

# Proxmox Virtual Environment (PVE)

功能強大、可取代 VMware  
ESXi 之緣起第一章

台大計中網路組  
游子興

# 先說一下 VMware 的壞話

---

# VMware 鐵粉之使用經歷

- \* V5.5 ~ v8.0
  - \* 2013~ 2024 超過十年使用經驗
- \* V6.7 區網雲端節能計畫
  - \* 13台 Hosts with vCenter(5台 v-SAN)
  - \* 最高 150+VMs，共 19個使用單位
  - \* 最長連續工作 36Hr 進行災難復原...
- \* V7.x V8.x 區網雲端大數據 + 區網網頁 + 個人電腦
  - \* 16台+ Host without vCenter (Free License)

# VMware

## 嚴格的硬體支援清單且淘汰快速

### \* RAID

- \* 僅支援硬體RAID 卡 (不支援主機板/軟體式)
- \* 硬體RAID卡 Server 最低價 Dell R750 \$157,614 (政府採購網)

### \* 網卡

- \* Realtek 網卡從未被正式支援
- \* V6.x
  - \* 100Mbps皆不支援、僅支援 1G 以上網卡
- \* 僅支援至 6.x
  - \* Broadcom BCM5751、BCM5721
  - \* Intel E1G42ET Dual Port (82576)
- \* 僅支援至 7.x
  - \* Intel Ethernet Server Adapter X520-DA2

### \* CPU

- \* IBM Server x3650 無法升級 v6.7
- \* ASUS BP1AF無法升級 v7
- \* Dell Precision Tower 5810、OptiPlex 990 無法升級 v8

# VMware ESXi

## 已有半年未推出更新 Patch

\* <https://esxi-patches.v-front.de/>

### ESXi Version 8.0 History.

Jump to release date: [2023-09-21 \(Update 2\)](#) | [2023-07-27 \(Update 1c\)](#) | [2023-06-01 \(Update 1a\)](#) | [2023-04-18 \(Update 1\)](#) | [2023-03-30 \(ESXi 8.0c\)](#) | [2023-02-14 \(ESXi 8.0b\)](#) | [2022-12-08 \(ESXi 8.0a\)](#) | [2022-10-11 \(GA\)](#) |

### ESXi Version 7.0 History.

Jump to release date: [2023-09-28 \(Update 3o\)](#) | [2023-07-06 \(Update 3n\)](#) | [2023-05-03 \(Update 3m\)](#) | [2023-03-30 \(Update 3l\)](#) | [2023-02-21 \(Update 3k\)](#) | [2023-01-31 \(Update 3j\)](#) | [2022-12-08 \(Update 3i\)](#) | [2022-09-01 \(Update 3g\)](#) | [2022-07-12 \(Update 3f\)](#) | [2022-06-14 \(Update](#)

# VMware 之近況

- \* 2022 博通Broadcom宣佈以 610億美元(溢價50%) 併購 VMware (Broadcom 市值 2200億)
  - \* 博通已分別收購軟體公司 CA Technologies，賽門鐵克企業端安全業務
- \* 2023/11 收購完成後一周裁減VMware人力，據稱超過2,000人。未經證實的消息指出博通以不續工作簽證為手段，迫使外籍工程師主動離職。
  - \* [https://www.reddit.com/r/vmware/comments/1b1sah4/broadcom\\_silent\\_layoffs\\_email\\_to\\_vmware\\_about\\_40/?rdt=61195](https://www.reddit.com/r/vmware/comments/1b1sah4/broadcom_silent_layoffs_email_to_vmware_about_40/?rdt=61195)
- \* 2023/12 取消現有合作夥伴關係，只接受下單量大的經銷商作為新的合作夥伴
  - \* <https://www.ithome.com.tw/news/160549>
  - \* 峰儀業務 Cors 最近提到 VMware 已經不提供教育版 License
- \* 2024/01 終結56項VMware產品，包括 vSphere Hypervisor，以簡化產品、減少研發成本
  - \* <https://www.thestack.technology/vmware-is-killing-off-56-products-including-vsphere-hypervisor-and-nsx/>
- \* 2024/01 VMware 宣布不再單獨銷售永久授權，未來改採用訂閱制
  - \* <https://blogs.vmware.com/cloud-foundation/2024/01/22/vmware-end-of-availability-of-perpetual-licensing-and-saas-services/>
- \* 2024/02 End Of General Availability of the Free vSphere Hypervisor (ESXi 7.x and 8.x)
  - \* [https://kb.vmware.com/s/article/2107518?lang=en\\_US](https://kb.vmware.com/s/article/2107518?lang=en_US)
- \* 2024/03 以38億美元將VMware 遠端存取技術用戶端運算部門 (End-User Computing, EUC) 出售給私募基金 KKR，包含企業平臺Workspace ONE整合終端管理 (Unified Endpoint Management, UEM)、旗艦桌機及應用程式虛擬化平臺 Horizon
  - \* <https://www.ithome.com.tw/news/161589>

