White Paper

EMC CLARiiON Virtual Provisioning

Applied Technology

Abstract

This white paper discusses the benefits provided by Virtual Provisioning. It describes how EMC® CLARiiON® implements thin 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 thin LUNs.

August 2008

Part Number H5512

Table of Contents

Executive summary

Virtual Provisioning enables organizations to reduce storage costs by increasing capacity utilization, simplifying storage management, and reducing application downtime. Virtual Provisioning also helps companies reduce power and cooling requirements and delay capital expenditures.

The EMC® CLARiiON® CX4 series introduces thin LUN technology that builds on CLARiiON's virtual LUN capabilities and seamlessly integrates with existing CLARiiON management and replication software. With CLARiiON Virtual Provisioning, you can choose between traditional LUNs, metaLUNs, and thin LUNs. The ability to nondisruptively migrate data to different LUN and disk types allows you to deploy the best solution for your changing application and business requirements without incurring downtime.

CLARiiON's thin LUN technology also supports features such as hot sparing, proactive sparing, and the ability to migrate data between thin LUNs, traditional LUNs, and metaLUNs without incurring application downtime. This flexibility separates CLARiiON Virtual Provisioning from typical thin provisioning implementations.

Introduction

One of the biggest challenges facing storage administrators is balancing how much storage space will be required by the various applications in their data centers. Administrators are typically forced to allocate space based on anticipated storage growth. They do this to reduce the management expense and application downtime required to add storage later on. This generally results in the overprovisioning of storage capacity. Overprovisioning leads to higher costs; increased power, cooling, and floor space requirements; and lower capacity utilization rates. Even with careful planning, it may be necessary to provision additional storage in the future. This may require application downtime depending on the operating systems involved.

To address these concerns, the CLARiiON CX4 series introduces thin LUN technology. This technology works with CLARiiON's metaLUNs and traditional LUNs to provide powerful, cost-effective, flexible solutions. CLARiiON thin LUNs present more storage to an application than is physically available. Storage managers are freed from the time-consuming administrative work of deciding how to allocate disk drive capacity. Instead, an array-based mapping service builds and maintains all of the storage structures based on a few high-level user inputs. Disk drives are grouped into storage pools that form the basis for provisioning actions. Physical storage is automatically allocated only 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. When additional physical storage space is required, disk drives can be nondisruptively added to the central storage pool. This reduces the time and effort required to provision additional storage, and avoids provisioning storage that may not be needed.

You manage thin LUNs using the same Navisphere® Manager GUI and Secure CLI commands that you use to manage traditional LUNs and metaLUNs. CLARiiON replication products, Navisphere Analyzer, and Navisphere Quality of Service Manager work seamlessly across thin LUNs, traditional LUNs, and metaLUNs.

This white paper discusses:

- How thin LUNs deliver space efficiency and make it easier to provision storage
- How CLARiiON implements provisioning for thin LUNs
- 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 thin 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:
% Full = Allocated capacity / usable pool 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.

LUN migration — A CLARiiON feature that dynamically migrates data to another LUN or metaLUN without disrupting running applications. LUN migration has many uses, for example:

- To change the type of disk drive (for example, from economical SATA to faster FC, or vice versa)
- To select a RAID type that better matches the current performance or availability requirements
- To re-create a LUN with more disk space

After migration is complete, the destination LUN assumes the identity (World Wide Name and other IDs) of the source LUN, and the source LUN is destroyed. LUN migration can be tuned by setting it to one of four speed settings: Low, Medium, High, or ASAP.

MetaLUN — A collection of traditional LUNs that are striped or concatenated together and presented to a host as a single LUN. A single metaLUN can be striped across any number of disk spindles, providing a much wider range of performance and configuration options. Additional LUNs can be added dynamically, allowing metaLUNs to be expanded on the fly.

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 zero 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.

