

EMC CLARiiON Virtual Provisioning

Applied Technology

Abstract

This white paper discusses the benefits provided by Virtual Provisioning. It describes how EMC® CLARiiON® implements Virtual Provisioning using thin pools and thin LUNs, and how storage space is allocated and monitored. The paper also discusses which application types are most suitable for Virtual Provisioning.

June 2008

Copyright © 2008 EMC Corporation. All rights reserved.

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED “AS IS.” EMC CORPORATION MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

For the most up-to-date listing of EMC product names, see EMC Corporation Trademarks on EMC.com

All other trademarks used herein are the property of their respective owners.

Part Number H5512

Table of Contents

Executive summary	4
Introduction	4
Audience	4
Terminology	5
Business requirements	6
What is Virtual Provisioning?	6
Thin pools.....	7
What is a thin pool?	7
Monitoring, adding, and deleting pool capacity.....	8
Virtual LUNs	9
What is a virtual LUN?	9
Attributes	10
Architecture and features.....	10
Virtual Provisioning and applications.....	11
General guidelines	11
Oracle.....	12
Microsoft SQL Server.....	12
Microsoft Exchange	13
VMware	13
Traditional LUNs versus thin LUNs.....	13
LUN migration	14
CLARiiON Virtual Provisioning phased implementation.....	14
Conclusion	14
Appendix A: Thin pool properties	16
Appendix B: Thin LUN properties	17

Executive summary

Virtual Provisioning enables organizations to reduce storage costs by increasing capacity utilization, simplifying storage management, and reducing downtime. Related benefits include reducing power and cooling requirements and delaying capital expenditures. EMC® CLARiiON® Virtual Provisioning integrates with existing CLARiiON storage management and replication software.

Most CLARiiON software products operate the same on thin LUNs and traditional LUNs. This includes the ability to migrate thin LUNs to traditional LUNs and vice versa. The flexibility inherent in CLARiiON Virtual Provisioning enables customers to use the best approach for their specific application and operating requirements. It adds a new dimension to tiered storage in the array, without disrupting organizational processes.

Introduction

It is challenging for storage administrators to forecast how much storage space will be required by the various applications in their data centers. Administrators typically allocate space based on anticipated storage growth. They do this to reduce the need to add storage later on while applications are running. This can result in overprovisioning storage capacity. Overprovisioning leads to higher costs and increased power, cooling, and floor space requirements. Even with careful planning, it may still be necessary to provision additional storage in the future, which may require application downtime depending on the operating systems involved.

The term *Virtual Provisioning* applies to EMC's family of provisioning technologies. Virtual Provisioning builds on "thin provisioning" functionality, which presents more storage to an application than is physically available. Additional physical storage is allocated automatically when writing new data blocks. Thin provisioning improves storage capacity utilization and simplifies storage management by presenting an application with sufficient capacity for an extended period of time. This reduces the time and effort required to provision additional storage and avoids provisioning storage that may not be needed.

CLARiiON Virtual Provisioning enables customers to use the same Navisphere® Manager GUI and Secure CLI commands on thin LUNs as they have been using on traditional LUNs. CLARiiON replication products, Navisphere Analyzer, and Navisphere Quality of Service Manager work seamlessly across thin LUNs and traditional LUNs. Using virtual LUN migration, users can migrate thin LUNs to traditional LUNs and vice versa. This flexibility is what separates CLARiiON Virtual Provisioning from typical thin provisioning implementations. Throughout the paper, the term virtual LUN denotes a CLARiiON thin LUN.

This white paper discusses:

- How Virtual Provisioning delivers space efficiency and makes it easier to provision storage.
- How CLARiiON implements Virtual Provisioning.
- How thin LUNs and thin pools compare with traditional LUNs and RAID groups.
- The types of file systems and applications that are best suited for Virtual Provisioning.

Audience

This white paper is intended for IT planners, storage architects, administrators, and others involved in evaluating, managing, operating, or designing CLARiiON storage systems.

Terminology

The following terminology appears in this white paper:

% Full — The percentage of pool capacity that is currently allocated. It is calculated using this formula:
$$\% \text{ Full} = \text{Allocated capacity} / \text{usable capacity}$$

% Full Threshold — A parameter that is set by the user. The system generates an alert when this threshold is exceeded.

Allocated capacity — The amount of actual physical thin pool space that is currently allocated for thin LUNs.

Available capacity — The amount of actual physical thin pool space that is currently *not* allocated for thin LUNs.

Consumed capacity — The amount of the thin pool that has been reserved and/or allocated for all the thin LUNs in the pool.

MetaLUN — A collection of traditional LUNs offered to a host as a single LUN.

Oversubscribed capacity — The amount of usable pool capacity configured for thin LUNs that exceeds the actual usable pool capacity.

RAID group — A type of storage pool that contains a set of disks on which traditional LUNs and metaLUNs can be created.

Storage pool — A general term used to describe RAID groups and thin pools. In the Navisphere Manager GUI, the storage pool node contains RAID groups and thin pool nodes.

