

# Thin provisioning: особенности применения

В публикации представлены особенности реализации технологий, связанных с т.н. механизмом “thin provisioning” – “тонкого выделения ресурсов” и преимущества дополняющих решений, позволяющих существенно повысить их эффективность.

## Введение

Thin Provisioning – функциональность, которая одной из первых была реализована компанией ZPAR<sup>1)</sup>, и в настоящее время поддерживается всеми ведущими (и не только) производителями систем, включая: HDS, EMC, HP, Sun Microsystems, IBM, Dell и др. HDS называет ее “Hitachi Dynamic Provisioning”, HP и Sun Microsystems – “Thin Provisioning”, EMC – “Virtual Provisioning” (используется в составе EMC DMX и Clarion). IBM поддерживают thin provisioning в составе своих продуктов IBM SVC и IBM XIV. На high-end моделях СХД от основных вендоров thin provisioning поддерживается уже достаточно давно, в основном, с 2006 г. На midrange моделях СХД – с 2008 г. и в течение 2009 г. эта функциональность будет включена почти у всех производителей СХД среднего класса.

Основная идея thin provisioning в том, что физическое выделение ресурсов для LUN происходит по мере необходимости или при достижении определенного уровня (watermark), а не сразу, как это обычно делается (рис. 1). Идеально объем физически выделенного пространства под LUN может соответствовать объему записанных на него данных. Однако хост сразу видит максимальный размер LUN и работает со всем объемом как с реальным. При таком выделении удастся существенно сэкономить на

объемах СХД и, соответственно, снизить объем ИТ-бюджета за счет отсутствия необходимости закупать объемы хранения “впрок” и, по сути, “омертвлять” денежные средства на длительное время, никак не используя их. При существующей сегодня в среднем 40%-й утилизации ресурсов хранения, общая потребность в них при использовании технологии thin provisioning может быть, соответственно, снижена до 60%.

Дополнительным преимуществом thin provisioning является существенное снижение затрат на администрирование LUNs, связанное с постоянным их мониторингом, увеличением размера и др.

Однако у технологии thin provisioning есть свои “подводные камни”, которые могут снизить возможную эффективность ее применения. В частности, например, система хранения “не видит” удаленных блоков файлов, и один раз записанные эти блоки воспринимаются как занятые в дальнейшем в течение всего времени. Это приводит к двум негативным последствиям: во-первых, удаленные блоки не могут быть возвращены в общий пул для использования в составе других LUNs или хотя бы восприниматься как свободные в составе своего LUN; во-вторых, например, при миграции LUN с СХД, не имеющей thin provisioning на СХД с поддержкой данной технологии, приходится на новом LUN выделять физическую память, равную и больше “старому” LUN, что сводит на нет все преимущества технологии thin provisioning.

В качестве примера другого “камня” может, например, служить особенность некоторых “недружественных” к thin provisioning файловых систем выделять под записываемые блоки память “кусками” (chunks), размер которых может быть от нескольких килобайт до нескольких мегабайт.

И в случае, если блок данных при записи существенно меньше блока (chunk), выделяемого под него системой хранения, то это может приводить к существенному перерасходу ресурсов (например, блок данных размером 8 Кбайт, а chunk – 1 Мбайт).

## Функциональные особенности продуктов Symantec для thin provisioning

Отмеченные проблемы и ряд других, возникающих при использовании тех-

нологии thin provisioning, можно решить за счет информации, которую передает хост/файловая система хоста СХД об удаленных блоках LUNs. Такую интеграцию на основе API стала предлагать, например, компания Symantec. В середине октября 2008 г. она анонсировала расширение к семейству продуктов Veritas Storage Foundation (SF), которые позволяют достичь максимального эффекта от использования технологии thin provisioning в системах хранения. В настоящее время эти расширения поддержива-

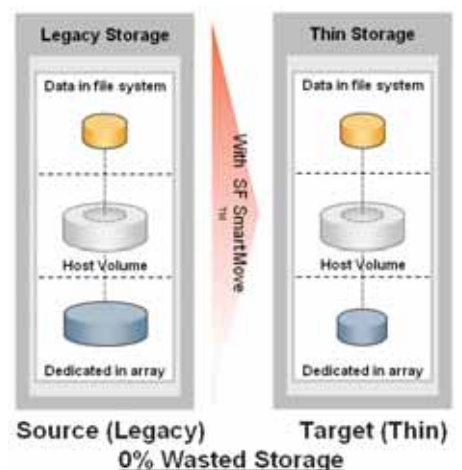
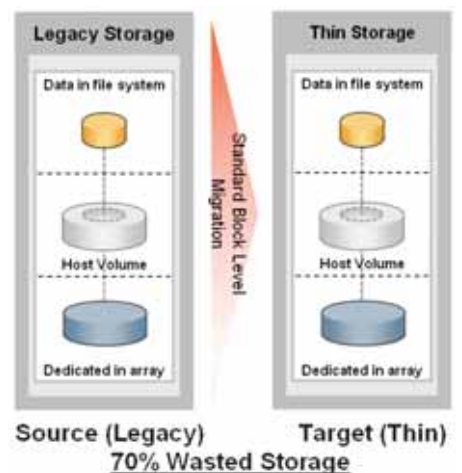


