Albireo SANblox

Albireo SANblox™: Red Hat® Enterprise Linux® Solution Profile

redhat.

TECHNOLOGY PARTNER

Permabit Technology Corporation
One Alewife Center, Suite 410
Cambridge, MA 02140 USA

Phone: (+1) 617.252.9600
Fax: (+1) 617.252.9977
info@permabit.com
www.permabit.com

Copyright © 2015 Permabit Technology Corporation. All Rights Reserved.
072315

PERMABIT

Summary

Permabit's Albireo technology allows companies to extend their existing storage infrastructure and leverage the benefits of data reduction. The Permabit Albireo SANblox™ appliance provides inline deduplication, compression, and thin provisioning to servers utilizing new or existing Fibre Channel (FC) attached storage.

This solution profile reviews best practices for implementing Red Hat® Enterprise Linux® and SANblox with leading enterprise storage arrays. As a Red Hat Partner, Permabit has worked with leading storage providers to develop operational guidelines for storage arrays fronted by the SANblox appliance that are being used by servers or desktops running the Red Hat Enterprise Linux operating system.

Target Audience

This document is intended for:

- Enterprise storage administrators
- SAN administrators
- Red Hat Enterprise Linux administrators
- Linux application administrators

Conventions

This Guide uses the following font conventions.

- **`Bold Courier`** indicates text that you must type into the terminal window (i.e., a SANblox command)
- *`Italic Courier`* indicates values you must provide in a command you enter (e.g., the name of a specific host)
- `Normal Courier` indicates text displayed in a terminal window (i.e., a response to a command you entered)

Permabit, HIOPS, and HIOPS Compression are registered trademarks, and the Permabit logo, Albireo logo, Permabit Enterprise Archive, SANblox, and Scalable Data Reduction are trademarks of the Permabit Technology Corporation, protected by the laws of the United States and other countries, and may not be reused or reproduced without permission. All other products or services mentioned may be covered by registered trademarks, trademarks, service marks, or product names as designated by the companies who market those products.

Introducing SANblox

SANblox is a complete, ready-to-run data reduction appliance that combines high performance and massive scalability with low resource overhead to deliver a powerful, cost effective solution for existing and new Fibre Channel SAN environments. SANblox applies both inline data deduplication and compression to data as it is passed from client systems through to designated LUNs on a storage array. The result is typical data reduction rates of 6:1 in mixed environments with rates as high as 50:1 for some VDI use cases. In environments that are able to leverage cache, these data reduction rates can lead to substantial performance benefits as more data is able to reside on high speed storage. Since nearly all flash storage is >25X faster at random I/O than a 15K RPM hard drive, the benefits from a larger flash-based cache or as a separate tier in hybrid storage environments can be substantial.

SANblox uses Permabit's Albireo data deduplication and HIOPS Compression™ software. Deduplication is a technique for reducing the consumption of storage resources by eliminating multiple copies of duplicate data. Instead of writing the same data more than once, each duplicate block is detected and recorded as a reference to the original block. Deduplication is distinct from compression in that it operates across much larger datasets than compression. The benefits of these two technologies are additive when combined, yielding greater possible savings than either on its own.

Installation is simple. Rack and cable the SANblox units, then provision a set of LUNs on an array. SANblox will aggregate the array LUNs into a global data reduction pool which can be carved up and exported as "SANblox-optimized" LUNs.

As seen in Figure 1, SANblox acts as a gateway for existing thick or thin-provisioned LUNs from a FC-attached storage array. The SANblox appliance is delivered as two 1U hardware units that are intended to be paired in an active-passive configuration for high availability with transparent failover. Data reliability is assured because all data is immediately written to back-end block storage with no write caching. Should the secondary system lose connectivity with the primary, it will take over operation in under 60 seconds.

To make it easy to manage, SANblox includes the SANblox Dashboard a graphical user interface (GUI), as well as a command-line interface (CLI) that integrates smoothly with existing management frameworks, user interfaces, and toolsets.