Subscribed capacity — The total amount of thin LUN user capacity configured in the pool. This number can be greater than the usable pool capacity. The usable pool capacity can be expanded so that it can accommodate the subscribed user capacity as more space is allocated for the thin LUNs in the pool.

Thin LUN — A logical unit of storage where physical space allocated on the storage system may be less than the user capacity seen by the host server.

Thin pool — A group of disk drives used specifically by thin LUNs. There may be 0 or more thin pools on a system. Disks may be a member of no more than one thin pool. Disks that are in a thin pool cannot also be in a RAID group.

Threshold alert — An alert issued when the % Full Threshold has been exceeded.

Total user capacity — The total capacity seen by all hosts using a thin pool.

Traditional LUN — A logical unit of storage that can span a number of disks on a storage system but looks like a single disk or partition to the server. The amount of physical space allocated is the same as the user capacity seen by the host server.

Usable pool capacity — Pool capacity measured as raw capacity minus overhead (RAID overhead and mapping overhead).

User capacity — This is also referred to as reported capacity. This is the size of the thin LUN as it appears to the host. This term also applies to traditional LUNs, where allocated capacity equals user capacity.

Business requirements

Organizations, both large and small, need to reduce the cost of managing their storage infrastructure while meeting rigorous service level requirements and accommodating explosive storage capacity growth.

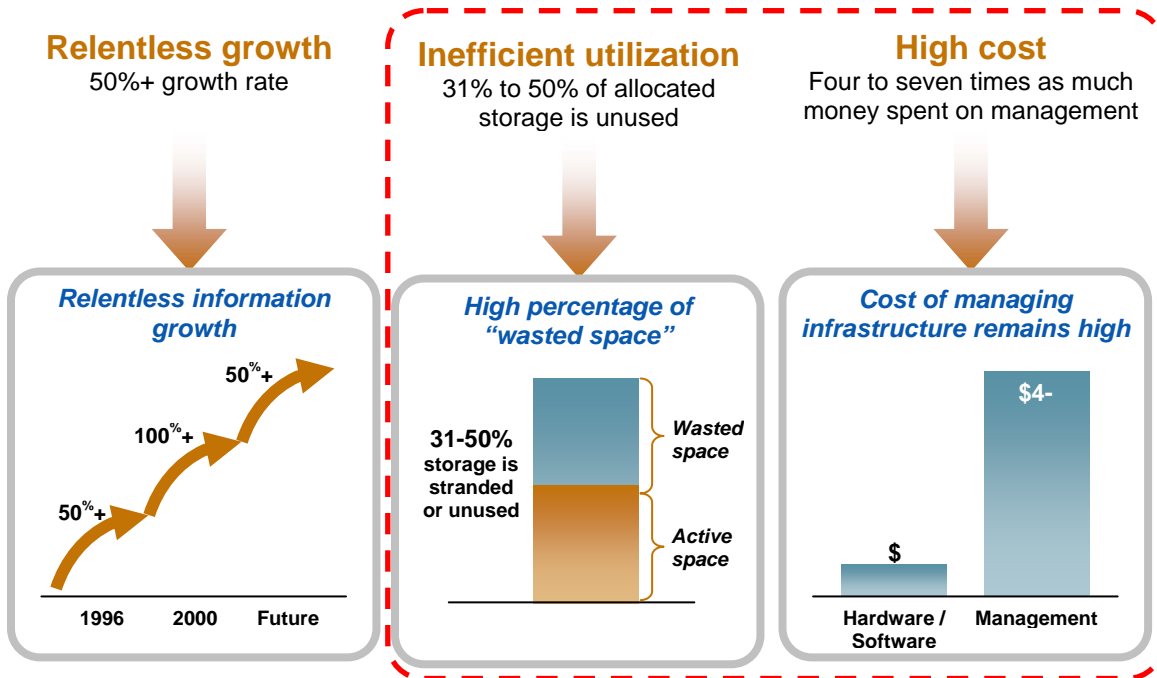


Figure 1. Information management challenges

Several business objectives have drawn increasing focus:

- *Reducing the cost of storage administration*
“Ease of use” initiatives span multiple aspects of storage processes, including staff training, initial storage provisioning, the addition of new storage, and the management and monitoring of storage systems. Virtual Provisioning reduces the ongoing staff time required to repeatedly add storage capacity.
- *Maximizing the utilization of storage assets*
Organizations need to accommodate growth by drawing more value from the same or fewer storage resources. Operational efficiency remains an ongoing challenge, as organizations often overallocate storage to applications to reduce the risk of outage and the need to reprovision later on.
- *Reducing capital expenditures and ongoing costs*
Virtual Provisioning reduces capital costs by delivering storage capacity on-demand. Ongoing costs are reduced because fewer disks consume less power and cooling, and less floor space.

What is Virtual Provisioning?

Storage provisioning is the process of assigning storage resources to meet the capacity, availability, and performance needs of applications. With traditional provisioning, the amount of storage allocated to an application is equal to the amount of physical storage that is actually allocated for that application on the storage system. Figure 2 and Figure 3 on page 7 show the differences between traditional and virtual provisioning.

