TECH GUIDE

InfiniFlash™ System Configuration and Tuning Guide for NexentaStor™

SanDisk®
a Western Digital brand

Western Digital Technologies, Inc.
951 SanDisk Drive, Milpitas, CA 95035

www.SanDisk.com

Table of Contents

Executive Summary

This guide is intended for use when installing the InfiniFlash System with NexentaStor. The following detail and screenshots were captured using Supermicro® X10 Servers, other server installations may vary slightly.

1. Configuring BIOS Settings

1.1 Configure IPMI

Set the IP of the IPMI so that you can access the remote console

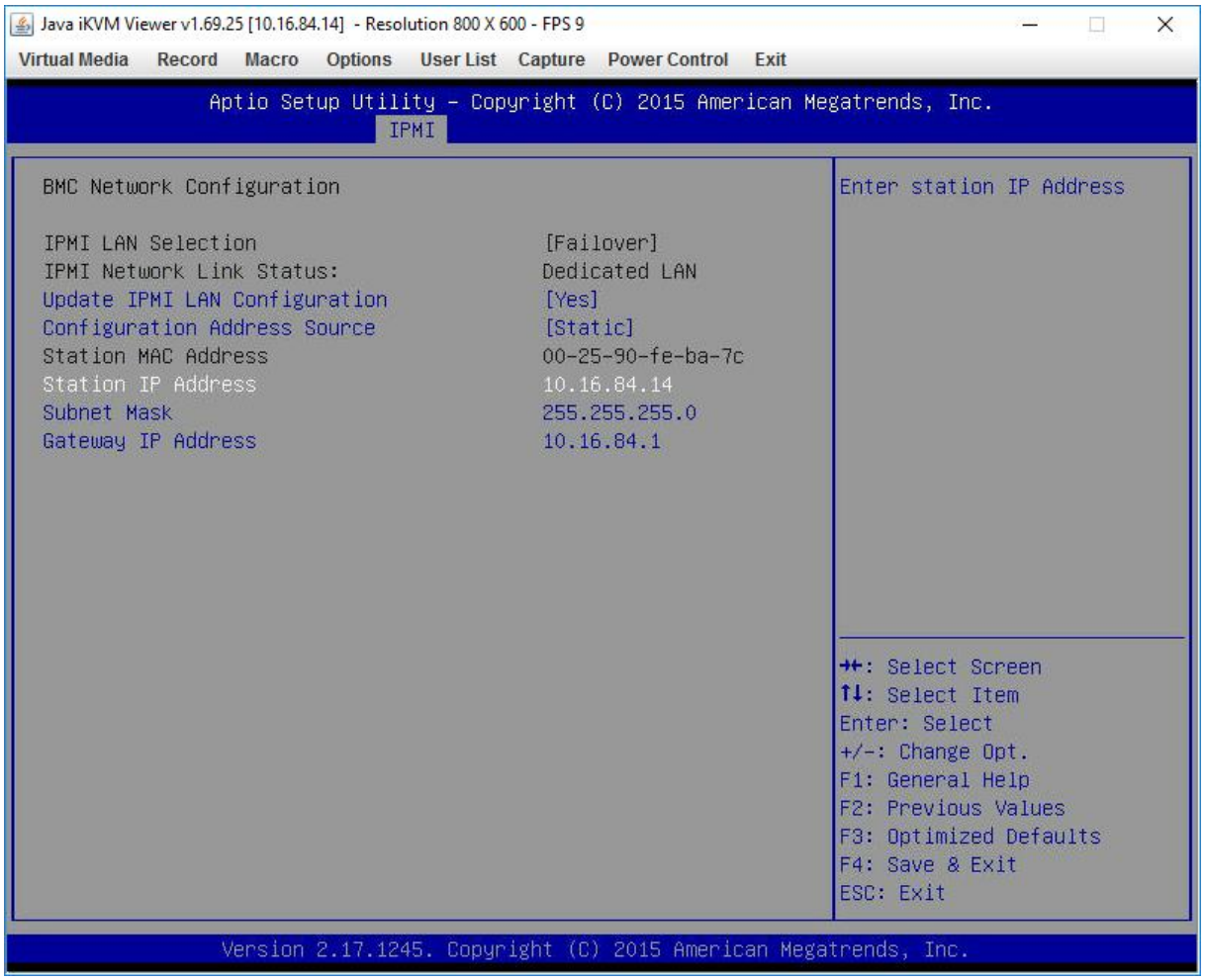

Figure 1. IPMI Configuration

1.2 Hyper-Threading and Virtualization

Disable the settings as shown below in the following section:

Advanced –> CPU Configuration -> Hyper-Threading

Advanced –> CPU Configuration -> Intel Virtualization Technology

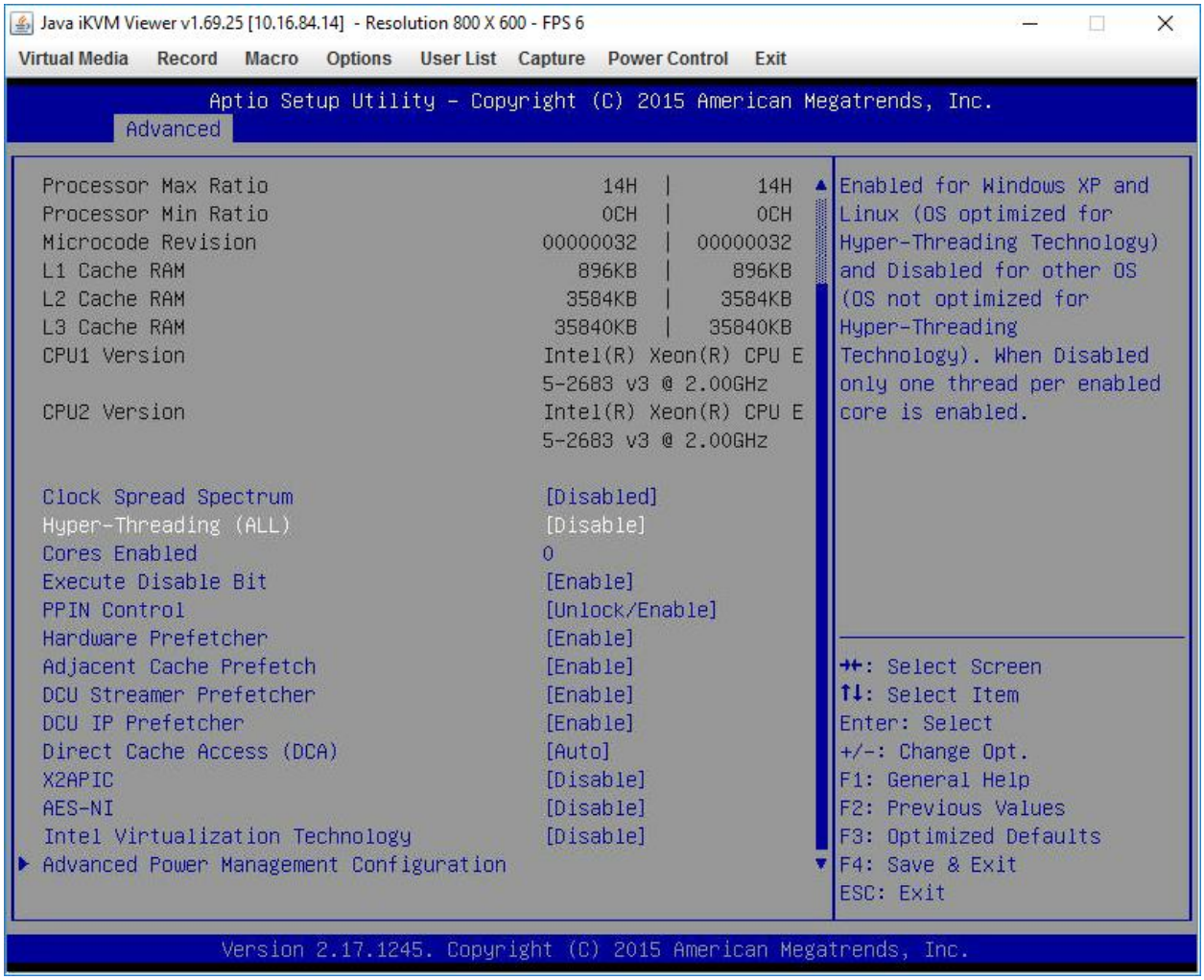

Figure 2. Hyper-Threading and Virtualization Configuration

1.3 Disable USB 3.0

Disable USB 3.0 support:

Advanced –> Chipset Configuration –> South Bridge –> USB 3.0 Support

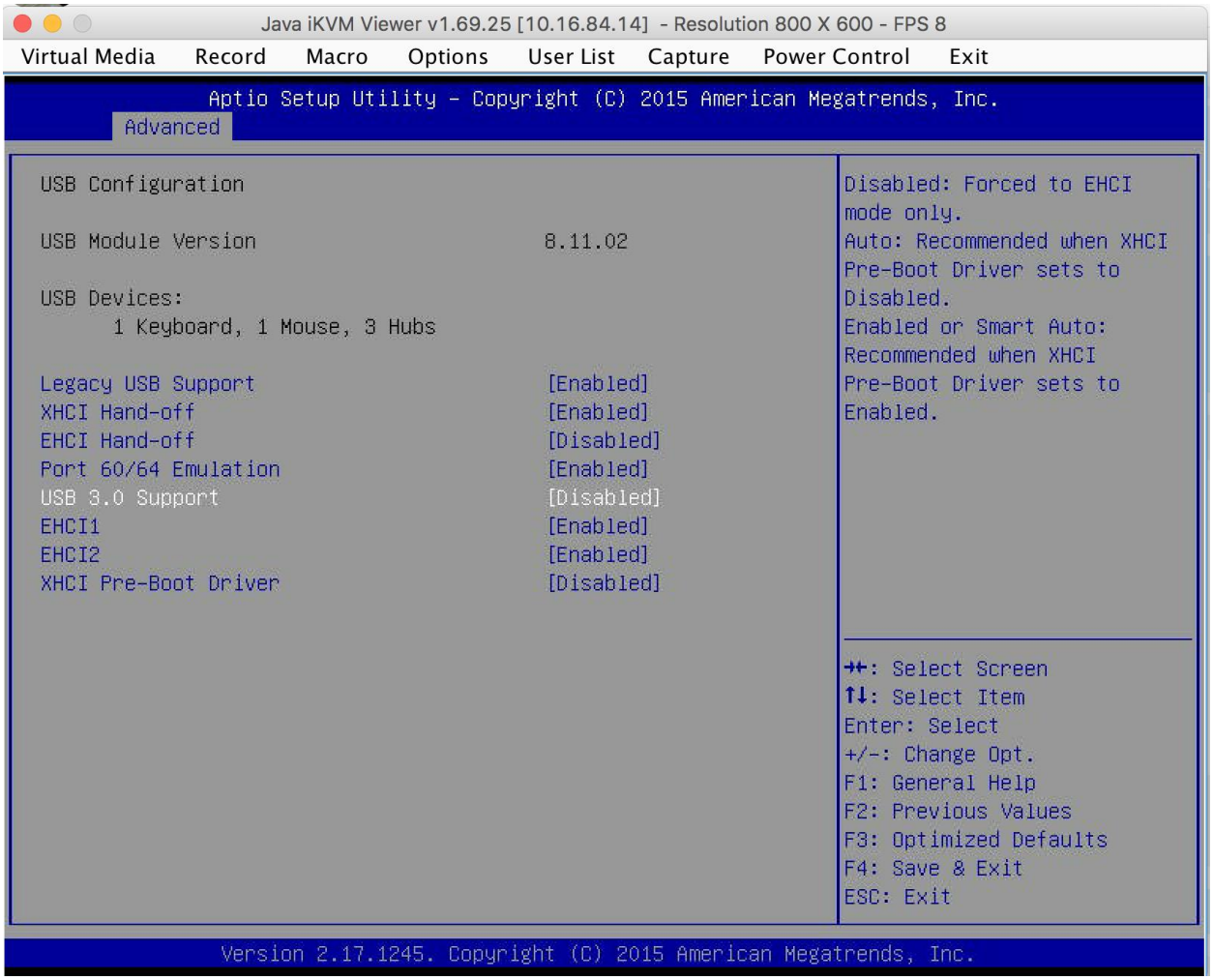

Figure 3. USB Configuration

1.4 CPU Power States

Configure the CPU power states:

Advanced –> CPU Configuration –> Advanced Power Management Configuration->Energy Performance BIAS Setting

