OPEN

Compute Project

Storage:
Example of Open Storage Ethernet Device Implementation

Rev 3

Author: Alvin Cox, Seagate Technology LLC

Author: Asghar Riahi, Seagate Technology LLC

1 Scope

This document illustrates an example implementation of the Storage Device with Ethernet Interface in an existing system that uses SAS expanders to communicate to SAS or SATA hard drives. The reference system is the Open Vault Storage System defined in Open Vault Storage Specification v0.8. Only substitution of the storage device and associated expander cards is covered in this example. Although enclosure management depends on data provided by the storage devices, implemention of enclosure management is not addressed in this document, but is necessary for the system to function properly.

2 Contents

1 Scope 2

2 Contents 2

3 Table of Figures 2

4 Open Storage with Ethernet Device (OSED) 3

5 OSED Storage System 4

5.1 Open Vault Storage System Overview 4

5.2 Open Vault Storage System Using OSED 7

6 FAQ 11

7 Revision History 13

3 Table of Figures

Figure 1 Storage device with Ethernet interface (OSED) 3

Figure 2 Open Vault Storage System 4

Figure 3 HDD tray diagram 5

Figure 4 SEB diagram 6

Figure 5 Connection between SEB and Winterfell 6

Figure 6 HDD tray diagram using OSEDs and NEBs 7

Figure 7 NEB diagram 8

Figure 8 Connection between NEBs and Rack 10G Switch 9

Figure 9 Open Vault Storage System with OSEDs and NEBs 10

4 Open Storage with Ethernet Device (OSED)

The storage device with Ethernet interface provides a common configuration for the implementation of the Ethernet interface on a storage device. It is intended to support large configurations of storage devices in data center applications. Other applications may find the features of this device advantageous where direct communication to the device through the Ethernet interface is desired. Figure 1 shows the storage device and connector. For additional details see "Storage device with Ethernet Interface" specification at this URL under IC Review section: http://www.opencompute.org/wiki/Storage/Dev/.

The following section illustrates how the storage device with Ethernet interface (OSED) may be used to replace SAS or SATA drives in the Open Vault Storage System to reduce the system complexity and allow increased overall capacity.