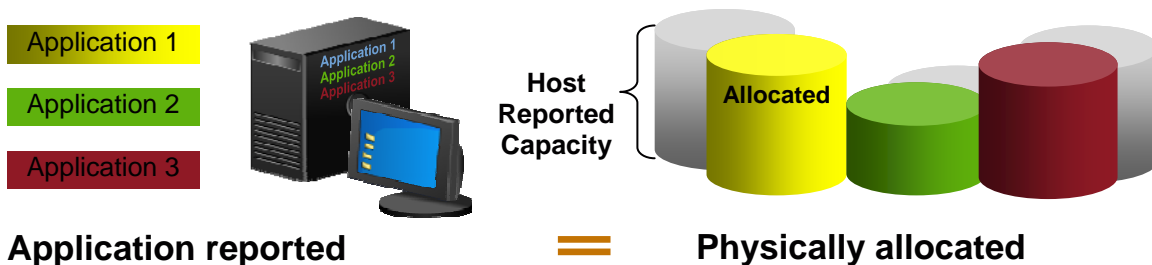


Figure 2. Traditional storage provisioning

With thin provisioning, *user capacity* (storage perceived by the application) is *larger* than the actual allocated space on the storage system. This simplifies the creation and allocation of storage capacity. The provisioning decision is not bound by currently available physical storage, but is assigned to the server in a capacity-on-demand fashion from a shared storage pool. The storage administrator monitors and replenishes each storage pool, not each LUN.

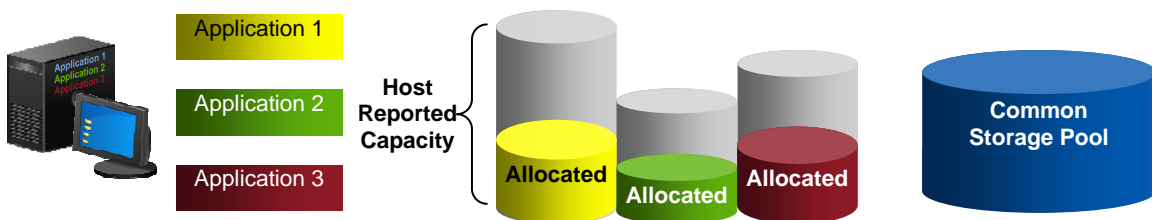


Figure 3. Virtual (thin) provisioning

With traditional provisioning, the storage administrator works with business people who forecast the amount of storage they will require. There is a tendency for these forecasts to be inflated. In some companies, an application administrator may monitor storage space and ask the storage administrator to provision additional storage. The storage administrator must rely on timely and accurate communications from various applications people to effectively manage storage space utilization.

With Virtual Provisioning, the storage administrator can configure more storage to a server than is actually allocated on the storage system. The storage administrator then monitors the actual physical storage space and adds disk drives to the pool as required.

In either case, storage utilization must be carefully monitored. CLARiiON provides the storage administrator with specific monitoring and alerting capabilities to help manage these risks.

Thin pools

What is a thin pool?

A thin pool is a collection of disks that are dedicated for use by virtual (thin) LUNs. A thin pool is somewhat analogous to a CLARiiON RAID group. Thin pools can contain a few disks or as many disks as a storage system can contain, whereas RAID groups are limited to 16 disks. Thin pools are:

- Simple to create – requiring only three user inputs:
 - Pool Name: For example, “Test & Dev Pool 2”
 - Resources: Number of disks or capacity

- Protection level: RAID 5 or 6
- Flexible – they can include:
 - One to many pools per storage system
 - Small pools or very large pools
 - Any Fibre Channel or SATA disk drive
- Simple to modify – you can:
 - Expand the pool size by adding drives to the pool
 - Contract the pool size by removing drives from the pool
- Simple to monitor, due to:
 - % Full Threshold Alerts
 - Storage capacity reports

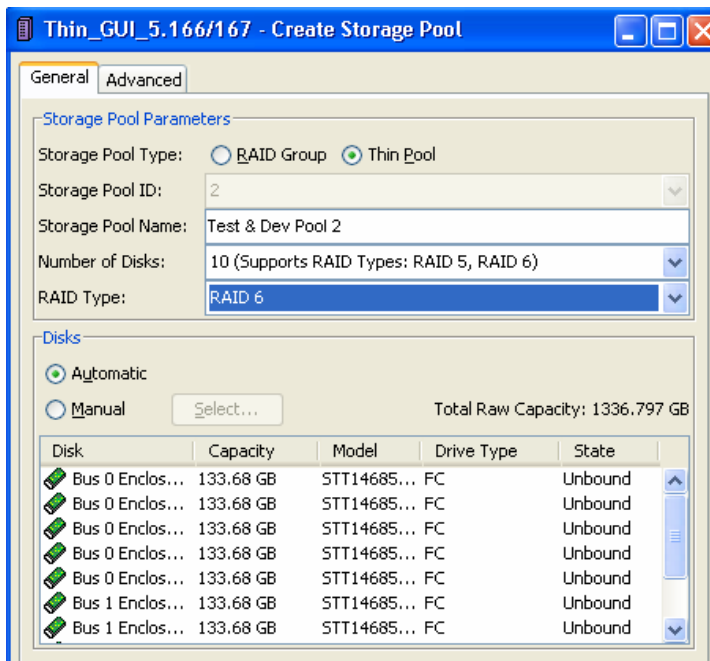


Figure 4. Create Storage Pool dialog box

Monitoring, adding, and deleting pool capacity

Usable pool capacity is the total physical capacity available to all LUNs in the pool. Allocated capacity is the total physical capacity currently assigned to all thin LUNs. Subscribed capacity is the total host reported capacity supported by the pool. The value for % Full Threshold (Allocated Capacity/Usable Pool Capacity) is used to trigger an alert. The initial alert value is user settable. Increasingly serious alerts will be sent on each successive 5% increment. The last two built-in thresholds will upgrade the alert severity to **Critical**.

