White Paper

EMC CLARiiON Virtual Provisioning

Applied Technology

Abstract

This white paper discusses the benefits of Virtual Provisioning™ on EMC® CLARiiON® CX4™ storage systems. It provides an overview of this technology, and describes how Virtual Provisioning is implemented on the CLARiiON CX4.

August 2010

Part Number H5512.4

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 to reduce power and cooling requirements and reduce capital expenditures.

EMC® CLARiiON® Virtual Provisioning provides pool-based storage provisioning by implementing pool LUNs that can be either *thin* or *thick*. Thin LUNs provide *on-demand* storage that maximizes the utilization of your storage by allocating storage as it is needed. Thick LUNs provide high and predictable performance for your applications. Both types of LUNs benefit from the ease-of-use features of pool-based provisioning. Pools and pool LUNs are also the building blocks for advanced data services such as Fully Automated Storage Tiering (FAST) and compression.

CLARiiON's Virtual Provisioning technology supports features such as hot sparing, proactive sparing, and the ability to migrate data between thin LUNs, thick LUNs, or traditional LUNs without incurring application downtime. The ability to nondisruptively migrate data to different LUN and disk types provides the best solution for meeting your changing application and business requirements without incurring 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 incurred when they need to add more storage as their business grows. This generally results in the overprovisioning of storage capacity, which then 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, FLARE® release 28.5 introduced Virtual Provisioning to the CX4™ with *thin LUN* technology. CLARiiON thin LUNs can 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 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 *pools* that form the basis for provisioning actions. Physical storage is automatically allocated only when new data blocks are written.

With FLARE release 30, we are introducing a new type of LUN called a *thick* LUN. Thick and thin LUNs can share the same pool, so they have the same ease-of-use benefits of pool-based provisioning. However, unlike a thin LUN, a thick LUN is fully allocated on creation. A thick LUN provides better performance than a thin LUN. However, a thick LUN does not provide the flexibility of overprovisioning and on-demand storage like a thin LUN does.

Virtual Provisioning also simplifies storage management tasks by providing the ability to expand pool LUNs (thick and thin LUNs) with a few simple clicks. The underlying pools can also be expanded by adding drives nondisruptively when additional physical storage space is required. This reduces the time and effort required to provision additional storage, and avoids provisioning storage that may not be needed.

Thin LUNs were made available, with a license, for CX4 storage systems in FLARE release 28.5. Thin LUNs and thick LUNs are now available, as a free feature, for CX4 storage systems running FLARE release 30. This free Virtual Provisioning feature requires a software enabler. For more information about features available with different FLARE releases, refer to the *Phased implementation for CLARiiON Virtual provisioning* section.

With Virtual Provisioning, CLARiiON users now have the flexibility of choosing either virtually provisioned *pool LUNs* (thick LUNs and thin LUNs) or CLARiiON's original LUNs (traditional LUNs and metaLUNs) for each application running on the CLARiiON.

You can manage pools and pool LUNs using the new Unisphere™ GUI or Secure CLI commands. CLARiiON replication products, Navisphere® Analyzer, and Navisphere Quality of Service Manager work seamlessly across pool LUNs. EMC RecoverPoint also supports the replication of CLARiiON pool LUNs.

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:

RAID group — A set of disks on which traditional LUNs and metaLUNs can be created.

Traditional LUN — A logical unit of storage created on a RAID group. The amount of physical space allocated is the same as the user capacity seen by the host server. Traditional LUNs cannot be created on a pool; they are always created on a RAID group.

MetaLUN — A collection of traditional LUNs that are striped and/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 to a metaLUN dynamically, allowing metaLUNs to be expanded on the fly.

Pool — A group of disk drives for configuring pool (thick and thin) LUNs. There may be zero or more pools in a storage system. Disks can only be a member of one pool, and they cannot also be in a RAID group.

Pool LUN – A logical unit of storage created on a pool. A pool LUN can be either a thin LUN or a thick LUN

Thick LUN — A type of pool LUN where physical space allocated is equal to the user capacity seen by the host server.

Thin LUN — A type of pool LUN where physical space allocated can be less than the user capacity seen by the host server.

LUN migration — A CLARiiON feature that dynamically migrates data to another traditional LUN, pool LUN, or metaLUN without disrupting running applications.

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

Consumed capacity — For a pool, this is the space currently used by all LUNs in the pool. For a thin LUN, this is the physical space used by the LUN. For a thick LUN, this is the host-visible capacity used by the LUN.

Allocated capacity — See “Consumed capacity.”

Oversubscribed capacity — The amount of user capacity configured for pool LUNs that exceeds the physical capacity in a pool.

Subscribed capacity — The total amount of capacity configured for thin LUNs in the pool. This number can be greater than the available user capacity. The user capacity can be expanded by adding drives to the pool.

User capacity — The total amount of physical storage capacity in the pool that is available for pool LUNs. This is also referred to as “usable capacity.” It is measured as raw disk capacity minus overhead (RAID overhead and mapping overhead). For a pool LUN, this is the size of the LUN as it appears to the host. For pool LUNs, this is sometimes called *host visible* capacity.

% Full — The percentage of pool capacity that is currently consumed. It is calculated using this formula:

% Full = Consumed capacity / User capacity

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

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

Thin friendly – A term that is frequently used for file systems and applications that do not preallocate all of the storage space during initialization. This term is also used for file systems that reuse deleted space before consuming additional storage. Both of these features improve capacity utilization in thin provisioning.

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.

