本文档主要整理了 iSCSI 的基本知识、创建及删除的一些命令及截图，还有关于 vmfs 系统里面的一些概念。

对于小小实习生而言，还是花了一点时间来了解 Esxi 的 iSCSI 的结构，先上图吧。

一、iSCSI SAN 存储器虚拟化架构

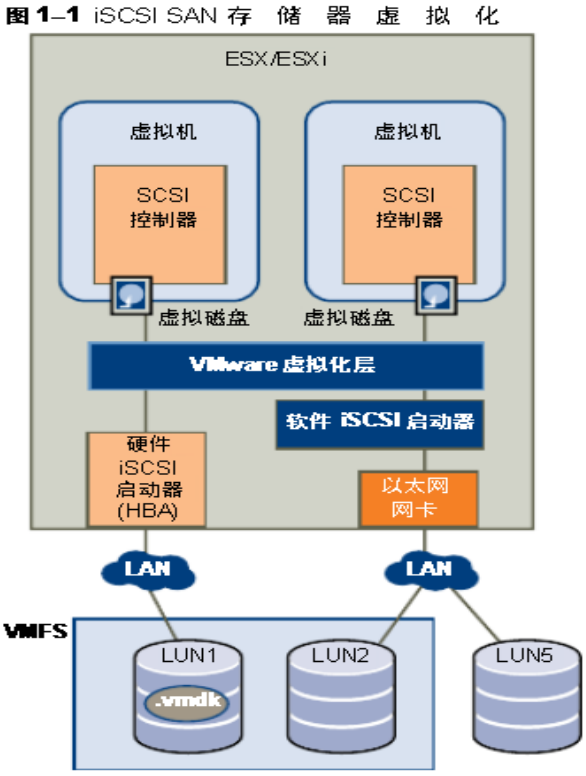


图 1 iSCSI SAN 存储器虚拟化

关于上图的解释，在 iSCSI 里面，虚拟机只是 register 到了 host 上面，并没有直接安装到 Host。它的安装文件 vmdk 是在 storage Lun 里面的 datastore 里。Esxi 与 storage Lun 之间通过网线（在同一个局域网 LAN 下面）相连接，它们在通信之前需要先发现 lun。

在跑测试的过程中，常常会因为找不到 Lun 而 failed,所以发现 lun 的过程很重要。重新装完新版本的 build 后，需要重新发现 Lun，才可以创建 datastore 等操作。

二、ESX/ESXi 支持 iSCSI 的三种方式

A，普通网卡，靠软件实现 iSCSI 封装，成本低，会占用较多的主机资源。
B，采用带有 iSCSI 封装功能的网卡，如 Broadcom 5709 等，它可以分担一部分 iSCSI 的封装操作，它的工作需要 VMware 的虚拟化软件进行配合。

C，全功能的 iSCSI HBA 卡，与 FC HBA 卡类似，可实现从 I/O 到 iSCSI 到 TCP/IP 的全部处理过程，成本较高，但性能好，基本不占用主机资源。

