Questions and Answers about QThin

Version 1.2
April 2016

Copyright

Copyright@2004~2016, Qsan Technology, Inc. All rights reserved. No part of this document may be reproduced or transmitted without written permission from Qsan Technology, Inc.

Trademarks

All products and trade names used in this manual are trademarks or registered trademarks of their respective companies.

Qsan Technology, Inc.

4F., No.103, Ruihu St.,
Neihu Dist., Taipei City 114,
Taiwan (R.O.C.)

Tel: +886-2-7720-2118
Fax: +886-2-7720-0295

Email: Sales@Qsan.com
Website: www.Qsan.com

Introduction

The QThin - Qsan Thin provisioning feature is implemented after firmware version 3.0.0 on the AegisSAN LX series. In this document, we will introduce some questions and answers which you might ask about the QThin and help you to get more understanding about QThin. Please read it carefully before using any thin provisioned RG on the production unit.

QThin RAID Design

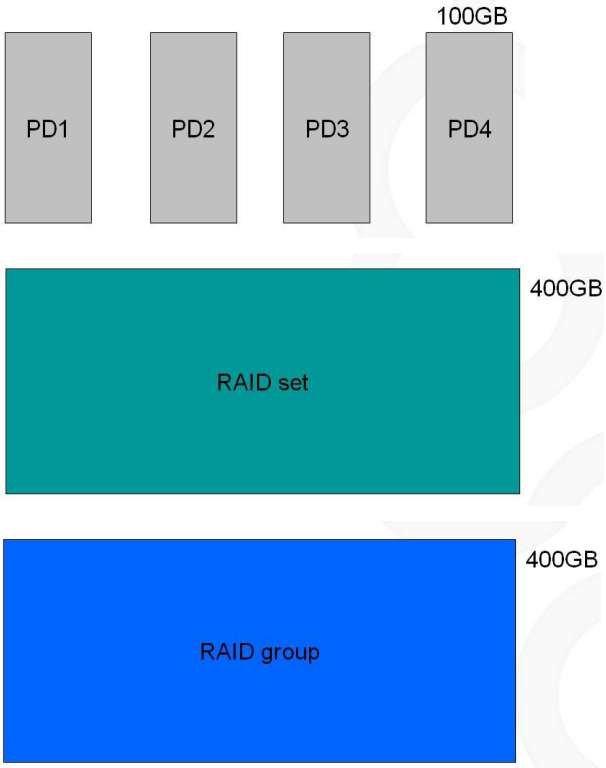
Question:

What is the relationship among the PD, RAID set, RG, VD and Granularity (the space unit for allocation and zero reclamation)?

Answer:

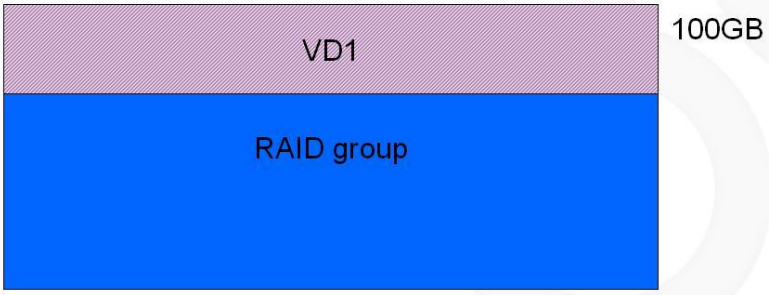
Here we would like to introduce the relationship with a simple example as below:

1. Creating a thin provisioned RG (RAID level 0) with 4 PD (100GB per each), and actually, a RG is created with a RAID set of RAID 0 with 4 disks.

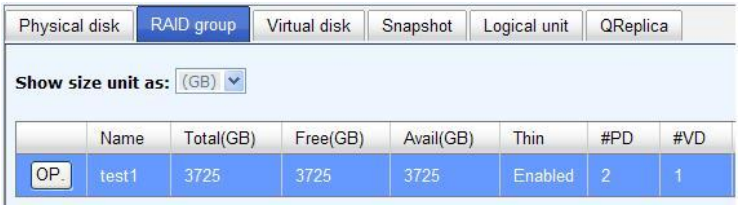


Copyright@2004~2013 Qsan Technology, Inc. All Rights Reserved.