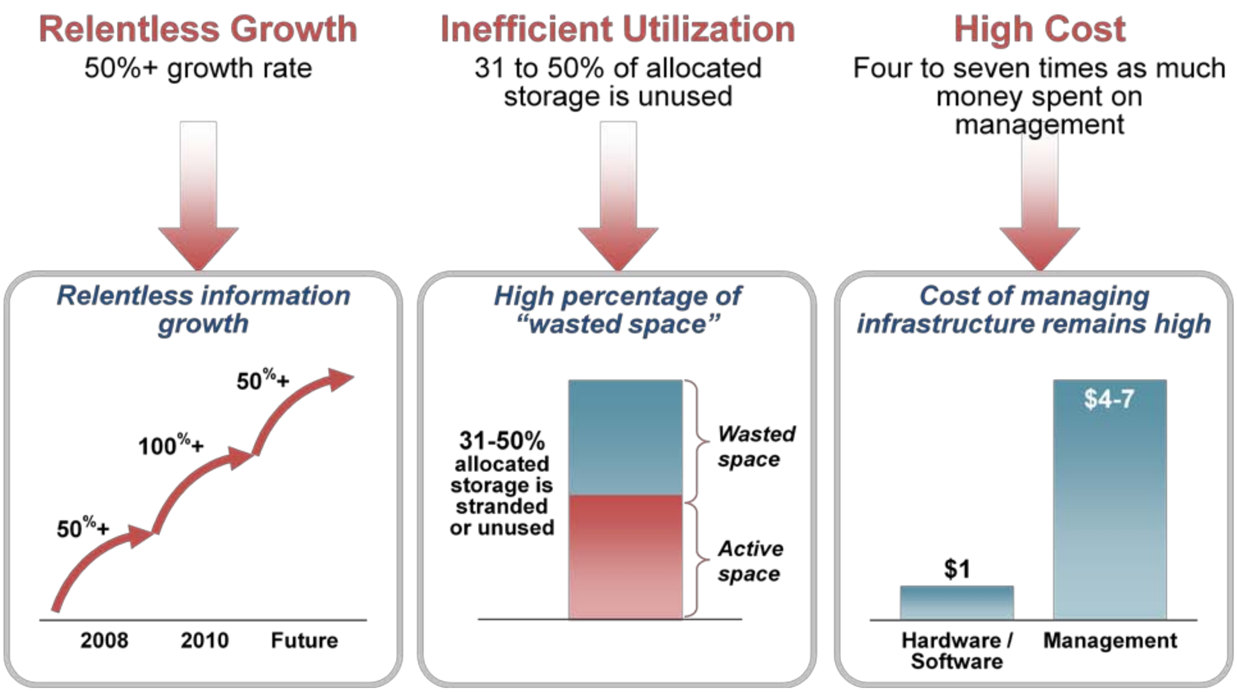

Figure 1. Information management challenges

Several business objectives have drawn increasing focus:

- *Reducing 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.

- *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 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 simplifies the process of adding storage.

Comparing traditional storage provisioning and Virtual Provisioning

Storage provisioning is the process of assigning storage resources to meet the capacity, availability, and performance needs of applications.

With traditional storage provisioning, you create a RAID group with a particular RAID protection level and a certain number of drives. RAID groups are restricted to a single drive type and to a maximum of 16 drives. When LUNs are bound on the RAID group, the host reported capacity of the LUNs 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 provisioning and Virtual Provisioning.

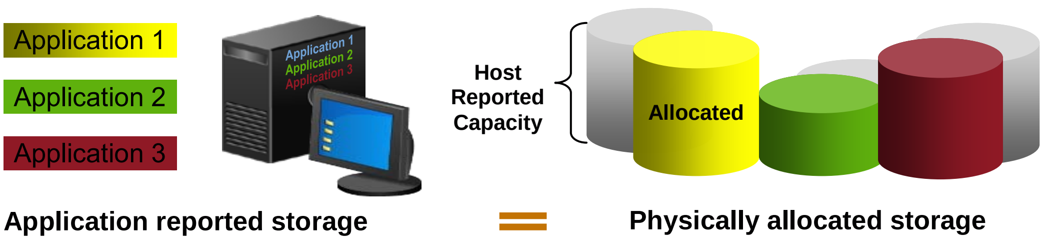

Figure 2. Traditional storage provisioning

With traditional provisioning, the storage administrator needs to carefully carve out the storage for an application based on the amount forecasted by the application administrator. 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. If the application requires more storage than forecasted, then storage can be expanded using metaLUN technology. Creating and maintaining metaLUNs requires additional planning and effort.

With Virtual Provisioning, you can create a pool with a particular RAID protection level and potentially all the drives in the storage system (except vault drives and hot spares). Pools can be homogeneous (having a single drive type) or heterogeneous (having different drive types). Heterogeneous pools are the building blocks for data efficiency services like FAST, which automatically places the most frequently used data on faster performing drives and places less active data on lower performing drives.

You can create thin LUNs or thick LUNs on the pools. With thin LUNs, the *host visible 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 pool. With thick LUNs, all the storage is reserved at the time of creation so the host reported capacity is equal to the actual storage allocated. Thick LUNs are higher performing than thin LUNs, and should be used for applications where performance is more important than space savings.

Pools and pool LUNs can be expanded on the fly any time additional storage is required. It is a completely nondisruptive procedure and can be achieved with a few simple clicks.

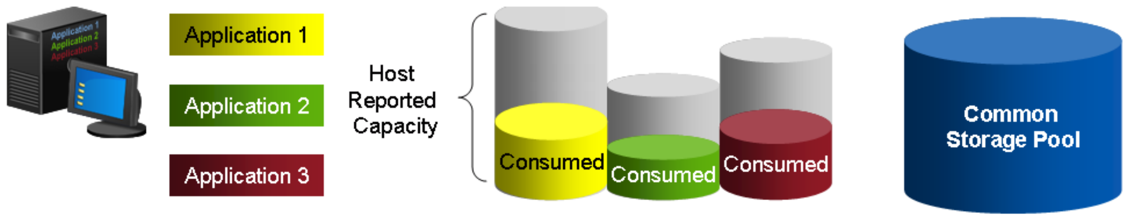

Figure 3. Virtual Provisioning

Virtual Provisioning provides the flexibility of choosing between thick or thin LUNs based on the application requirements.

Pools

A pool is somewhat analogous to a CLARiiON RAID group, which is a physical collection of disks on which logical units (LUNs) are created. Pools are dedicated for use by pool (thin and thick) LUNs. Pools can contain a few disks or hundreds of disks, whereas RAID groups are limited to 16 disks. Because of the large number of disks supported in a pool, pool-based provisioning provides the same benefits of metaLUN striping across many drives, and, unlike metaLUNs, it requires no planning and management effort. Pools, like RAID groups, support a single RAID protection level that can be RAID 5, RAID 6 or RAID 1/0 (with FLARE release 30). Pools can be homogeneous (having a single drive type) or heterogeneous (containing different drive types).

Homogeneous pools

The concept of homogeneous pools was introduced with the first release of Virtual Provisioning in FLARE release 28.5. With release 28.5, these homogeneous pools can be composed of either Fibre Channel or SATA disk drives. With FLARE release 28.7, homogeneous pools can also be composed of Flash drives. Homogeneous pools support a single RAID level. Until FLARE release 30, it was EMC’s best practice recommendation to use homogeneous pools. Homogeneous pools are still recommended for applications with similar performance requirements, and when you do not intend to use advanced data services like FAST, which is introduced in FLARE release 30.

