napp-it

napp-in-One

ESXi Virtual Server + virtualized ZFS Storage Server NAS/SAN

User‘s Guide
Setup and First steps

published: 2023-Apr-23 (c) napp-it.org

Content:

1. About Napp-in-One
2. Hardware requirements
   „best to“ with new Intel Optane
3. Setup ESXi (use newest)
4. Setup the napp-it storage VM
5. OS barebone Installation
6. Manuals, Help and infos
7. Remote Management
8. Napp-it Web-UI
9. ZFS Pools
10. ZFS Filesystems
11. SMB Server
12. User and Groups/ Active Directory
13. NFS Server
14. iSCSI/ FC server

15. Data Scrubbing
16. Data Snapshots/ Backup

17. Operational settings
18. Security
19. Tuning

20. ZFS vCluster in a Box

21. Other manuals

1. about Napp-in-One

Napp-in-one is an approach to combine ESXi, the leading VM environment with a minimalistic and resource efficient ZFS storage for general filer use and to store virtual machines on ZFS with snaps and online replication.
As napp-in-one is based on an enterprise class Solarish operating system, it includes the whole set of NAS/SAN features like iSCSI/FC/NFS/S3 or replication webmanaged by napp-it either based on native ZFS (Oracle Solaris) or Open-ZFS in its native environment (free Solaris fork OmniOS or OI) with bootenvironments and the „it just works" experience after setup and updates or recovery recovery.

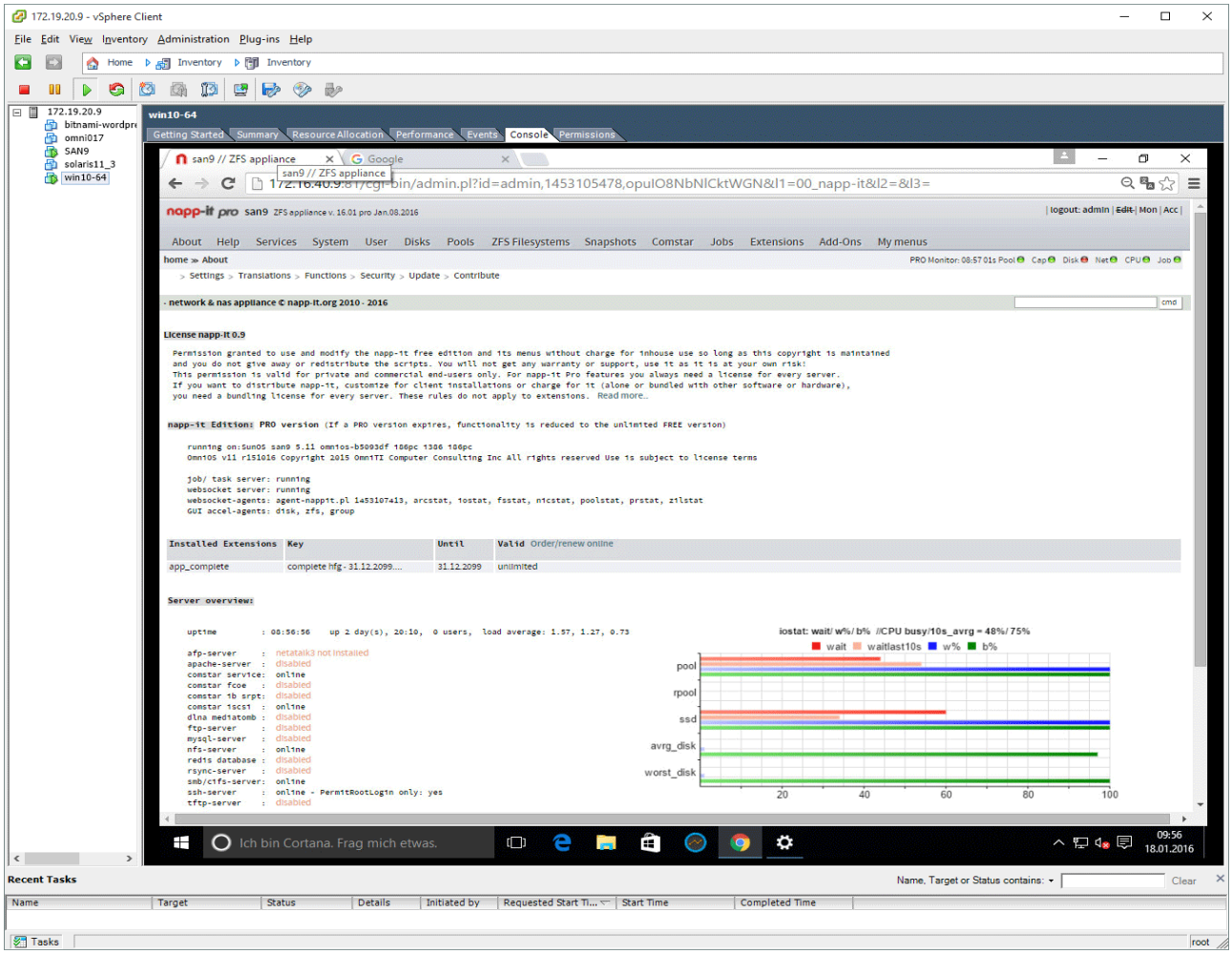
The VM environment

In a server environment you offer services like directory services, databases, webservices and desktop or application virtualisation in a VM environment. VMware, the leading vendor for virtual environments offers ESXi the tiniest type-1 hypervisor as a commercial product or for free with a reduced feature set regarding HA and storage (This is why you want ESXi+ZFS). ESXi comes with the smallest footprint and best support for any guest operating systems from BSD over Linux, OSX, Solaris to Windows. It is often simply the fastest. It is initially installed or recovered after a crash within minutes. Management can be done remotely by its free webconsole.

The Storage environment

The revolutionary ZFS filesystem offers a new level of data security with checksums and CopyOnWrite with snapshots and bitrot protection/repair. Solaris where ZFS comes from, offers a unique integration of ZFS within the operating system combined with its own ZFS embedded kernelbased NFS and SMB services, the Comstar FC/iSCSI stack and Crossbow virtual networking - best integrated by one source/vendor Sun/Oracle.

Napp-in-One integrates this to a ready to use „Just Enough Storage OS" solution with webmanagement. You can download and setup or recover after a crash within minutes as you use it for storage only. All other server services that require special configurations are VMs on safe ZFS storage with backups and versioning.



Napp-in-One Storage VM on ESXi

2. Hardware requirements

While there are many options, (see VMware HCL), the following is suggested:

1. A serverclass system from Dell, HP or Lenovo that is certified for VMware
This starts from the affordable HP or Dell server line up to 19" storage systems or

2. A configurable storage server with a serverline mainboard from Asrock or Supermicro

Best is to check the following:
- Use a server class mainboard with ipmi and a CPU with ECC support
- Use a Xeon or AMD CPU and at least 16 GB ECC RAM

Critical and limiting for ESXi is mostly the nic and disk adapter.
- prefer an LSI HBA without raid functionality (referred as IT mode) ex LSI/ BroadCom 9x00-8i
- prefer Intel Nics, I would prefer 10G-BaseT ones like the Intel X540-T1 (1x 10g) or X540-T2 (2 x 10G)
  or the Intel X520 with an SFP+ interface but you can start with1G.

A good current mainboard, start searching here
- http://www.supermicro.com
- http://www.asrockrack.com/general/products.asp#Server

Criteria:

- You need a bootmedia for ESXi, this can be a M.2 disk or an Sata SSD/disk (boot and local datastore)
- You need a second disk controller for storage in passthrough mode (Sata, SAS or NVMe)

- As you want to use a ZFS storage server, you must connect disks to your storage VM,
best with a dedicated HBA/Sata in pass-through mode as this offers a barebone alike performance.
ZFS storage from the storage VM is available for ESXi via NFS when the storage VM is up and running.

see http://www.virten.net/2015/10/usb-devices-as-vmfs-datastore-in-vsphere-esxi-6-0/
where you can use the new ESXi webclient to reformat the USB stick/disk (prefer SSD via an USB case)
https://labs.vmware.com/flings/esxi-embedded-host-client

Hardware/ mainboard setup
In mainboard bios settings, enable vt-d or iommu (pass-through of real hardware to VMs) and set Sata to AHCI.
The vt-d option is only available with server class mainbaords and some CPUs (prefer a Xeon).

Attention: Only some CPUs are capable of ECC and vt-d, mainly Xeons and some lowcost Intel CPUs. AMD has often ECC support even with desktop boards. Sometimes you need a AMD Pro cpu (mainly with integrated graphics).
In general prefer server class boards and CPUs.

About All in One

When I came up with the All in One idea more than 10 years ago (ESXi server with a full featured virtualized ZFS storage appliance VM in one box), many declared a virtualized storage server as a stupid idea. In the meantime the All in One idea is quite common with a wide range of storage appliances not only with a minimalistic Solaris/OmniOS but also on BSD or Linux after the following post: https://b3n.org/freenas-9-3-on-vmware-esxi-6-0-guide/

2.1 M.2 NVMe or The new Optane is a game-changing technology

In the past a typical AiO setting was a server system with an Sata bootdisk for ESXi and the napp-it storage VM on the local Sata datastore. All other disks including disks for Slog and L2ARC were connected on an SAS HBA pass-through mode. To make your VMs crash resistent you enabled sync-write with an optional and dedicated Slog device.

M.2 and new Intel Optane NVMe starting with the 900P makes a whole different setup possible that is faster, more flexible and cheaper. Unlike traditonal Flash, the Optane is adressed like RAM so there is no need for Trim, Garbage Collection or erase prior write like with current Flash disks. This dramatically reduces latency (from the 20-30us of normal Flash) to 10us while iops go high from 80k to 500k. This allows a sync write performance of a ZFS pool that is near to the performance with sync=disabled.

What are the consequenses for a „best of" AiO setup?

Basically the most important thing is:
Use an M.2 bootdisk (ESXi boot and local datastore for the storage VM)
Optionally use an Intel Optane NVMe (128GB) bootdisk with vdisks for Slog and L2Arc
Use passthrough for ZFS storage (Sata, SAS or NVMe). For VM storage prefer a Flash based pool.