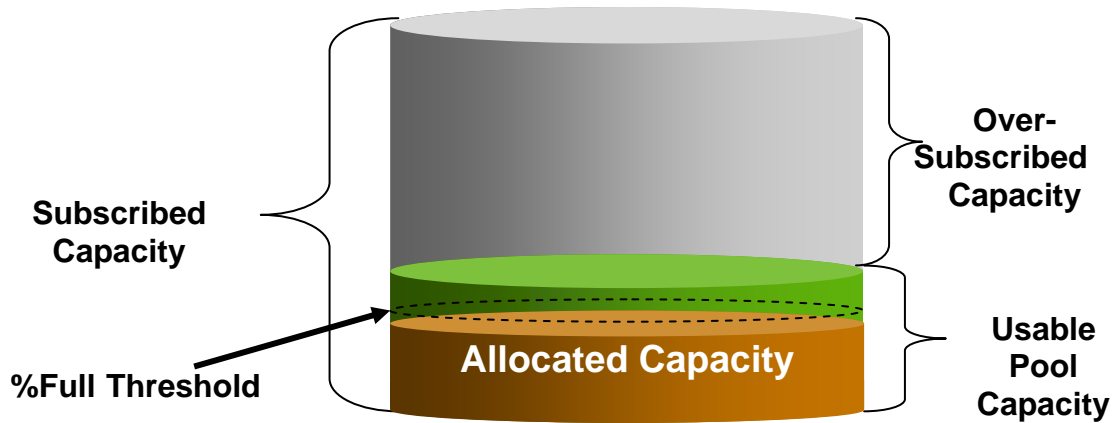


Figure 5. Pool % full threshold

Adding drives to the pool nondisruptively increases available usable pool capacity for all thin LUNs in the pool. Allocated capacity is reclaimed by the pool when LUNs are deleted. There is no need to defrag.

Reserved capacity can be set at the LUN level if desired ($Consumed\ capacity = Allocated\ capacity + Reserved\ capacity$). Reserved capacity can be guaranteed to one or more thin LUNs, but is not allocated. It is reserved space that is not available for use by other LUNs in the thin pool. Reserved capacity can be reclaimed nondisruptively. The ability to reserve capacity for one or more thin LUNs will be added in the second phase of CLARiiON Virtual Provisioning (explained later in this paper).

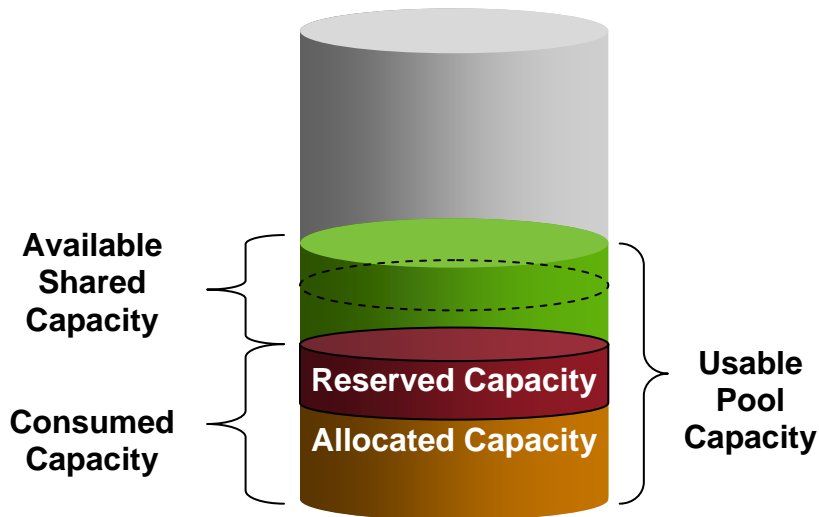


Figure 6. Reserved capacity

Virtual LUNs

What is a virtual LUN?

A CLARiiON virtual (thin) LUN is similar to a traditional LUN in many ways. CLARiiON customers use the same Navisphere Manager GUI operations and Secure CLI commands on virtual LUNs as they have been using on traditional LUNs. Existing CLI scripts continue to work, and all user-oriented FLARE[®]

functions work the same, including underlying data integrity features, local replication, and LUN properties information. Exceptions include MirrorView™ and SAN Copy™, which are not supported for thin LUNs in Phase 1. Navisphere Taskbar wizards maintain a similar look and feel. Virtual LUNs work with applications such as Navisphere Analyzer and Quality of Service Manager. Features such as hot sparing and proactive sparing operate in the same way. It is also possible to migrate a traditional LUN to a virtual LUN and vice versa. This flexibility separates CLARiiON Virtual Provisioning from typical thin provisioning implementations.