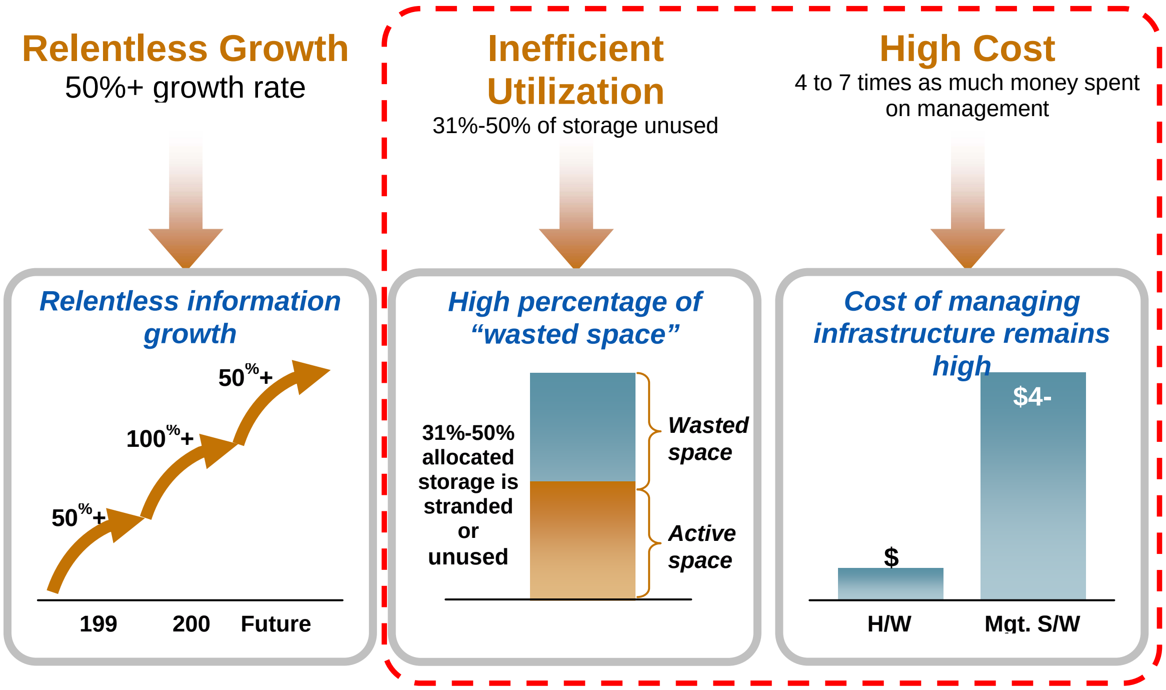

Figure 1. Information management challenges

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.

Comparing traditional LUNs and thin LUNs

Storage provisioning is the process of assigning storage resources to meet the capacity, availability, and performance needs of applications. With traditional storage provisioning, the host reported capacity is equal to the amount of physical storage capacity allocated. The entire amount of physical storage capacity must be present on day one, resulting in low levels of utilization. Recovering underutilized space remains a challenge. Figure 2 and Figure 3 show the differences between traditional and thin provisioning.

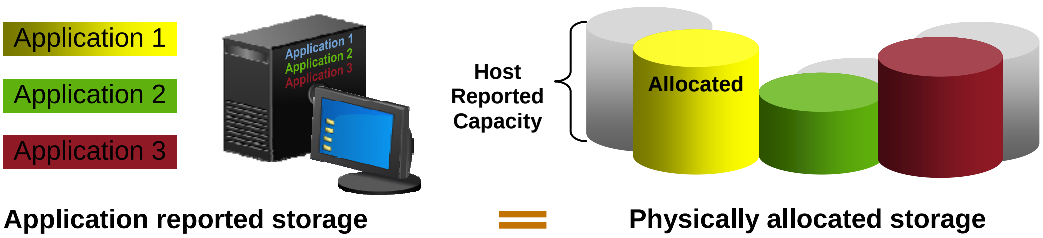

Figure 2. Traditional storage provisioning

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 thin provisioning, the *host reported 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. Thin LUNs can be sized to accommodate growth without regard for currently available assets. Physical storage 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.