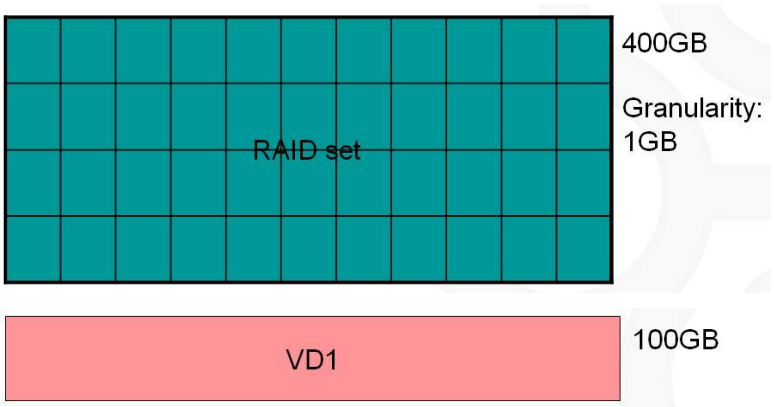
2. Creating a VD1 (100GB) on the RG which is of course a thin provisioned VD, the system will create a VD which does not occupy any physical capacity at the beginning.



The storage will tell the server (host) that there is a 100GB VD1 can be used, and pretending the 100GB VD1 has been allocated on the RG. But in fact, the spaces on the RG will not decrease until reading or until any data is really written into the VD1 from the server (host). Such as the below illustration, there is a VD created on a thin provisioned RG, but the Free space will not be used before writing any data into the VD.

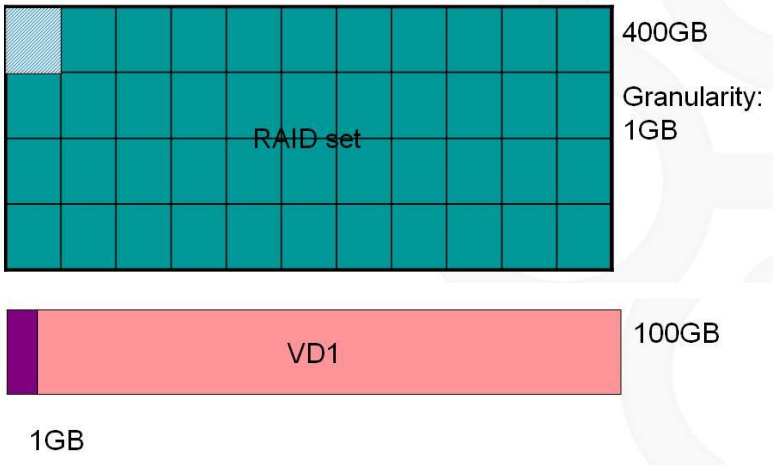


3. When you try to write data into the VD1 from the server (host), the system will allocate available space from the RG to the VD1 to store the data.

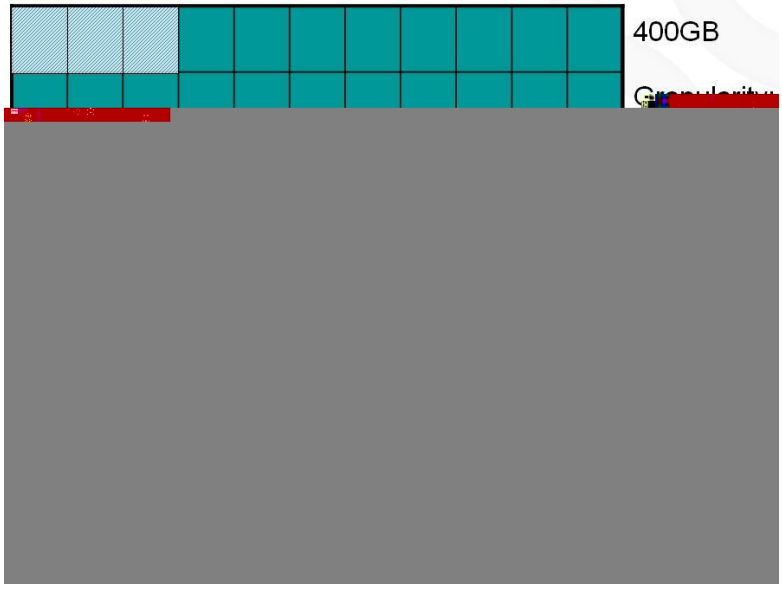


The RAID set will be divided by the Granularity (1 GB) which is the minimum size while allocating the space to a VD.