# 功能強大、友善硬體支援 Proxmox Virtual Environment (PVE)

---

# Proxmox Virtual Environment (PVE)

- \* Linux Debian + QEMU/KVM + LXC(Linux Container)
  - \* QEMU(Quick Emulator)
  - \* KVM(Kernel-based Virtual Machine)
- \* 硬體支援
  - \* 相容於 Linux Debian Kernel
    - \* 支援 Realtek 網卡
    - \* 支援 100Mbps 網卡
    - \* 支援舊型 PCI 介面網卡 (主流: PCI-E/PCI Express)
  - \* PCI Passthrough 限制少
    - \* 支援 USB Mouse/Keyboard
    - \* 支援 USB Audio/Video (USB Camera)



# Proxmox VE 特色

- \* Cluster 中控台
  - \* 不需安裝額外軟體 (vCenter Appliance)
    - \* 節省後續維運升級之困擾
  - \* 每台 Node 皆可當成中控台
    - \* 避免 vCenter 當機或所在 Host 當機
- \* Migrate: (Like VMware vMotion)
  - \* Node 主機搬移
  - \* Storage 搬移
- \* HA 高可用性
- \* 內建 Email Notification
  - \* 不需設定 SMTP Server，內建寄信機制
  - \* 待確認
    - \* Received: from davisyoupc.cc.ntu.edu.tw (HELO pve32.cc.ntu.edu.tw) ([140.112.3.82])
    - \* by mailgw2.ntu.edu.tw with ESMTP; 01 Apr 2024 15:32:22 +0800
    - \* Received: by pve32.cc.ntu.edu.tw (Postfix, from userid 0)

# Proxmox VE 特色

- \* VM Clone
  - \* Full Clone
- \* VM Template
  - \* Link Clone
    - \* 以共用範本的虛擬磁碟為基礎產生新的 VM，新 VM 僅有差異存在新的虛擬磁碟，可節省容量，而且複製速度極快。
    - \* 應用: VMware Horizon/VDI
- \* Proxmox Backup Server
  - \* 增量備份、資料壓縮、重複資料刪除
  - \* 備份 VM Disk 內容檢視，不需還原即可取出檔案

# Proxmox VE Storage

---

# Support Storage Type

- \* Proxmox VE

- \* [https://pve.proxmox.com/pve-docs/pve-admin-guide.html#chapter\\_storage](https://pve.proxmox.com/pve-docs/pve-admin-guide.html#chapter_storage)

- \* All storage types which have the “Snapshots” feature also support “Thin Provisioning”.

- \* Support Block or File(.raw .qcow2 .vmdk) format

- \* VMware ESXi

- \* VMFS5、VMFS6

- \* iSCSI、NFS

- \* Support File Format(.vmdk) Only

**Table 2. Available storage types**

Description	Plugin type	Level	Shared	Snapshots	Stable
ZFS (local)	<code>zfspool</code>	both <sup>1</sup>	no	yes	yes
Directory	<code>dir</code>	file	no	no <sup>2</sup>	yes
BTRFS	<code>btrfs</code>	file	no	yes	technology preview
NFS	<code>nfs</code>	file	yes	no <sup>2</sup>	yes
CIFS	<code>cifs</code>	file	yes	no <sup>2</sup>	yes
Proxmox Backup	<code>pbs</code>	both	yes	n/a	yes
GlusterFS	<code>glusterfs</code>	file	yes	no <sup>2</sup>	yes
CephFS	<code>cephfs</code>	file	yes	yes	yes
LVM	<code>lvm</code>	block	no <sup>3</sup>	no	yes
LVM-thin	<code>lvmthin</code>	block	no	yes	yes
iSCSI/kernel	<code>iscsi</code>	block	yes	no	yes
iSCSI/libiscsi	<code>iscsidirect</code>	block	yes	no	yes
Ceph/RBD	<code>rbd</code>	block	yes	yes	yes
ZFS over iSCSI	<code>zfs</code>	block	yes	yes	yes