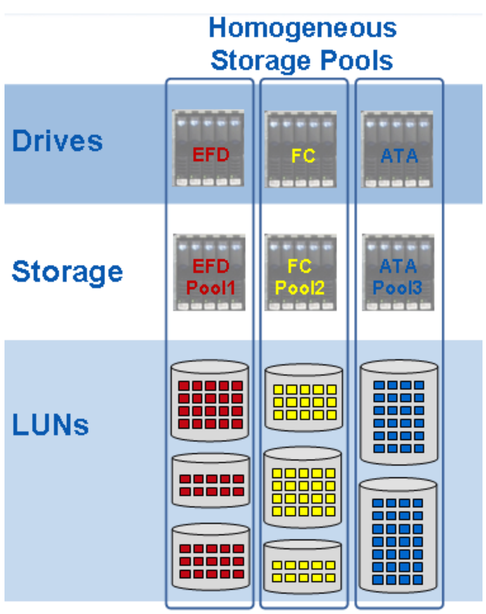

Figure 4. Homogeneous storage pools

Heterogeneous pools

FLARE release 30 introduces heterogeneous pools, in which a pool can consist of different types of drives such as Flash, FC, and SATA drives. Heterogeneous pools, like homogeneous pools, support a single RAID protection level.

Heterogeneous pools are the foundation for FAST, which facilitates automatic data movement to appropriate drive tiers depending on the I/O activity for that data. The most frequently accessed data is moved to the highest tier (Flash drives) in your pool for faster access, medium activity data is moved to FC drives, and low activity data is moved to the lowest tier (SATA drives). For more information on FAST, refer to the *EMC CLARiiON FAST – A Detailed Review* white paper on Powerlink®.

There can be a maximum of three tiers in a heterogeneous pool, based on the three drive types. Heterogeneous pools can also have two drive types and still leverage FAST. FAST does not differentiate tiers by the drive speeds.

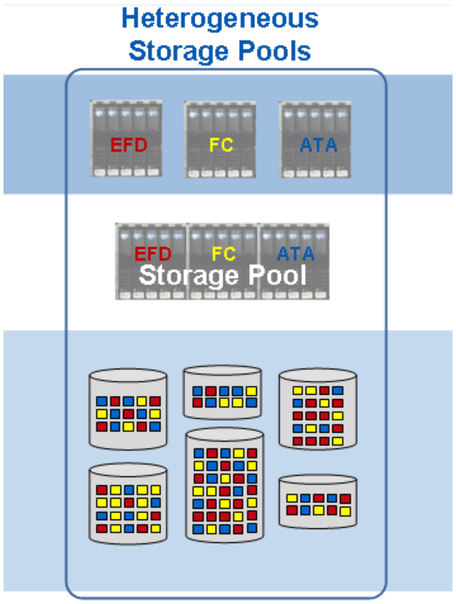

Figure 5. Heterogeneous storage pools

Attributes

Pools are simple to create because they require only three user inputs:

- Pool Name: For example, “Application Pool 2”
- Resources: Number and type of disks
- Protection level: RAID 5, 6, or 1/0

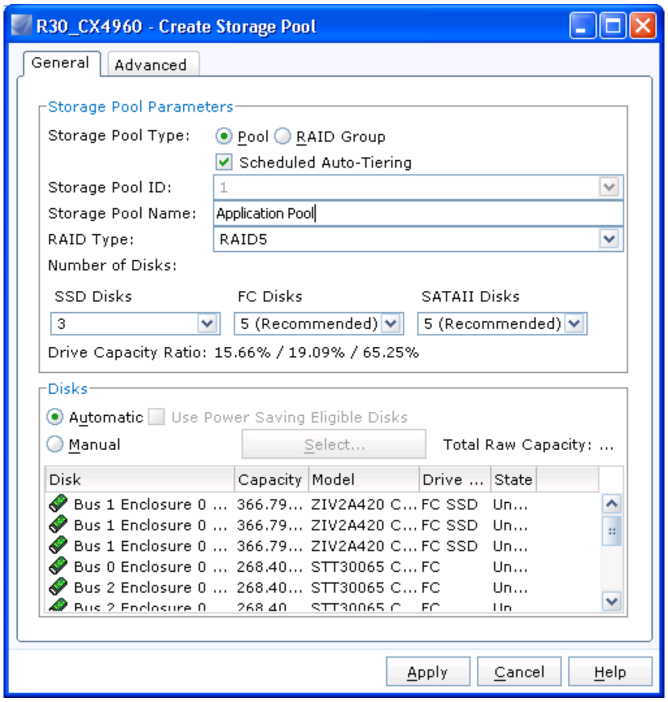

Figure 6. Create Storage Pool dialog box

A CLARiiON CX4 can contain one or many pools. RAID 5 is the default selection for pools and is recommended for most situations. The smallest pool size is three drives for RAID 5, four drives for RAID 6, and two drives for RAID 1/0. However, recommended pool sizes are multiples of five for RAID 5, multiples of eight for RAID 6, and multiples of eight for RAID 1/0. These drive counts are chosen based on the best practices to ensure performance and availability. For very large pools, RAID 6 may be a better choice. For best practice recommendations, refer to the *EMC CLARiiON Best Practices for Performance and Availability* white paper on Powerlink.

With FLARE release 30, pools can be as large as the maximum number of drives (except vault drives and hot spares) allowed per system type. For example, a CX4-960 can contain 955 drives in a single pool or between all pools. Vault drives (the first five drives in a storage system) cannot be part of a pool, so Unisphere dialog boxes and wizards do not allow you to select these drives. Large pools cannot be created in a single operation. Depending on the system type, pools can be created by the maximum allowed drive increment and then need to be expanded until you reach the desired number of drives in a pool. Once completely created, you can create LUNs on it. For example, to create an 80-drive pool on a CX4-120, you need to create a pool with 40 drives and then expand the pool with another 40 drives. Pools can also be expanded at a later time if more storage is needed. The maximum allowed drive increments for different system types are shown in Table 1.

Table 1. Drive increments for CX4 models

CX4 model	Maximum allowed drive increments
CX4-120	40
CX4-240	80
CX4-480	120
CX4-960	180

More detailed information on the limits for different FLARE releases, such as disks per pool and pools per system, is provided in <u>*Appendix A: Virtual Provisioning limits.*</u> Also refer to the *EMC Virtual Provisioning Release Notes* for current information.