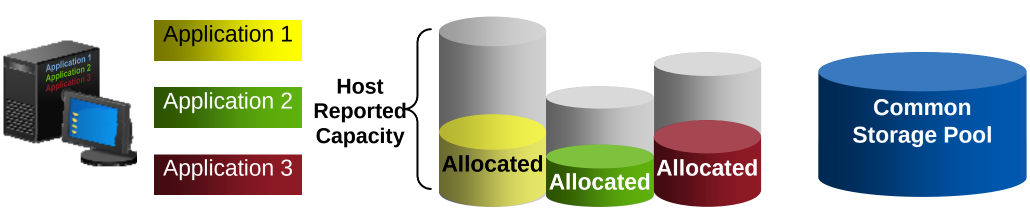

Figure 3. Thin provisioning

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

A thin pool is a collection of disks that are dedicated for use by thin LUNs. A thin pool is somewhat analogous to a CLARiiON RAID group. Thin pools can contain a few disks or hundreds of disks, whereas RAID groups are limited to 16 disks. Thin pools are simple to create because they require only three user inputs:

- Pool Name: For example, “Test & Dev Pool 2”
- Resources: Number of disks
- Protection level: RAID 5 or 6

Thin pools are flexible. They can consist of any supported Fibre Channel or SATA disk drive. CLARiiON can contain one or many thin pools per storage system. The smallest pool size is three drives for RAID 5 and four drives for RAID 6. However, EMC recommends a minimum of five drives for RAID 5 and eight drives for RAID 6. CX4-480 and CX4-960 will support hundreds of drives per thin pool. Actual limits such as disks per pool and thin LUNs per system will be available when Virtual Provisioning becomes available. Thin pools are also easy to modify. You can expand the pool size by adding drives to the pool and contract the pool size by removing drives from the pool. Thin pools are simple to monitor using the “% Full Threshold Alert” setting and storage capacity reports.

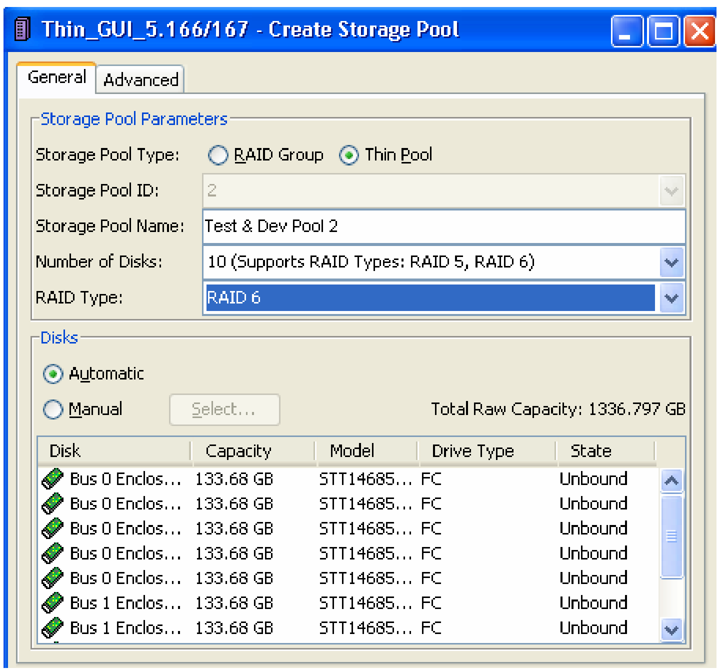

Figure 4. Create Storage Pool dialog box

It is a best practice to limit each thin pool to one type of disk drive. Fibre Channel and SATA drives should be deployed in separate pools. Drives within a thin pool should be rated at the same rpm and be the same size. It is best to provision thin LUNs only to applications that can tolerate some variation in performance. Although thin LUNs are internally striped, performance may vary. A misbehaving application will impact other applications whose LUNs occupy the same spindles within a thin pool. Use traditional LUNs for applications with stringent performance requirements. Use metaLUNs for high-throughput applications that can benefit from spreading I/O over a large number of disk drives striped over multiple RAID groups

Monitoring, adding, and deleting thin pool capacity

Usable pool capacity is the total physical capacity available to all thin 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**.