# 參考資料

- \* <https://ithelp.ithome.com.tw/users/20118848/ironman/4678>



**Jason Cheng (節省哥)** (jasoncheng7115)  
IT邦研究生 3 級 · 點數 2265  
節省工具箱

0 72 9

# ZFS

## (Zettabyte File System)

### 簡介

---

# ZFS

## (Zettabyte File System)

- \* ZFS was initially launched with Sun Microsystems(2010 被 Oracle 併購), and later it was a part of Linux distributions and macOS.
- \* ZFS is not a filesystem but a volume manager.
  - \* 推翻磁碟區的概念，無分割區管理
- \* ZFS does not require defragmentation and there isn't a tool to do it.
- \* jail -- The system administration utility which allows launching of processes within a jail environment.



# Support OS & 廠商

- \* Support OS

- \* Oracle Solaris、Linux (Ubuntu 16.04 LTS+)、FreeBSD、macOS
- \* <https://en.wikipedia.org/wiki/ZFS>

- \* NAS Storage 廠商

- \* QNAP

- \* QuTS hero 系列
- \* 為什麼 QNAP 企業級 NAS 選用 ZFS 檔案系統
- \* <https://www.qnap.com/solution/qnap-zfs/zh-tw/>

- \* Infortrend

- \* EonNAS Pro 200、3000 Series

- \* TrueNAS (FreeNAS)

- \* <https://www.truenas.com/>

- \* Unraid

- \* <https://unraid.net/zfs-pools-rc3>

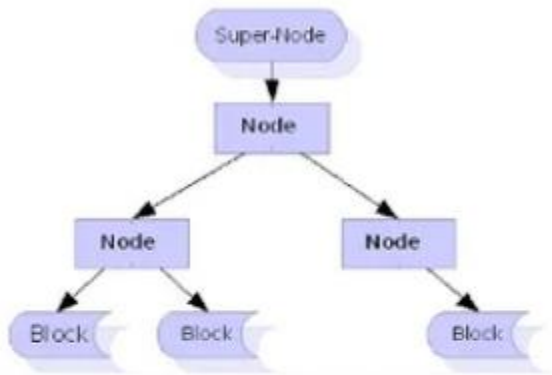
# ZFS Features

- \* 128 Bit 定址: 高達 16EB 的容量
- \* 支援軟體式 RAID 組態
  - \* RAID 0 (Striped)、RAID 1 (Mirror)
  - \* RAID 5 (RAID Z1)、RAID 6 (RAID Z2)
  - \* RAID 10、RAID 50、RAID 60
  - \* Triple Mirror: 3 個資料鏡射保護
  - \* Triple Parity (RAID Z3): 即可提供 3 個冗余容錯保護
  - \* ※不建議使用硬體 RAID Card
- \* WORM (Write Once, Read Many)
  - \* 寫入後不能作任何修改，只能夠讀取
  - \* 避免資料被變更，確保存檔的完整性。可讓機敏行業的企業能符合相關法規。
- \* COW(Copy On Write) 寫入時複製
  - \* 對資料進行修改時，不會將原有資料覆蓋，而是寫入新的空間，完成後更新 Metadata，將指標指向新的區塊，最後才把舊區塊刪掉

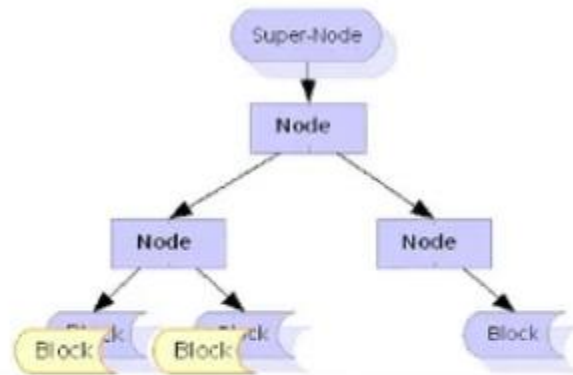
# COW (Copy On Write)