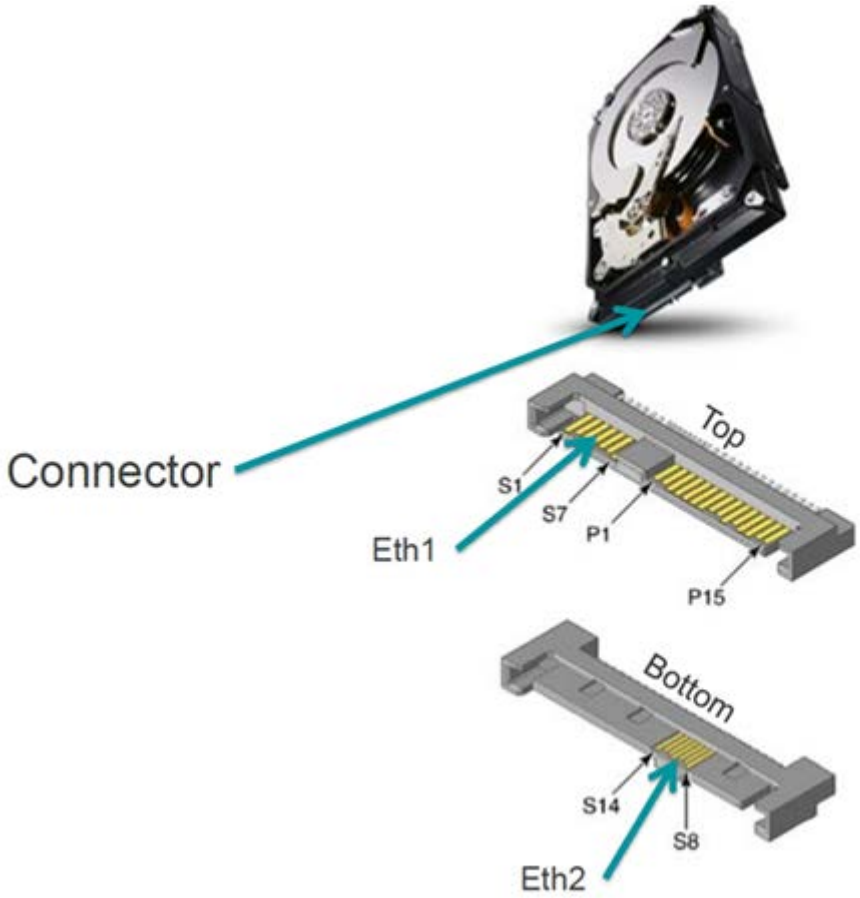

Figure 1 Storage device with Ethernet interface (OSED)

5 OSED Storage System

5.1 Open Vault Storage System Overview

The configuration of the Open Vault Storage System defined in Open Vault Storage Specification v0.8 is illustrated in Figure 2. The maximum storage capacity is achieved when there are 8 Knox storage slots per one Winterfell server. This provides a total number of storage devices available in the Open Vault Storage System equal to 480 (16 *30 = 480). The storage rack is attached to the host system through Ethernet cables connected to the 10G Switch of the storage rack.

Figure 2 Open Vault Storage System

The section labeled Knox illustrated in Figure 2 is referred to as an Open Vault Storage Unit in the Open Vault Storage System. The Open Vault Storage Unit is a 2U-30HDD storage enclosure, consisting of two identical 1U high HDD trays with fifteen 3.5" HDDs and slots for two SAS expander boards on each, one fan control board, and six redundant fan modules mounted externally in the rear of the chassis (see Figure 3). Each HDD tray contains one drive plane board, one power transition board and at least one SAS expander board. Within the rack, each SAS expander board (SEB) (see Figure 4) is connected with Mini SAS cables to a Winterfell server (see Figure 5). The Winterfell server provides a bridge function to convert the SAS interface of the HDD tray to Ethernet which is then connected to the 10G Switch in the top of the rack.