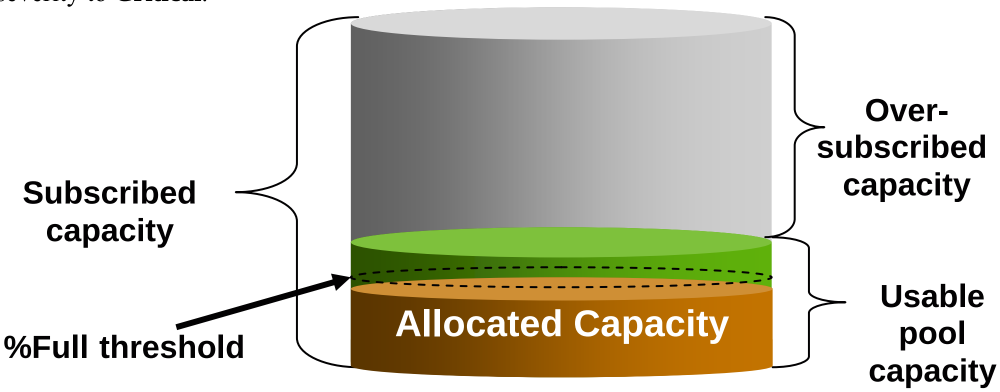

Figure 5. Pool % full threshold

Adding drives to the pool nondisruptively increases 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.

Thin LUNs

A CLARiiON thin LUN is similar to a traditional LUN in many ways. CLARiiON customers use many of the same Navisphere Manager GUI operations and Secure CLI commands on thin LUNs and traditional LUNs. All user-oriented FLARE® functions work the same, including underlying data integrity features, hot sparing, LUN migration, local replication, and LUN properties information. Thin LUNs are available as choices in Navisphere Taskbar wizards, Navisphere Analyzer, and Navisphere Quality of Service Manager. Features such as hot sparing and proactive sparing operate in the same manner. It is also possible to migrate a traditional LUN (or metaLUN) to a thin LUN and vice versa. This flexibility separates CLARiiON thin provisioning from typical thin provisioning implementations.

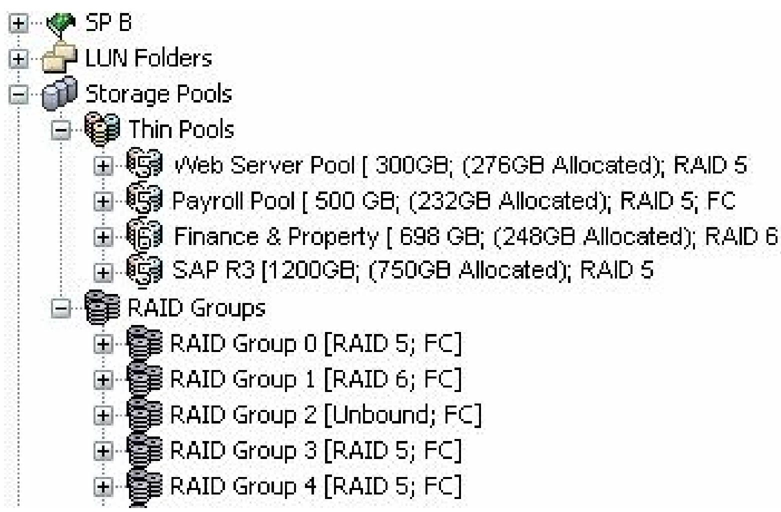

Figure 6. Storage pools in Navisphere Manager

Attributes

The primary difference between traditional and thin LUNs is that thin LUNs can consume less physical space on the storage system. Thin LUNs are simple to create, with three inputs:

- Thin pool name
- Host to provision
- Amount of host visible capacity

Thin LUNs are easy to use because the system automatically manages the drives within the pool according to CLARiiON best practices. The Mapping Service distributes data evenly throughout the pool for optimal performance. Thin LUNs are easy to manage because Navisphere Manager and Navisphere CLI commands work the same for thin LUNs as they do for traditional LUNs. New property screens and reports are available to obtain information about thin LUNs.

Architecture and features

Specialized software known as the Mapping Service manages the placement and use of the physical storage used by thin LUNs. Data is written to 8K chunks (*extents*) and is optimally packed. This makes configuring thin 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 monitor LUN space utilization using the Allocation Limit to control a runaway process.