寫入時複製

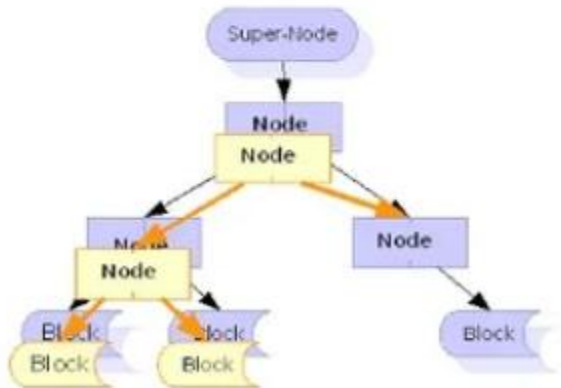
1. Initial State



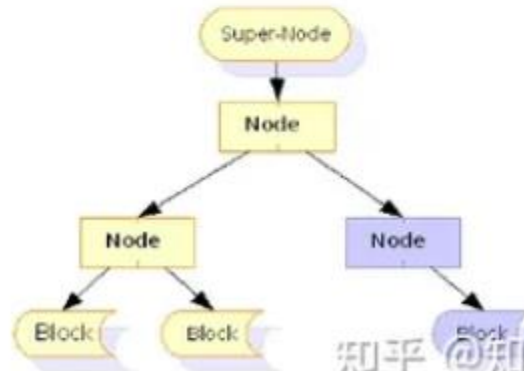
2. Copying of Nodes



3. Creating new Nodes



4. Write new Super Node



# ZFS Features

## \* Snapshot and Clone

- \* 因為 COW 之特性，建立快照時幾乎不佔額外空間且非常快速，因此能更密集地建立快照，且支援更多的快照數量

# Snapshot

## ZFS vs. VMware

- \* 在 QNAP QuTS hero 及 QES 作業系統

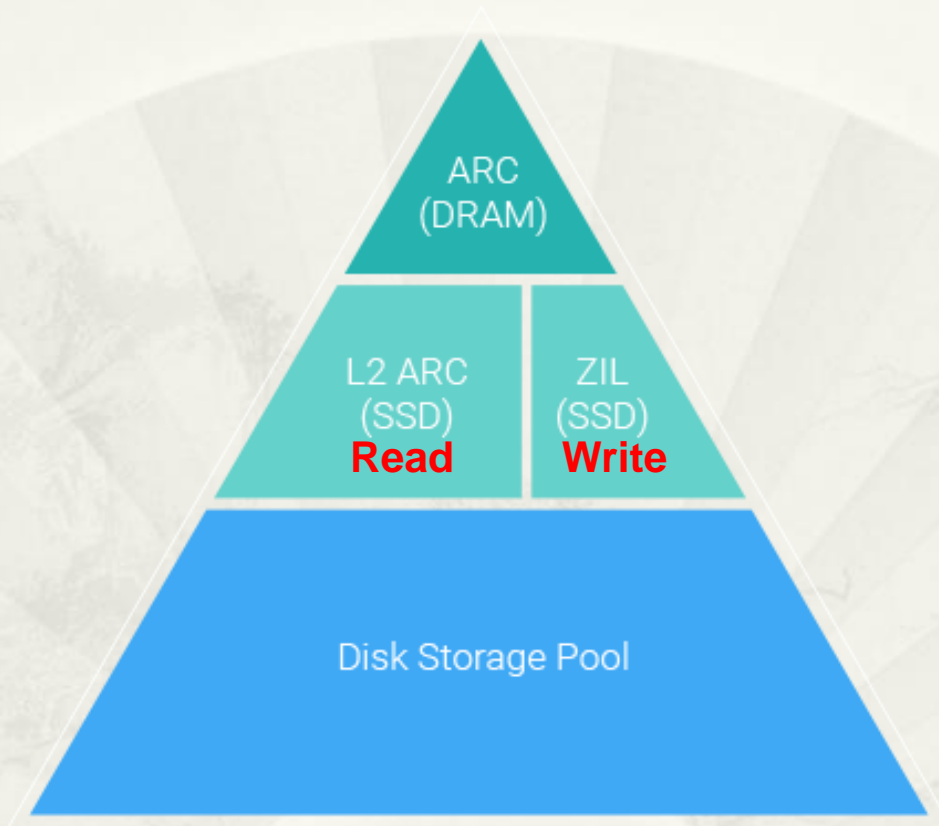
高達65536張的快照(Snapshot)  
每小時拍 1 張快照，1 天 24 張，可維持長達 7 年的多版本保護