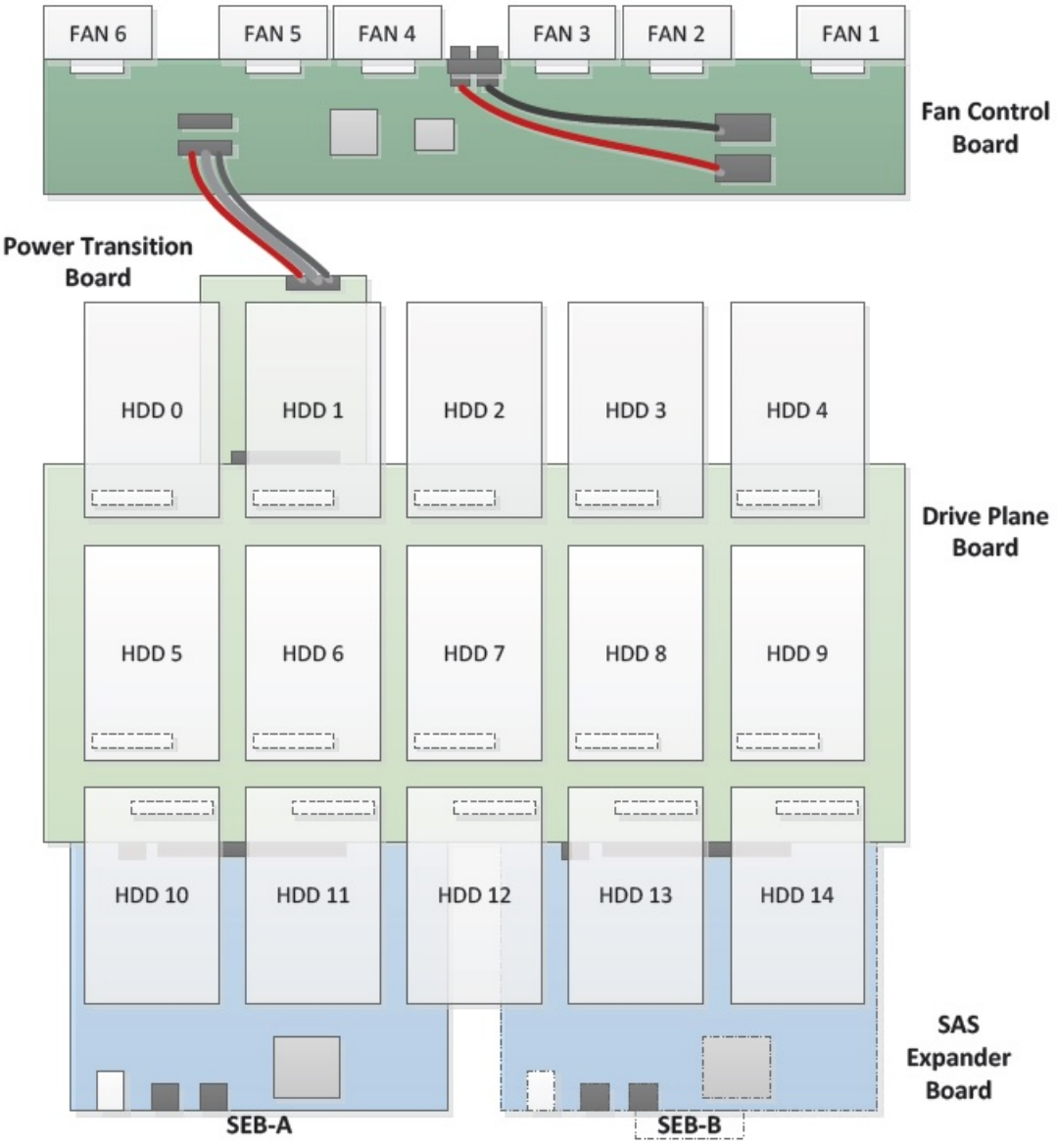

Figure 3 HDD tray diagram

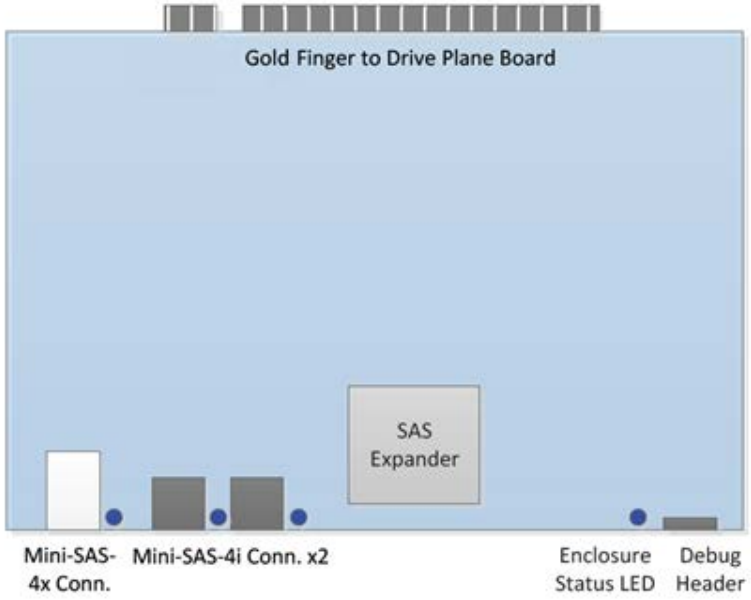

Figure 4 SEB diagram

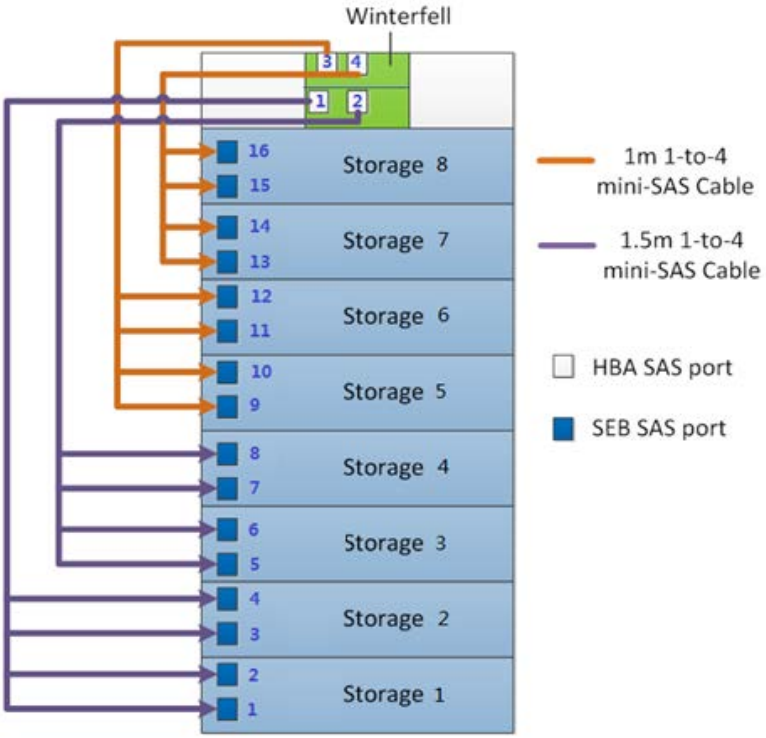

Figure 5 Connection between SEB and Winterfell

5.2 Open Vault Storage System Using OSED

The Open Vault Storage System may be simplified by replacing the SAS/SATA HDDs with OSEDs and changing the SEB to a Network Expander Board (NEB). When OSEDs (i.e., Ethernet storage devices) are used with an NEB, the NEB board may be connected directly to the Ethernet 10G Switch in the top of the rack, eliminating the need for the Winterfall to provide a SAS-to-Ethernet translation. The HDD tray diagram using OSEDs and NEBs is shown in Figure 6.