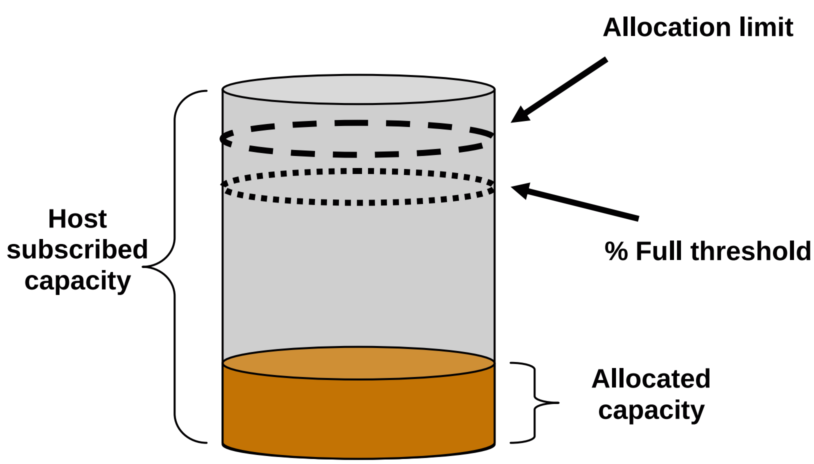

Figure 7. LUN Threshold and Allocation Limit

Storage administrators can monitor thin pools and add additional storage to them as required. In addition, they can reserve up to the user (host visible) 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.

Reserved capacity is guaranteed to a particular thin LUN but not allocated. *Consumed capacity = Allocated capacity + Reserved capacity*. Reserved capacity is not available to any other thin LUN. When additional space is allocated, any remaining reserved capacity is consumed first. Unused reserved capacity can be reclaimed nondisruptively because it has not been allocated. The ability to reserve capacity for one or more thin LUNs will be added in the second phase of CLARiiON Virtual Provisioning, This is discussed later in this paper.

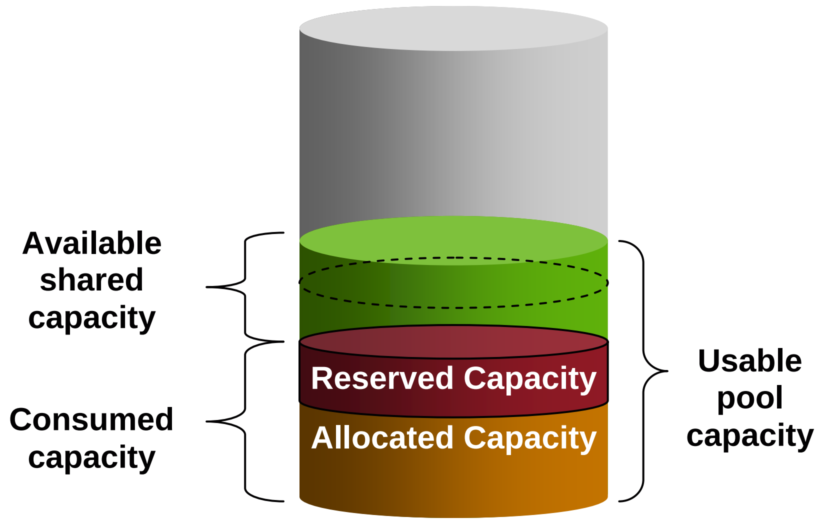

Figure 8. Reserved capacity

Using thin LUNs with applications

Thin LUN 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.

- *Thin Provisioning "friendly" environments*

  To improve capacity utilization with file systems, use thin provisioning only when files are *not* frequently added or deleted. Many file systems do not efficiently reuse the space associated with deleted files, which reduces the capacity utilization benefits of Thin Provisioning.

  You should carefully weigh the space consumption characteristics of databases before using thin provisioning. 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, thin provisioning can still improve ease of use and improve performance via wide data striping in some cases. But capacity utilization will depend on the allocation methods used by specific file systems and databases.