For disk based pools

You want an Optane as Slog and propable L2ARC due the read ahead persistent caching option on L2Arc.
You can add the Optane in passthrough mode or you can use the Optane as ESXi bootdisk and add a small 10-20 GB vdisk as Slog, optionally another vdisk (max 5xRAM) as L2Arc

Disk Pool	8k sync/unsync /s	random sync/unsync/s	seq sync/unsync/s	dd sync /unsync /s
no slog	520K / 1.9M	1.6M / 65.8M	41,8M/ 1024M	283M/ 939M
Optane Slog	1.6M / 1.9M	39.4M/ 68.4M	512M / 1023M	849M/961M
SSD Pool	8k sync/unsync /s	random sync/unsync/s	sequ sync /unsync /s	dd sync/unsync/s
no slog	1.5M/ 1.9M	16M / 50.2M	341M/ 1023M	423M/ 806M
Optane Slog	1.6M / 1.9M	38.2M/50.2M	512M/ 1023M	731M/ 806M
Optane Pool	8k sync/unsync /s	random sync/unsync/s	sequ sync /unsync /s	dd sync/unsync/s
one 900p	1.6M/ 1.9M	32M / 75M	511M/ 1023M	711M/ 1.1G

see some benchmark values with a 20G Slog as vdisk on a Optane 900P

My suggested AiO setup now

- use an bootdisk > 100GB (M.2 NVMe or Optane NVMe) to boot ESXi
- use the free space on the bootdisk as local datastore and place the napp-it storage VM onto
- Use an LSI SAS HBA or Sata or NVMe in pass-through mode for your datapool
- prefer Flash disks with powerloss protection for VM storage

for diskbased pools.
- add a 20 G vdisk on the Optane datastore to the napp-it storage VM and use as Slog for your datapool
- add a vdisk for L2ARC (around 5x and no more than 10 x size of RAM) and enable read ahead

SSDs especially desktop ones suffer on steady write or mixed read/write workloads. For an Optane, steady write or simultanious read/write workload is uncritical.

read
http://napp-it.org/doc/downloads/optane_slog_pool_performane.pdf
https://forums.servethehome.com/index.php?threads/optane-nvme-for-slog-pooldisks-or-all-in-one-via-vdisk-on-omnios.17596/#post-169543

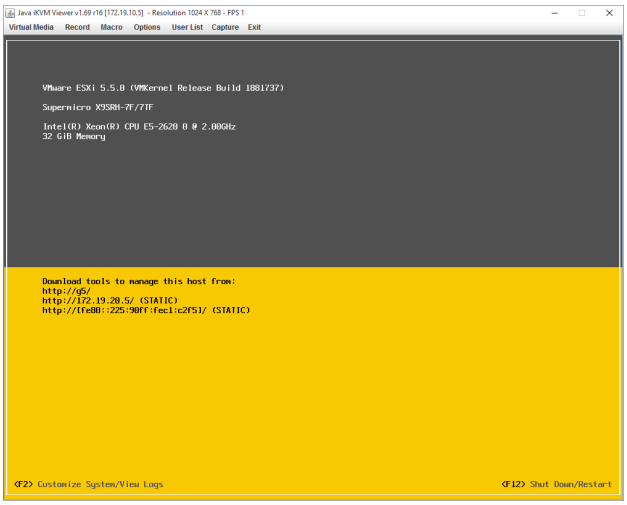
3. Setup ESXi

3.1 Download ESXi

The „VMware Vsphere Hypervisor ESXi" iso installer can be downloaded for free after registration from https://my.vmware.com/web/vmware/downloads ESXi comes with a 60day trial key that includes all commercial features. Download the free Key from the registration page to enable all free features without a time restriction or use your commercial keys. Use the Rufus tool to create a bootable USB installer stick.

Boot the installer and install ESXi to an M.2 or an Sata SSD/Disk (100GB min). During setup you must enter a root password, select a management nic and ip settings.

After setup, you see the local ESXi status screen that allows to edit network settings of the management interface and a reboot or shutdown

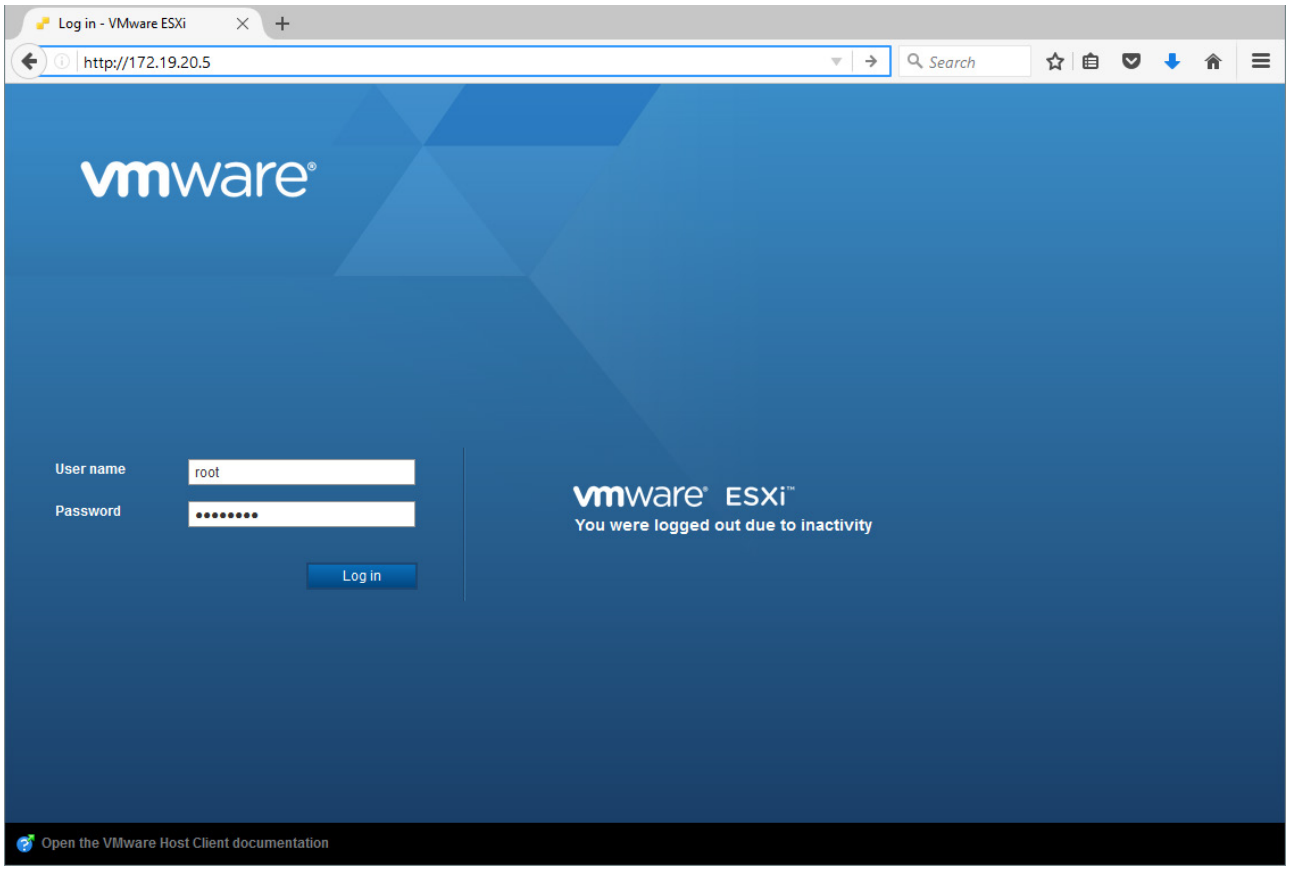


ESXi console (console monitor or IPMI)