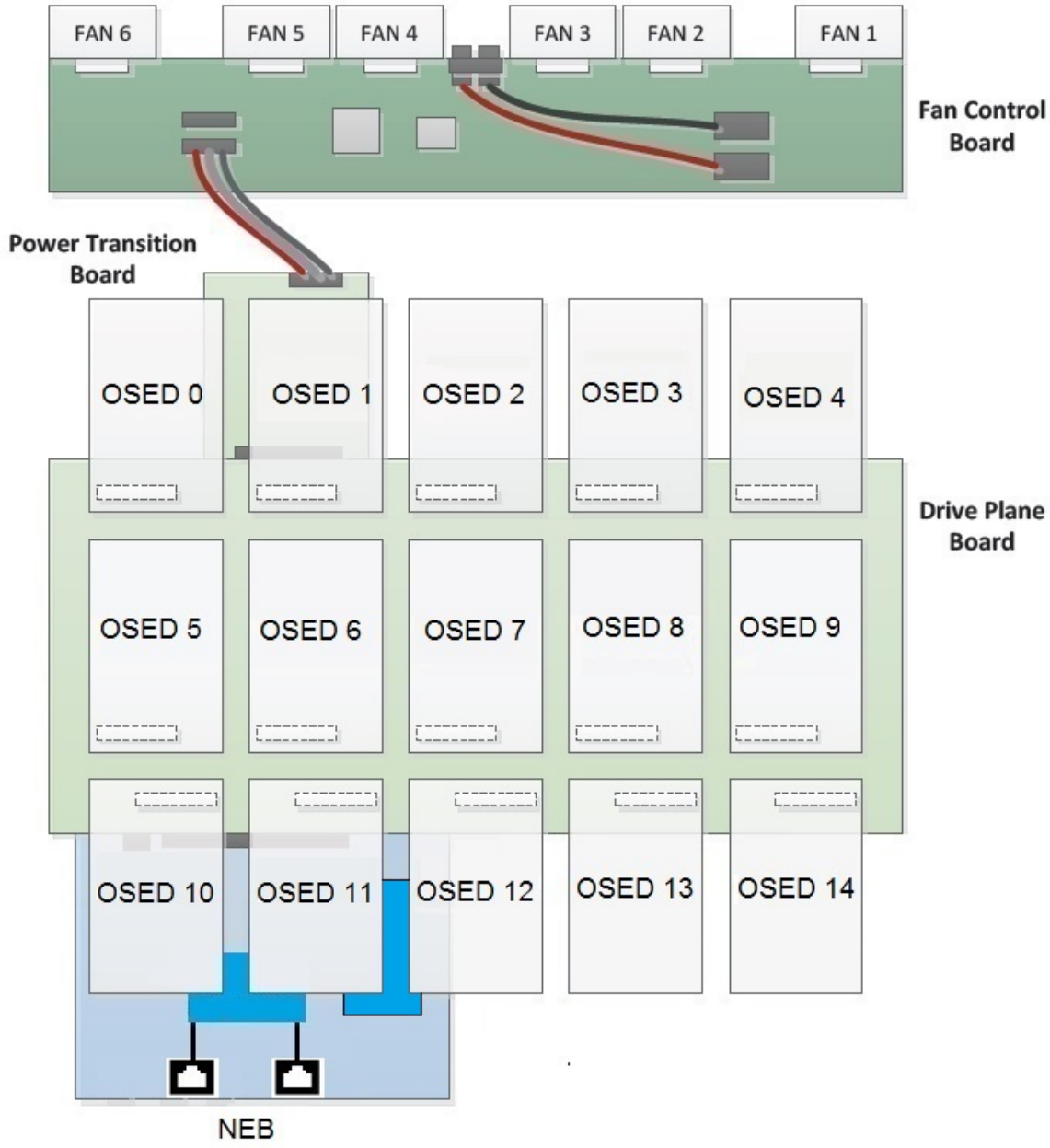

Figure 6 HDD tray diagram using OSEDs and NEBs

The same drive plane board may be used for either the SAS/SATA or OSED storage devices. The drive plane board should be populated with only one type of drive (i.e., SAS/SATA or OSED) since the expander board is unique to the type of interface. Figure 7 shows the NEB diagram. Two RJ45 jacks are used to connect each tray in the Knox to the 10G Switch in the top of the rack. There is an Ethernet Switch on the NEB that is used to redirect the data traffic to individual OSEDs on the drive plane board. The enclosure status LED and debug header are not illustrated in the diagram as implementation of enclosure management is not addressed in this example.