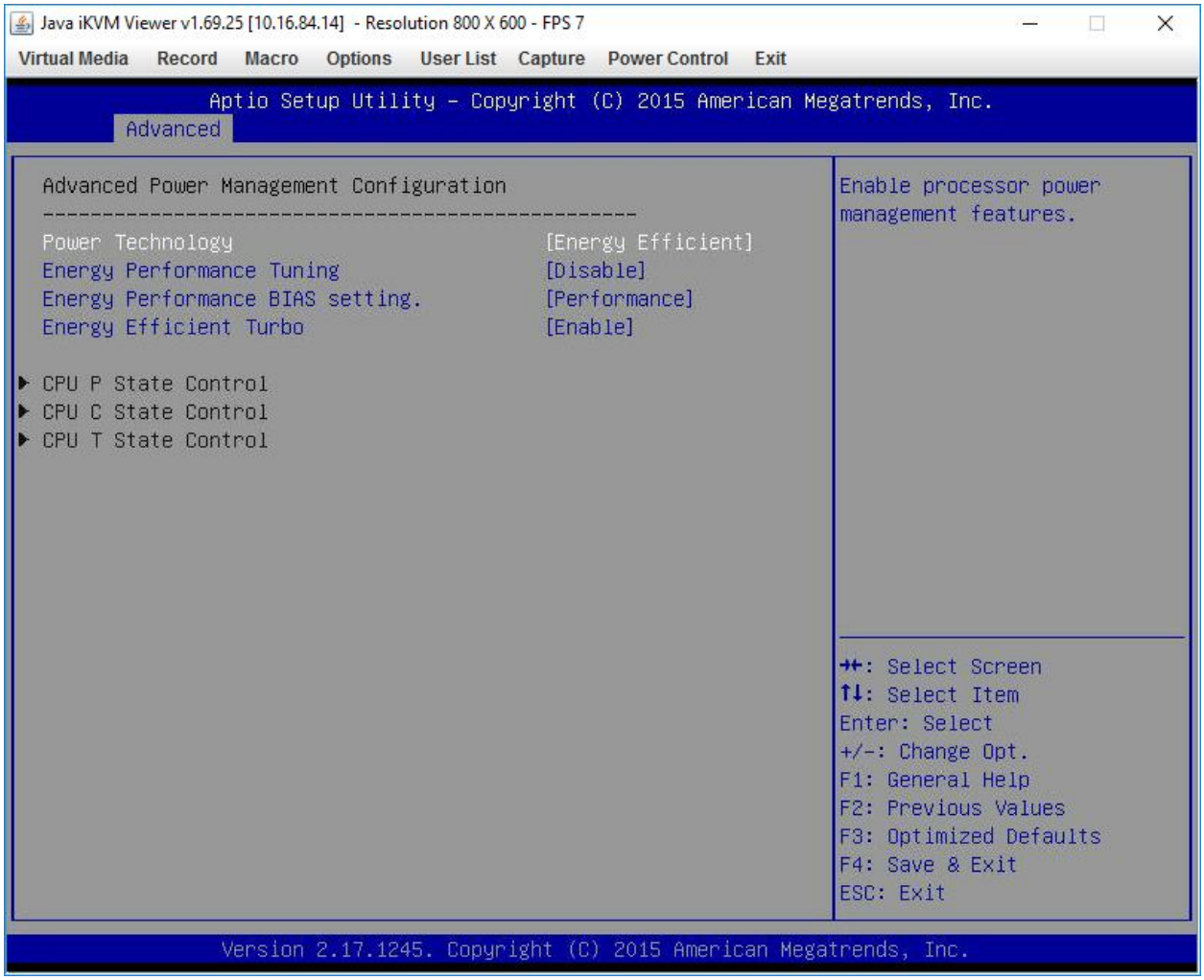

Figure 4. CPU Power State Configuration

1.5 AHCI Setting

Advanced – SATA Configuration

Configure SATA as AHCI as shown below.