- \* <https://www.qnap.com/solution/qnap-zfs/zh-tw/>
- \* Best practices for using VMware snapshots in the vSphere environment (1025279)
  - \* <https://kb.vmware.com/s/article/1025279>

vmware<sup>®</sup> CUSTOMER CONNECT  
by Broadcom

- Maximum of 32 snapshots are **supported** in a chain. However, for a better performance use only 2 to 3 snapshots.
- Do not use a single snapshot for more than 72 hours.

# Multi Layer Cache



<https://hk.xfastest.com/64819/xf-news-qnap-new-qts-hero-qts-hero-introduction/2/>

# Multi Layer Cache 硬碟效能實測

\* WD 8TB @PVE (ZFS)

	Read (MB/s)	Write (MB/s)
SEQ1M Q8T1	3467.12	310.69
SEQ1M Q1T1	1335.94	552.06
RND4K Q32T1	214.10	120.99
RND4K Q1T1	60.09	56.34

\* Seagate IronWolf Pro 20TB



Task name	Read mB/s	Write mB/s
Sequential Max (128K) Q32T1	281.058	285.001
Random Max (128K) Q1T1	31.053	93.795
Random 4K Q1T1	1.131	3.363
Random 4K Q32T1	3.591	3.458

# ZFS Features

- \* Online Data Compression
  - \* 支援壓縮速度最快的 LZ4 演算法，不會對系統造成太大負擔
- \* Online Data Deduplication
  - \* Block Level 運作的功能
  - \* 寫入數據前會先計算數據特徵指紋，若相同的資料已經存在，只會更新 Metadata，而不會再次寫入相同的資料。
  - \* 應用情境
    - \* VM 環境: 很多 VM 作業系統皆相同
    - \* Mail Server Storage : Email to 很多人、附件檔案很大
- \* 上述兩者皆支援 Enable/Disable at any time
  - \* A pool can contain any mix of compression and non-compression data, coexisting.
  - \* A pool can contain any mix of deduplicated and non-deduplicated data, coexisting.



# Deduplication 應用案例

- \* Email 1MB to 30,000 在校師生
- \* Storage Size: 1MB \* 30,000 = 30 GB

敬賀端午佳節安康如意 ; Dragon Boat Festival Greetings from the President



陳文章

收件者 校內訊息

# 端午節

祝福您和家人安康如意

Wish you and your family a  
nice Dragon Boat Festival!

Dr. Wen-Chang Chen  
President, NTU

陳文章 敬賀

國立臺灣大學 National Taiwan University



# Demo1

# TrueNAS

---

# Create 3 Different Datasets

Details for  
SMB

Add Zvol

**Dataset Details**

Type: FILESYSTEM  
Sync: STANDARD  
Compression Level: OFF  
Enable Atime: OFF  
ZFS Deduplication: OFF  
Case Sensitivity: OFF  
Path: NAS/SMB

Details for  
SMB\_LZ4

Add Zvol

**Dataset Details**

Type: FILESYSTEM  
Sync: STANDARD  
Compression Level: Inherit (LZ4)  
Enable Atime: OFF  
ZFS Deduplication: OFF  
Case Sensitivity: OFF  
Path: NAS/SMB\_LZ4


Details for  
SMB\_Dedup





Add Zvol

**Dataset Details**

Type: FILESYSTEM  
Sync: STANDARD  
Compression Level: OFF  
Enable Atime: OFF  
ZFS Deduplication: ON  
Case Sensitivity: OFF  
Path: NAS/SMB\_Dedup