Рис. 2. При использовании продукта SmartMove, входящего в состав Veritas Storage Foundation, для миграции данных между традиционными СХД и с поддержкой thin provisioning удается полностью исключить неиспользуемые (удаленные) блоки LUN.

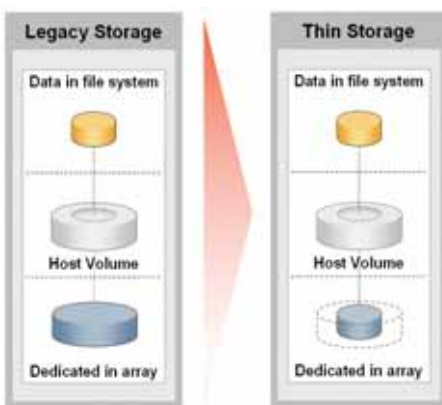
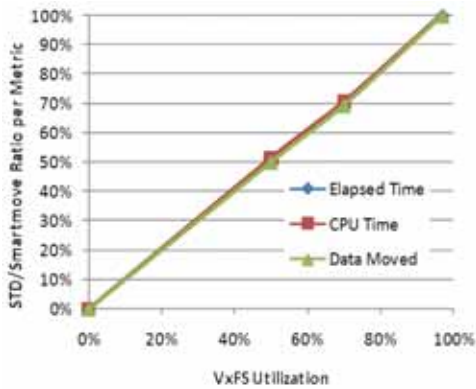


Рис. 1. При использовании технологии thin provisioning выделение физических ресурсов для LUN происходит по мере необходимости.

<sup>1)</sup> Hough, Geoffrey and Sandeep Singh. “ZPAR Thin Provisioning.” ZPAR Incorporated, June 2003, <http://www.zpar.com/documents/ZPAR-tp-wp-01.2.pdf>



**Рис. 3.** Как показало тестирование (по данным Symantec) наблюдается практически линейная зависимость между процентом использования файловой системы и объемом системных ресурсов, требующихся для выполнения операций миграции/зеркалирования.

ются в составе семейства моделей Hitachi TagmaStore USP и, соответственно, в ее OEM-продуктах от HP и Sun. В течение 1 кв. 2009 г. она будет поддерживаться и в моделях среднего уровня — HDS AMS2000.

В составе вышеуказанного анонса Symantec представила два продукта. Первый — Veritas Thin Reclamation API — новый API, позволяющий автоматизировать возврат освобождающегося пространства (например, после удаления) LUN в общий пул для систем хранения, поддерживающих механизм thin provisioning. Необходимо заметить, что, если LUN управляется VxFS (файловая система SF), то при операциях записи автоматически в первую очередь ищется освободившееся пространство, и только после его использования занимают новые блоки.

Также в состав новой версии Storage Foundation включен продукт SmartMove — решение, позволяющее автоматически исключать неиспользуемое пространство LUN при он-лайн миграции данных LUN с традиционных СХД на СХД, поддерживающих thin provisioning (рис. 2). Данное решение работает со всеми дисковыми массивами, предлагающими thin provisioning, и поддерживает Windows-, Unix- и Linux-платформы на серверах с файловыми системами.

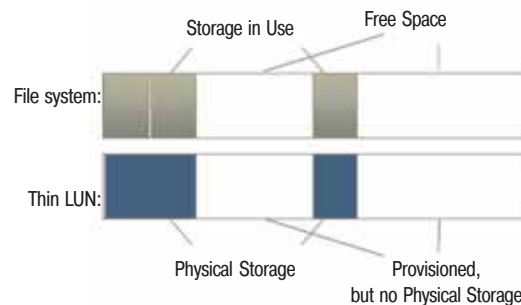
При этом, как показало тестирование (по данным Symantec), наблюдается практически линейная зависимость между процентом использования файловой системы и объемом системных ресурсов, требующихся для выполнения операций миграции/зеркалирования (рис. 3).

