American Megatrends

www.ami.com

Whitepaper

06 03 2005

Thin Provisioning

Overview

American Megatrends, Inc.
6145-F Northbelt Parkway
Norcross, GA 30071

TRADEMARK AND COPYRIGHT ACKNOWLEDGMENTS

FOR ADDITIONAL INFORMATION

Call American Megatrends at 1-800-246-8600 for additional information. You can also visit us online at ami.com.

LIMITATIONS OF LIABILITY

In no event shall American Megatrends be held liable for any loss, expenses, or damages of any kind whatsoever, whether direct, indirect, incidental, or consequential, arising from the design or use of this product or the support materials provided with the product.

LIMITED WARRANTY

No warranties are made, either express or implied, with regard to the contents of this work, its merchantability, or fitness for a particular use. American Megatrends assumes no responsibility for errors and omissions or for the uses made of the material contained herein or reader decisions based on such use.

REVISION HISTORY

06/03/2005	Preliminary release

THIN PROVISIONING

THE NEED FOR THIN PROVISIONING

Given the decreasing cost of storage and the increasing cost of administration, the bigger problem in storage today is, unsurprisingly, not storage capacity but storage capacity management.

When an administrator takes a storage system fresh out of a box, he or she is presented with several terabytes of disk space, often with the capacity to add several more terabytes later. Waiting to use this capacity are several users, each of whom may be running various applications requiring dedicated storage.

Initially setting up the system, the administrator is faced with the unpleasant task of making several decisions that must be taken in what is essentially an ad-hoc manner, without any strictly scientific basis or precedent. One decision is that of partitioning the total storage capacity amongst various volumes for different users and applications. Except for the most tightly-controlled of cases, there is simply no way for an administrator to accurately estimate how much storage capacity a particular volume will need. Consequently, estimates are often wildly exaggerated, and the situation more often than not boils down to guesswork, causing space to run out on some volumes while others are sitting on huge amounts of unused space. This space is called allocated-but-unutilized capacity, and is the bane of *exact provisioning* or *hard provisioning*.

Another related problem relates to the fact that an administrator would prefer to over-estimate rather than under-estimate storage capacity, and as a result, he or she would most likely tend to purchase more capacity than is expected to be used. A significant fraction of this capacity may not be used even in one or two years. The administrator has, therefore, purchased several terabytes of space to serve the needs of users for several years in the future, often at an exorbitant price, and two or three years later, finds that most of the space is still unused, and prices of disks have dropped by such a huge amount that the space could now be bought for a tiny fraction of what was initially spent on it.

Exact provisioning is, therefore, a bad idea, and the market demands a better solution than that for a problem that should not have existed in the first place. While the storage industry cannot as of now provide tools that will allow an administrator to estimate capacity more accurately, it can, however, protect the administrator from having to make ill-informed, unnecessary, and dangerous decisions. The mechanism by which this is done is known as ‘thin provisioning’.

Thin provisioning is a mechanism by which administrators do not need to hard-partition their physical capacity into various volumes at the beginning – they may create volumes that have a logical capacity in excess of what is present physically, and then grow their physical capacity as the need arises. Volumes do not ‘own’ space that is allocated to them at startup; rather, they take space out of a common pool as and when they need it, thus sharing the storage capacity on an as-needed basis rather than an ad hoc basis.

There are three major benefits of thin provisioning: avoiding allocated-but-unutilized storage capacity, amortizing storage cost, and providing OS-independent support for capacity expansion. These three advantages are discussed in detail in the sections below.