- *"General purpose" performance requirements*

  Thin provisioning is appropriate for applications in which some performance variability can be tolerated. Some workloads will see performance improvements from wide striping with thin 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)*

  Thin provisioning can be an effective way to improve 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 thin provisioning, provided their environments meet the previously outlined critirea.
- *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 writes 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 thin 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 thin 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 thin provisioning context. First, a minimal number of thin extents are allocated from the thin pool when a VMware file system is created on thin LUNs. 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 writes to previously uninitialized blocks are performed. The VMware kernel provides a number of allocation mechanisms for creating thin disks, and not all of them are thin-provisioning friendly. The "eagerzeroedthick" format is not ideal for use with thinly provisioned devices. The "thin" allocation policy is somewhat thin provisioning friendly. However, the risk of exceeding the thin pool capacity is much higher when thin disks are allocated using this policy, since oversubscription of 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:
- 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 thin provisioning with thin pools for:
- The best space efficiency
- Easy setup and management
- Minimal host impact
- Energy and capital savings
- Applications where space consumption is difficult to forecast

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

LUN migration

Thin LUNs can be the target or source of LUN migration operations. LUNs can be moved between or within storage pools and or RAID groups. Thin LUNs can be moved to a different storage pool that offers different characteristics to rebalance array performance, or to take advantage of newly available capacity. For example, thin LUNs can be migrated to a traditional fully allocated LUN if an application in R&D is deployed to production. Traditional LUNs can be migrated to a thin LUN structure if necessary but the physical storage space will remain fully allocated.

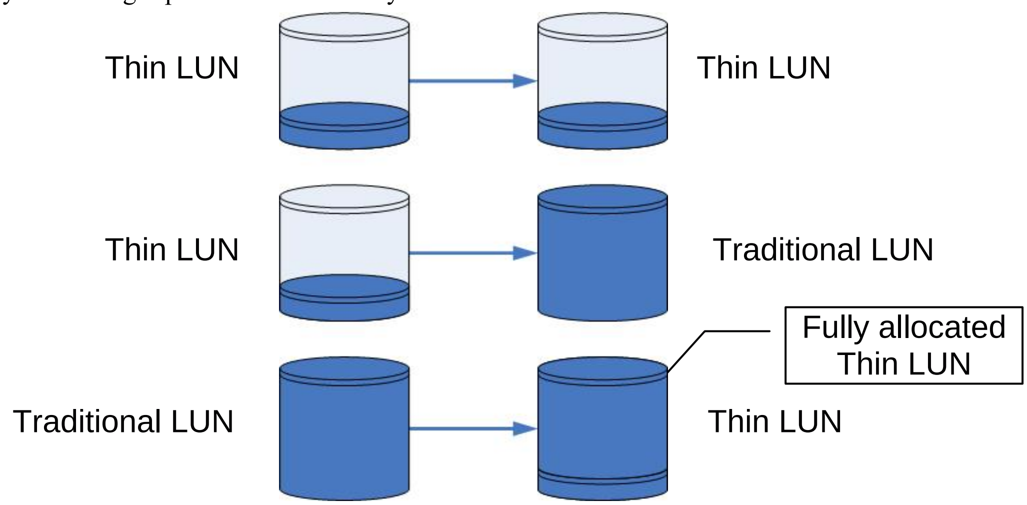

Figure 9. LUN migration

Phased implementation for CLARiiON thin provisioning

CLARiiON thin provisioning will be implemented in two phases. Phase 1 will include basic thin pool and thin LUN structures as well as the ability to expand a thin pool with additional disk drives. Phase 1 will support thin LUN technology including LUN migration and support for SnapView™ snapshots and clones. Navisphere Analyzer and Navisphere Quality of Service Manager will also be supported.

Phase 2 will include support for array-based remote replication: MirrorView®/S, MirrorView/A, and SAN Copy™. Phase 2 will provide the ability to shrink a thin pool by withdrawing disk drives, create and track thin LUN space reservations, and expand thin LUNs. Additional reports detailing thin pool and thin LUN characteristics and relationships will be made available via Taskbar-based reporting

Conclusion

When implemented appropriately, thin 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 Thin Provisioning integrates well with existing management and business continuity technologies, and is an important advancement in capabilities for CLARiiON customers. CLARiiON thin provisioning:

- 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

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	• 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	• 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	• 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