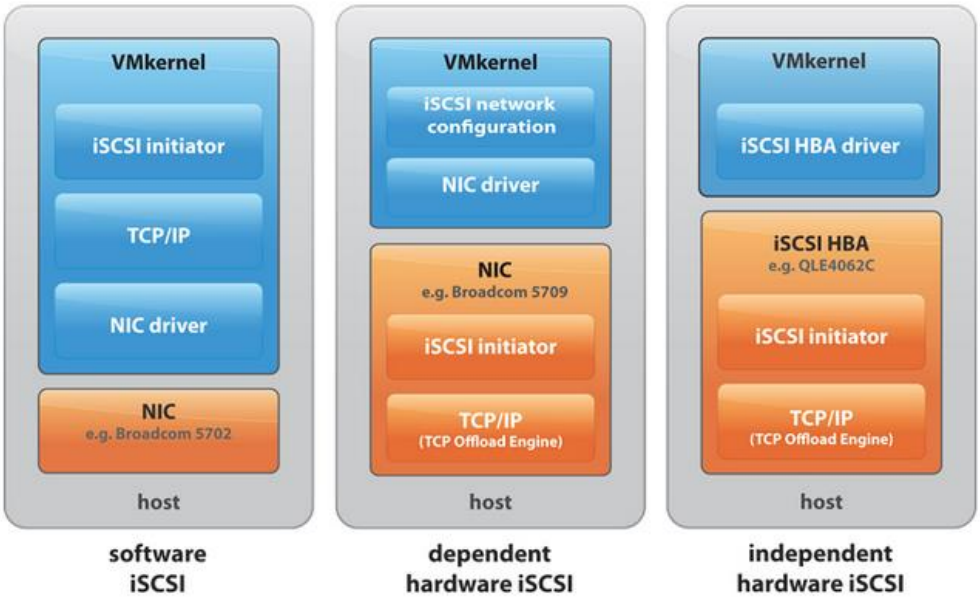


图 2 三种 iSCSI 结构

科普了上面 iSCSI 的知识，接下来对于 iSCSI 环境的搭建会涉及到 vswitch,uplink,portgroup,vmk,vmknic，vmhba.快速了解这个过程的方法是先通过图形化界面配置 iscsi，了解了大致流程后，再结合下面的命令行熟悉下每个命令及在图形化界面的变化。

程序中也有相应的代码，运行 case 的时候会自动搭建 iSCSI 环境。

下面主要是 Hardware iSCSI set up & remove:

- **Set up**

1. Check the mapping between vmhba and vmnic

esxcli iscsi physicalnetworkportal list

显示 vmhba 与 Vmnic 之间的映射关系（手动创建 iscsi 环境前，需要先通过此命令查看映射关系，否则无法正确配置的）

```
[root@sin2-pekhwe-iscsi-006:~] esxcli iscsi physicalnetworkportal list
Adapter  Vmnic   MAC Address        MAC Address Valid  Current Speed  Max Speed  Max Frame Size
-------  ------  -----------------  -----------------  -------------  ---------  --------------
vmhba38  vmnic6  00:10:18:e9:98:20               true           1000       1000            1500
vmhba39  vmnic7  00:10:18:e9:98:22               true           1000       1000            1500
vmhba40  vmnic4  00:10:18:f0:05:90               true              0          0            1500
vmhba41  vmnic5  00:10:18:f0:05:92               true              0          0            1500
vmhba45  vmnic0  d4:ae:52:ea:09:18               true           1000       1000            1500
vmhba45  vmnic1  d4:ae:52:ea:09:1a               true           1000       1000            1500
vmhba45  vmnic2  d4:ae:52:ea:09:1c               true           1000       1000            1500
vmhba45  vmnic3  d4:ae:52:ea:09:1e               true           1000       1000            1500
vmhba45  vmnic6  00:10:18:e9:98:20               true           1000       1000            1500
vmhba45  vmnic7  00:10:18:e9:98:22               true           1000       1000            1500
vmhba45  vmnic4  00:10:18:f0:05:90               true              0          0            1500
vmhba45  vmnic5  00:10:18:f0:05:92               true              0          0            1500
vmhba34  vmnic0  d4:ae:52:ea:09:18               true           1000       1000            1500
vmhba35  vmnic1  d4:ae:52:ea:09:1a               true           1000       1000            1500
vmhba36  vmnic2  d4:ae:52:ea:09:1c               true           1000       1000            1500
vmhba37  vmnic3  d4:ae:52:ea:09:1e               true           1000       1000            1500
```

图 3 vmhba 与 Vmnic 之间的映射关系

2. Create vSwitch

esxcli network vswitch standard add -v iSCSI

3. Bind uplink

esxcli network vswitch standard uplink add -u vmnic2 -v iSCSI

4. Create port group on vSwitch

esxcli network vswitch standard portgroup add -p iscsi_vmk -v iSCSI

5. Create vmknic on port group

esxcli network ip interface add -i vmk123 -p iscsi_vmk

6. Configure IP on vmknic

esxcli network ip interface ipv4 set -i vmk123 -t dhcp

7. Bind your vmknic with vmhba

esxcli iscsi networkportal add -A vmhba34 -n vmk123

8. Start to discovery your staroge array(10.111.13.104:3260)

esxcli iscsi adapter discovery sendtarget add -A vmhba34 -a 10.111.13.104:3260

9. Rescan your vmhba

esxcfg-rescan vmhba34

之后用 esxcfg-scsidevs –A ,查看刚刚的 vmhba 卡是否发现 storage Array,如果一切配置妥当，但是还是没有反应的话，可以在 iSCSI 上添加 session

esxcli iscsi session add -A vmhba34.