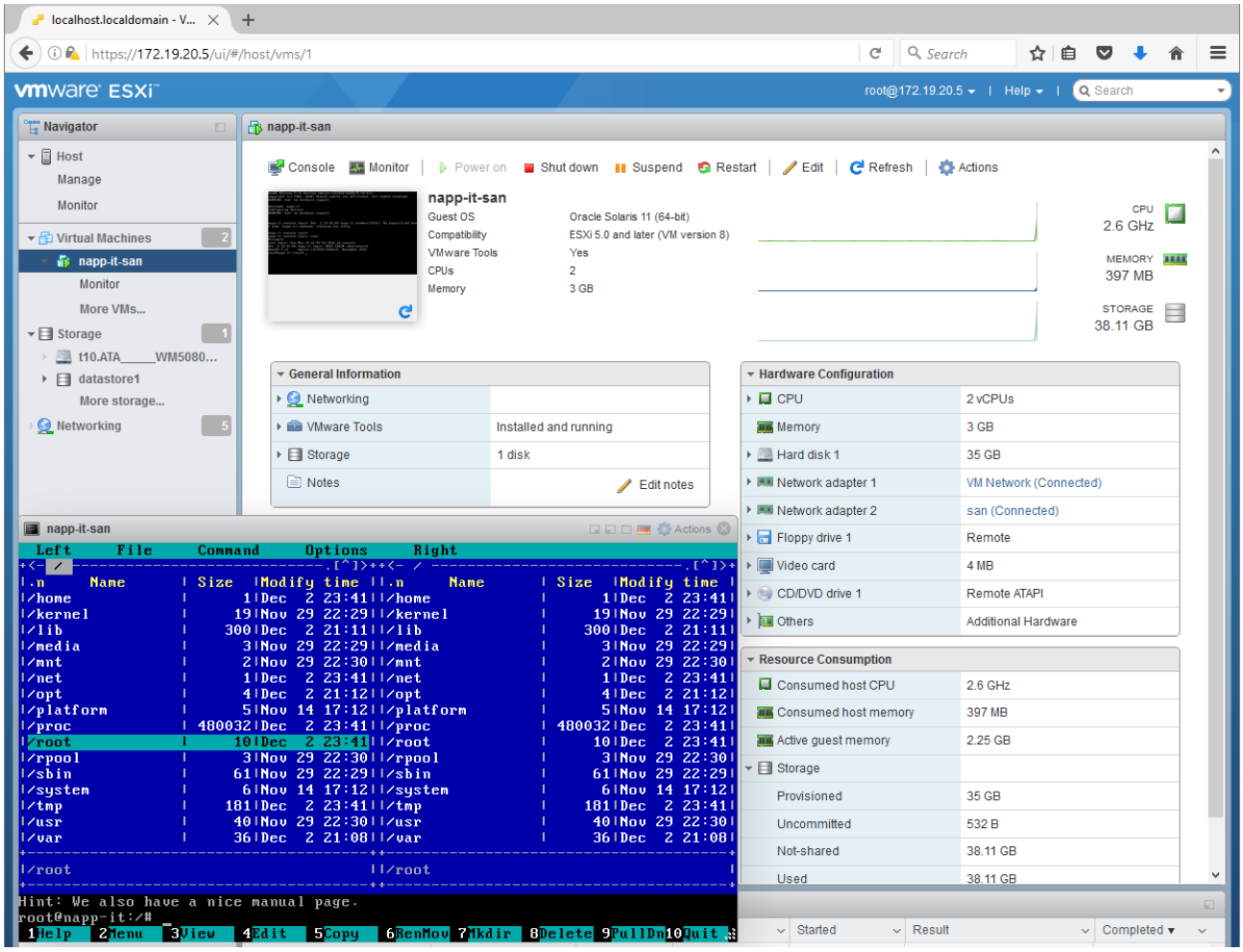
3.2 Start Management Interface

ESXI can be managed via your browser.

Start your browser and connect to the ip of ESXI: http://esxi-ip



3.3 Webmanagement Interface for ESXi



3.4 Setup ESXi

- Replace 60 day trial key with an enterprise key or a free key
- Deploy the napp-it ZFS server template to local datastore (minimalistic OmniOS, free download from napp-it.org)
- Assign at least 4GB RAM and 1 vcpu
- Add Sata/SAS disk controller or NVMe in passthrough mode

or

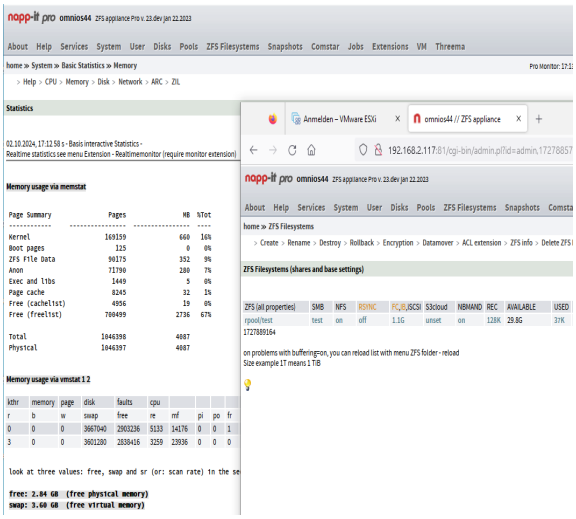
- Create a folder iso on your local datastore and upload OS isos to this folder
- create a new VM (ex Solaris 64 bit), bootdisk > 30GB, 4 GB RAM and the iso file as DVD bootdevice
- boot OS installer and setup, optionally install VMware tools (already included in OmniOS)

3.5 Setup Storage VM

After the ZFS server template is deployed and OmniOS started
Start your webbrowser and open http://omnios_ip:81

- Create a datapool ex tank
- create a filesystem ex nfs
- set NFS and SMB share to on in menu ZFS filersystems

- import this filesystem in ESXi as an NFS v3 storage.
- use this NFS datastore for all VMs beside the storage VM
- use Windows SMB for copy/move VMs or
  ZFS snaps=Windows previous versions



efficient napp-it Storage VM with 4 GB RAM

4.0 Backup the napp-it storage VM

First of all, you do not need a backup of the storage VM if you follow the suggestion that you should not install any services beside storage that require complex configurations. In such a case, you only need to reimport the original template or your modified one, then import the datapool and re-add the VMs to inventory.

As the storage VM is on local datastore, you have the the following backup options:
- Online Backup: You need a backup tool for ESXi what mostly work only with a commercial ESXi license

- Offline Backup (recommended)
  Shutdown the Storage VM and export the VM as a template in main menu
  File >> Export >> Export OVF template

  Such a template is quite small in size and can be reimported with all current settings, see 3.4 or 4.0

4.1 Recovery of the napp-it storage VM

Basically you can simply reimport the OVA template that you have created under 4.5 or you re-deploy the default napp-it OVA template and import the old datapool. If you have added users for SMB you must recreate the users and the jobs. If you have enabled a napp-it autojob „backup napp-it" you can restore all napp-it settings when you restore the content of /var/web-gui/_log and /var/web-gui/_my from your datapool/backup_appliance folder. You can use napp-it restore or WinSCP from Windows to copy these files back.

Again: You should not add services that require configuration or a complex recovery to the base/storage VM

Use VMs for any services beside basic storage. With ESXi you have the free choice between BSD, OSX, Linux, Solaris and Windows. Put the VMs on ZFS storage (shared NFS storage delivered by napp-it) as this offers performance and fast access from your desktop via SMB, versioning with unlimited snaps and manual online backup/clone from snaps (accessable via Windows Previous Versions) or automated online storage replication to a second backupsystem with zfs send.

4.6 Backup and Recovery of additional VMs (any VM beside the storage VM)

offline backup options (VM is down)
- save the VM folder on NFS or export as a template
- create a ZFS snap and replicate the snap to another location

online backup options (VM is running)
- use a commercial tool like VEEAM
- create simple ZFS snaps and replicate to another location with the risk of a corrupted VM.
- create ZFS snaps with embedded ESXi hotmemory snaps and replicate to another location, see
https://forums.servethehome.com/index.php?threads/napp-it-zfs-server-on-omnios-solaris-news-tips-and-tricks.38240/#post-367124

restore/ add to inventory
- Restore a VM folder on NFS (Windows SMB copy or previous version) or restore the NFS filesystem via ZFS rollback or replication. When you have access to an unregistered VM ex on NFS, navigate to the VM folder with the ESXi file-browser and use a right-mouse-click on the .vmx file to register (add to inventory) the VM. The VM is off.

Recovery of VMs from ZFS snaps with embedded ESXi hot memory snap (save restore of a running VM)
This requires that you power down all VMs, then restore the VM folder to a ZFS snap state with embedded ESXi hot memory snaps (ex via Windows previous versions) or start a ZFS rollback of the whole NFS filesystem.

You can open the VM folder on NFS via the SMB share to check if snapshots are available. ESXi will not see the snapshots until you reboot ESXi, so reboot now.
After reboot you can restore the ESXi hot memory snap. The VM is then online with the state of backup time.

5. ZFS Storage Server Management via napp-it

5.1 Manuals

read the manuals from http://napp-it.org/manuals/index_en.html like

- http://www.napp-it.org/doc/downloads/napp-it.pdf
- Basic System Administration Guide (Oracle Solaris 11 Express)
- Advanced System Adminstration Guide (Oracle Solaris 11 Express)
- ZFS Administration Guide (Oracle Solaris 11 Express)

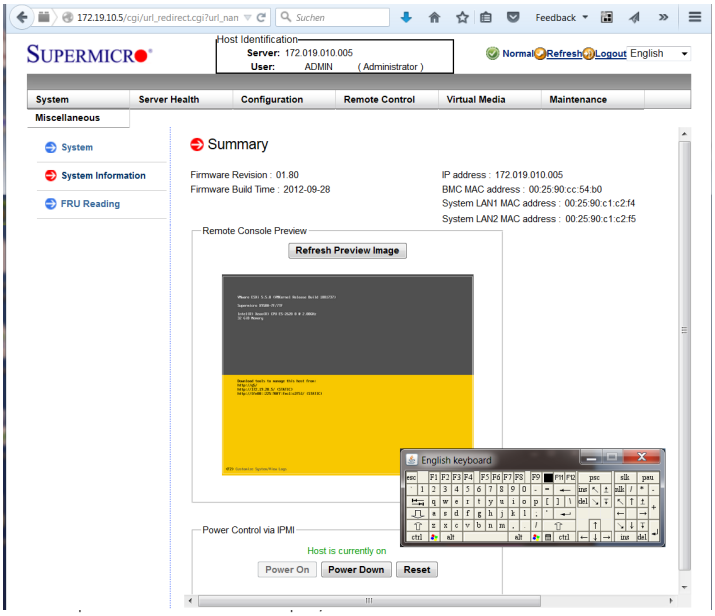
7. Remote management

A Server can be managed remotely, use these tools

7.1 IPMI

IPMI ist a must have for a server: https://en.wikipedia.org/wiki/Intelligent_Platform_Management_Interface

IPMI is a remote management microcontroller on serverclass hardware like Supermicro mainboards that ends with a „-F„ . You can connect the microcontroller remotely with a webbrowser even when the server is in a power-off state. Functions are mainly power on/off/reset, a remote console/keyboard and the ability to mount ISOs like a lokal CD/DVD drive.