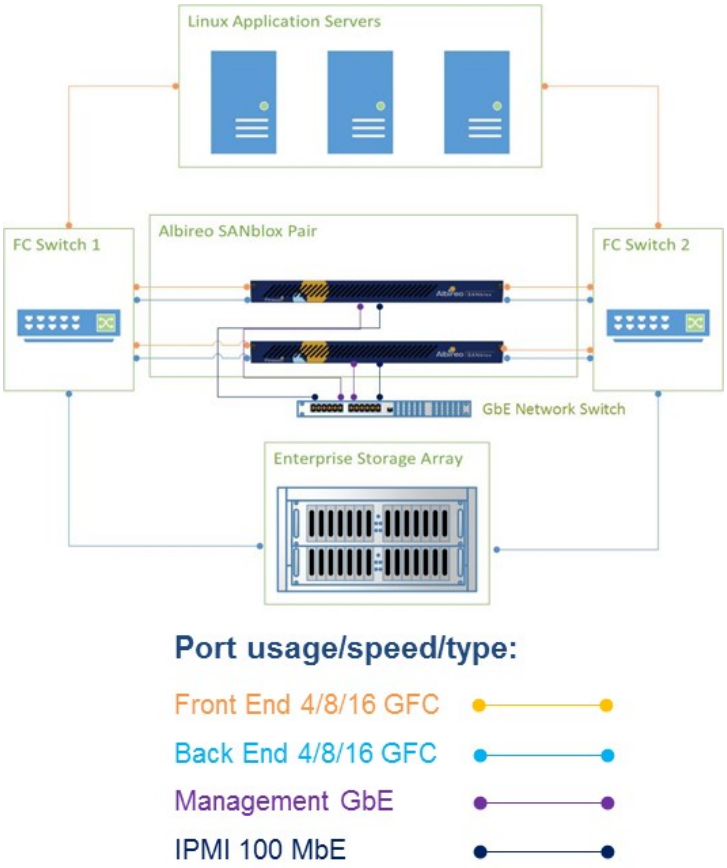


Figure 1: Basic SANblox Deployment Topology

Red Hat Enterprise Linux Overview

As the world's leading enterprise Linux platform, Red Hat Enterprise Linux has been designed for deployment on physical hardware, in virtual machines, and in the cloud while delivering reliability, security and performance.

Red Hat enables data centers to support a broad range of application environments from the small business to the enterprise. Key areas that will be explored further in this document include:

- **Storage**

  The default file system provided in the latest release is XFS, which is highly scalable (up to 500 TB) and robust in terms of data integrity. In addition, Red Hat Enterprise Linux delivers advanced network and clustered file systems such as NFS, CIFS, and GFS2. Block level data management functionality is also built in to Red Hat Enterprise Linux, including logical volume management, software RAID, snapshot capabilities, and backup tools — many of which are enabling enterprise organizations today to meet business continuity and disaster recovery goals for large volumes of data.

- **Availability**

  The Pacemaker cluster manager and GFS2 can be used to achieve five-nines availability in enterprise environments. In addition, Red Hat provides the Red Hat High Availability Add-On which extends reliability, availability, and serviceability for both applications and storage.

- **Virtualization**

  Red Hat Enterprise Linux Server incorporates KVM virtualization capabilities. Red Hat Enterprise Linux can serve as a host for virtual guests and can run as a guest on hypervisors supplied by Red Hat Enterprise Virtualization, VMware, and Microsoft Hyper-V. For many workloads, performance as a guest is nearly the same as performance on a physical machine.

- **Containers**

  Linux containers is Red Hat Enterprise Linux combine lightweight application isolation with flexible image-based deployment. They include the application dependencies so that an application can run across hosts, including Red Hat Enterprise Linux, Red Hat Atomic Host, and OpenShift by Red Hat. Containers combine a small footprint with minimal overhead an simplified maintenance to reduce application deployment costs.

Red Hat Enterprise Linux and SANblox

Benefits of Using Red Hat Enterprise Linux with SANblox

SANblox delivers data reduction technologies that enhance the storage savings already provided with the thin-provisioning capabilities provided by Red Hat Enterprise Linux. SANblox data deduplication, compression, and thin provisioning offer substantial storage savings by delivering data reduction to enterprise storage arrays. Because SANblox is implemented as a separate appliance with its own dedicated processor and memory, none of these features consume additional processing resources on the enterprise array or have any performance impact on the array's data management capabilities. In addition, they do not require any configuration changes on Red Hat Enterprise Linux servers. By using the data reduction technologies of SANblox with Red Hat Enterprise Linux, you can increase array hardware use and reduce acquisition and operational costs.

While deduplication and compression both reduce the amount of consumed storage, these savings remain hidden from Red Hat Enterprise Linux administrators because the Linux view of the storage is at the SANblox LUN layer, and SANblox LUNs represent their provisioned capacity. Red Hat Enterprise Linux can also be configured as an intermediate gateway that delivers advanced CIFS and NFS-based file sharing for SANblox-enabled enterprise storage.

SANblox Thin Provisioning

Traditional (thick) storage provisioning pre-allocates storage and assigns it to a server or VM. Administrators typically overprovision storage to avoid running out of storage and to avoid application downtime associated with expanding provisioned storage. Thin provisioning allows administrators to oversubscribe, delivering storage on demand. Thin provisioned storage is treated as a shared resource pool and only consumed as individual VMs require it. Sharing increases the usage rate of storage by avoiding allocation of capacity that is not needed.