Copyright@2004~2013 Qsan Technology, Inc. All Rights Reserved.



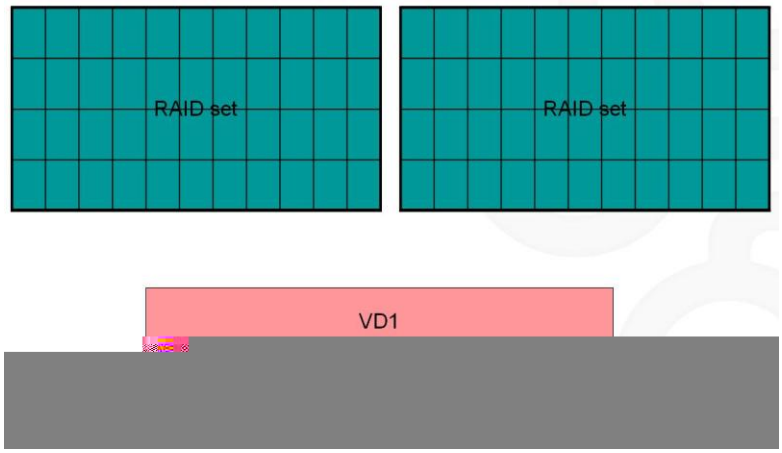
And it will also occupy one Granularity (1 GB) even the written data is less than 1 GB.



Even only 2.5 GB of data was written into the VD, there is still 3GB of physical space on the RG allocated to the VD.

4. And what if there are two or more RAID sets in the RG, and then creating two or more VDs? Which one of the RAID sets will the data be allocated to?

Copyright@2004~2013 Qsan Technology, Inc. All Rights Reserved.



The system has its own policies (the health status of the RAID sets, the available capacity of the RAID and etc.) to judge which one of the RAID sets is proper to be allocated to a VD while space is demanded as the data changes. And you may observe which one of the RAID sets had been allocated with the data from the RG page on the web GUI when you clicked on the RG. For example, please see the below illustration, the space on the RAID set 2 is allocated more than on the RAID set 1.

Compatibility

Question:

Can I change the fat provisioned RG/VD to thin provisioned RG/VD, or change the thin provisioned RG/VD back to the fat provisioned RG/VD?

Answer:

The RG cannot be changed from fat to thin provisioned, or the other way around. You have to backup the data and destroy the RG to create the type of RG (fat or thin) to restore the data back.

Copyright@2004~2013 Qsan Technology, Inc. All Rights Reserved.

Question:

If I am using the dual-controller, do I still need to schedule a down-time for upgrading the controller firmware from 1.x.x to 3.x.x?

Answer:

Yes, after upgrading the controller firmware to the 3.x.x version, please perform a cold reboot (shutdown and then power on) to the system. Otherwise, the system might have some abnormal behaviors during the reboot process. Due to the system perform rebooting on one controller after another which will cause the controllers with firmware 1.0.x and 3.0.x exist on the system at the same time and cause the problem.

Capacity

Question:

What is the maximum number or capacity of the thin provisioned RG and VD can be included in the system (such as AegisSAN LX)?

Answer:

Please see the below comparison form between fat provision and thin provision:

RG / Max.	Fat Prov. RG	Thin Prov. RG
PD - system	192	192
VD - system	4096	4096
PD - RG	64	256

Copyright@2004~2013 Qsan Technology, Inc. All Rights Reserved.

There are items on the "Max." column, such as "PD - system", it means the maximum number of PD that can be supported in a system. And for the "Space - RAID set", it represents the maximum capacity that can be supported in a RAID set. And so on.