# Test 1. Compression

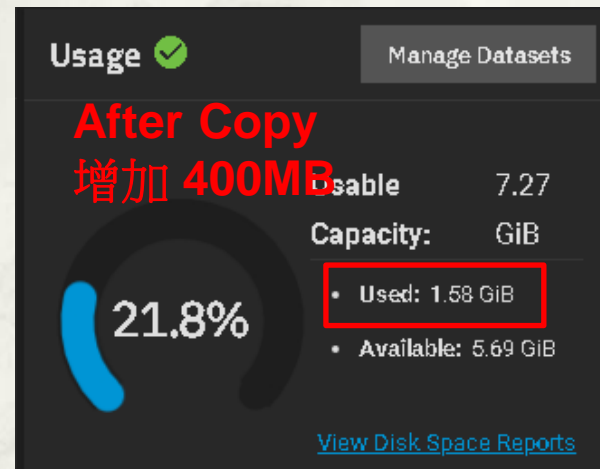
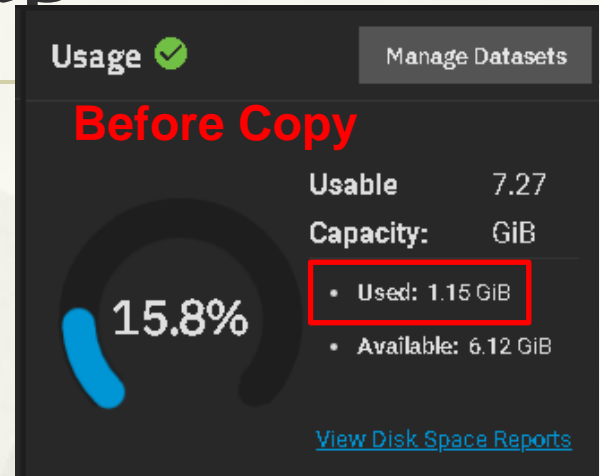
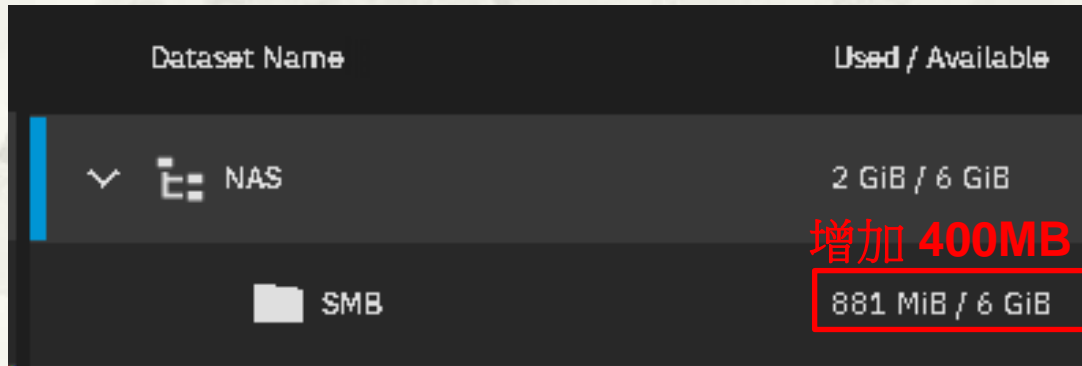
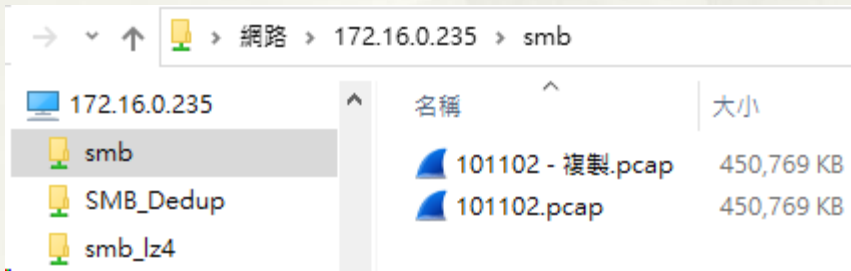
名稱	大小
 101102.pcap	450,769 KB

Dataset Name	Used / Available
▼  NAS	1 GiB / 6 GiB
 SMB	441 MiB / 6 GiB
 SMB_Dedup	441 MiB / 6 GiB
 SMB_LZ4	280 MiB / 6 GiB