Array (backend) LUNs used for SANblox can be thick or thin provisioned LUNs. SANblox automatically provides 10:1 thin provisioning. Use array-based thin provisioning if a higher ratio is desired, or to allow for the expansion of physical storage. For example, if you initially allocate 25 TB of thick array storage, SANblox will present that as 250 TB. Alternatively, if you are using array-based thin provisioning, you could physically consume just 2.5 TB of storage, logically present that as 25 TB of storage, and SANblox will then present it as the same 250 TB. A major advantage of this second approach is that the array has room to allocate physical storage as needed.

Red Hat Enterprise Linux and SANblox Albireo Deduplication

Enterprise organizations deploy multiple instances of applications on Red Hat Enterprise Linux to compartmentalize needs by department and to scale out processing across multiple servers.

Red Hat Satellite allows administrators to rapidly deploy desktop or server software to physical or virtual servers. A template consists of one or more system images, which can include an operating system, applications, and system patches/updates. Deployment of an image involves copying these files to create a new independent server or VM, which introduces duplicate data for each image deployed. SANblox uses Permabit's Albireo data deduplication technology to eliminate this duplicate data, increasing storage utilization. Application images and data stored across SANblox LUNs share physical blocks.

SANblox deduplication is enabled globally across the storage pool, and data reduction rates will vary based on the data blocks in common across the pool. To achieve the highest data reduction rates, Permabit recommends using a shared SANblox appliance to store images that utilize a common operating system and applications. In multi-SANblox deployments, you might have separate SANblox units for VDI (desktops) and for your server farms.

Red Hat Enterprise Linux and SANblox HIOPS Compression

While deduplication is an ideal data reduction technique for VMs (because of their high rates of redundancy), compression plays a crucial role in realizing savings for application data. SANblox provides inline compression using its HIOPS technology, which operates at a 4 KB block level and is optimized to work in conjunction with the Albireo deduplication feature of SANblox to deliver high performance. When a block of data is sent down from Red Hat Enterprise Linux, SANblox uses Albireo deduplication to check if that block has been seen before. If it has, then it is deduplicated and no compression operation is performed. If it hasn't, then the block is compressed and packed together with other compressed blocks into a single 4 KB block. The performance of sequential reads is dramatically improved by HIOPS Compression because a single physical 4 KB I/O request can read in multiple compressed blocks. Data reduction from HIOPS Compression in Red Hat Enterprise Linux application environments is typically in the 2X to 5X range.

Deployment

Red Hat Enterprise Linux can be deployed on new or existing storage systems fronted by SANblox. SANblox is deployed as a pair of 1U nodes (servers) which must be racked and cabled. SANblox will consume a total of 8 Fibre Channel switch ports (four per switch across two fabrics) and 4 Ethernet switch ports (ideally divided between two switches).

Once configured, the basic procedure for deploying storage for Red Hat Enterprise Linux is as follows:

1. Provision backend storage for SANblox
2. Create the storage pool
3. Provision frontend SANblox LUNs
4. Configure Red Hat Enterprise Linux hosts

Users of Red Hat Linux Virtualization can seamlessly migrate applications to SANblox LUNs, after which the space used on the original LUNs can be reclaimed. This can also be achieved through the use of tools like DRDB or by setting up LVM mirroring to a SANblox device.

Provisioning Backend Storage for SANblox

SANblox uses fabric-attached LUNs for creating its optimized storage pools. A single storage pool will use all available backend LUNs presented to it by the SAN fabric.

Array backend LUNs provisioned to a SANblox storage pool are aggregated together into a striped configuration to allow SANblox to balance load across all constituent devices. As a result, all LUNs in a storage pool must be the same size. The aggregate size of backend LUNs must be at least 1 TB, and should be zoned and mapped to the backend ports of both members of the SANblox HA pair.

The backend ports of SANblox act as SCSI initiators to the backend storage, and you will need to configure switch zoning and storage array host mapping (LUN masking) to access LUNs from those ports.

To verify that the LUNs are correctly mapped to the device, first run this command to force SANblox to rescan the fabric:

```
$ sanblox backend-storage-lun-scan
```

Then retrieve the LUN list to see which LUNs are recognized:

```
$ sanblox backend-storage-lun-list --verbose
 - Backend Storage LUN:
       Name: 360060e8010591e3005701a730000004b
       Path: /dev/mapper/360060e8010591e3005701a730000004b
       Resolved Name: dm-2
       Resolved Path: /dev/dm-2
       Size: 500G
       Vendor: ARRAY_VENDOR
       Product: PRODUCT_CODE
       Used by:
 - Backend Storage LUN:
       Name: 360060e8010591e3005701a730000004c
       Path: /dev/mapper/360060e8010591e3005701a730000004c
       Resolved Name: dm-0
       Resolved Path: /dev/dm-0
       Size: 500G
       Vendor: ARRAY_VENDOR
       Product: PRODUCT_CODE
       Used by:
```

In this example, two LUNs, each 500 GB, have been mapped to the SANblox device. Omitted from this output is the configuration LUN, which will also be displayed from the command. Repeat these commands on the passive node and verify that the list of LUNs is the same.