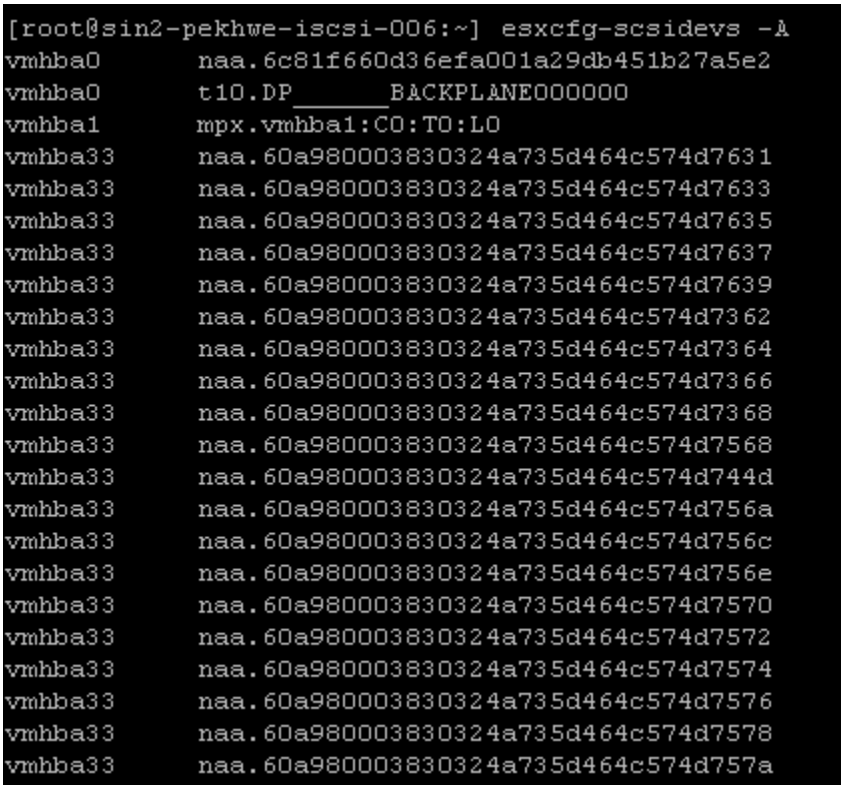


图4 最终发现的luns

列出的这些Lun是该vmhba下的所有的Lun（即创建了datastore的Lun和空白的Lun）另外在运行VMFS的时候，需要选择一个空白的Lun，这样不会把之前的datasore清除掉，给其他测试带来麻烦(因为Create VMFS本身是一个创建VMFS和新的datastore的过程，它会清除之前的设置的)

- Remove iSCSI configuration

关于ISCSI的删除：

只有在vmhba 与vmknic解绑的情况下才可以进行其他的remove操作，有时候需要强制删除才可以解绑iSCSI port(因为有session在所以无法解绑)

具体删除步骤如下：

1.esxcli iscsi adapter discovery sendtarget remove-A vmhba34 -a 10.111.13.104:3260

2.esxcfg-rescan vmhba34

3.esxcli iscsi networkportal remove -A vmhba34 -n vmk123 –f true(其中 f 代表强制删除，force)

4.esxcli network ip interface remove-i vmk123 (or esxcli network ip interface remove -p iscsi_vmk)

5.esxcli network vswitch standard portgroup remove -p iscsi_vmk -v iSCSI

6. esxcli network vswitch standard uplink remove -u vmnic2 -v iSCSI

7.esxcli network vswitch standard remove -v iSCSI

Tips:

1.How to check vmknic ipaddress?

esxcfg-vmknic -l

2.How to check the vSwitch?