ALLOCATED-BUT-UNUTILIZED STORAGE CAPACITY

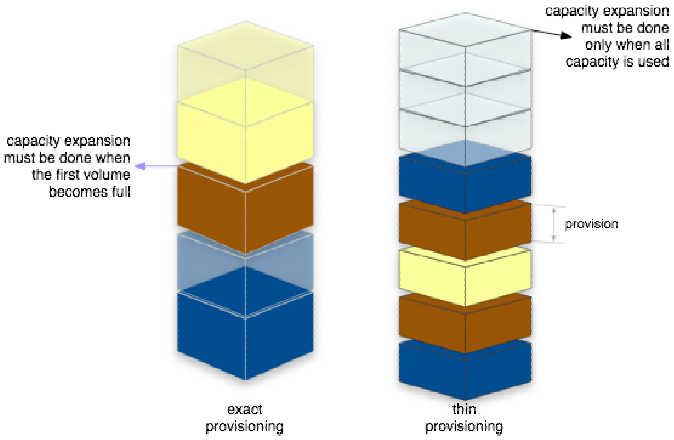
An administrator's decision to hard-partition the available storage space amongst various volumes of different sizes is made difficult, because the way in which volumes are allocated and utilized governs how much space is wasted, and how soon capacity expansion will need to be done.

When storage is exact-provisioned into volumes, the bottleneck for the system becomes the volume that is utilized the most, and which therefore fills up first. As soon as any volume reaches its capacity limit, capacity will need to be expanded, because the volume cannot use unutilized space that has been allocated to other volumes.

Thin provisioning avoids this problem by not allocating space for the volumes up front, but rather, waiting until a write request demands additional space before allocating it for a particular volume. In this way, a volume may be allocated space flexibly from anywhere in the storage system, and the utilization of space grows at a rate that is governed by the cumulative size of all volumes, rather than the rate at which the fastest, smallest volume grows.

The following figure shows a storage server that is being used by three volumes. Exact provisioning will necessitate expansion when the smallest volume becomes full; thin provisioning will allow the administrator to defer expansion until all the storage capacity is exhausted.

FIGURE 1: THIN PROVISIONING AVOIDS ALLOCATED-BUT-UNUTILIZED CAPACITY



AMORTIZING STORAGE COST

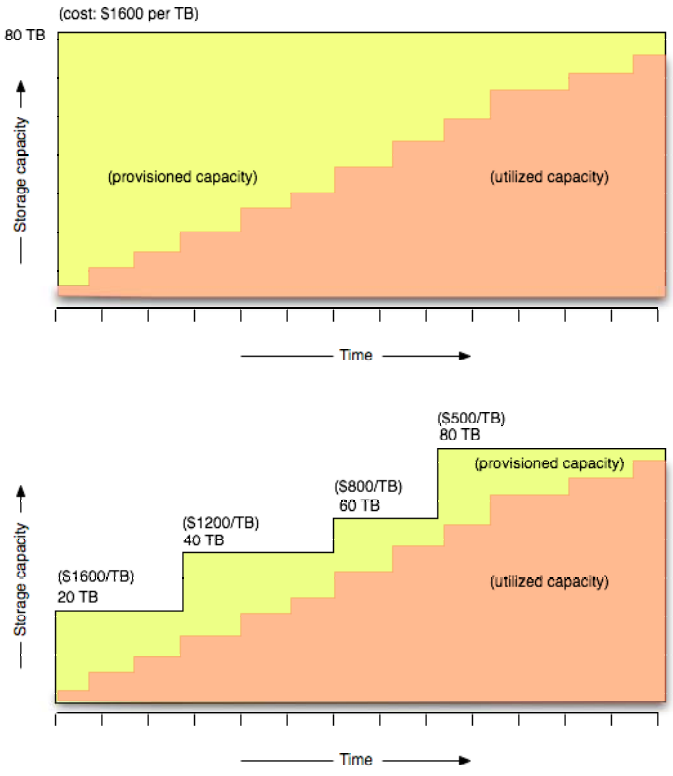
When an administrator purchases storage space to meet predicted demand, he or she is traditionally required to commit to purchasing space that must satisfy the needs of users over a period of several years. Since this space is purchased up front, a large fraction of it will remain unused for several years, while users slowly use up more and more space.

Cost of storage is amongst the most volatile in the IT industry, with prices per gigabyte falling by as much as one half every year. If an administrator purchases a large amount of space up front, instead of purchasing space as and when capacity demand arises, he or she is buying storage at a far higher price than what is on the market when the demand for it is actually present.