Multipath I/O is automatically configured for supported storage arrays.

Creating the Storage Pool

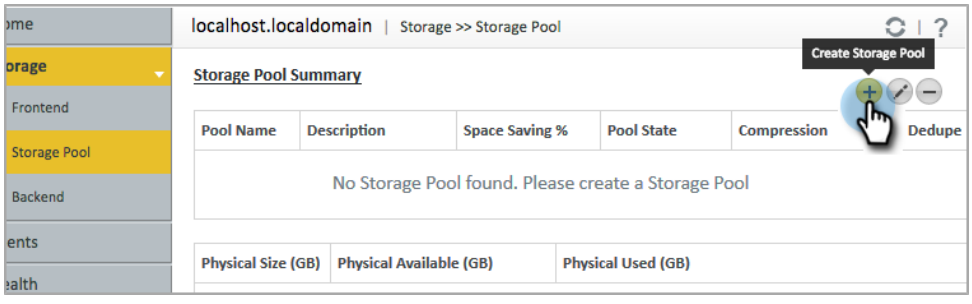
Once backend LUN mapping is in place, you're ready to create an optimized SANblox storage pool. A storage pool stripes backend thin-provisioned storage LUNs together and optimizes the data written to them.

Frontend LUNs provisioned out of a storage pool share the same backend storage. A storage pool is also considered a 'deduplication domain,' so identical blocks written to different frontend LUNs will be identified and deduplicated.

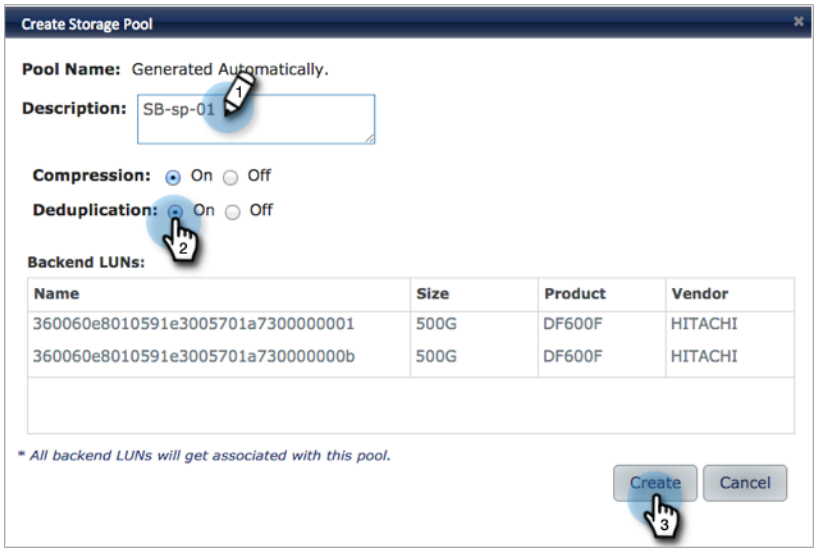
All backend LUNs should be considered as a unit when performing array-based snapshots, backup, and replication. If multiple backend storage LUNs are aggregated together, then they must all be snapshotted simultaneously. Refer to your storage array documentation to see how to take consistent snapshots of a set of volumes.

To create the storage pool, log in to the SANblox Dashboard and perform the following commands:

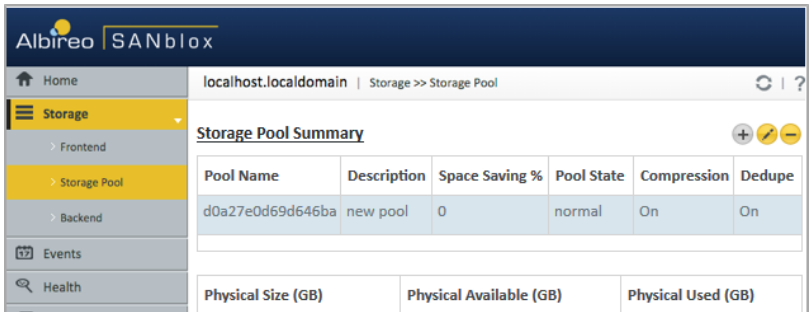
1. From the **Storage > Storage Pool** page, click the **Create Storage Pool** button in the top right.



2. Verify that backend LUNs listed in the dialog are correct (this is important), enter a **Description**, and click **Create**.



The operation may take a few moments to complete, after which the Storage Pool page displays the new pool.