В составе Storage Foundation с версии 5.0 имеется продукт — Dynamic Storage Tiering (DST — в более ранних версиях продвигался как Quality of Storage Services; см. SN № 3/28, 2006 — “Dynamic Storage Tiering — уровневое хранение от Symantec”), предназначенный для построения многоуровневых решений хранения на основе VxFS файловой системы и совершенно прозрачный для пользователей, приложений, БД, а также политик резервного копирования/восстановления. В интеграции с технологией thin provisioning DST может быть ис-

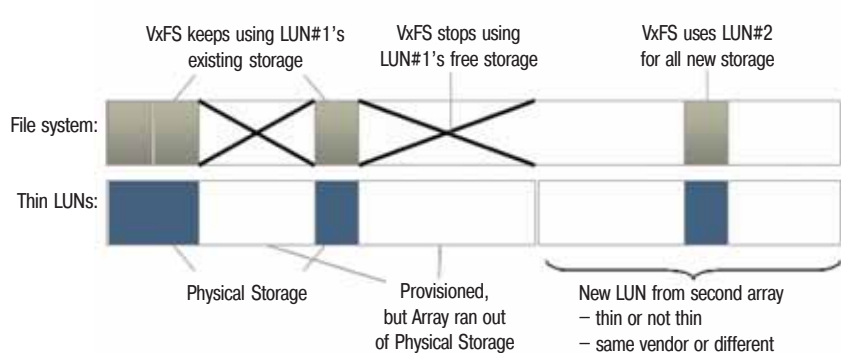
пользован для защиты от возможных ошибок, возникающих при необходимости расширения LUN и отсутствия возможности добавления физических ресурсов к массиву, на котором размещается LUN (рис. 4). В этих случаях DST обеспечивает метод, позволяющий LUN оставаться в доступности, пока не появится возможность расширения thin provisioning массива.

Это основано на возможности VxFS использовать множество томов, что позволяет снять традиционные ограничения по развертыванию файловой системы на единственном томе. В момент (или ранее) превышения физического размера thin provisioning LUN, чтобы продолжить выполнение приложения, к файловой системе (VxFS) добавляется второй том, размещенный на другом дисковом массиве (рис. 5). Далее, на основе DST-политик все новые записи перенаправляются на второй том, в то время как при чтении и изменении данных операции осуществляются с первым томом.

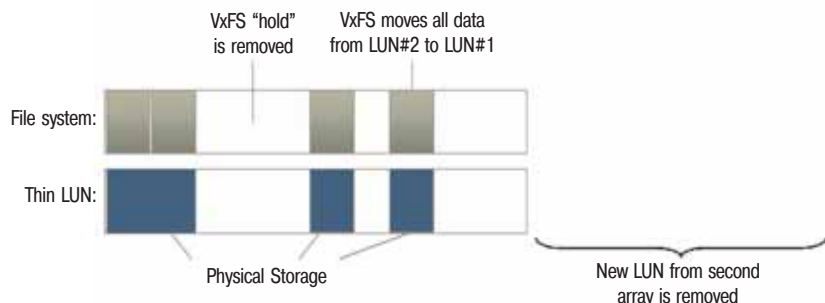
Когда физическая память добавлена к первому тому, второй том может быть “ejected” — VxFS автоматически переместит все данные со второго тома на первый и затем удалит второй том (рис. 6).



**Рис. 4.** Thin provisioning LUN достигает предела физического объема СХД.



**Рис. 5.** Добавляется второй LUN на другом дисковом массиве. DST-политики все новые записи перенаправляют на второй том, в то время как при чтении и изменении данных операции осуществляются с первым томом.



**Рис. 6.** Физическая память добавлена к первому тому. VxFS перемещает все данные со второго LUN на первый и затем удаляет второй LUN.

При оценке эффективности файловой системы, используемой для thin provisioning в ряде случаев может иметь значение и тот факт, что большинство блочных файловых систем при создании файловой системы сразу создают метаданные (inodes). В противоположность этому, например, VxFS создает их по мере необходимости, что также положительно сказывается на эффективности thin provisioning.

## Заключение

Эффект от использования дополнительных решений к технологии thin provisioning на СХД существенно зависит от приложений. Поэтому при их развертывании следует всесторонне проанализировать все особенности продуктивных приложений, которые планируется развернуть на этих СХД, а также оценить все требования и возможности по дальнейшему управлению ими и ресурсами СХД.

Базы данных в основном предраспределяются на месяцы и даже годы. В этих случаях с учетом того, что, например, для Oracle, начиная с версии 10g опция storage reclamation включена в состав базовой функциональности, эффект от использования Veritas Thin Reclamation при миграции минимальный или вообще отсутствует. Для других приложений со средней или высокой интенсивностью удаляемых файлов эффект от дополнительных решений может составлять десятки процентов и более.