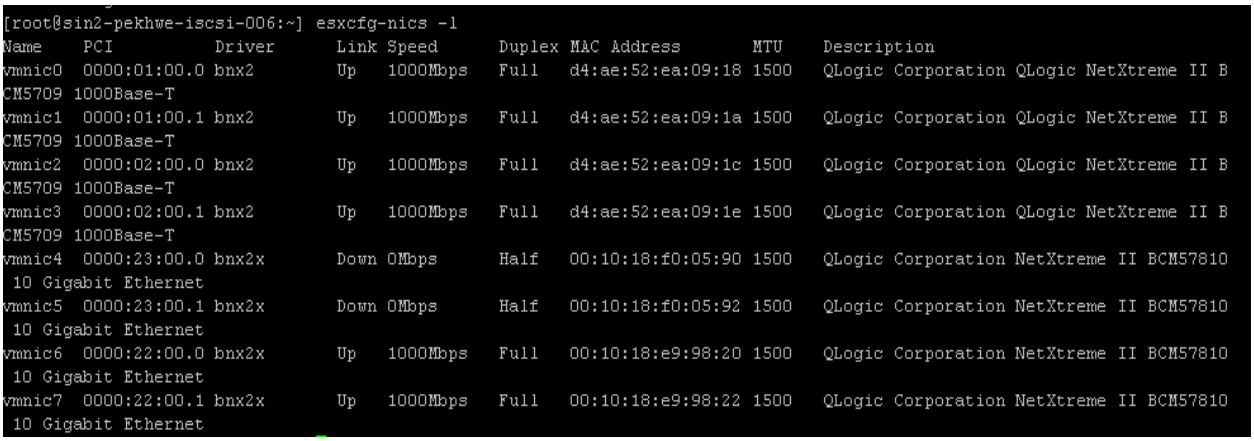
esxcfg-vswitch -l

3.How to check the vmhba?

esxcfg-scsidevs -a

4.How to check the vmnic?

esxcfg-nics –l



```
[root@sin2-pekhwe-iscsi-006:~] esxcfg-nics -l
Name    PCI          Driver      Link Speed      Duplex MAC Address       MTU    Description
vmnic0  0000:01:00.0 bnx2        Up   1000Mbps   Full   d4:ae:52:ea:09:18 1500   QLogic Corporation QLogic NetXtreme II B
CM5709 1000Base-T
vmnic1  0000:01:00.1 bnx2        Up   1000Mbps   Full   d4:ae:52:ea:09:1a 1500   QLogic Corporation QLogic NetXtreme II B
CM5709 1000Base-T
vmnic2  0000:02:00.0 bnx2        Up   1000Mbps   Full   d4:ae:52:ea:09:1c 1500   QLogic Corporation QLogic NetXtreme II B
CM5709 1000Base-T
vmnic3  0000:02:00.1 bnx2        Up   1000Mbps   Full   d4:ae:52:ea:09:1e 1500   QLogic Corporation QLogic NetXtreme II B
CM5709 1000Base-T
vmnic4  0000:23:00.0 bnx2x       Down 0Mbps      Half   00:10:18:f0:05:90 1500   QLogic Corporation NetXtreme II BCM57810
 10 Gigabit Ethernet
vmnic5  0000:23:00.1 bnx2x       Down 0Mbps      Half   00:10:18:f0:05:92 1500   QLogic Corporation NetXtreme II BCM57810
 10 Gigabit Ethernet
vmnic6  0000:22:00.0 bnx2x       Up   1000Mbps   Full   00:10:18:e9:98:20 1500   QLogic Corporation NetXtreme II BCM57810
 10 Gigabit Ethernet
vmnic7  0000:22:00.1 bnx2x       Up   1000Mbps   Full   00:10:18:e9:98:22 1500   QLogic Corporation NetXtreme II BCM57810
 10 Gigabit Ethernet
```

图 5 vmnic 信息（各个 nic 的网速可知）

5.How to check the devices on specified vmhba?

esxcfg-scsidevs -A/-m/-c

- esxcfg-scsidevs –A

  是已经匹配到的所有的 lun,不论使用与否。

- esxcfg-scsidevs –m

  显示的是使用中的 lun

(经过上述 2 条命令比对，发现空白的 lun)

