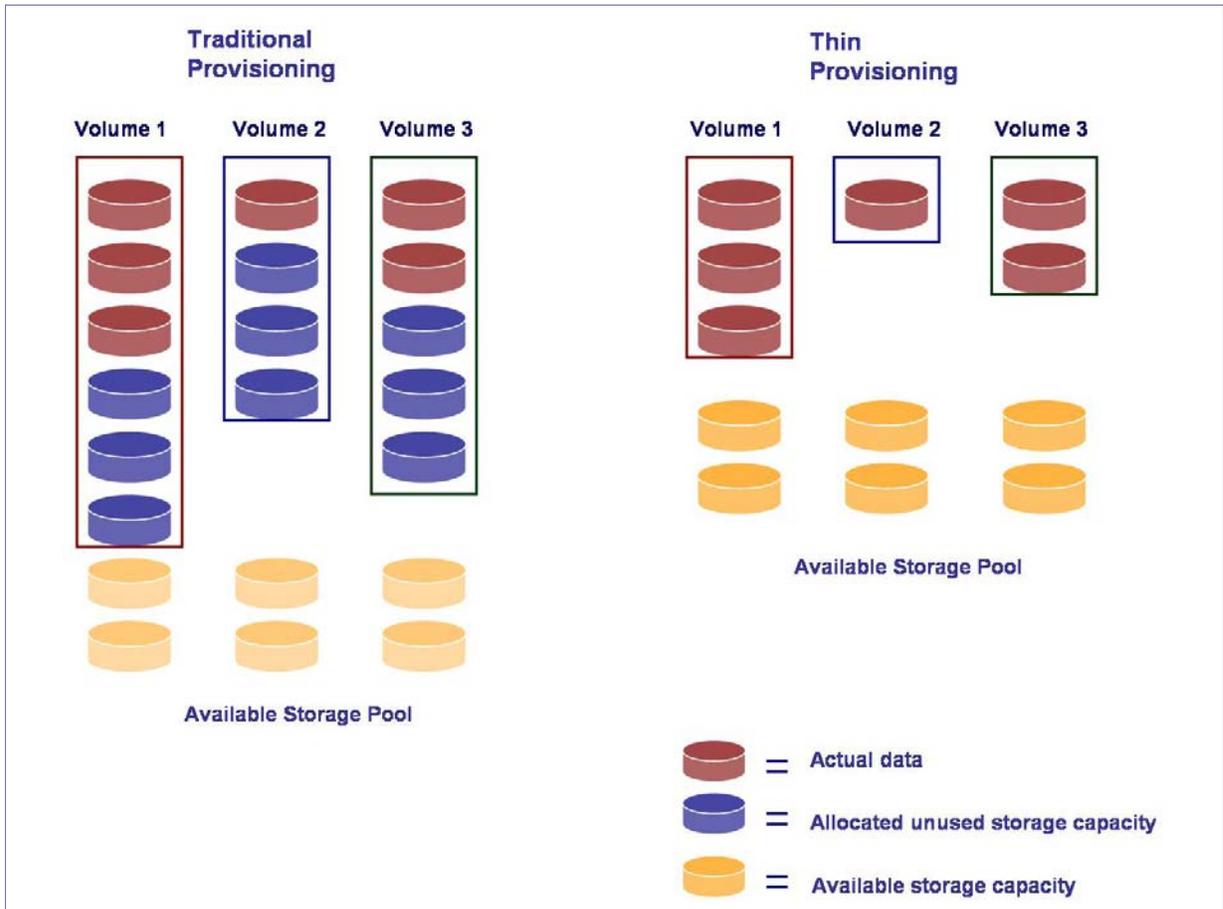
Instead of anticipating the behavior of users and applications with a predetermined amount of capacity, system administrators instead add capacity as needed.   This results in requiring less storage initially and for the life of the storage system.

Thin Provisioning and Copy Services

Another example of how thin provisioning reduces cost and actually improves efficiency is through copy services that are thin provisioning aware.   Traditionally, local and remote mirroring solutions copy entire volumes, which include actual data and allocated but unused capacity as well. If a volume contains 2 TB with 100 GB of actual data, the local and remote mirroring solutions copy all 2 TB. Customers will often make more than one copy of primary data in order to run tests, perform data mining, or for disaster recovery purposes. This can result in three or four or more copies of volumes with over 8 TB of capacity required for only 400 GB of actual data. Then multiply that number by dozens or even hundreds of volumes. The cost savings, impact on resources and performance, and storage administration reduction can be substantial. **In this example, the amount of capacity required is 20 times less using thin provisioning versus traditional methods**.

Objections and Responses

The critics of thin provisioning are, not surprisingly, the storage vendors that do not support it. Here are some of the objections and ESG's answers to them:

Objection: Thin provisioning supports bad practices of over-allocation.

ESG's View: This is untrue, since system administrators can still enforce hard limits using thin provisioning. One of the major benefits of thin provisioning is to prevent over-allocation.

Objection: The storage administrator is essentially deceiving the application.

ESG's View: The application is being "deceived." However, we see no issue with this. It is not uncommon for various technologies to spoof applications.

Objection: Thin provisioning requires the system administrator to second-guess the application users' requirements.

ESG's View: This is fundamentally wrong, since thin provisioning takes away the guesswork that is required by traditional provisioning.

Objection: Additional management must be provided to know when to increase actual allocation and provision more storage.

ESG's View: System administrators are alerted when to add more storage capacity. However, this is a simple process. Thin provisioning actually reduces managemen, by simplifying the task of storage provisioning.