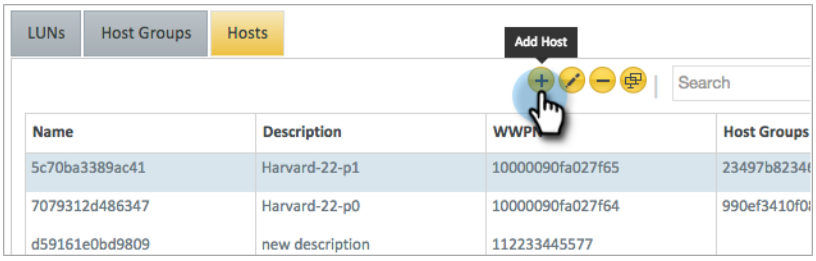
Provisioning Frontend SANblox LUNs

Once the storage pool has been configured, the next step is to provision LUNs and map them for use by Red Hat Enterprise Linux servers.

Adding Hosts

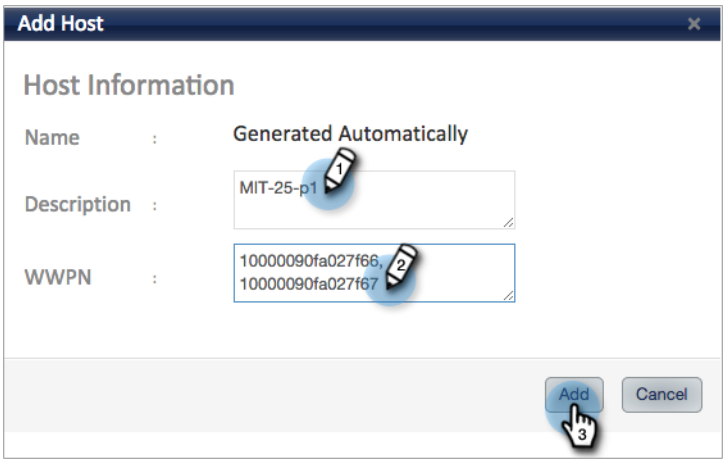
First, you must configure SANblox with the hosts you plan to access.

1. From the **Storage > Frontend** page, click the **Hosts** tab.



2. For each physical host running Red Hat Enterprise Linux Server:

   a. Click the **Add Host** button.



   b. Enter a **Description**.

      The Name of the host is generated automatically.

   c. Enter the host **WWPNs** separated by a comma.

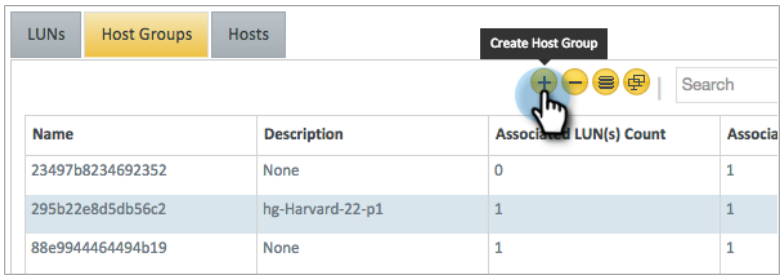
      Refer to your HBA or operating system documentation for instructions on locating the WWPNs for the storage adapters installed on your hosts.

   d. Click **Add**.

Creating a Host Group

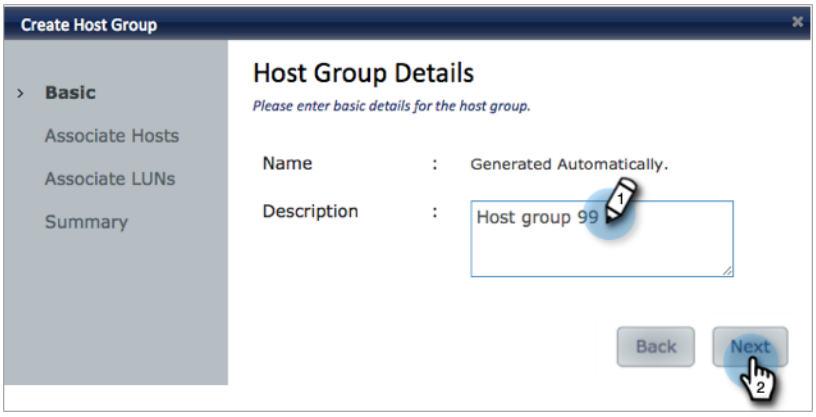
Once your Red Hat Enterprise Linux hosts have been configured, it's time to add them into a host group so that they can be associated with one or more frontend SANblox volumes. A host group is a collection of one or more hosts that require the same access rights to a LUN or group of LUNs.