- esxcfg-scsidevs –a

  所有 vmhba 的状态，不论 online 或者是 unbound

```
[root@sin2-pekhwe-iscsi-006:~] esxcfg-scsidevs -a
vmhba38 bnx2i             unbound   iscsi.vmhba38                          QLogic NetXtreme II iSCSI Adapter
vmhba39 bnx2i             unbound   iscsi.vmhba39                          QLogic NetXtreme II iSCSI Adapter
vmhba0  megaraid_sas      link-n/a  unknown.vmhba0                         (0000:05:00.0) LSI / Symbios Logi
vmhba1  ahci              link-n/a  sata.vmhba1                            (0000:00:11.0) ATI Technologies I
vmhba40 ahci              link-n/a  sata.vmhba40                           (0000:00:11.0) ATI Technologies I
vmhba41 ahci              link-n/a  sata.vmhba41                           (0000:00:11.0) ATI Technologies I
vmhba42 ahci              link-n/a  sata.vmhba42                           (0000:00:11.0) ATI Technologies I
vmhba43 iscsi_vmk         online    iscsi.vmhba43                          iSCSI Software Adapter
vmhba32 bnx2i             unbound   iscsi.vmhba32                          QLogic NetXtreme II iSCSI Adapter
vmhba33 bnx2i             online    iqn.2014-09.com.vmware:iscsi-006:vmhba35QLogic NetXtreme II iSCSI Adapter
vmhba34 bnx2i             online    iqn.1998-01.com.vmware:sin2-cpdnet-dhcp-10-111-13-207:1825694306:36QLogic
vmhba35 bnx2i             unbound   iscsi.vmhba35                          QLogic NetXtreme II iSCSI Adapter
vmhba36 bnx2i             unbound   iscsi.vmhba36                          QLogic NetXtreme II iSCSI Adapter
vmhba37 bnx2i             unbound   iscsi.vmhba37                          QLogic NetXtreme II iSCSI Adapter
```

图 5 Vmhba online or unbound(有时候还会 offline)

6.esxcli iscsi networkportal list

查看 vmhba 的详细情况，包括 ip 地址等。（有时候 case failed 是因为虚拟机找不到 ip 了）

另外可以通过 ls /vmfs/devices/disks/ 查看所有的 disk(naa 与 vml)

之所以有下面的附录，是因为初来乍到遇到了好多新名词，为了理解无误，还是找官方确认了下，帮助小伙伴们理解。

Appendix:

1.VMFS Volume Manager Terminology

This appendix defines key terms used to describe components of a logical volume.

Datastore A formatted file system that is layered on top of either a VMFS volume with block-based storage (iSCSI and FC) or a mount point for NFS storage. It is a shared storage resource in which VMhomes, virtual disks, and VM objects are stored.

LUN A single block storage allocation presented to a server. This logical unit number is the unique identification a host has assigned to a given block device resource (disk) it finds when it scans the storage array network. The term disk is often used interchangeably with LUN. From the perspective of an ESX host, a LUN is a single unique raw storage block device or disk.

ESX LUN number The host-based number assigned to the storage unit as seen be the ESX host addressing the storage. In most cases this can be set by the array and should be consistent across a set of ESX hosts that share a common pool of storage resources.

Disk serial number A unique number assigned to a LUN by the storage array for identification by hosts. This number is also called a SCSI disk ID, SCSI ID, LUN ID, or LUN UUID.

NAA disk ID National Address Authority disk identification that is presented by most storage arrays.

VMFS volume extent VMFS volumes can be made up of one or more disks (LUNs). Each disk (LUN) is called a VMFS extent. In many cases there is a 1:1 ratio of LUNs to VMFS volumes. However, some VMFS volumes have many extents, and in some rare cases several VMFS volumes might exist on a single LUN.

VMFS volume manager The process that scans, sorts, and manages status, membership, and protection of underlying components that are presented to an upper layer as a singe resource.

VMFS volume signature A unique identification assigned to a given storage resource and used by VMFS to identify the disk as unique. It is a combination of the LUN number and the disk serial number.

VMFS volume A collection of storage resources managed as a single shared resource. In most cases the VMFS volume contains a single LUN. In those cases the datastore and the VMFS volume are identical. However, in some cases the VMFS volume might span two or more LUNs and be composed of multiple extents.

2. 猜猜我是谁

vm vmx vmdk vmxf virtual disk 一对对孪生兄弟。

vmx 虚拟机配置文件

.vmdk VM 的虚拟磁盘特性

vmxf 其他虚拟机配置文件

*-flat.vmdk：虚拟机的原始磁盘文件，包含整个虚拟机镜像。
*.nvram：虚拟机的 BIOS 信息。
*.vmdk：虚拟硬盘的描述文件，与-flat.vmdk 文件相关联。主要用于指定虚拟硬盘的接口类型，扇区，簇等信息，也用于定义是否启用“thin Provision”

*.vmx：虚拟机的配置文件，包含虚拟机所有虚拟硬件的信息和设置。如电源，网络，内存等。

*.vmxf：配置虚拟机聚合时，扩展的虚拟机配置文件。

详见 http://jackiechen.blog.51cto.com/196075/210492/ ESX 虚拟机文件列表详解

https://wiki.eng.vmware.com/HWE/Drivers/TestBed/Singapore 新加坡的机器的信息