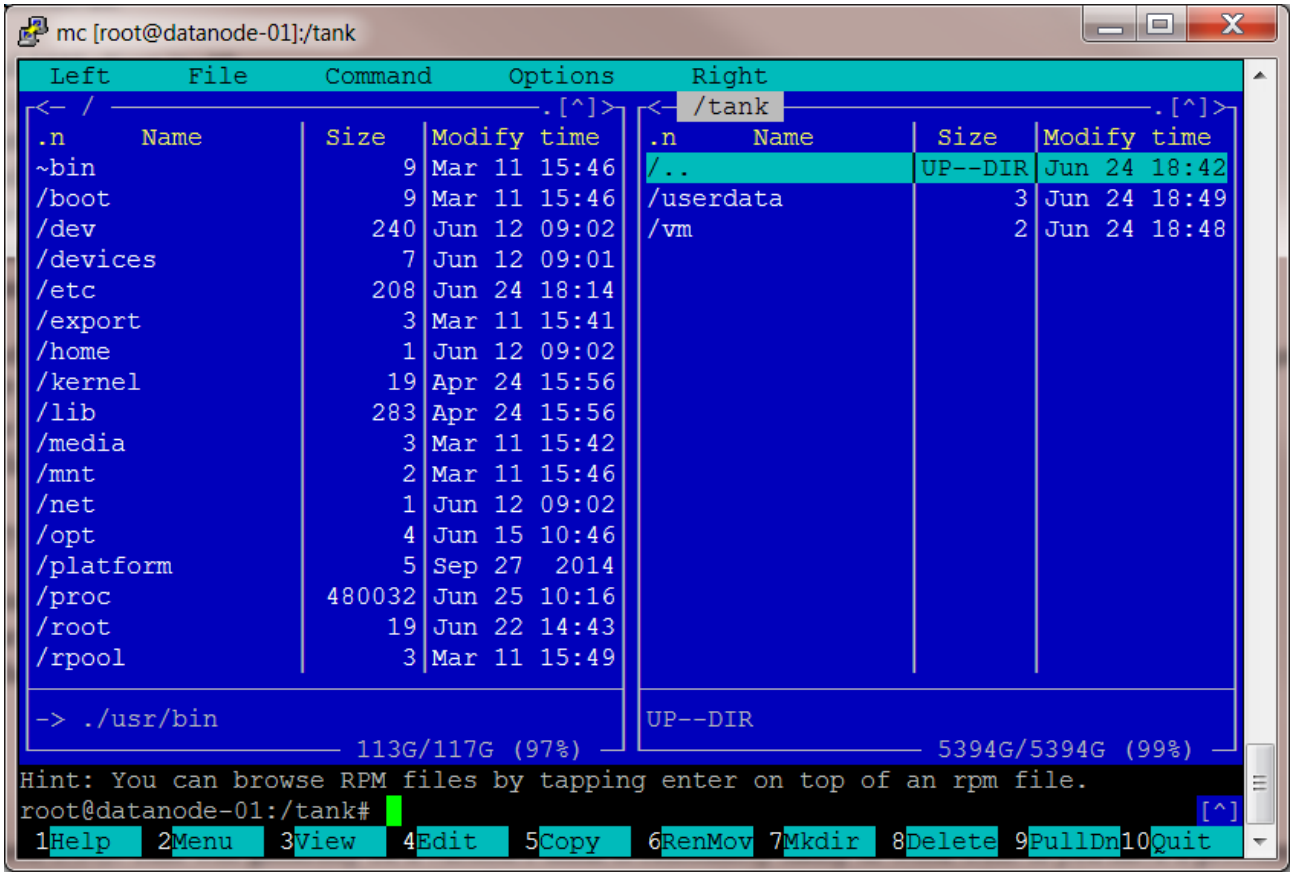


You can enable IPMI and its ip adress in your mainboard bios. It comes with a dedicated network port so you can connect with a dedicated and isolated management network. As an option, you can use your regular Lan port (insecure). IPMI requires a current Java (free download from www.java.com). For security reasons, you must allow the ip of your server (ex https://172.19.10.5) for java applets.

SuperMicro default IPMI user/pw (you should change that)
user: ADMIN
pw: ADMIN

7.2 Remote Console via Putty

http://www.chiark.greenend.org.uk/~sgtatham/putty/download.html



Daily management is done via the napp-it Web-GUI or vsphere. Some tasks require console access. This can be done locally or remotely via Putty, a free Windows application. Download and run - no installation required.

To use Putty, you must enable SSH on OmniOS. This is the case per default but per default only regular users can login, not root. So you must either create a regular user than can login. After this you can gain admin permissions wit a su command. Other option is to enable remote root access in the napp-it Web-GUI in menu „Services >> SSH >> allow root". As this can be a security problem, you should disable remote root afterwards with menu „Services >> SSH >> deny root"

Tips:

You can copy/ paste CLI commands with a „right mouse click" into the Putty Window.
The same is the case when you mark text within the Putty console.

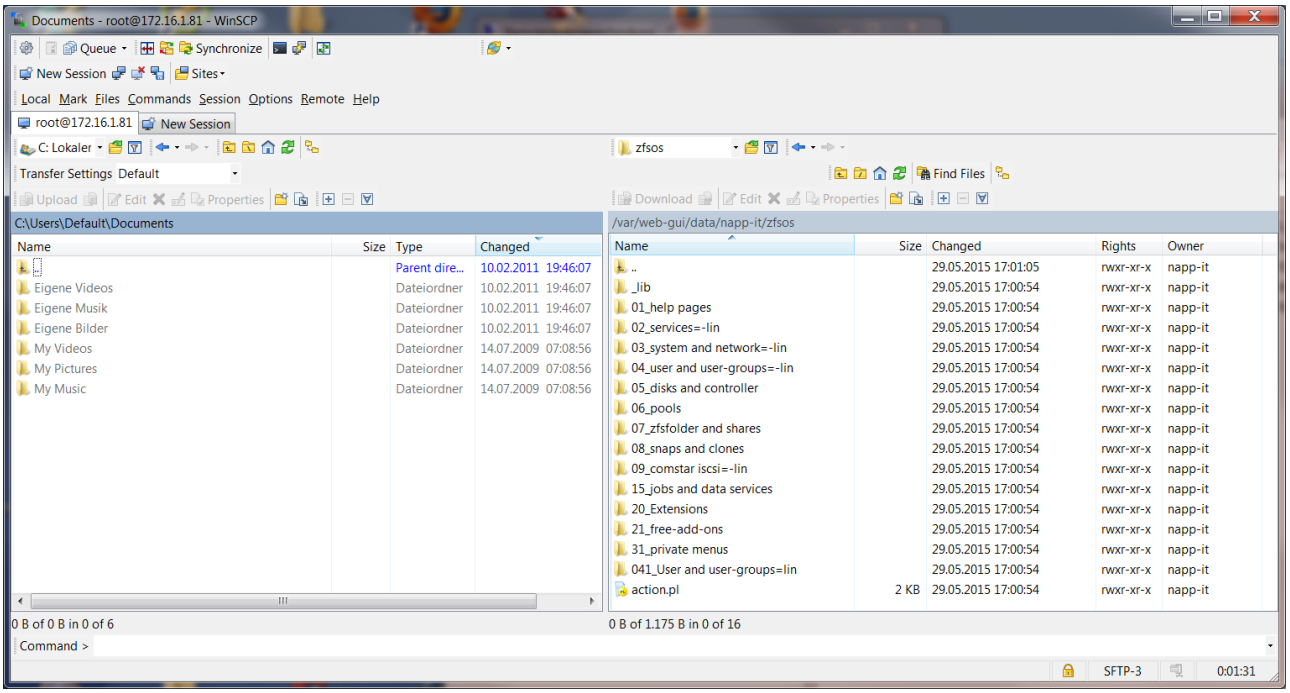
Midnight commander, a console filebrowser that runs on OmniOS with an optional usermenu is installed automatically by napp-it to do local file management (copy/move/delete/edit). This is the fastest way to copy/move files as it is done locally and not over your network.

You can start midnight commander when you enter „mc" at console and quit with F10.

If Midnight commander is showing wrong characters, you can either set a proper environment variable or call Midnight commander directly with a LANG envronment ex (German), start mc like:
LANG=de_DE mc

7.3 Remote Filemanagement/ Fileediting on Windows with WinSCP

http://winscp.net/eng/download.php



WinSCP is a „must have" tool.

WinSCP is a free Windows application that allows

- upload/download files (binary/text/auto) like a ftp client but encrypted
- edit /find files on OmniOS (you can use different editors for differen filetypes)
- delete/copy/move files (not as fast as Midnight Commander as files must be transferred encrypted over LAN)
- check/modify Unix permissions and ownership

To use WinSCP you must enable SSH on OmniOS. This is the case per default but per default only allows that regular users can login, not root. An option is to enable remote root access in the napp-it Web-GUI in menu „Services >> SSH >> allow root". As this can be a security problem, you should disable remote root afterwards with menu „Services >> SSH >> deny root" .

Tips:

When you connect as root, you have full permissions to edit all files on OmniOS including systemfiles. This makes Unix magagement a lot easier as you can manage remotely from Windows and do not need to use ancient editors like vi.