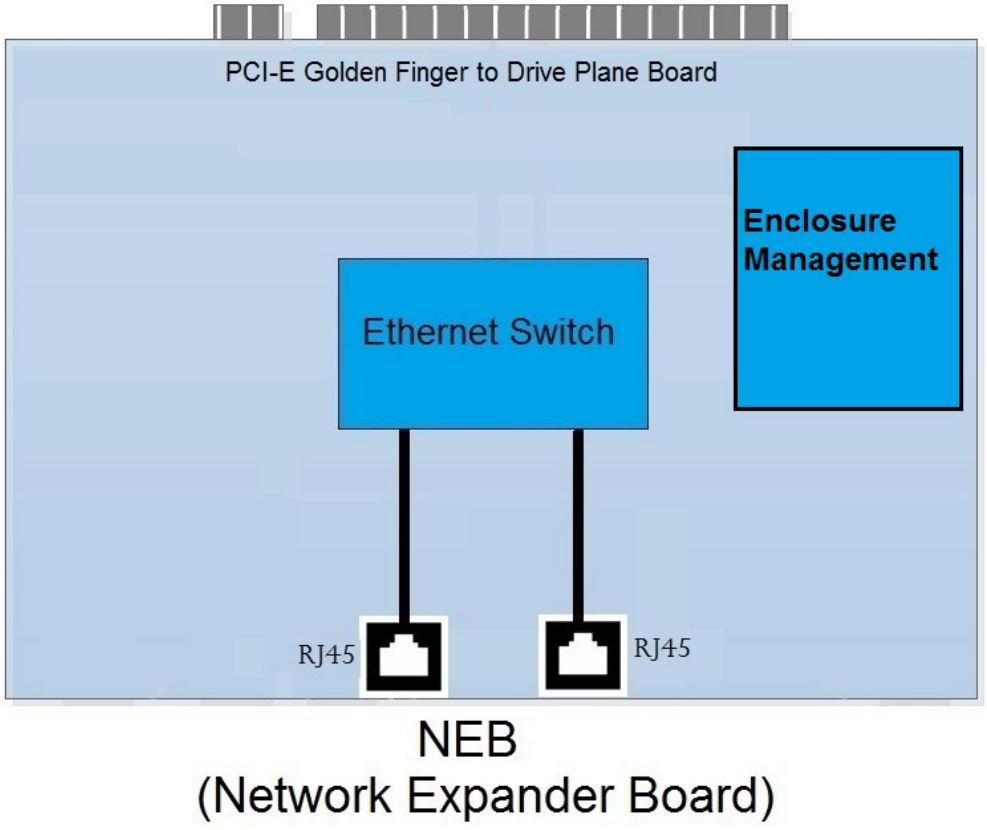

Figure 7 NEB diagram

The NEB is directly connected to the 10G Switch in the top of the rack rather than to the intermediate Winterfell server. The connection diagram is shown in Figure 8. Figure 9 shows how two additional 2U slots are available in the rack where the Winterfell servers were located. These slots may be used for two additional Open Vault Storage Units when OSEDs and NEBs are used instead of SAS/SATA HDDs and SEBs, allowing a total number of storage devices available in the Open Vault Storage System to equal 540 Storage Devices (18 *30 = 540).

The original configuration supported up to 480 Storage Devices. With OSEDs in the rack the maximum storage increases by 12.5% as compared to SAS/SATA HDDs for Storage Devices of equal capacity.