When creating a pool, different types of disks (EFD, FC and SATA) can be of different sizes. However, to maximize space utilization, in each pool all disks of a particular type should be the same size. When using different capacity drives to create a pool, it is best to create the pool in stages. For example, if you have ten 450 GB FC drives and five 300 GB FC drives, first create the pool selecting only the ten 450 GB drives, and then expand the pool by adding the other five 300 GB drives.

Monitoring, adding, and deleting pool capacity

User capacity is the total physical capacity available to all LUNs in the pool. *Consumed capacity* is the total physical capacity currently assigned to all pool LUNs. *Subscribed capacity* is the total host reported capacity supported by the pool. Oversubscribed capacity is the amount of user capacity configured for LUNs that exceeds the physical capacity in a pool. Thin LUNs with larger user capacity than the physical capacity can result in an oversubscribed pool.

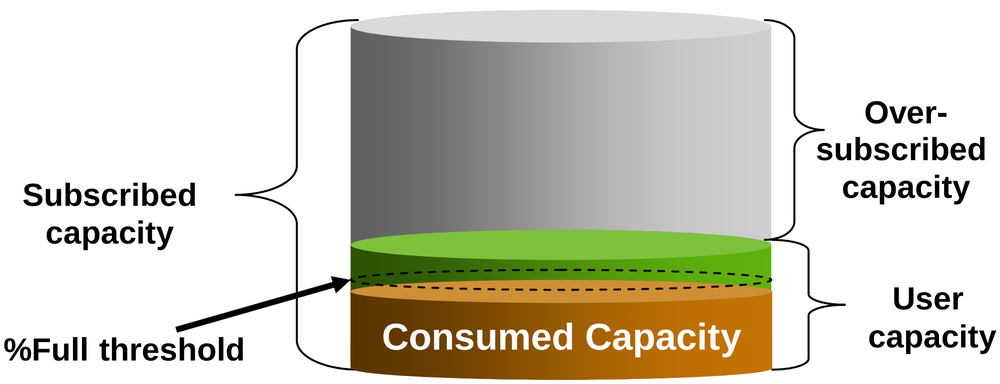

Figure 7. Pool % full threshold

Pools are monitored using the **% Full Threshold Alert** setting and storage capacity reports. You can specify the value for **% Full Threshold** (Consumed capacity/User capacity) when a pool is created. It can be changed at any time. It ranges from 1% to 84%. When **% Full Threshold** is exceeded, an alert is triggered at the *warning* severity level. When the **%Full** reaches 85%, the pool issues a built-in alert at the *critical* severity level. Beyond 85%, the built-in alert mechanism continues to track the actual **%Full** value as the pool continues to fill.

Alerts can be monitored via the **Alerts** tab in Unisphere. In Unisphere's Event Monitor wizard, you can also select the option of receiving alerts via e-mail, a paging service, or an SNMP trap. Table 2 lists the thresholds for different types of alerts.

Table 2. Threshold alerts

Threshold Type	Threshold Range	Threshold Default	Alert Severity	Side-effect
User settable	1%-84%	70%	Warning	None
Built in	n/a	85%	Critical	Clears user settable alert

Allowing consumed capacity to exceed 90% of user capacity puts you at the risk of running out of space and impacting all applications using LUNs in the pool.

The **Storage Pool Properties** dialog box shows parameters such as subscribed capacity, oversubscribed capacity, percent subscribed, and consumed capacity.

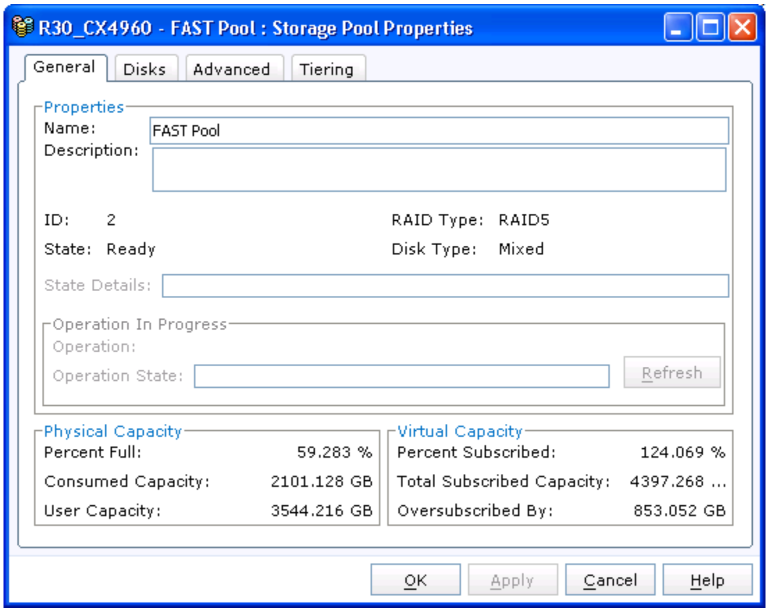

Figure 8. Storage Pool Properties dialog box

A pool can be expanded by adding drives to it. It is a completely nondisruptive operation, and the increased capacity can be used by LUNs in that pool. Consumed (allocated) capacity is reclaimed by the pool when LUNs are deleted. Oversubscribed pools can run out of space; hence it is a best practice to ensure that a monitoring strategy is in place and you have the appropriate resources to expand the pool.

Expanding pools with only a few disks as compared to the original number of disks in the pool can impact performance for the applications running on the pool; hence it is our best practice recommendation to expand pools by approximately the same number of disks as originally present in the pool.

Pool LUNs

Pool LUNs are created within pools. A pool LUN can be either thick or thin. A CLARiiON pool LUN is similar to a traditional LUN in many ways. Many of the same Unisphere operations and Secure CLI commands can be used on pool LUNs and traditional LUNs. Most user-oriented FLARE functions work the same way, including underlying data integrity features, hot sparing, LUN migration, local and remote replication, and LUN properties information. Pool LUNs are available as choices in Unisphere Tasks, Unisphere Analyzer, and Unisphere Quality of Service Manager. Features such as hot sparing and proactive sparing also work with pool LUNs. It is also possible to migrate a traditional LUN (or metaLUN) to a pool LUN and vice versa.

Thin LUNs

Thin LUNs were introduced with the first release of Virtual Provisioning in FLARE release 28.5. The primary difference between a thin LUN when compared to traditional LUNs and thick LUNs is that thin LUNs present more storage to an application than is physically allocated. Presenting storage that is not physically available avoids overprovisioning the storage system and underutilizing its capacity. Thin LUNs consume storage as needed from the underlying pool.