8. First steps with the napp-it Web-Gui

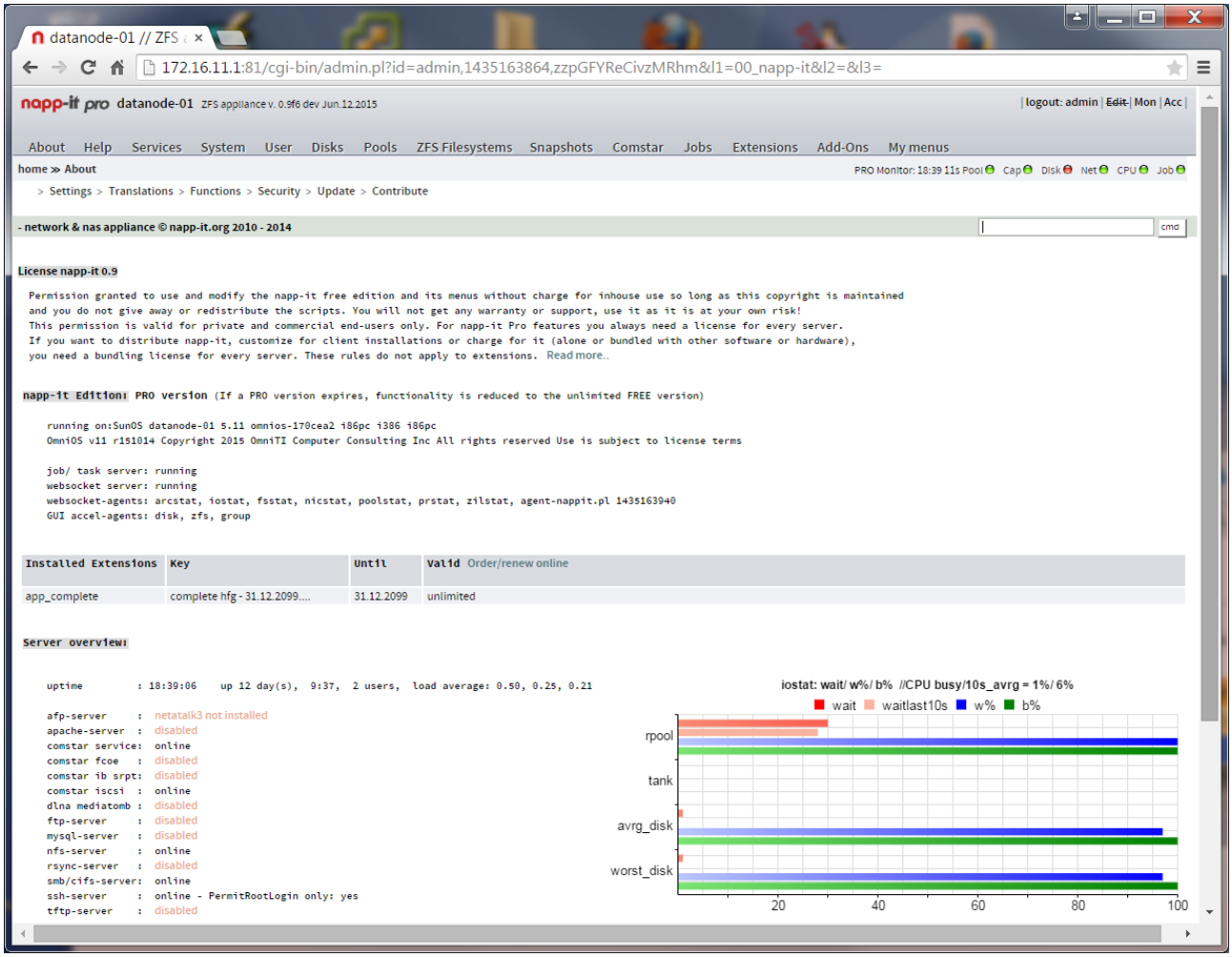
Use your browser to manage napp-it: http://serverip:81 example

http://192.168.1.1:81

If you are unsure about your ip, enter the following console command

ifconfig -a

If you start napp-it the first time, you are asked to setup napp-it passwords and email.



Napp-it welcomes you with the above startscreen that shows basic infos about your OS/ napp-it release, the state of services and your current iostat. After an initial setup, napp-it comes with a 30 day evaluation of Pro features. This includes realtime monitoring of single appliances and appliance groups, WWN enclosure management, remote replication, advanced ACL management and an improved GUI performance due background acceleration agents.

Napp-it Pro includes support options like access to developer or bugfix releases or email-support for the complete edition. After the 30 day evaluation, you can continue to use napp-it free without a time or capacity restriction -even commercially- with all features that are needed to manage a ZFS storage appliance.

Napp-it free is not crippleware or a product that is limited in essential features. It is sufficient for many cases. It is a stable state of napp-it that is updated from the dev release from time to time. If you want to support napp-it or use Pro features or require immediate access to bugfix releases commercially or as a homeoffer, check http://napp-it.org/extensions/quotation_en.html

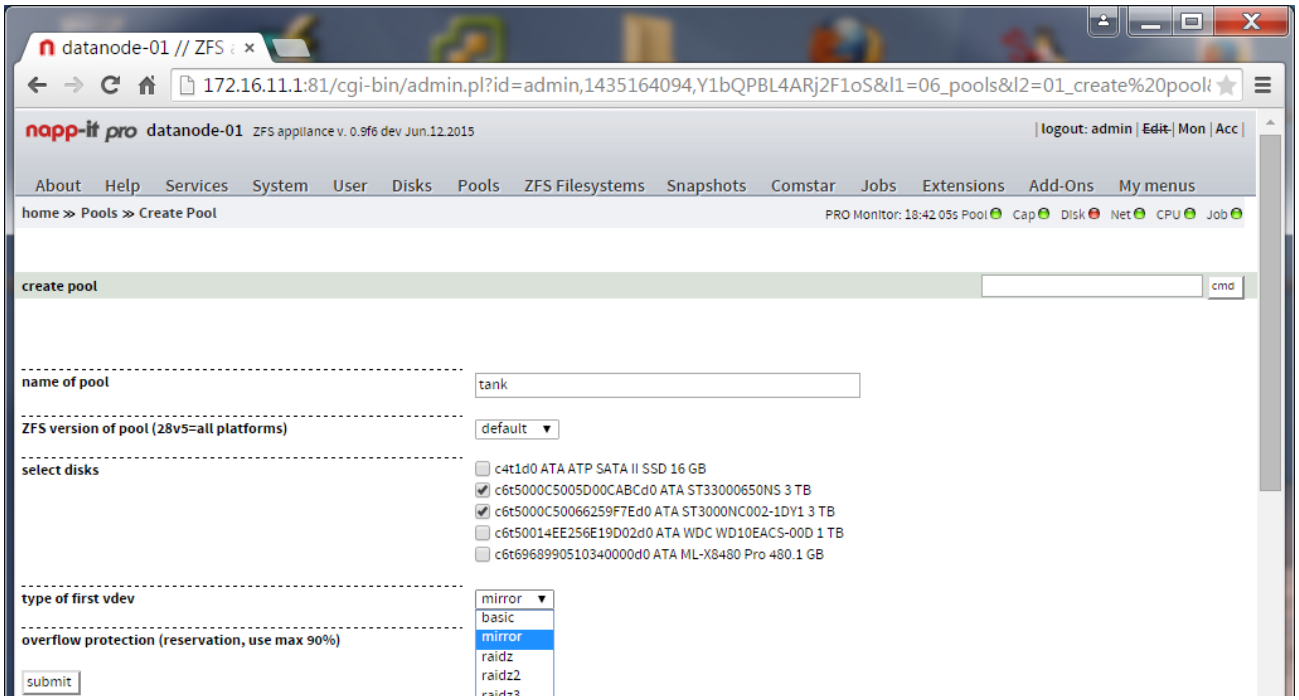
9. Create a ZFS datapool

From other systems like Windows, you know disks where you can create one or more partitions with a fixed size. You can combine single disks or partitions to a Raid that is treated like a single disk. It is possible to increase a partition up to disk or Raid-array size. But you cannot span a partition afterwards over multiple disks or raid-arrays without destroying the old partition. If a disk is full, you must create a new one and copy over data.

ZFS allows a more flexible handling of disc capacity with a concept that is called storage virtualization. Base of this is a storage pool. Unlike a disk or conventional raid array, the size of a ZFS pool can grow dynamically. You start with a new pool that is build from a Raid-Array example a ZFS Raid-1 or Raid-Z without Raid problems like the write hole problem of conventional Raid. If you need more capacity, you add more Raid-Arrays to build storage up to the Petabyte range as this is the real design goal of ZFS.

Similar to oldstyle partitions, you create ZFS filesystems on your pool but unlike old partitions, you do not set a size of a filesystem as it can grow dynamically up to the poolsize. If you increase the pool, the additional capacity is immediately available to all filesystems. You can limit capacity with quotas and ensure with reservations.

create a ZFS pool with menu Pools >> create Pool



- name your pool (ex tank)
- select version (only needed for compatibility mainly with Oracle Solaris and pool v28/5)
- select disks that you want to use for your first raid-array/ vdev (ex two 3 TB disks in a mirror)
- enable overflow protection (a 10% pool reservation that limits usable capacity to prevent a full/slow pool)
  You can reduce/ delete the reservation in menu ZFS Filesystems at any time

click submit and your pool „tank„ is created. Details about the pool: see menu Pools

Pool	VER	RAW SIZE/ USABLE	ALLOC	RES	FRES	AVAIL zfs [df -h/df -H]	DEDUP	FAILM	EXP	REPL	ALT	GUID	HEALTH	SYNC	ENCRYPT	ACTION	ATIME
rpool	-	149G/ **143.9GB**	26.8G	-	-	**113G** [114G /122G]	1.00x	wait	off	off	-	162800115664189026	ONLINE	standard	n.a.	clear errors	off
tank	-	5.44T/ **5.8TB**	427K	-	543G	**5.27T** [5.3T /5.8T]	1.00x	wait	off	off	-	16273592907840458286	ONLINE	standard	n.a.	clear errors	-

info: RAW poolsize does not count redundancy, usable/available size is from zfs list, df -h displays size as a power of 1024 wheras df -H displays as a power of 1000

click on poolname to list all pool-properties or on a property to modify.