1. From the **Storage > Frontend** page, click the **Host Groups** tab and then click the **Create Host Group** button.

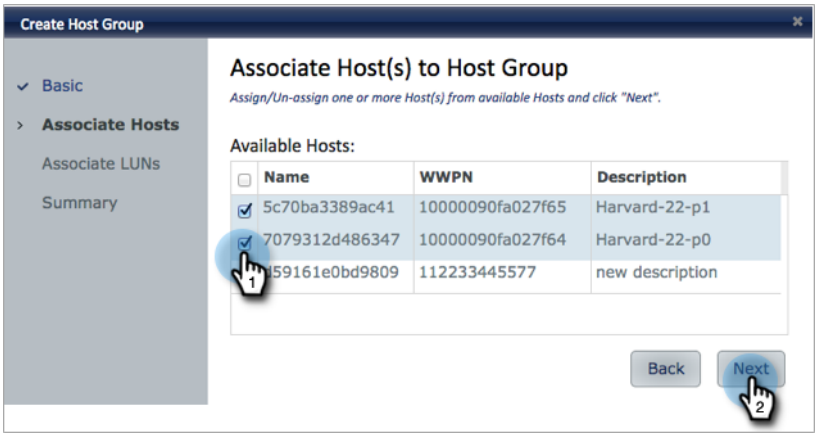


2. Enter a **Description** and click **Next**.

   The Name of the host group is generated automatically.



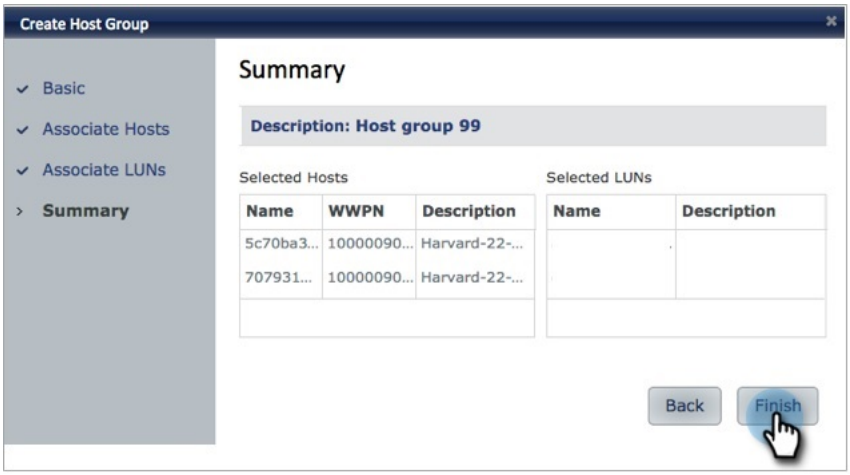
3. Select the hosts to associate with your group and click **Next**.



4. Click **Next** to skip the Associate LUNs page.

   Normally you could choose the LUNs to associate with your host group at this point, but you have not yet configured your LUNs.

5. Review the summary for the host group you are about to create. If you need to make changes, click **Back**. Otherwise, click **Finish.**

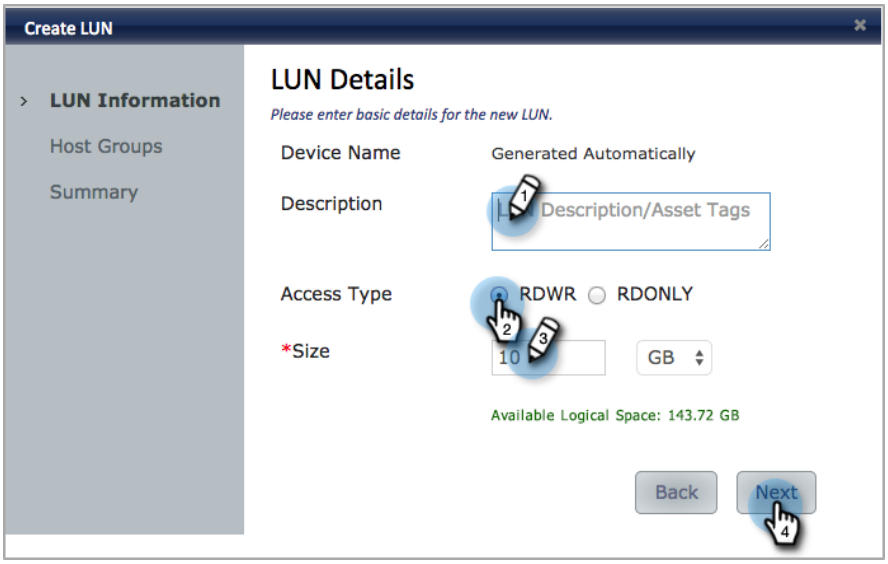


Creating LUNs

1. From the **Storage > Frontend** page, click the **LUNs** tab and then click the **Create LUN** button.

2. Enter a **Description** and ensure that the LUN is configured for **RDWR** (read/write) access.

   The Device Name is generated automatically.