Thin provisioning allows the administrator to create volumes that are, logically, of the size that an administrator foresees that a user will need for several years, without having to purchase the physical capacity required for that space during initial setup. As capacity slowly depletes, the administrator may purchase more space – often at dramatically lower prices.

The following figure shows the cost difference between exact provisioned and thin provisioned storage. Since thin provisioned storage allows capacity to grow far more slowly, the cost of storage is amortized over various years, instead of being bought with a lump sum up front.

FIGURE 2: AMORTIZING COST OF STORAGE IN A THIN PROVISIONED SYSTEM



OS-INDEPENDENT SUPPORT FOR CAPACITY EXPANSION

While there are traditional ways in which capacity may be expanded in a storage system, all of them invariably rely upon the operating system's support for dynamically expandable volumes in order to work. Typically, when capacity expansion is required, the administrator adds capacity, makes it visible to the OS on the client side, and allows the OS to use this additional capacity to expand a volume that it has configured.

Several operating systems provide this capability, but there are a substantial number of file systems that do not allow expansion without data loss. Thin provisioning allows the use of these file systems in places where they would traditionally be disallowed due to their limitations.

This is because, in a thin provisioned storage system, it is possible to create volumes that may exceed the actual disk capacity that is available by a large factor. Hence, an administrator may create as large a volume as he or she wishes on a disk array, expose it to the operating system (which will format it with a file system), and, as the operating system utilizes more and more capacity, expand capacity below the volume invisible to the OS.

This form of capacity expansion removes various limitations that operating systems may impose on volumes. For example:

- Microsoft Windows supports capacity expansion using the concept of Dynamic disks. However, dynamic disks have several significant limitations; for example, disks that are discovered via iSCSI may not be configured as dynamic volumes, and clustering is incompatible with dynamic disks in current versions of Windows.
- Most operating systems require human intervention to expand dynamic disks, making it inconvenient for an administrator to do so on systems with a large number of volumes. Thin provisioning can alleviate this problem.
- Thin provisioning also removes limitations that the RAID stack may impose on capacity expansion. For example, some RAID stacks may allow capacity expansion to only take the form of new logical disks, and not permit newly added drives to be used to expand existing logical disks. Since thin provisioning uses a map-based approach to expansion, it may be used to expand volumes without RAID intervention, thereby providing the benefit of spindles without needing re-striping.
- In storage systems with multiple volumes, each volume may host a different file system. Thin provisioning offers a consistent and simplistic view of capacity growth across all the file systems and volumes.
- As a result of OS independence of capacity expansion, any storage system that uses thin provisioning may also have a much smaller test matrix, since it need not be tested with every Operating System that it needs to be tested with.

The ability to expand capacity on a physical layer without affecting the logical layer is, therefore, one of the foremost advantages of thin provisioning.

COLLATERAL BENEFITS OF THIN PROVISIONING

Thin provisioning has several benefits over exact provisioning, as described above. However, the infrastructure required to implement thin provisioning also greatly facilitates a number of additional features, providing collateral benefits that are either very difficult or impossible without it.

One such feature is called *provision tracking*. Provision tracking involves keeping track of which provisions have been written to and which have not, thereby being able to distinguish meaningful data from junk data on the disks. Provision tracking can be exploited in several applications, such as capacity expansion, replication, rebuilds, snapshots etc. Since these operations do not need to be performed on junk data, and need to be performed only on data that was written previously, this knowledge will allow all of these operations (and many others) to be performed significantly faster.

Provision tracking is a powerful idea, and may also be exploited to provide some other features that were typically beyond the scope of storage systems. For example, by appropriately interfacing provision tracking with a file system such as XFS, it is possible for a storage system to be aware of file deletions and system reconfigurations. By performing this, it is possible for a storage system to recover space, perform housekeeping, and provide far better utilization and performance to higher layers.

Thin provisioning is, therefore, the solution to several existing industry problems, and the gateway to providing several highly desirable features. AMI, along with a growing number of vendors in the storage industry, is actively engaged in developing and supporting thin-provisioning based storage systems.