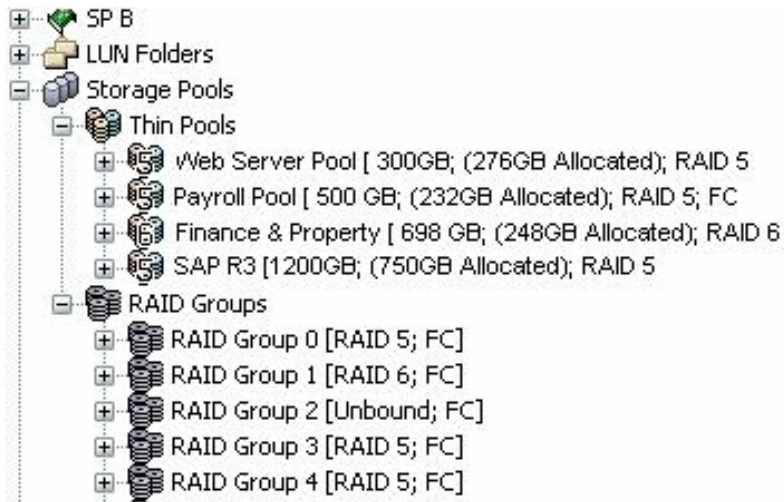


Figure 7. Storage pools in Navisphere Manager

Attributes

The primary difference between traditional and virtual LUNs is that virtual LUNs can consume less physical space on the storage system. Virtual LUNs are:

- Simple to create, with three inputs:
 - Thin pool name
 - Host to provision
 - Amount of host visible capacity

- Simple to manage with LUNs that can expand or shrink

- Easy to use because the system automatically:
 - Manages the drives within the pool according to CLARiiON best practices
 - Distributes data throughout the pool for optimal performance
 - Maps data in 8K chunks (LBA extents) for maximum space efficiency

- Simple to monitor with:
 - Allocated capacity
 - Reports

Architecture and features

Specialized software known as the Mapping Service manages the placement and use of the physical storage used by virtual LUNs. Data is written to 8K chunks (*extents*) and is optimally packed. This makes

configuring virtual LUNs easy, because the underlying software makes all the decisions about how to lay out actual storage blocks across the disks in a thin pool. Less experienced storage administrators will benefit from not having to be directly involved in the details of configuring storage. The Mapping Service performs these functions adhering to performance best practices.

Storage administrators can manage LUN space utilization with the following thresholds:

- Optional % Full Threshold Alerts
- Optional Allocation Limit to control a runaway process

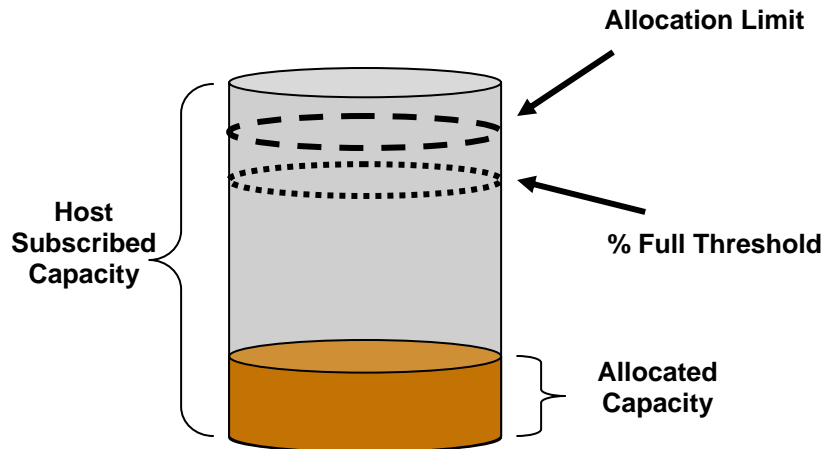


Figure 8. LUN Threshold and Allocation Limit

Storage administrators can monitor thin pools and add additional storage to them as required. In addition, they can reserve up to the host capacity of the thin LUN. The criteria for reserving thin LUN storage include the following:

- The storage system must have available shared storage.
- Space is guaranteed but not necessarily allocated.
- Space is consumed (allocated) on a “use-first” basis.
- Space can be nondisruptively returned to the pool at any time.