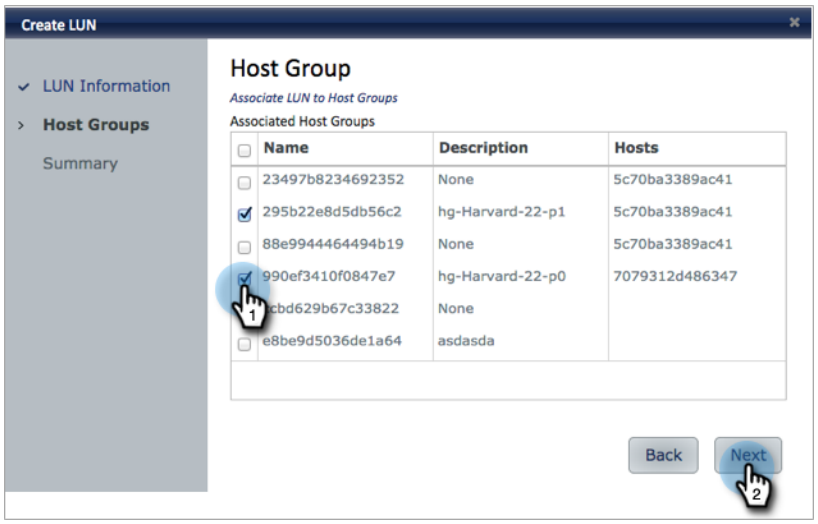


3. Enter the appropriate **Size** for this LUN, and then click **Next**.

   The available space is noted in green text. Enter a size in GB. For smaller LUNs, you can select MB as the unit of measure and enter any multiple of 4 MB.

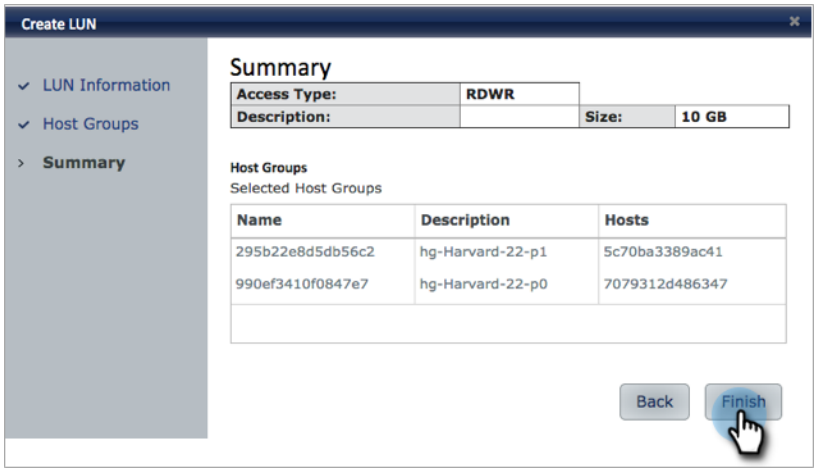
4. Select one or more host groups to associate with the LUN.

   SANblox restricts LUN visibility to only those hosts that are explicitly associated with a frontend volume to prevent unintended hosts from gaining access to SANblox LUNs.



5. Click **Next**.

6. Review the summary for the LUN you are about to create. If you need to make changes, click **Back**. Otherwise, click **Finish**.



Configuring Red Hat Enterprise Linux

Connecting a SANblox LUN to a Red Hat Enterprise Linux client requires two configuration steps, installing the HBA/drivers and configuring multipath mapping.

Installing HBAs and Drivers

The process to install the HBA (host bus adapter) and drivers is out of the scope of this document and will vary depending on the manufacturer of the HBA. This section assumes you have installed the HBA and loaded the initiator functionality to facilitate discovering the LUNs that are discoverable by the client. Discovery also is dependent on correct FC switch zoning and the corresponding host group configuration on SANblox.

1. List the block devices and find the SANblox devices. The following command can be used to list the Permabit SANblox devices and their associated SCSI block device name (e.g., sdx). In this example, we created a single 50G Frontend Storage LUN on SANblox, which is in an HA configuration (two paths per active/passive node).

```
# for i in /sys/block/sd*; do \
  echo -n "$(basename $i)"; \
  echo -n "$(cat $i/device/vendor) "; \
        cat $i/device/model; done
sda ATA     WDC WD1002FBYS-0
sdi PERMABIT SB100
sdj PERMABIT SB100
sdk PERMABIT SB100
sdl PERMABIT SB100
```

2. Verify that the list contains the expected paths.

   In the example, four block devices were discovered by the initiator. The number of devices you will have is determined by the total number of active and passive paths from the SANblox to the client, which may vary due to zoning and SANblox HA configuration. At least two devices should be discovered, one active and one unavailable.

Configuring Multipath

To benefit from the throughput available by utilizing multiple paths to storage, or HA (High Availability) capabilities, multipath must be configured on the client.