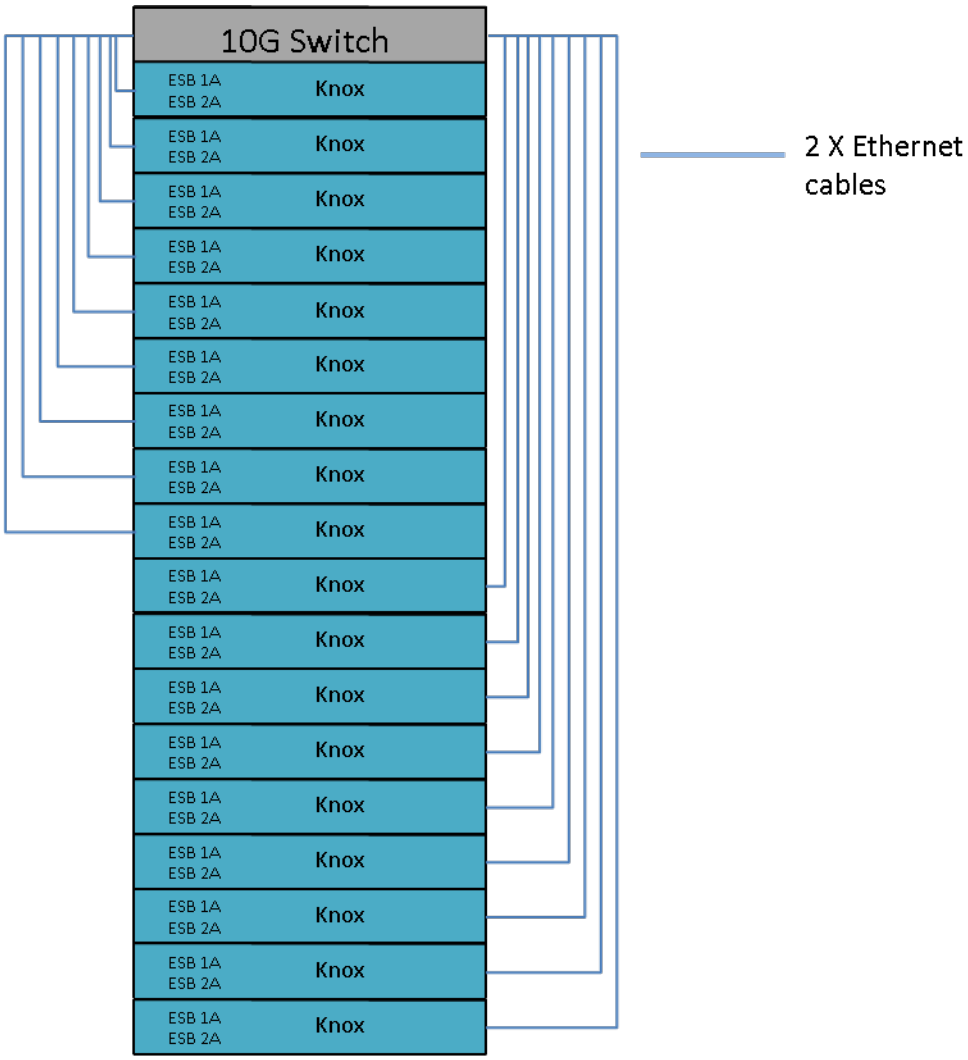

Figure 8 Connection between NEBs and Rack 10G Switch

Figure 9 Open Vault Storage System with OSEDs and NEBs

6 FAQ

1. Q: Is there any software needed (Driver's etc.) If so should define?

   A: No the system uses standard Ethernet communication protocols: The SEB (SAS expander board) provides enclosure management services to the chassis. Propose a solution to replace this function.

2. Q:The SEB (SAS expander board) provides enclosure management services to the chassis. Propose a solution to replace this function.

   A: The existing enclosure management may be duplicated or an alternate solution may be implemented per the system design requirements. This example focuses only on the OSED storage interface. (We are looking at the broader application of the OSED)

3. Q:The SAS expander has been qualified to operate with 6G SAS signals. Signal integrity from OSEDs have not been qualified.

   A: Since the OSED is using Gigabit Ethernet, there are no issues with the existing infrastructure with regards to signal integrity.

4. Q: How do estimated power budgets of NEBs compare to the SEBs? Can the Knox support the power requirement?

   A: The Ethernet switching can be implemented with less power than the SAS expander.

5. Q: Why RJ45? Can other interfaces be used?

   A: We chose RJ45 because It is a common Ethernet connector. Any connector that supports Ethernet signaling may be used at the discretion of the implementer.

6. Q: Propose a sample switch IC for use. Does the switch require a management processor? What type?

   A: Since the switch is used to forward the TCP/IP packets to a specific IP address (OSED storage device) the complexity of the switch is at the discretion of the implementer.

7. Q: Do bandwidth bottlenecks exist in this system? Where are they?

   A: There is no bottleneck in this system,

8. Q: I would like to see more information on the NEB. Here are a few things that I thought of.

a. Q: Does a reference design exist yet?

   A: Not for the Knox box but other implementations do exist. See: http://rnt.de/en/bf_storageobject.html

b. Q: Is it a switch, or a NIC, or could it be both?

   A: It is both

c. Q: What are the details behind its configuration and management? Does it include any sideband or debug console interface?

   A: The details of the configuration is at the discretion of the implementer.

d. Q: How would it integrate with a larger network?

   A: It will be integrated like any other switch

e. Q: How are the drives presented on the network, behind the NEB?

   A: Each drive has two IP addresses.

      i. Individual L2/L3 devices with their own MACs and IPs, visible everywhere?

         A: Yes

      ii. Hidden behind an application proxy service?

         A: No

      iii. As devices behind a NAT or PAT?

         A: No

For additional FAQ regarding the Storage device with Ethernet interface refer to wiki FAQ:
https://developers.seagate.com/display/KV/FAQ

7 Revision History

Version	Date	Author	Changes
1	04/03/2014	Alvin Cox, Asghar Riahi	Initial Public Version
2	08/13/2014	Alvin Cox, Asghar Riahi	Add revision history
3	08/20/2014	Alvin Cox, Asghar Riahi	Add FAQs