Thick LUNs

FLARE release 30 introduces a new type of pool LUN called a thick LUN. Thick LUNs and thin LUNs can share the same pool. When a thick LUN is created, all of its capacity is reserved in the pool for use by that LUN. Thick LUNs are better performing than thin LUNs, but they do not provide the capacity utilization benefits that a thin LUN provides. The performance of a thick LUN is comparable to a traditional LUN and it also provides the ease-of-use benefits of pool-based provisioning. Thick LUNs are the default choice when creating LUNs in a pool.

Attributes

The maximum size for a pool LUN is 14 TB. Pool LUNs are simple to create, with four inputs:

- Pool information
- Amount of host visible user capacity
- Pool LUN name (optional)
- Thin or not

The default option will create a thick LUN. A thin LUN can be created by selecting the **Thin** checkbox, as shown in Figure 9.

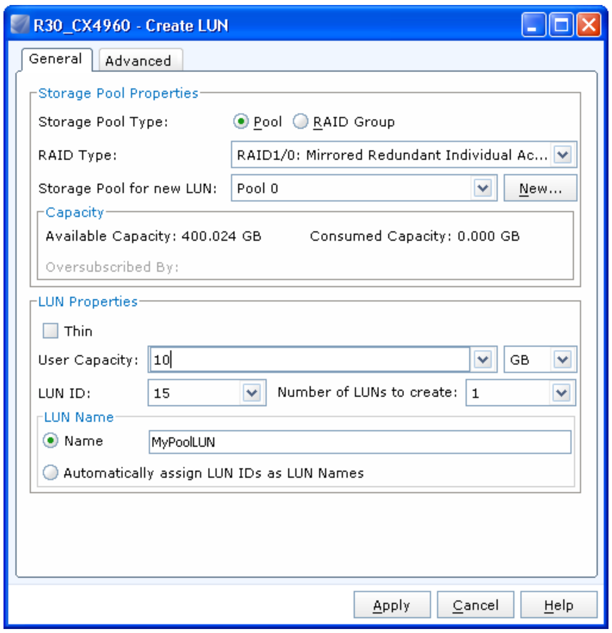

Figure 9. Create LUN dialog box

Pool LUNs are easy to use because the system automatically manages the drives within the pool according to CLARiiON best practices. The Virtual Provisioning software distributes data evenly throughout the pool. Pool LUNs are easy to manage because Unisphere and Secure CLI commands work the same for pool LUNs as they do for traditional LUNs. New property screens and reports are available to obtain information about pool LUNs.

There are only a few restrictions with pool LUNs and they are enforced through Navisphere:

- A pool LUN cannot be used for a clone private LUN (which contains maps for clones).
- A pool LUN cannot be used for a write intent log for MirrorView™ operations.
- A pool LUN cannot be used as a component of a metaLUN.
- A thin LUN cannot be used in the reserved LUN pool (which reserves space for snapshots), but thick LUNs can be used.

Expanding and shrinking pool LUNs

With FLARE release 30, EMC introduces the capability to expand a pool LUN. Pool LUN expansion can be done with a few simple clicks, and the expanded capacity is immediately available. Pool LUN expansion does not use metaLUN technology. It is a much simpler task than creating a metaLUN, because you do not need to create a destination LUN first, and you can expand the LUN to any size as long as the pool can support that capacity. For a thick LUN, the pool must have enough storage for the expansion to succeed, whereas for a thin LUN the storage does not need to be available. It is important to note that you cannot expand a pool LUN if it is part of a replication or LUN-migration operation.

FLARE release 30 also introduces the capability of *shrinking* a pool LUN. This capability is only available for LUNs that are served up to Windows Server 2008. (This capability was introduced for traditional LUNs in FLARE release 29). The shrinking process involves two steps:

1. Shrink the file system from Windows Disk Management.
2. Shrink the pool LUN using a command window and the DISKRAID utility. The utility is available through VDS Provider, which is part of the EMC Solutions Enabler package.

The new LUN size is shown as soon as the shrink process is completed. A background task is triggered to reclaim the deleted or shrunk space and return it back to the pool. Once the task is completed, the reclaimed space can be used by all other LUNs in that pool.

Architecture of pool LUNs

When you create a thin LUN or a thick LUN, it will initially consume 2 GB from the pool. Some of the 2 GB is used for LUN metadata, and the remaining space is available for incoming writes. For a thick LUN, the remaining user capacity of the LUN is also reserved in the pool. For a thin LUN, no space is reserved after the initial 2 GB. Because all of the user-capacity is reserved upon creation of a thick LUN, the mapping requirements for a thick LUN are much less compared to a thin LUN, which is why thick LUNs perform better than thin LUNs. As new writes come into a thin LUN or a thick LUN, more physical space is allocated in 1 GB slices. This storage-on-demand algorithm ensures that at least a 1 GB slice is available at all times. LUN metadata is also embedded within these additional 1 GB slices.

Data is written to thin LUNs in 8k chunks and optimally placed within the 1 GB slices. On thick LUNs, data is written within the 1 GB slices in a contiguous fashion. This automated data placement makes configuring pool LUNs easy, because the Virtual Provisioning software makes all the decisions about how to lay out actual storage blocks across the disks in a pool while adhering to performance best practices. Less-experienced storage administrators benefit from not having to be directly involved in the details of configuring storage.

Best practices

EMC recommends that you do *not* change the default owner of the pool LUN once it is provisioned. This can impact performance because the underlying private structures, which provide storage for the pool LUN, are still controlled by the original SP. It is a best practice to balance the pool LUNs on both SPs when you create the pool LUNs. The option to choose SP ownership is under the **Advanced** tab in the **Create LUN** dialog box. If you must change the SP ownership after the LUN has been created, LUN migration can be leveraged to migrate the LUN to a new LUN with the desired SP.

It is important to consider the metadata overhead *before* provisioning pool LUNs for your applications. For a pool LUN, you can estimate the overhead using the following simple formula. Note that this is an approximation; the actual overhead depends on the write pattern of your application.

Metadata Overhead (in GB) = LUN Size (in GB units) * .02 + 3 GB

This formula approximates the storage consumed from the pool if you write to the entire LUN. For example, if you have a 200 GB pool LUN that gets fully consumed over a period of time, based on the above formula, it will utilize approximately 207 GB from the pool.

Use cases for traditional, thick, and thin LUNs

It is important 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 between thin, thick, and traditional LUNs.