Extend a ZFS datapool

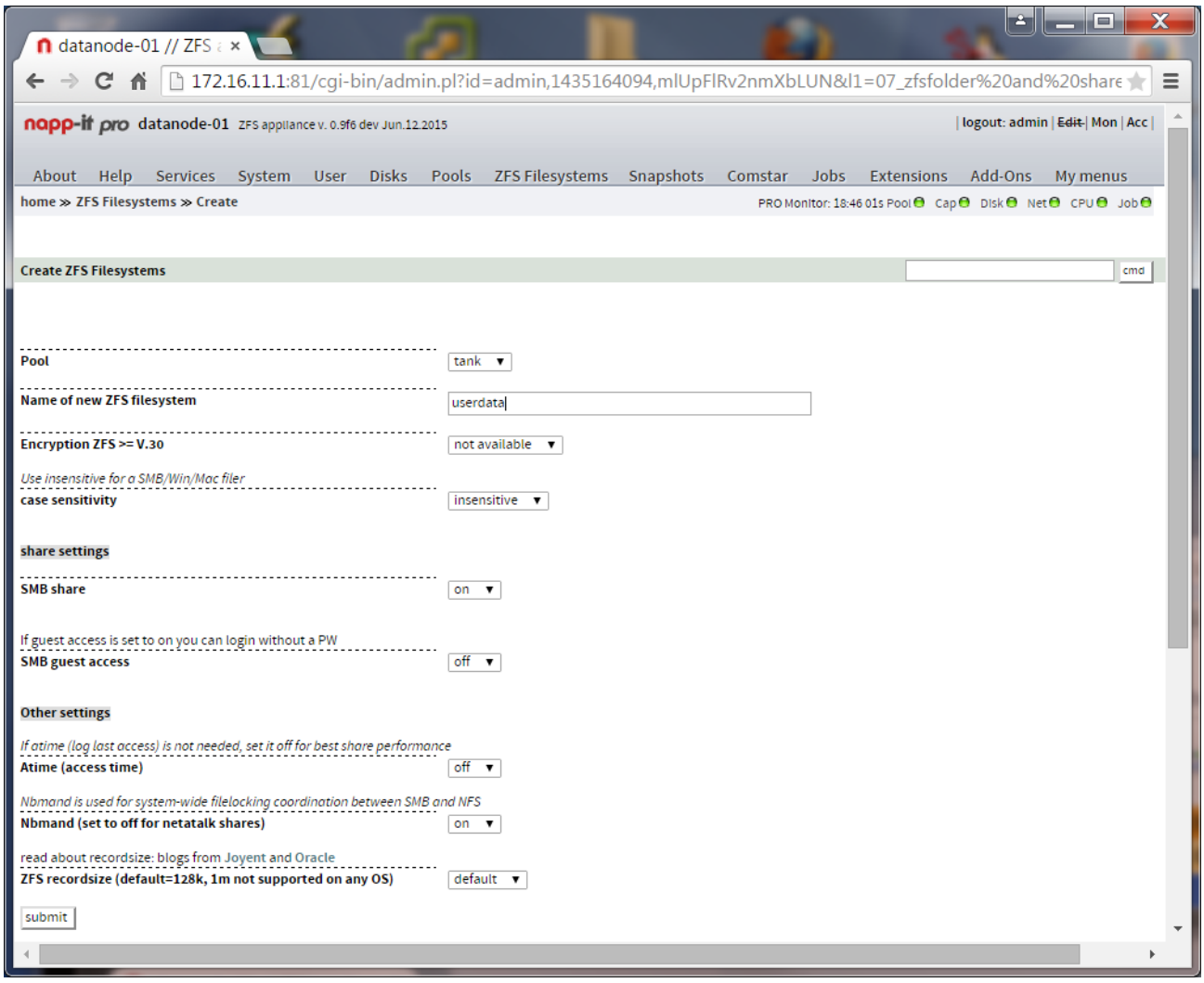
If you want to increase capacity, use menu Pools >> extend pool and add another Raid-array (ZFS call it vdev) ex a mirror or raid-Z. ZFS will stripe data over all vdevs to increase not only capacity but performance.

10. Create a ZFS filesystem

From other systems like Windows, you know partitions that you can format in FAT32 or NTFS. This is similar to OmniOS and ZFS with the difference, that you always format to ZFS and that the size of a filesystem can grow dynamically up to poolsize. You can limit the available capacity of a filesystem with quotas and ensure with reservations. This is called storage virtualization.

Basically it is enough to create a single filesystem and use traditional folders below to organize your data. But as every filesystem can have different ZFS properties, can be replicated and has its own snapshots, it is common to use as many filesystems as you like, up to thousands (example one filesystem per user).

create a ZFS pool with menu ZFS Filesystems >> create

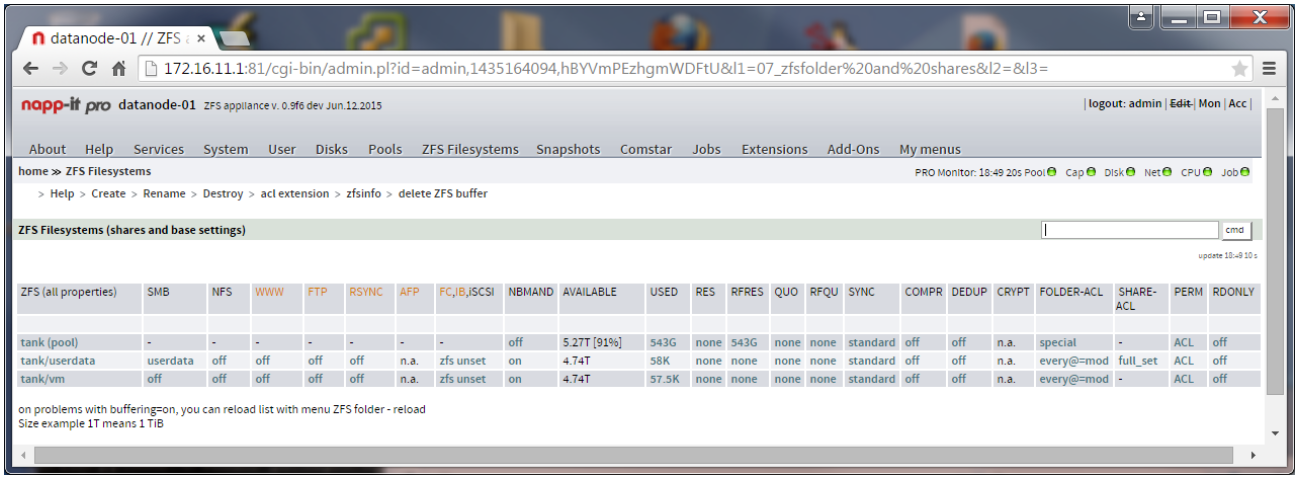


- select your pool (ex tank)
- name your new filesystem (ex userdata)
- select case sensitivity (Unix is case sensitive, Windows not - for a SMB server use the „Windows-behaviour")
- set immediate SMB sharing on or off
- other settings like atime, nbmand and recordsize

click submit and you have created a filesystem, optionally with SMB sharing enabled. You can now connect from Windows as user root as you do not have created other users yet. Default permission is everyone=modify.

Create more filesystems ex vm when needed (ESXi datastore)

Menu ZFS filesystems



Most settings about share and filesystem properties are ZFS filesystem properties that can can be set/controlled in this menu. You can click on an editable setting (they are blue coloured) to modify. Examples:

enable/ disable a SMB share:
click in the row of a filesystem example tank/userdata to the entry under the column SMB

enable/ disable a NFS share
click in the row of a filesystem example tank/vm to the entry under the column NFS

enable/ disable a iSCSI share (a ZFS volume as a blockdevice)
click in the row of a filesystem example tank/vm to the entry under the column iSCSI

set a quota for a filesystem
click in the row of a filesystem example tank/userdata to the entry under the column QUO or RFQU

set a reservation for a filesystem
click in the row of a filesystem example tank/userdata to the entry under the column RES or RFRES

enable sync write for a filesystem
click in the row of a filesystem example tank/userdata to the entry under the column SYNC
Sync write setting affects data security. Off means fast cached writes but last 5s are lost on a powerloss.

enable LZ4 compress for a filesystem
click in the row of a filesystem example tank/userdata to the entry under the column COMPR

enable dedup for a filesystem
click in the row of a filesystem example tank/userdata to the entry under the column DEDUP
Warning: dedup works poolwide. With low RAM this can dramatically reduce performance.

set/reset ACL for files and folders
click in the row of a filesystem example tank/userdata to the entry under the column Folder-ACL
Reset ACL is free. Other features are part of an extension. You can set ACL via Windows in napp-it free.

list all ZFS filesystem properties
click on the filesystem name ex tank/userdata

11. SMB/ CIFS Server

SMB/CIFS is a filesharing protocol from the Windows world. It is widely used on any platform.
Even Apple switched to SMB in their newer OSX releases as the default sharing protocol.

On OmniOS/ Solaris you have two options for an SMB server. One is SAMBA that is available on any Linux/ Unix system. The other is the Solaris CIFS server that is available on Solaris based systems only and is the de facto standard SMB server there.

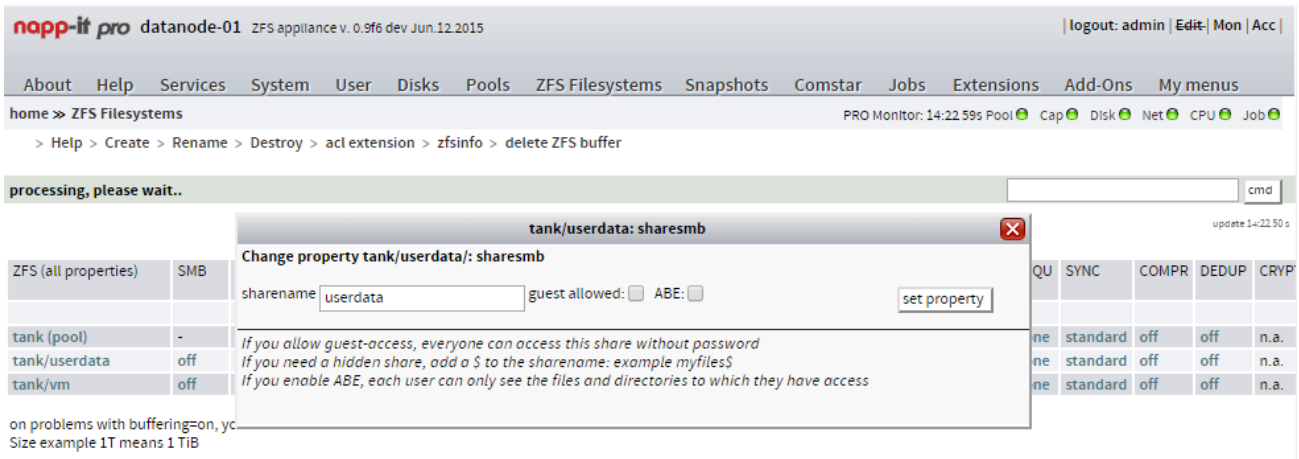
If one compare SAMBA with Solaris CIFS you will find many features in SAMBA that are not available in Solaris CIFS. But Solaris CIFS has some advantages that are not in SAMBA, mainly because SAMBA must run on any X-System with any filesystem. Some of these features are killer features as they affect easyness, performance or Windows compatibility like:

some Advantages of SAMBA over Solaris CIFS

- same server on any X-system
- can act as AD server
- a lot of sharing options
- nested shares/ shares independent from ZFS filesystems
- permissions are based on Unix UID/GID/ Posix ACL, this is a plus if you work mainly in a Unix world
https://en.wikipedia.org/wiki/Access_control_list

some Advantages of Solaris CIFS over SAMBA (used by napp-it)

- fully integrated in ZFS as a filesystem property, easy handling via zfs set command
  there is no configuration file, enable it and set permissions as file/share attribut.
- multithreaded and fast
- integration of ZFS snaps as „Windows previous version"
- manageable via Windows management console (connected users, open files, share level permissions)
- share and file/folder level permissions (Windows server alike)
- permissions are based on NFS4 ACL. allow/ deny with inheritance settings.
  They work very similar to Windows NTFS
- Windows SID as extended ZFS attribute. This allow a move/backup of data in a Windows AD
  environment between servers where permissions are preserved.



enable SMB sharing in menu ZFS Filesystems, click on off in the row of a filesystem under SMB

11.1 SMB related settings (Solaris CIFS)

SMB Service

The SMB service is started automatically when you enable a share.
Some modifications (like share level ACL) require a service restart. This is done automatically by napp-it.