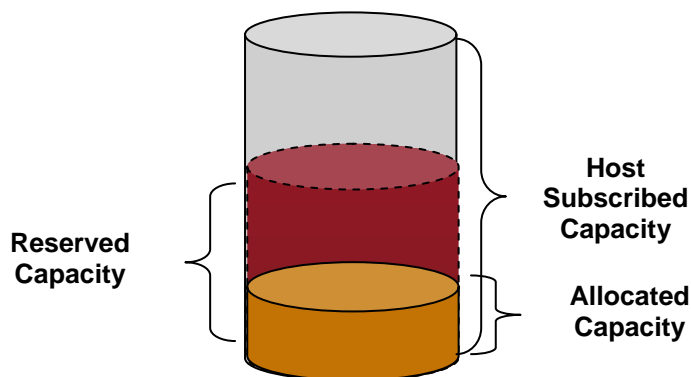


Figure 9. Reserved capacity

Virtual Provisioning and applications

General guidelines

Virtual Provisioning is most appropriate for applications that have the following characteristics:

-
- *Predictable and controllable growth*
Applications with well-understood capacity growth requirements will help organizations avoid unexpected consumption of all space in the pool, and enable accurate initial provisioning of thin devices.
 - *Virtual Provisioning “friendly” environments*
To improve capacity utilization with file systems, you should apply Virtual Provisioning only when files are *not* frequently added or deleted. Many file systems do not reuse the space associated with deleted files, which reduces the capacity utilization benefits of Virtual Provisioning.
You should carefully weigh the space consumption characteristics of databases before using Virtual Provisioning, to improve their capacity utilization. Some databases preallocate space and write zeros to it. This preallocated, but unused, space cannot be shared in a thin pool, reducing or eliminating the capacity utilization benefits.
Independent of these factors, Virtual Provisioning can still improve ease of use and improve performance via wide striping in some cases. But capacity utilization will depend on the allocation methods used by specific file systems and databases, as outlined previously.
 - *“General purpose” performance requirements*
Virtual Provisioning is appropriate for applications in which some performance variability can be tolerated. Some workloads will see performance improvements from wide striping with Virtual Provisioning. However, when multiple thin devices contend for shared spindle resources in a given pool, and when utilization reaches higher levels, the performance for a given application can become more variable. Navisphere Quality of Service Manager can be used to manage resource contention within the pool as well as between LUNs in different thin pools and RAID groups.
 - *Environments that need flexible provisioning (for example, test and development)*
Virtual Provisioning can be an effective means of improving ease of use and capacity utilization for lower storage tiers such as test and development.
 - *Document repositories*
Document repositories with rapidly rising capacity requirements can benefit greatly from the improved capacity utilization offered by Virtual Provisioning, provided their environments meet the previously outlined criteria.
 - *Software development/source code*
Many organizations will see an opportunity to lower TCO by improving ease of use and capacity utilization for storage associated with software development, because these development activities can usually tolerate some level of performance variability.

Oracle

Traditional database files as well as Oracle ASM include an Auto Extend feature that can take advantage of thin provisioning. Without Auto Extend, using CREATE DATABASE with traditional database files would cause Oracle to write zeros to all blocks of a tablespace file. Oracle DBAs can elect to use traditional LUNs in small RAID groups for log files and place the Oracle database volumes in a thin pool. This places log and database volumes on separate disks (a best practice) while still providing the benefits of Virtual Provisioning for the database volumes.

Microsoft SQL Server

Microsoft SQL Server 2005 introduced new functionality that altered behavior during the database creation phase. Earlier versions of SQL Server fully initialized all data file and transaction log file components, writing to every page in all database files and log files. With SQL Server 2005, the database creation phase no longer requires initializing all data files if *Instant File Initialization* can be utilized by the SQL Server environment. EMC recommends deployment of Microsoft SQL Server 2005 (and later) with Instant File Initialization functionality when using Virtual Provisioning. Earlier versions of SQL Server that do not

include Instant File Initialization will result in full extent allocations for the physical files located on thin devices.

Microsoft Exchange

By default, Exchange 2007 database files start between 2 MB and 4 MB in size, and incrementally grow by 2 MB as additional space is needed. This auto-extend behavior is efficient from a thin pool perspective, as only space immediately needed by the database file is allocated. While the default database file extension size of 2 MB (256 pages each of 8K) can be changed in the Active Directory, there is no reason to make any modifications to the default behavior.

VMware

The VMware Virtual Machine File System (VMFS) has some useful characteristics when viewed in a Virtual Provisioning context. First, a minimal number of thin extents is allocated from the thin pool when a VMware file system is created on virtually provisioned devices. The amount of storage required to store the file system metadata is a function of the size of the thin device. VMware's file system does not write all of its metadata to disks on creation. The VMware file system formats and uses the reserved area for metadata as requirements arise.

Using the “zeroedthick” allocation method in VMware, the storage required for the virtual disks is reserved in the datastore, but the VMware kernel does not initialize all the blocks. The blocks are initialized by the guest operating system as write activities to previously uninitialized blocks are performed. The VMware kernel provides a number of allocation mechanisms for creating virtual disks, and not all of them are Virtual Provisioning friendly. The “eagerzeroedthick” format is not ideal for use with virtually provisioned devices. The “thin” allocation policy is somewhat Virtual Provisioning friendly. However, the risk of exceeding the thin pool capacity is much higher when virtual disks are allocated using this policy, since the oversubscription to physical storage occurs at two independent layers that do not communicate with each other.

The VMware **cp copy** command is thin friendly. DRS, VMotion, and “cold” VM migration are unaffected. VM Clones and Templates are problematic. VM Cloning fully allocates all blocks. There is currently no workaround for this. VMware Templates also allocate all blocks. The workaround is to shrink VMDKs before creating a template and use the “Compact” option.