Use RAID groups and traditional LUNs:

- When milliseconds of performance are critical
- For the best and most predictable performance
- For precise data placement
- For physical separation of data
- When you are not as concerned about space efficiency

Use thin LUNs with 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
- Automated storage tiering
- Applications with moderate performance requirements
- Taking advantage of data efficiency services like compression and space reclamation

Use thick LUNs with pools for:

- Applications that require very good performance
- Easy setup and management
- Automated storage tiering
- Minimal host impact

Thin LUN space reclamation

Space reclamation is a new feature in FLARE release 30 that allows the system to free up allocated storage that is not used. This feature works in conjunction with LUN migration and SAN Copy™ when migrating existing volumes to thin LUNs.   Unused storage is often locked up by applications that are either no longer important to organizations or do not need the storage that was originally allocated to them. With the space reclamation feature, you can nondisruptively migrate existing volumes of your applications to FLARE release 30 thin LUNs and reclaim the unused space. Since thin LUNs only consume storage to which data is written, all the allocated but unused storage is returned to the pool that can be used by other LUNs in the pool. The process is completely transparent and so you can move your applications without requiring any downtime.

Space reclamation is automatically invoked when performing a SAN Copy pull operation on source volumes from older CLARiiON systems, Symmetrix® storage systems, and supported third-party systems to a thin destination LUN running FLARE release 30. Space reclamation is also invoked when you perform a LUN migration to move an existing traditional LUN or thick LUN to a thin LUN within the same array, (running FLARE release 30). The software detects zeros at 8k chunk granularity. For example, it will only migrate 8k chunks with data in them. If the 8k chunk is filled with zeros, it will not be migrated to a thin LUN

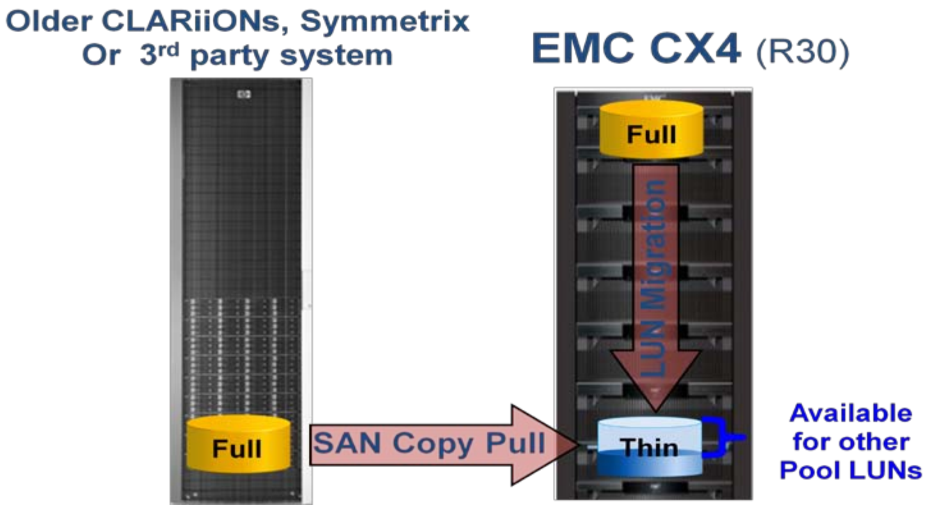

Figure 10. Space reclamation

Replication

CLARiiON Virtual Provisioning supports local and remote replication. When replicating a thin pool LUN, the *host visible* capacity, and not the *consumed* capacity, is used to determine which LUNs are eligible for replication. For example, assume you plan to replicate a thin LUN to another thin LUN. The thin LUN has a *host visible capacity* of 100 GB and *consumed capacity* of 20 GB. The destination thin LUN has to be at least 100 GB to be eligible for the replication operation. Thin-to-thin replication provides optimal space efficiency for your replicas because only the utilized or consumed space is replicated. To provide even more flexibility, EMC also supports replication between pool LUNs, traditional LUNs, and metaLUNs.

Local replication

For local replication, SnapView snapshots and clones are both supported on thin LUNs since FLARE release 28.5. SnapView will support thick LUNs starting with FLARE release 30. SnapView clones support replication between thick, thin, and traditional LUNs. When cloning from a thin LUN to a traditional LUN or thick LUN, the physical space of the traditional/thick LUN must equal the host visible capacity of the thin LUN. This will result in a fully allocated thin LUN if the traditional LUN/thick LUN is reverse-synchronized. Cloning from traditional/thick to thin LUN will result in a fully allocated thin LUN as the initial synchronization will force the initialization of all the subscribed capacity.

Remote replication

Remote replication of thin LUNs, MirrorView/S, MirrorView/A, and SAN Copy are all supported starting with FLARE release 29. In FLARE release 30, remote replication on thick pool LUNs is also supported.

MirrorView

You need FLARE release 29 or later on the primary and secondary storage systems to use MirrorView for remote replication of thin LUNs. To replicate thick LUNs with MirrorView, you need FLARE release 30 running on the primary system. The secondary storage system in this case has to be at least FLARE release 29. When mirroring a thin LUN to another thin LUN, only consumed capacity is replicated between storage systems. This is the most beneficial for initial synchronizations. Steady state replication is similar, since only new writes are written from the primary storage system to the secondary system. When mirroring from a thin LUN to a traditional or thick LUN, the thin LUN host visible capacity must be equal to the traditional LUN capacity or thick LUN’s user capacity. Any failback scenario that requires a full synchronization from the secondary to the thin primary image causes the thin LUN to become fully allocated. When mirroring from a thick LUN or traditional LUN to a thin LUN, the secondary thin LUN will be fully allocated. With MirrorView, if the secondary image LUN is added with the no initial sync option, the secondary image retains its thin attributes. However, any subsequent full synchronization from the traditional LUN or thick LUN to the thin LUN, as a result of a recovery operation, causes the thin LUN to become fully allocated. For more information on using pool LUNs with MirrorView, see the white paper *MirrorView Knowledgebook* on Powerlink.

SAN Copy

In SAN Copy sessions, pools LUNs can be source and/or destination LUNs. The support for thin LUNs was added in FLARE release 29. Thick LUNs will also be supported for SAN Copy operations starting with FLARE release 30. When all the systems involved in SAN Copy operation are running FLARE release 29 and later, and the source and destination LUNs are thin, only the consumed capacity is copied; the unused or zeroed space is not copied. This is true for a full-copy session, initial synchronization, and a bulk copy of an incremental session. Incremental updates are similar because only data that has changes is copied.