LZ4 壓縮結果

# Test 2. Dedup

\* Copy to SMB again



```
root@truenas[~]# zpool list
```

NAME	SIZE	ALLOC	FREE	CKPOINT	EXPANDED	FRAG	CAP	DEDUP	HEALTH	ALTROOT
NAS	7.50G	1.15G	6.35G	-	-	0%	15%	1.00x	ONLINE	/mnt

**Before Copy**

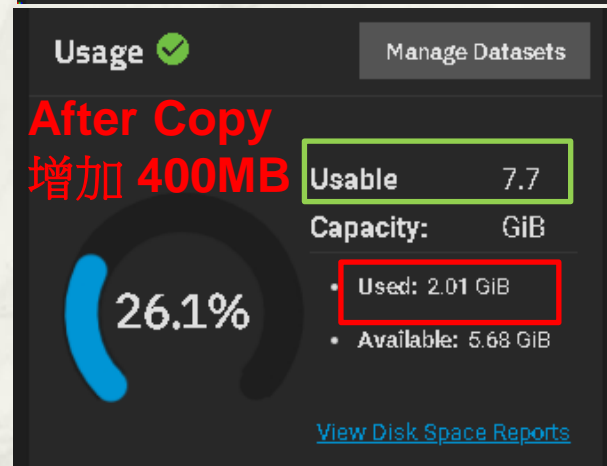
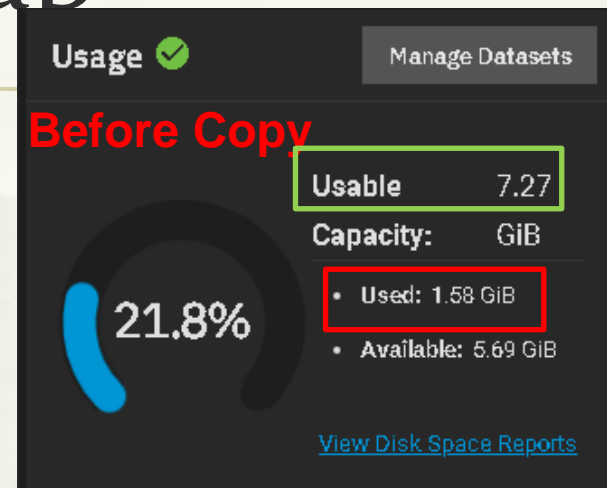
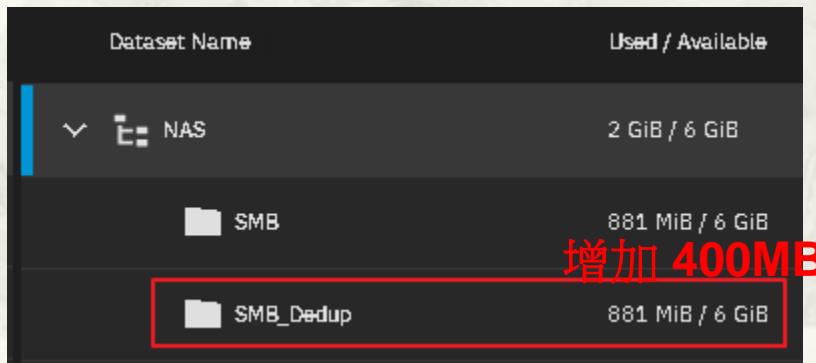
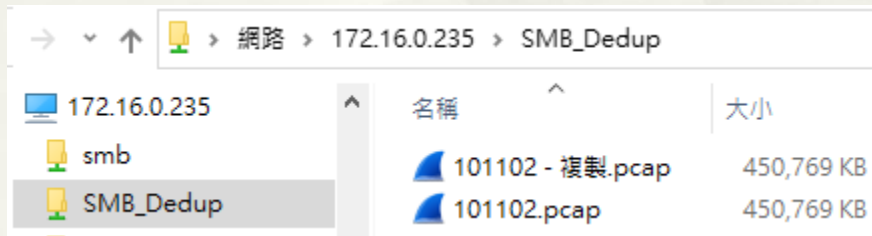
```
root@truenas[~]# zpool list
```