On problems with the SMB server or if you are in a AD Domain that was temporarely unavailable,
it may be needed to restart the service manually in menu „Services >> SMB"

If you import a pool with shares enabled and SMB service disabled, you may get a warning that the SMB service is not enabled. You can ignore as the service is started automatically or after a share off/on.

SMB Share On

As SMB Sharing is a in OmniOS and ZFS integrated property of filesystem, you can enable a share in menu ZFS Filesystems when you click on off in the row of a filesystem under SMB with the following options:

- sharename: The share is visible to a client like Windows under this name
  if you add a „$" to the name, the share is hidden, example userdata$
  To connect such a hidden share, you must connect from Windows like \\datanode-01\userdata$
- guest allowed: You do not need to login with a name and password to access the share (ex from Windows)
- ABE (access based enumeration): Only files and folders are visible where you have permissions

SMB Share off

To disable a share, click on the sharename in the row of the filesystem and set sharesmb = off

SMB permissions

In contrast to other Unix services, Solaris CIFS uses Windows alike NFS4 ACL with permission inheritance, not traditional Unix permissions like 755 or Posix ACLs (https://en.wikipedia.org/wiki/Access_control_list).
This is the reason why you should not set Unix permissions like 755 on files/folders that are shared over SMB as this would delete ACL inheritance settings that are not know in traditional Unix.

Always use ACL to set permissions on Solaris. As tradititional Unix permissions are a subset of the ACL possibilities, they are reduced automatically to fit the ACL permissions.

As ZFS is a Unix filesystem, it must use Unix UID and GID as file security attributes. Solaris CIFS additionally store Windows Security ID's (SID) as extended ZFS attributs. They are used by the CIFS server only and allows file movements/ backups where Windows NTFS alike permissions were preserved - does not matter what UID a user has. This is an advantage especially in an AD environment.

When you create a new ZFS filesystem with napp-it, the default permission is set to
root = full access
everyone@ = modify

This allows that any user can connect a SMB share with read/write permissions as default.
If you do not create new users, only root has (full) access to regular SMB shares at the moment unless you do not had enabled the guest option that allows a connect without login.

ACLs on files and folders

Each file and folder on OmniOS has an owner (root or the creator), Unix permissions (traditional Unix permissions like 755) and NFS4 ACL permissions.

If you enter for example at console

/usr/bin/ls -V /var/web-gui/napp-it

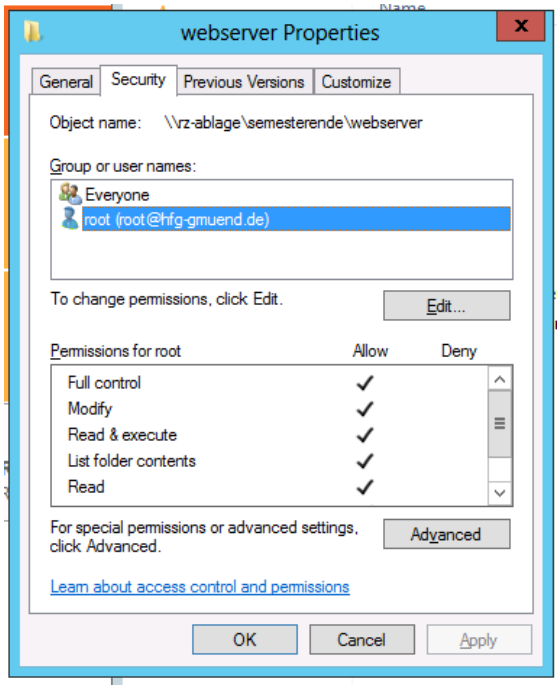
you may get as a result

```
root@datanode-01:/root# /usr/bin/ls -V /var/web-gui/_my
total 2
drwxr-xr-x   2 napp-it  root         2 Mar 19 15:00 wwwroot
             owner@:rwxp--aARWcCos:-------:allow
             group@:r-x---a-R-c--s:-------:allow
          everyone@:r-x---a-R-c--s:-------:allow
drwxr-xr-x   3 napp-it  root         3 Mar 19 15:00 zfsos
      user:2147484183:rwxp-DaARWc--s:fd-----:allow
             owner@:rwxp--aARWcCos:-------:allow
             group@:r-x---a-R-c--s:-------:allow
          everyone@:r-x---a-R-c--s:-------:allow
```

Owner (napp-it), group (root), Unix permissions like drwxr-xr-x and ACLs like owner@:rwxp--aARWcCos:-------:allow or auser ACL are shown.
Windows SID informations are not shown here as they are used in the CIFS server only.

ACLs can be assigned to OmniOS/AD users, OmniOS/SMB or AD groups or as a trivial ACL to owner@, group@ or everyone@ to be compatible to traditional Unix permissions.

You can modify ACL permissions with the /usr/bin/chmod command, per Windows or per napp-it ACL extension. Modifying ACL via CLI command is really stupid. Especially with napp-it free, you can use Windows (beside Home editions) to modify permissions. To do this, you can login to the SMB share from Windows as user root. A right-mouse click >> Property on a file or folders opens the Windows property dialog where you can select Security. Set Permissions just like you would do on a real Windows server.



Good to know:

User root or the owner/creator have always full access, you cannot lock them out - even if permission is not set explicitely. This is normal on Unix and different to Windows (and a boon for any admin doing backups)

If you set ACL on a folder, they are per default inherited to newly created files and folders unless you set „inherit to this folder only" The user that is logged in is the owner of new files and folders (with full permission).

You can override this behaviour with the ZFS property acl inheritance = discard or restricted (default is pass-through).

Windows processes first deny rules then allow. Solaris processes them in their order where the first matching rule is relevant. To set correct deny rules, use napp-it/ ACL extension

ACLs on shares

ACL on shares was introduced by Windows to restrict access independently and additionally to permissions on files and folders, mainly to restrict access without the need to modify file attributes.

On Solaris, a share control file is used that is created when you enable a share ex if you enable a SMB share share tank/userdata, you find the share control file as /tank/userdata/.zfs/shares/userdata

You can set share level ACL via napp-it/ ACL extension, or remotely via Windows server management (You must connect a share/ server management as a user that is a member of the SMB admin group.). Napp-it can restore Share ACL when you re-enable a share.

examples:

File ACL: full access
Share ACL: readonly

real permission: readonly

File ACL: readonly
Share ACL: full

real permission: readonly

File ACL: full
Share ACL: user: paul=full

real permission: paul= full, no other user allowed

The default share-level ACL is: full access
(only file attributes are relevant)

Modifications on share level ACL require that you disable/enable a share or restart the SMB service to take effect.

ZFS Properties aclinherit and aclmode

From http://docs.oracle.com/cd/E36784_01/html/E36835/gbaaz.html#scrolltoc

„ aclinherit

– Determine the behavior of ACL inheritance. Values include:

discard – For new objects, no ACL entries are inherited when a file or directory is created. The ACL on the file or directory is equal to the permission mode of the file or directory.

noallow – For new objects, only inheritable ACL entries that have an access type of deny are inherited.

restricted – For new objects, the write_owner and write_acl permissions are removed when an ACL entry is inherited.

passthrough – When property value is set to passthrough, files are created with a mode determined by the inheritable ACEs. If no inheritable ACEs exist that affect the mode, then the mode is set in accordance to the requested mode from the application.

passthrough-x – Has the same semantics as passthrough, except that when passthrough-x is enabled, files are created with the execute (x) permission, but only if execute permission is set in the file creation mode and in an inheritable ACE that affects the mode.

The default mode for the aclinherit is passthrough (napp-it only).

aclmode

– Modifies ACL behavior when a file is initially created or controls how an ACL is modified during a chmod operation. Values include the following:

discard – A file system with an aclmode property of discard deletes all ACL entries that do not represent the mode of the file. This is the default value.

mask – A file system with an aclmode property of mask reduces user or group permissions. The permissions are reduced, such that they are no greater than the group permission bits, unless it is a user entry that has the same UID as the owner of the file or directory. In this case, the ACL permissions are reduced so that they are no greater than owner permission bits. The mask value also preserves the ACL across mode changes, provided an explicit ACL set operation has not been performed.