SAN Copy supports migrations between thin, thick, and traditional LUNs. With the space reclamation feature in FLARE release 30, you can do a SAN Copy pull operation from a non-thin source on an older release CLARiiON or any supported third-party array to a FLARE release 30 thin destination and reclaim the zeroed space from the non-thin source.

LUN migration

LUN migration moves data from a *source* LUN to a *destination* LUN (of the same or larger size) within a single storage system. This migration is accomplished without disruption to the applications running on the host. Pool LUNs can be the target or source of LUN migration operations. LUN migration can enhance performance or increase disk utilization for your changing business needs and applications by allowing the user to change LUN type and characteristics, such as RAID type or size, while their production volume remains online. Pool LUNs can be moved between pools or to a traditional LUN in another RAID group.

When migrating a thin LUN to another thin LUN, only the consumed space is copied. When migrating a thick LUN or traditional LUN to a thin LUN with FLARE release 30, the space reclamation feature is invoked and only the consumed capacity is copied. If you are migrating a traditional LUN on a prerelease 30 system to a thin LUN, the thin LUN will end up being fully allocated. Migration from thick LUNs to traditional LUNs or another thick LUN and from traditional LUN to thick LUNs is also permitted.

PowerPath Migration Enabler

EMC PowerPath® Migration Enabler (PPME) is a host-based migration tool that enables nondisruptive or minimally disruptive data migration between storage systems or between logical units within a single storage system. The Host Copy technology in PPME works with the host operating system to migrate data from the source logical unit to the target. With PPME 5.3, the Host Copy technology supports migrating virtually provisioned devices. When migrating to a thin target, the target's thin-device capability is maintained. This can be very useful when migrating data from older systems with non-thin volumes to new thinly provisioned storage systems. Host Copy migrations for thin LUNs are supported on Windows, Linux, AIX, and Solaris hosts connected to CLARiiON arrays running FLARE version 28.5 or later.

RecoverPoint

Replication is also supported through EMC RecoverPoint. Continuous data protection (CDP) and continuous remote replication (CRR) support replication for thin LUNs, thick LUNs, and traditional LUNs. When using RecoverPoint to replicate to a thin LUN, only data is copied; unused space is ignored so the target LUN is thin after the replication. This can provide significant space savings when replicating from a non-thin volume to a thin volume. When using RecoverPoint, we recommend that you not use journal and repository volumes on thin LUNs. For more information on using RecoverPoint, see the white paper *EMC RecoverPoint/SE for the CLARiiON CX4* on Powerlink.

Using thin LUNs with applications

Due to the storage-on-demand feature of thin LUNs, not all application environments are well suited to thin LUNs. In general it is a best practice to use *thin-friendly* applications, that is, applications that do not preallocate all of the storage space during initialization and that reuse deleted space before consuming additional storage. The following are guidelines for using thin LUNs with the most common applications

Oracle

Traditional database files as well as Oracle ASM include an Auto Extend feature that can take advantage of thin LUNs. 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 LUNs for the database volumes. For best practices planning for database deployments, refer to the *Leveraging EMC CLARiiON CX4 Virtual Provisioning for Database Deployment* white paper available on Powerlink

Microsoft SQL Server

Microsoft SQL Server 2005 introduced new functionality that changed how databases are created. Earlier versions of SQL Server fully initialized all data file and transaction log file components by 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. EMC recommends using Microsoft SQL Server 2005 (or later) with Instant File Initialization functionality when using thin LUNs.

Microsoft Exchange

By default, a new Exchange 2007 database is between 2 MB and 4 MB in size, and grows by 2 MB as additional space is needed. This auto-extend behavior is efficient from a thin LUN perspective, as only space immediately needed by the database file is allocated. Exchange environments that most effectively benefit from Virtual Provisioning have a large space allocated for mailboxes and are slow to fill that space.

File systems

When creating a file system on a thin LUN, you need to consider how much metadata is written to the thin LUN. An inefficient file system, which writes a lot of metadata to the LUN, causes thin devices to become fully allocated more quickly than with an efficient file system. For this reason, an efficient file system is thin friendly.

Another thin-friendly feature is the ability to effectively reuse deleted space. If you delete a file, the file system knows it can use those blocks to store new data. The underlying CLARiiON thin LUN, however, is not aware of this. When you create a new file on the host file system, depending on the file system, it may or may not use the space freed up by a deleted file. If it writes the new file in the previously freed-up area, then the thin LUN will not consume more space from the pool. However, if it writes in a previously unwritten area, then the thin LUN will consume more space from the pool.

With NTFS, some *operations* are thin friendly. For example, a NTFS format does not preallocate the physical space. Instead, it creates file system metadata that only consumes a few gigabytes of metadata on the thin LUN. NTFS writes new data on the LUN and updates its own metadata accordingly. However, NTFS is not very effective when it comes to reusing deleted space.

Other file systems like Linux ext2 and ext3, Solaris ZFS and UFS, and Symantec VxFS do not preallocate all the storage and also effectively reuse deleted space before allocating new space and thus work nicely with thin LUNs. For the latest list of file systems supported with CLARiiON Virtual Provisioning, check the *EMC Support Matrix* at https://elabnavigator.emc.com/ on Powerlink.

Veritas Storage Foundation

Veritas Storage Foundation, by Symantec, enhances the value of thin-provisioned storage environments by optimizing migrations and enabling storage reclamation. CLARiiON thin LUNs can take advantage of the *SmartMove* and *Thin Reclamation* capabilities in Storage Foundation to help customers capitalize on their existing storage investment and minimize costs.

SmartMove

Veritas SmartMove enables efficient online migrations from thick storage to thin storage, and in the process, reclaims unused space. It works with CLARiiON thin LUNs running FLARE release 28.5 or later. It supports UNIX and Linux platforms with Storage Foundation 5.0 MP3, and Windows platforms with Storage Foundations 5.1. Tight integration between Veritas Volume Manager and Veritas File System allows SmartMove to identify the blocks deleted by an application. SmartMove uses this information to ensure that only the necessary blocks of data are moved during online migration.

Thin Reclamation API

Thin Reclamation API enables online reclamation of unused storage over time, so that thin-provisioned storage can remain thin. When data is deleted from a host file system, the physical storage stays allocated to a thin LUN. To reclaim this unused but allocated storage, Thin Reclamation API sends the SCSI WRITE_SAME command (with the UNMAP bit turned on) to inform the storage array that all physical storage for a certain block range can be reclaimed. With FLARE release 30, when CLARiiON receives SCSI WRITE_SAME commands from the Thin Reclamation API, it zeros out the specified block range in the thin LUN. Over time, a background process returns the zeroed space back to the pool where they can used by other applications.