1. Ensure that your Red Hat Enterprise Linux system has `device-mapper-multipath` installed.

2. In your multipath configuration (typically, `/etc/multipath.conf`), we recommend using the following configuration for SANblox in an HA configuration:

```
defaults {
       udev_dir /dev
       path_selector "round-robin 0"
       path_grouping_policy group_by_prio
       features "1 queue_if_no_path"
       prio alua
       failback manual
       no_path_retry 30
       user_friendly_names yes
       path_checker tur
}
```

3. After making that configuration change, run the following commands to refresh and list the multipath state:

```
# multipath –r
# multipath –ll
  mpathg (23638323034653230) dm-2 PERMABIT,SB100
  size=49G features='0' hwhandler='0' wp=rw
     |-+- policy='round-robin 0' prio=50 status=active
     | |- 6:0:2:0 sdj 8:144 active ready running
     | `- 7:0:2:0 sdl 8:176 active ready running
     `-+- policy='round-robin 0' prio=0 status=enabled
     |- 7:0:3:0 sdi 8:128 active ready running
     `- 6:0:3:0 sdk 8:160 active ready running
```

4. Verify the multipath configuration is correct based on your SANblox and LUN configuration.

   The example above shows two priority groups. These represent the active and passive group of paths to a single SANblox LUN. The active and passive groups are denoted by `status=active` and `status=enabled` respectively.

   There are multiple paths within each priority group, reflecting the paths to this LUN over both SANblox FC ports. The groups have a roundrobin I/O queuing policy and a higher priority on the active policy group to ensure the passive node will not receive I/O. The multipath device, `mpathg`, has been created at `/dev/mapper/mpathg` and `/dev/dm-2`.

Performance Considerations

Several factors contribute to optimizing Red Hat Enterprise Linux performance in a SANblox environment.

Properly configured SAN fabric switches introduce only minor latency because of their low latencies relative to servers, SANblox, and storage arrays. Make sure that the paths through the switch fabric are not saturated — that is, ensure that the switch fabric is not already running at its maximum throughput.

SANblox Performance

- For best performance, a SANblox storage pool containing Red Hat Enterprise Linux LUNs should be dedicated to applications with a similar workload. For example, mixing virtual desktop and Red Hat JBOSS Middleware applications on the same storage pool is not advisable.

Backend Array Performance

- Ensure that each LUN presented to SANblox by the backend array is on a RAID group that provides adequate performance for application hosts. Pay attention to activities and resource utilization of other LUNs in the RAID group, as they may prevent you from achieving your performance goals. In general, a RAID group containing the SANblox LUNs should not include LUNs used by other applications.
- Make sure read/write caching is enabled.
- Distribute the paths to the LUNs presented to SANblox among all available storage processors to provide optimal load balancing.

Red Hat Enterprise Linux Performance

- Permabit generally recommends using the no-op scheduler for fast storage.
- Tuned and tuned-adm provide a number of profiles designed to improve storage performance.
- Ensure all software is up-to-date with critical patches and recommended updates.
- Ensure hosts have the latest BIOS version, as well as other hardware devices to address any known supportability issues.
- Each server should have a sufficient number of HBAs to allow maximum throughput for all the applications/VMs hosted on the server for the peak period. I/O spread across multiple HBAs provides higher throughput and less latency for each application.
- To provide redundancy in the event of HBA failure, make sure the server is connected to a dual redundant fabric.

Conclusions

When introduced into a Red Hat Enterprise Linux SAN environment, SANblox delivers substantial data reduction benefits, reducing data footprint while enhancing performance. The ease of SANblox deployment and management makes this data reduction appliance an ideal solution for supporting Red Hat Enterprise Linux environments in new or existing Fibre Channel SANs.

This solution profile is intended to provide an overview of SANblox deployment in Red Hat Enterprise Linux environments. Additional guidance may be required to resolve user-specific deployment issues. For additional assistance, contact your reseller to speak with a SANblox implementation specialist.

Your feedback on this document is welcome. Send comments to marketing@permabit.com. Refer to “RHEL Server SSP” in the Subject of your email.

References

Miscellaneous References

- ESG: Permabit SANblox: Fitting More Bricks in Storage Buildings

  http://www.esg-global.com/blogs/permabit-sanblox-fitting-more-bricks-in-storage-buildings
- Wikibon: Permabit Throws a Data Reduction Lifeline to Legacy Arrays

  http://wikibon.org/wiki/v/Permabit_Throws_a_Data_Reduction_Lifeline_to_Legacy_Arrays
- Wikipedia: RAID Definitions and Explanations

  http://en.wikipedia.org/wiki/RAID

Permabit References

- Albireo SANblox Datasheet

  http://permabit.com/resources/sanblox-data-sheet
- Permabit Albireo SANblox FAQ

  http://permabit.com/resources/sanblox-faq

Red Hat References

- Red Hat Satellite Documentation

  https://access.redhat.com/documentation/en-US/Red_Hat_Satellite/
- Red Hat Enterprise Virtualization Documentation

  https://access.redhat.com/documentation/en-US/Red_Hat_Enterprise_Virtualization/
- Red Hat Enterprise Linux Documentation

  https://access.redhat.com/documentation/en-US/Red_Hat_Enterprise_Linux/