3PAR Thin Provisioning

3PAR is a pioneer and leader in creating and advancing thin provisioning-based solutions. Its implementation of thin provisioning is noteworthy for many reasons, including the following:

- The 3PAR InServ Storage Server is an Enterprise-class storage system that has a high level of scalability, performance, protection features, and reliability. InServ supports a clustered architecture, scaling from two to eight nodes in a single cluster.   The InServ provides a utility storage model, allowing system administrators to very simply and easily provision virtual storage systems to dedicate hardware resources to departments and/or applications. 3PAR InServ is made all the more compelling with thin provisioning support as well as these other attributes.

- The 3PAR volume copy technologies, including its Remote Copy (remote mirrors to other InServ storage systems for disaster recovery), Full Copy (full local volume copies), and Virtual Copy (differential copy-on-write snapshots), are fully thin-provisioning-aware.

- 3PAR maintains a single pool of raw capacity for all volumes. This means that only one common pool of raw available capacity must be purchased and managed, as opposed to multiple separate pools, one for each service level.

- The 3PAR thin provisioning implementation is fully automated. Whereas other implementations may require manual tracking, assignment, and service level configuration (i.e., disk type, RAID level, degree of striping, etc.) of the capacity to be over-allocated, 3PAR does this automatically.   This greatly simplifies administration of the thin provisioned environment, especially when users wish to provide multiple levels of service, which would otherwise require separate care and feeding.

- The 3PAR InServ is a massively parallel storage system that ensures a high level of performance levels for thin provisioned volumes as well.   Other storage systems may experience performance issues with the additional metadata redirection required by thin provisioning.

- The useable size of 3PAR thin provisioned volumes can be grown on demand.

- Thresholds and multiple warnings around capacity utilization can be set for 3PAR thin provisioned volumes, groups of thin provisioned volumes, and the InServ system as a whole. Alert and event notices are visible through e-mail, in the CLI & GUI management interfaces, and to 3PAR service and support.   This instrumentation contributes to the safety and ease-of-use of a thin provisioning approach.

- 3PAR thin provisioned volumes, by default, consume physical capacity in a balanced fashion across all disk drives, even as additional resources are added.

- 3PAR thin provisioned volumes can be dynamically converted from one RAID type to another or from one drive type (FC) to another (Nearline) with 3PAR Dynamic Optimization. This optimization can be performed online and non-disruptively.

ESG interviewed two 3PAR customers to get their insights on their actual production implementation of the InServ storage system and its thin provisioning technology.

Customer One: Leading Search Engine

The first 3PAR customer that we spoke to specifically sought out the InServ storage system for its thin provisioning capability. This company was using EMC storage and noticed that they were allocating capacity to a database that they never used. In fact, 50 percent of their storage system was made up of stranded unused capacity. On top of that, the full volume copies they made also replicated these empty blocks, adding insult to injury. After implementing 3PAR and its thin provisioning technology, this company has eliminated its stranded capacity.   Additionally, they have essentially eliminated provisioning administration tasks.   They are very

comfortable with thin provisioning and haven't had any issues or problems. Today they have 48 TB of capacity on 3PAR.

Customer Two: Leading Infrastructure Hosting Company

This company has over 250 TB of 3PAR storage capacity and provides storage services for its clients for different applications. They use thin provisioning in order to lower costs. They provide 32 GB of virtual storage to each of their clients, who typically don't need that much at first. Initially the capacity required by their clients is a small number of gigabytes. Thin provisioning enables the hosting company to keep costs significantly down. The company does not buy storage capacity in advance. In their view, buying capacity in advance is a waste, especially since storage prices go down consistently over time. This concept of just-in-time storage is enabled with thin provisioning.

ESG's View

Thin provisioning ranks at the top of the list of the most important storage system features, along with differential snapshots. The economic benefits and the reduction of storage provisioning tasks are reasons enough to evaluate storage systems that support thin provisioning, giving them a distinct advantage over products that only provide traditional storage provisioning methods. Thin provisioning is just too valuable for customers to ignore and over time as more customers become aware of its benefits, demand will drive support.

For years only a handful of storage systems supported differential snapshots versus other products that used full point-in-time copies. The vendors that once didn't support differential snapshots did all they could to create fear, uncertainty, and doubt, but eventually came around and supported it. And not surprisingly, these same vendors now spend a great deal of time expounding on the virtues of differential snapshots. ESG believes that the same chain of events will occur with thin provisioning. More and more storage vendors will support it, going from opponents to proponents over the next 24 months.

Thin provisioning is "just in time capacity" that essentially eliminates allocated but unused storage or stranded storage. It also greatly simplifies storage provisioning tasks, reducing administration costs. Customers should look at storage systems that support thin provisioning differently than those that do not when evaluating the initial capacity configuration, budgeting and planning capacity growth over time, and calculating the total cost of ownership.

3PAR is clearly the leader and pioneer of thin provisioning. Storage system vendors should embrace the new school of thought that it is important to provide value and not just capacity. 3PAR thin provisioning breaks the traditional model of selling as much capacity as possible but instead minimizes capacity requirements. This shows that 3PAR wants to provide its customers solutions and not the biggest box they can. ESG believes that a new wave in which vendors approach the needs of storage customers by putting more emphasis on minimizing capacity and maximizing efficiency. Thin provisioning is an excellent example of how storage systems vendors can provide this new wave of value to its customers. **The economic impact of thin provisioning can be substantial, saving literally thousands, tens of thousands, and potentially hundreds of thousands of dollars, depending on the scope of the environment.**