According to the above form, we want to remind you with the below notices:

1. Although the Max. PD number that the thin provisioned RG can support is 256; it is still limited by the Max. PD supported in the system (192) as well.
2. For the thin provisioned VD, once the total capacity of the thin provisioned VD on the system reach the limitation of 256 TB, you cannot create the thin provisioned VD or allocate the snapshot space on a thin provisioned VD anymore. For example, if you have created a thin provisioned VD with 128 TB and allocated 50 TB snapshot space on the VD, the rest capacity that can be used to create the thin provisioned VD will be limited in 78 TB.
3. The fat provisioned RG does not have the limitation in the above RED columns. Say if there were thin provisioned VDs with size of 256 TB in total on the system, the fat VD still can be created as long as there is still free space on the fat RG.
4. No matter you create a fat or thin provisioned VD, the capacity cannot be shrunk.

Failure Cases

Question:

How can I recover the thin provisioned RG/VD which has failed already?

Answer:

Currently, there is no way to restore a failed thin provisioned VD. It is highly recommended that users need to keep monitoring the health of the RAID and backup the data regularly, no matter it is fat or thin provisioned RAID.

Question:

Will the thin provisioned RG get failed when one of the RAID sets in the RG got the failed status?

Answer:

Yes, the RG will be in the failed status if one of the RAID sets in the RG got failed. But the RG will be restored to good status after removing the failed RAID set.

Copyright@2004~2013 Qsan Technology, Inc. All Rights Reserved.

Question:

Continue the question above. When one of the RAID sets in a thin provisioned RG is failed, will the VD on it get failed as well? Is the failed VD still available to be accessed?

Answer:

Yes, the VD will get failed as well if it has the space allocated from the failed RAID set. The failed VD still can be accessed, but we could not guarantee if the consistency of the data on the VD will be corrupted or not, since there are parts of data stored on the failed RAID set.

Question:

Why does a thin provisioned VD with size bigger than a fat provisioned VD rebuild faster than the fat provisioned VD sometimes?

Answer:

When a thin provisioned VD is rebuilding, it rebuilds only the space actually allocated to it. Say there is a 1TB thin provisioned VD but only 30GB physical space is allocated to it, then it will take less time in rebuilding compares to a fat provisioned VD with size of 100GB.

Zero Reclaim

Question:

What is the difference between enabling “Space reclaim” when creating the VD and performing “Space Reclamation” on the VD manually?

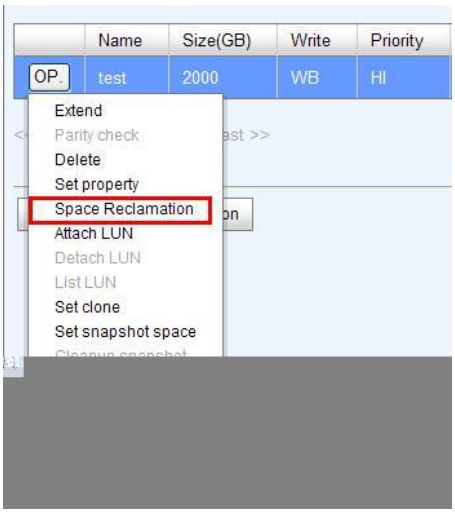
Answer:

Enabling “Space reclaim” option when creating the VD:

Copyright@2004~2013 Qsan Technology, Inc. All Rights Reserved.

Enabling this option means when the data is written into the VD, the system will check in real time if there is space could be reclaimed.

Performing the "Space Reclamation" function on the VD:



This function will check the whole data of the VD which allocated on the RAID sets in background, see if there are any spaces could be reclaimed.

Question:

Can the snapshot space on the thin provisioned VD be zero reclaimed?

Copyright@2004~2013 Qsan Technology, Inc. All Rights Reserved.

Summary

The above questions and answers might be helpful for you to save times when you tried to use and realize the QThin function. Due to the behaviors of thin provisioned VD are different comparing to the thin provisioned VD, please try to test much more scenarios for making sure that you are very familiar with each scenario of the QThin before you will store your valuable data into the QThin VD.

Reference

- Wikipedia - Thin provisioning
  http://en.wikipedia.org/wiki/Thin_provisioning

Copyright@2004~2013 Qsan Technology, Inc. All Rights Reserved.