Traditional LUNs versus thin LUNs

The key is to understand your application requirements and select the approach that meets your needs. If conditions change, you have the option of using CLARiiON LUN migration to migrate from thin LUN to traditional LUN or vice versa.

Use RAID groups and traditional LUNs in the following circumstances:

- When microseconds of performance matters!
- For the best and most predictable performance
- For precise data placement
- When you want one method for all applications and services
- When you are not as concerned about space efficiency

Use Virtual Provisioning with thin pools and thin LUNs for the following:

- The best space efficiency
- Easy setup and management
- Minimal host impact
- Energy and capital savings
- Applications where space consumption is difficult to forecast

It is important to remember that thin pools and thin LUNs offer many benefits, but they require careful monitoring and are not suited to every application.

LUN migration

LUN migration operations behave as shown in Figure 10.

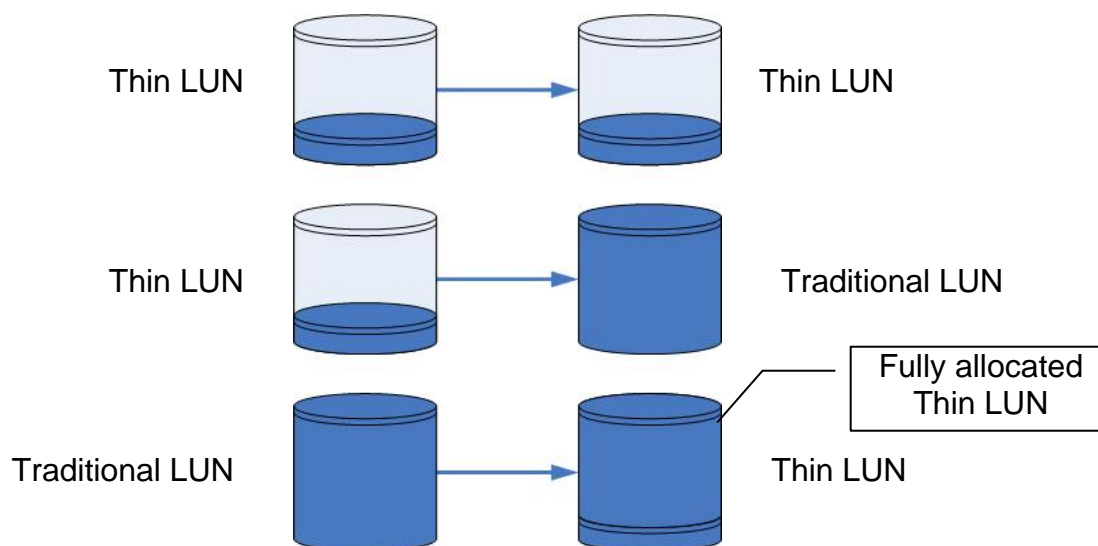


Figure 10. LUN migration

Migrating a traditional LUN to a thin LUN does not save space initially. But the new thin LUN can now be expanded (see Phase 2 below) so that configured user capacity exceeds allocated capacity, thereby adding some of the benefits of Virtual Provisioning in the process.

CLARiiON Virtual Provisioning phased implementation

CLARiiON Virtual Provisioning will be implemented in phases. The first two phases are planned as follows.

- Phase 1 will include LUN migration support as well as support for SnapView snapshots and clones. Navisphere Analyzer and Quality of Service Manager will be supported. Dynamic thin pool expansion will also be supported.
- Phase 2 will include support for MirrorView/S, MirrorView/A, and SAN Copy as well as thin pool reservations, thin pool shrink, and thin LUN expansion.

Conclusion

When implemented appropriately, Virtual Provisioning can be a powerful complement to organizations' processes and technologies for improving ease of use, enhancing performance, and utilizing storage capacity more efficiently. CLARiiON Virtual Provisioning integrates well with existing management and business continuity technologies, and is an important advancement in capabilities for CLARiiON customers. CLARiiON Virtual Provisioning does the following:

- Saves time:
 - Easy to create pools and LUNs
 - Easy to monitor and manage

-
- Reduces provisioning uncertainty:
 - Decisions are easy to modify
 - No impact on host servers
 - Reduces upfront investment and saves energy:
 - Highly space-efficient
 - Multiple applications share resources
 - Physical storage can be added as required
 - Builds on existing CLARiiON features:
 - Migration is supported between all types of LUNs
 - Thin replicas and mirrors
 - Navisphere Quality of Service Manager and Navisphere Analyzer
 - Navisphere Reports