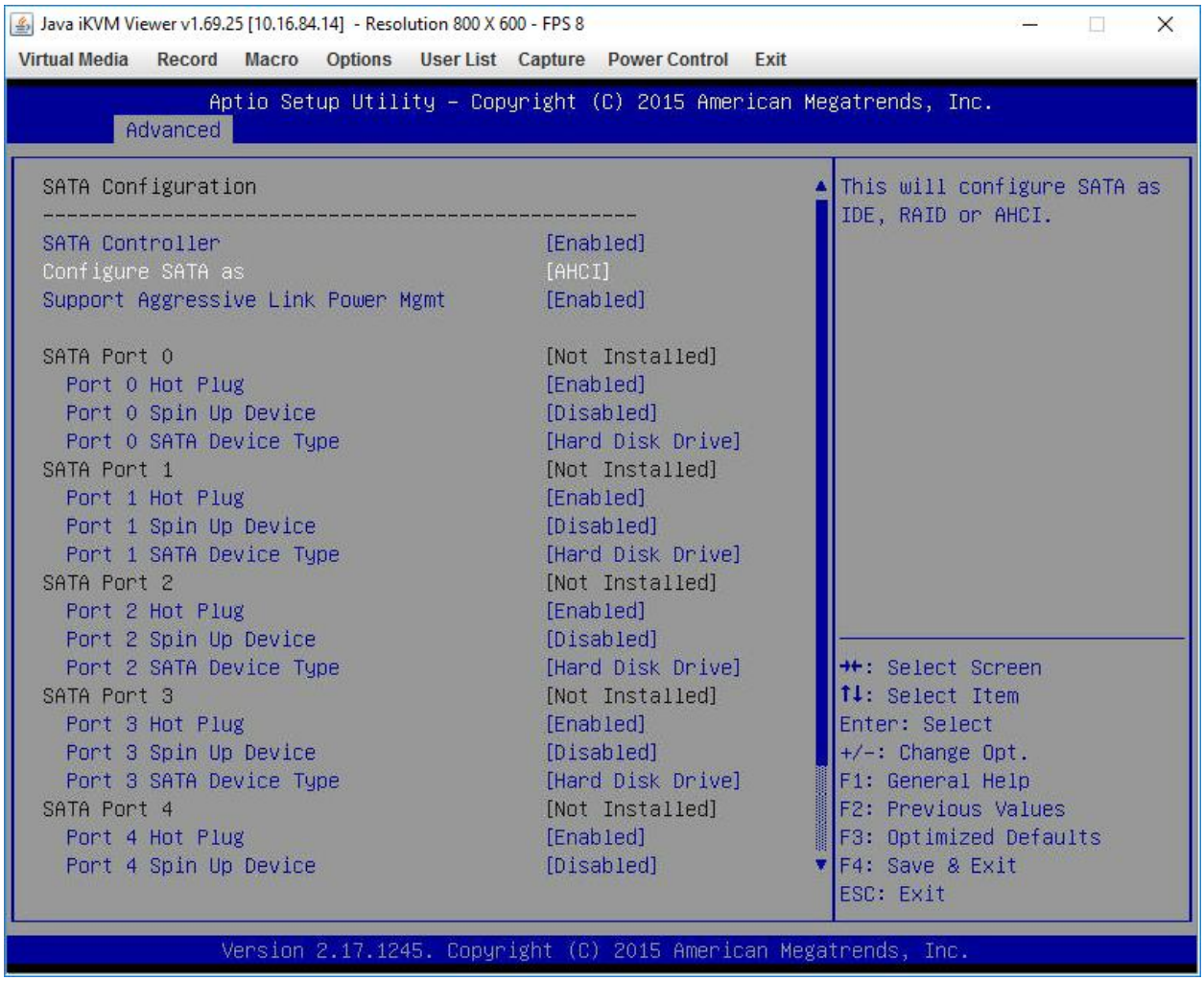

Figure 5. SATA Configuration

2. NexentaStor Tunables

2.1 Tuning Devices with Larger Queue Depths

You may increase the vdev queues if you know the device's queue depths. When deploying NexentaStor on InfiniFlash, add the lines below to the file

/etc/system:

set zfs:zfs_vdev_sync_read_max_active = 30

set zfs:zfs_vdev_sync_write_max_active = 30

set zfs:zfs_vdev_async_read_max_active = 9

set zfs:zfs_vdev_async_write_max_active = 30

set sd:sd_max_throttle=128

2.2 Generic Pool Tuning

The settings below will improve compression and reduce write inflation by reducing the metadata copies and by not modifying access time on a file read. This is done on a per-pool basis.

Run the zfs commands below in a bash root shell:

zfs set compression=lz4 <InsertPoolNameHere>

zfs set redundant_metadata=most <InsertPoolNameHere>

zfs set atime=off <InsertPoolNameHere>

2.2.1 InfiniFlash Recommendations for Pool Configurations

Nexenta recommends that pools configured on InfiniFlash comprise one of:

- 1+1 mirrors with a minimum record size of 16KB
- raidz1(4+1) vdevs with record size set to 64KB
- raidz2(8+2) vdevs with record size set to 128KB

Choosing between the options above will be a function of performance, storage efficiency, and redundancy requirements. More vdevs with a smaller record size will yield better IOPS performance with smaller random IOs.

Configuring the storage pool with 1+1 mirrors provides the best performance at the expense of usable capacity. The record size should be tuned based on the block size of the application in order to maximize application performance. Because InfiniFlash uses a native block size of 16k, the record size of the storage pool should always be 16k or higher to prevent write amplification.

2.3 Multi-Pathing Configuration

The mpxio driver will take the multiple paths to a disk drive and present them as a single, logical disk. The driver uses the configuration file /kernel/drv/scsi_vhci.conf to determine if a disk is supported by mpxio. Nexenta maintains a list of supported storage devices and their recommended settings, which can be downloaded from their website.

The current configuration file can be found here: http://www.nexenta.com/scsi_vhci.new

These are the commands for installing the file:

```
cd /kernel/drv/
wget http://www.nexenta.com/scsi_vhci.new
cp scsi_vhci.conf scsi_vhci.conf.orig
cp scsi_vhci.new scsi_vhci.conf
```

This will be an entry in /kernel/drv/scsi_vhci.conf for the InfiniFlash flash cards. If this line is not present, or if the latest configuration file cannot be downloaded, it can be entered manually by adding the following line near the end of /kernel/drv/scsi_vhci.conf:

```
"SANDISK S", "f_sym",           # SanDisk Infiniflash
```