NAME	SIZE	ALLOC	FREE	CKPOINT	EXPANDED	FRAG	CAP	DEDUP	HEALTH	ALTROOT
NAS	7.50G	1.58G	5.92G	-	-	0%	21%	1.00x	ONLINE	/mnt

**After Copy**  
增加 400MB

# Test 2. Dedup

\* Copy to SMB\_Dedup again



```
root@truenas[~]# zpool list
```

NAME	SIZE	ALLOC	FREE	CKPOINT	EXPANDSZ	FRAG	CAP	DEDUP	HEALTH	ALTROOT
NAS	7.50G	1.58G	5.92G			0%	21%	1.00x	ONLINE	/mnt

Before Copy

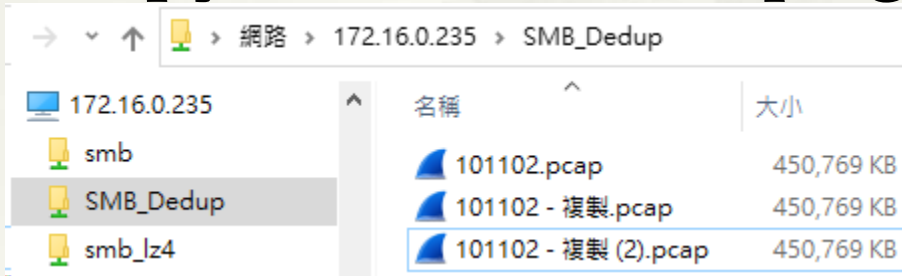
```
root@truenas[~]# zpool list
```

NAME	SIZE	ALLOC	FREE	CKPOINT	EXPANDSZ	FRAG	CAP	DEDUP	HEALTH	ALTROOT
NAS	7.50G	1.58G	5.92G			0%	21%	2.00x	ONLINE	/mnt

After Copy

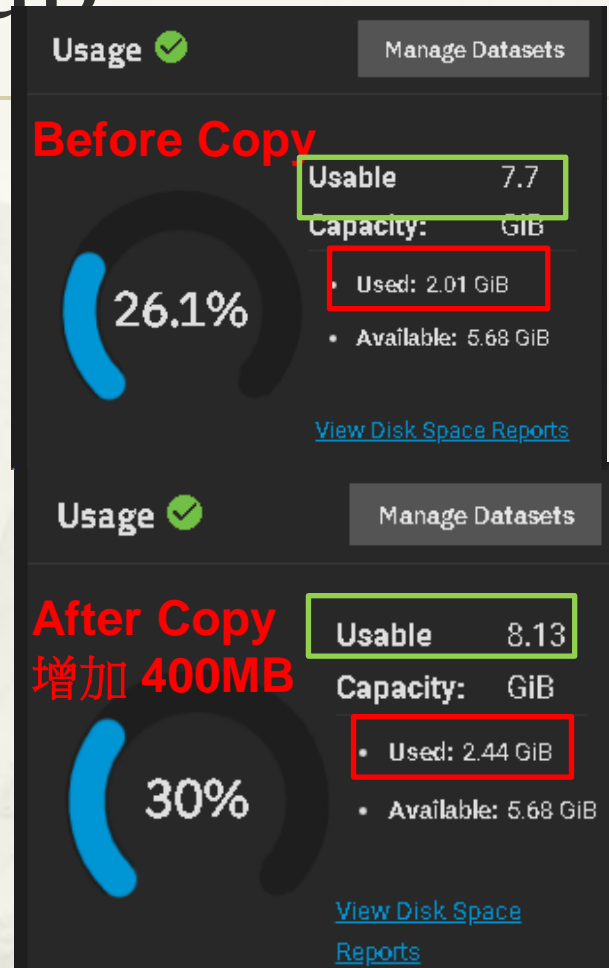
# Test 2. Dedup

\* Copy to SMB\_Dedup again



Dataset Name	Used / Available
NAS	2 GiB / 8 GiB
SMB	881 MiB / 6 GiB
SMB_Dedup	1 GiB / 6 GiB

增加 400MB



```
root@truenas[~]# zpool list
```


NAME	SIZE	ALLOC	FREE	CKPOINT	EXPANDSZ	FRAG	CAP	DEDUP	HEALTH	ALTROOT
NAS	7.50G	1.58G	5.92G	-	-	0%	21%	2.00x	ONLINE	/mnt

Before Copy

```
root@truenas[~]# zpool list
```

NAME	SIZE	ALLOC	FREE	CKPOINT	EXPANDSZ	FRAG	CAP	DEDUP	HEALTH	ALTROOT
NAS	7.50G	1.58G	5.92G	-	-	0%	21%	3.00x	ONLINE	/mnt

After Copy



簡報完畢  
謝謝