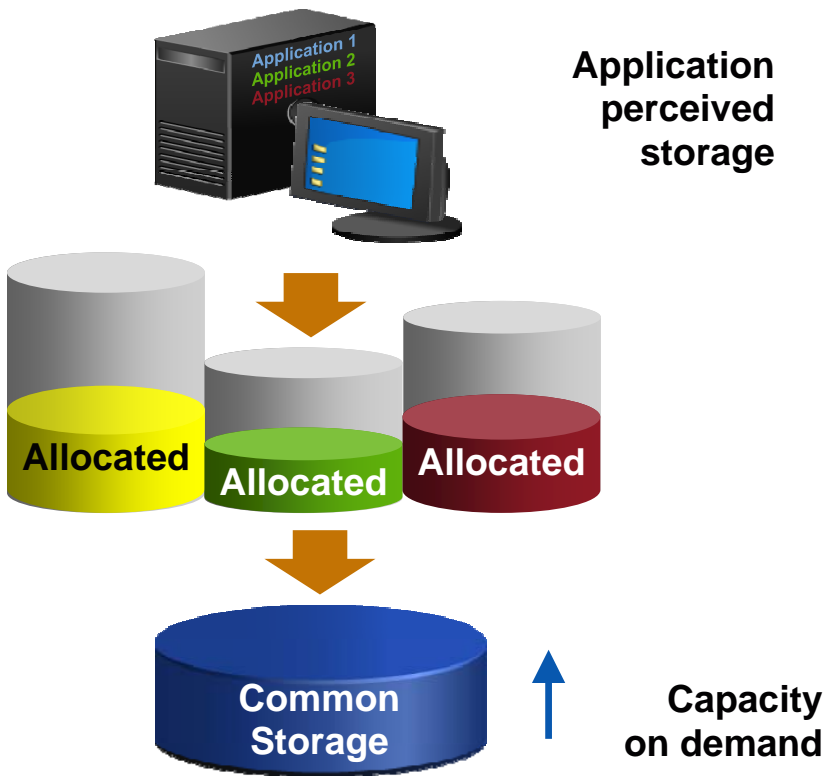


Figure 11. Thin provisioning

Appendix A: Thin pool properties

Thin pool property	Description	User editable?	Value range
Name	Friendly name of the thin pool that is also its unique ID (must be unique across all thin pools on the array) Max length: 255	Yes	ASCII, all printable characters (0x20 to 0x7E) with no leading or trailing spaces
Description	User text describing the thin pool Max length: 255	Yes	n/a
ID	Numeric ID of the thin pool (unsigned 32 bit) Note: This will display as a property but is only used internally. The Name is the ID.	No	0 to 4294967295
RAID Type	RAID protection level of the thin pool	Only at thin pool creation time	5, 6
Disk Type	Type of disks used by the thin pool	Disks selected at pool creation/expansion time	FC, ATA, SATA, Mixed
State	Current state of the thin pool	No	<ul style="list-style-type: none"> • Initializing • Ready • Faulted • Offline • Destroying
State Details	Optional additional information to describe the state of the pool	No	Status code from driver
Raw Capacity	Total amount of physical disk capacity in the pool	Entered at pool creation/expansion time	Total capacity of all physical disks in the pool
User Capacity	Total amount of storage capacity in the pool excluding RAID overhead	No	Raw Capacity minus RAID overhead
Allocated Capacity	Amount of user capacity that is allocated to thin LUNs	No	0 to User Capacity
Available Capacity	Amount of user capacity that can be allocated to thin LUNs	Indirectly by expanding pool	0 to User Capacity
Subscribed Capacity	Total amount of storage capacity subscribed to by all thin LUNs in the pool. This number can be larger than the pool capacity	Indirectly by creating/destroying thin LUNs	0 to Total User Capacity of all thin LUNs in the pool
Operation in Progress	Indicates that a long-running operation is in progress, and may show the percentage completed	No	<ul style="list-style-type: none"> • Initializing • Recovering • Destroying thin LUN

Appendix B: Thin LUN properties

Property	Description	User editable?	Value range
Name	Friendly name of the thin LUN (must be unique across all thin LUNs per array) Max Length: 255	Yes	ASCII, all printable characters (0x20 to 0x7E) with no leading or trailing spaces
ID	Numeric ID of the thin LUN	No	Shared with all LUN types
Unique ID	WWN	No	WWN
Thin Pool	Friendly name of the thin pool that contains this thin LUN	No	Valid thin pool name
RAID Type	RAID protection level of the thin LUN's thin pool	No	5, 6
Drive Type	Type of disks used by the thin LUN's thin pool	No	FC, ATA, SATA, Mixed
State	Current state of the thin LUN	No	<ul style="list-style-type: none"> • Initializing • Ready • Faulted • Offline
State Details	Optional additional information to describe the state of the thin LUN	No	Status code from driver
User Capacity	Size of the thin LUN that is to be presented to the host	Only at LUN creation time	Depends on physical disks in pool
Consumed Capacity	Amount of space of the thin LUN that has been used by user data	No	Depends on physical disks in pool and RAID type
Free Available Pool Capacity	Amount of space in the thin pool that is available for thin LUNs to consume	No	Depends on pool
LUN Type	Type of LUN	No	Thin
Auto Assignment	Enable the array to auto-assign the thin LUN to an SP	Yes	Yes/No
Current Owner	Current ownership of the thin LUN (same as traditional LUN)	No	SP A, SP B
Default Owner	Default ownership of the thin LUN (same as traditional LUN)	Yes	SP A, SP B
Allocation Owner	SP that owns the TLU allocation table	No	SP A, SP B
Alignment Offset	Offset value to align host access	Only at LUN creation time	Same as other LUN types