Virtual environments

For VMware environments, the Virtual Machine File System (VMFS) has many characteristics that are thin friendly. First, a minimal number of thin extents are allocated from the pool when a VMware file system is created on thin LUNs. Also, a VMFS datastore reuses previously allocated blocks that are beneficial to thin LUNs. When using RDM volumes, the file system or device created on the guest OS dictates whether the RDM volume is thin friendly.

With ESX 3.x, when creating a VMware virtual disk, the default type is “zeroedthick.” This is the recommended virtual disk type for thin LUNs. Using zeroedthick, the storage required for the virtual disk is reserved in the datastore, but the VMware kernel does not initialize all the blocks. In addition, features such as VMware DRS, Converter, and VMotion are thin friendly. VM Clones, Storage VMotion, Cold Migration, and Templates are not thin friendly. 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. Customers are advised to use these features on thick LUNs. For detailed information about implementing Virtual Provisioning with VMware, refer to the *EMC CLARiiON Integration with VMware ESX* white paper available on Powerlink

With ESX 4.0, CLARiiON thin LUNs can be configured as thick (zeroedthick) or thin. With either option, the VMware admin must monitor the consumed capacity on the VMFS datastore. vSphere provides an alert when the thresholds are reached. Also, with ESX 4.0, features such as VM Clones, Storage VMotion, Cold Migration, and Templates become thin friendly. This is because, unlike ESX 3.x, the zeroedthick or thin format remains intact on the destination datastore. In other words, the consumed capacity of the source virtual disk is preserved on the destination virtual disk and is not fully allocated.

The VMware kernel provides other mechanisms for creating virtual disks that are not thin friendly. The “eagerzeroedthick” format is not recommended for thin LUNs, because at creation it performs a write to every block of the virtual disk and results in equivalent storage use in the thin pool.

In Hyper-V environments, EMC recommends that you select the “dynamically expanding” option for virtual hard disk on the NTFS file system with CLARiiON thin LUNs; this preserves disk resources. When using pass-through, the file system or guest OS dictates whether the pass-through device will be thin friendly. For more information on using Hyper-V Server, see the whitepaper *Using EMC CLARiiON with Microsoft Hyper-V Server* on Powerlink.

Phased implementation for CLARiiON Virtual Provisioning

EMC is implementing CLARiiON Virtual Provisioning in phases. The initial phase was released in December 2008 (FLARE release 28.5). This release included basic pool and thin LUN structures as well as the ability to expand a pool by adding disk drives. Thin LUNs were supported for LUN migration and local replication using SnapView (snapshots and clones). Navisphere Analyzer and Navisphere Quality of Service Manager were also supported. Remote replication of thin LUNs was supported using RecoverPoint. With FLARE release 28.7, support for Enterprise Flash Drives in a thin pool was also added

In 2009, with FLARE release 29, support was added for array-based remote replication: MirrorView/S, MirrorView/A, and SAN Copy. Some of the limits have also been doubled with release 29. *Appendix A: Virtual Provisioning limits* lists the limits with different FLARE releases.

With FLARE release 30, the concept of heterogeneous pools and a new type of LUN in a pool called a thick LUN are introduced. The RAID 1/0 protection level and the ability to expand and shrink (Windows Server 2008) pool LUNs are included. Release 30 also introduces the space reclamation feature when migrating to thin LUNs. A single pool can now support all the drives in the system except vault drives. Refer to *Appendix A: Virtual Provisioning limits* for limits information.

Conclusion

When implemented appropriately, Virtual Provisioning can be a powerful complement to organizations' processes and technologies for improving ease of use 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:

- Saves time:
  - Easy to create pools and pool LUNs
  - Easy to expand pools and pool 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
  - SnapView snapshots and clones, MirrorView/S, MirrorView/A, SAN Copy
  - Unisphere Quality of Service Manager and Unisphere Analyzer
  - Unisphere Reports

References

- *EMC CLARiiON Best Practices for Performance and Availability — Applied Best Practices*
- *Leveraging EMC CLARiiON CX4 Virtual Provisioning for Database Deployment—Best Practices Planning*
- *MirrorView Knowledgebook: FLARE 29 – A Detailed Review*
- *EMC CLARiiON SAN Copy — A Detailed Review*
- *EMC RecoverPoint/SE for the CLARiiON CX4 — Applied Technology*
- *EMC CLARiiON Integration with VMware ESX — Applied Technology*

Appendix A: Virtual Provisioning limits

Table 3. Limits with FLARE release 28.5

System	Max. thin pools	Max. disks per thin pool	Total disks in all thin pools	Max. thin LUNs per pool	Max. thin LUNs in all pools
CX4-120	10	20	40	256	256
CX4-240	20	40	80	512	512
CX4-480	40	80	160	1024	1024
CX4-960	60	120	240	2048	2048

Table 4. Limits with FLARE release 29

System	Max. thin pools	Max. disks per thin pool	Total disks in all thin pools	Max. thin LUNs per pool	Max. thin LUNs in all pools
CX4-120	20	40	80	512	512
CX4-240	40	80	160	1024	1024
CX4-480	40	120	240	2048	2048
CX4-960	60	180	360	2048	2048

Table 5. Limits with FLARE release 30

System	Max. thin pools	Max. disks per thin pool	Total disks in all thin pools	Max. thin LUNs per pool	Max. thin LUNs in all pools
CX4-120	20	115	115	512	512
CX4-240	40	235	235	1024	1024
CX4-480	40	475	475	2048	2048
CX4-960	60	955	955	2048	2048

Appendix B: Guidelines for pools and pool LUNs[1]

Guidelines for pools

- Create or expand pools in increments of five drives for RAID 5 and eight drives for RAID 6 and eight drives for RAID 1/0.
- RAID 5 is recommended for most situations. For very large pools of drives, RAID 6 may be a better choice.
- Avoid running pools at near maximum capacity.
- Ensure that a monitoring strategy is in place for all thin pools.
- Expand pools by approximately the same number of disks as originally present in the pool.

Guidelines for pool LUNs

- Operating systems and applications are thin LUN friendly when:
  - The file systems efficiently reuse deleted space.
  - Applications reserve space but do not initialize all storage.
- Consider metadata overhead before provisioning thin LUNs.
- Balance your thin LUNs on both SPs and do not change the default owner once provisioned.

[1] For continuous updates to best practices concerning Virtual Provisioning please use the CLARiiON white paper *EMC CLARiiON Best Practices for Performance and Availability — Applied Best Practices*.