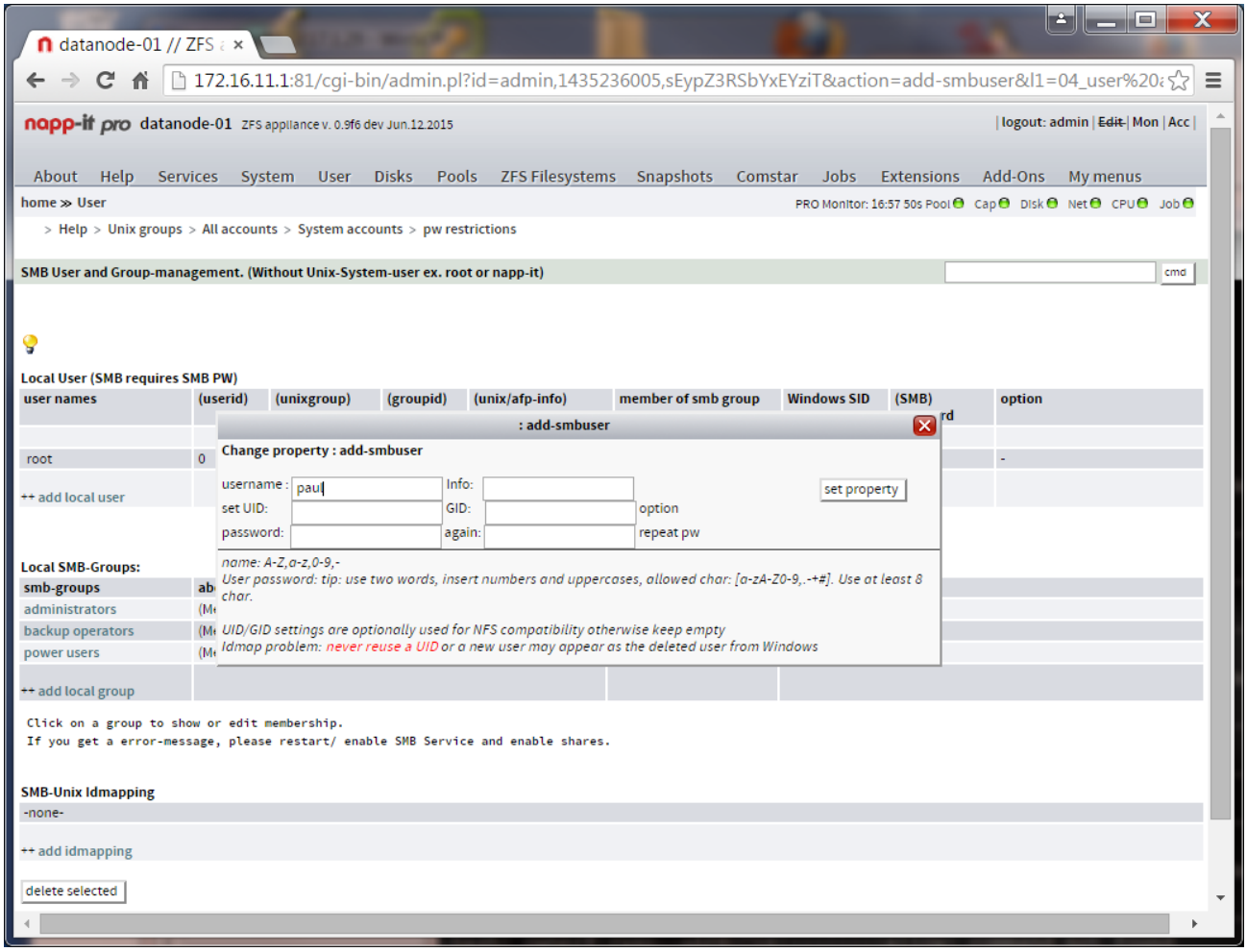
passthrough – A file system with an aclmode property of passthrough indicates that no changes are made to the ACL other than generating the necessary ACL entries to represent the new mode of the file or directory.

The default mode for the aclmode is pass-through (napp-it only). "

aclmode = restricted is added in Illumos/OmniOS to avoid permission
modifications with a chmod command (ex via NFS)

12. User and Groups

If you do not need to restrict access to a SMB share, you can enable guestaccess and you are ready.
If you want to restrict access, you can create users with napp-it menu „User"



When you create a user, you only need to enter a username and a password. This user is valid for SMB access and is a valid Unix user. You can assign a UID/GID for a new user when needed ex for NFS (optional).

Attention:

Windows groups and Unix groups behaves different. This is the reason why the Solaris CIFS server come with an own SMB group management that works independently from Unix groups.

If you need groups to restrict SMB access, you must do this with SMB groups.
Menu „User" allows to create SMB groups add add users to these groups.

There is an idmapping option Winuser -> Unixuser and Wingroup -> Unixgroup in Solaris/ OmniOS.
While a usermapping makes sense only in an AD-environment to map an AD user to a Unix user (never map a local Unixuser to a local Unixuser) you can map local SMB groups to local Unixgroups to achieve a similar permission behaviour within SMB and locally on Unix.

ACL settings for multi-user SMB access

some basic examples for File and Folder ACL settings

Goal:
- everyone can access and read files from a share like data (data is a filesystem below a pool named tank)
- everyone can modify files in data/common and below
- user paul is the only one to access data/paul and below

needed ACL settings:

folder /tank/data:
allow everyone@=readx (read and execute), no-inherit, this folder only

folder /tank/data/common:
allow everyone@=modify, inherit to folders and subfolders

folder /tank/data/paul:
allow paul=modify or full, inherit to folders and subfolders

Goal:
- everyone can access and read from a share like data (data is a filesystem below a pool named tank)
  read should be allowed only from folder /tank/data, not folders below
- everyone can create new folders but not files on data
- only the creator of a folder (=owner) has access to the new folder and below

needed ACL settings (aclmode must not restrict ownership when creating folders):

folder /tank/data:
allow everyone@=readx (read and execute), no-inherit, this folder only
allow everyone@=create_folder_set, inherit to folders and subfolders
allow owner@=modify or full, inherit to folders and subfolders

Goal:
- everyone can access and read from a share like data (data is a filesystem below a pool named tank)
- everyone can read files files on /tank/data/common and below
- members of SMB group „professors" are allow to modify /tank/data/common and below
- only members of SMB group „professors" are allowed to modify /tank/data/professors and below

needed ACL settings (aclmode must not restrict ownership when creating folders):

folder /tank/data:
allow everyone@=readx (read and execute), no-inherit, this folder only

folder /tank/data/common:
allow everyone@=readx, inherit to folders and subfolders
allow group:professors=modify, inherit to folders and subfolders

folder /tank/data/professors:
allow group:professors=modify, inherit to folders and subfolders

Active Directory: If you want to assign ACL to AD users, this may require that the AD user was logged in once to a SMB share.

12.1 Active Directory

With a few users especially if you use only one or two fileservers, a simple local user management is sufficient, simple and idiot proof. With more users, many servers you need to have the same credidentials on any machine. To achieve this you need a centralized user database like Ldap or Active Directory.

The Solaris CIFS server is prepared to be a member computer in an AD environment.
You can join a Domain with menu „Services >> SMB >> Active Directory

Enter:

Domainname ex	ex myuniversity.edu
IP-Adress of your AD-Master server	ex: 172.16.1.11
Domainadmin username:	admin
Domaiadmin password:	*****

When you click submit, OmniOS syncronizes time and sets DNS to your AD server and joins the domain.
If OmniOS lost connectivity to your AD server, you can disable/enable the SMB service or rejoin the domain.

Care about

Your napp-it Server can be either a member of a workgroup (use local user)
or a domain (can use either local or domain user).

If you switch from domain to workgroup-mode, remove all mappings with idmap remove -a
If you join a domain, you should create at least one idmapping (give domainadmin root permission):
Other mappings are not needed, replace domainadmin with your admin username.

```
idmap add winuser:domainadmin unixuser:root
```

Possible Problems:

If you get a „UNSUCCESSFULL" error

If you want to join a domain newly, please verify that your domain does not already have a
computer member with the name of your OmniOS server. In this case you get UNSUCCESSFULL

other reasons for UNSUCCESSFULL:
- wrong username/password

Sometimes UNSUCCESSFULL happens for unknown or timeout reasons.
- If you try again a second or third time it may work.

If you get a „INTERNAL ERROR"
- Check if SMB service is online (you must have enabled a SMB share)

If you get a „Failed to find any domain controller"
- check network, ip and DNS settings
- try another lmauth level (4 is ok for Windows 2012)

13. NFS Server

NFS is a filesharing protocol from the Unix world that is supported in NFS v3 and NFS v4.
Mostly NFS3 is used in secure environments where you mainly need performance as NFS3 lacks Authenthification (to login with name/pw) and Authorisation (no restrictions based on file permissions) beside some „good will" settings based on client ip and client UID.

You can enable/disable NFS in menu „ZFS filesystems" in the row of a filesystem under NFS as a filesystem property, similar to SMB sharing. Mostly you set on or off. Other option is to restrict access based on client ip or allow full acces independently from the client UID.

Enable NFS
set NFS = on

or instead on on something like
rw=@192.168.1.0/24,root=@192.168.1.0/24 tank/vm

Disable NFS
set NFS = off

14. iSCSI Server

iSCSI is not a multiuser filesharing protocoll like SMB or NFS. It offers blockstorage to a single client that is treated there like a local disk and formatted from the client with a filesystem. You need a cluster filesystem if you want to allow access from two clients simultaniously.

Originally, Sun implemented iSCSI sharing as a filesystem property like NFS and SMB. As iSCSI is mostly used in large and complex HA environments, this approach was replaced by COMSTAR, a enterprise framework to manage iSCSI and FC environments.

When should you use iSCSI/ FC

- when you need a non-ZFS filesystem like ext4, HFS+, NTFS or VMFS
  ex ESXi environments or a Windows Server based on ZFS blockstorage
- in HA environments with a setup similar to SAS multipath but with iSCSI datanode multipath
  to allow large capacity, high performance or remote installations (not limited by cable length)
- in HA environments with dataheads or clients (ex storageserver with services or a server) and datanodes

that provide their storage via iSCSI. This can be a simple datanode mirror or a raid-Z over datanodes.

Enable iSCSI via menu Comstar

1. create a Logical Unit (LU). This can be a ZFS volume, a file or a RAW disk
2. create a target (this is the part that you connect from a client)
3. create a target group with targets as members
4. add a view from your logical units to a target group to make them visible as a LUN.

Enable iSCSI via menu ZFS Filesystems

For smaller installations, Comstar is quite complex. Napp-it offers a way where you can enable iSCSI on a per filesystem way with a on/off switch in menu ZFS filesystems in the row of a filesystem under iSCSI.
If you enable iSCSI here, you create a logical unit, a target, a target group and a view in a 1:1 relation.
If you need more than on/off or basic settings, you can manage the targets with menu Comstar as well.

15. Data Scrubbing

With older filesystems like Ext4 or NTFS you have had the following problem.
You edit a file and you have a crash or powerloss while you write/ update the file. As these filesystems do inline modifications (modifiy parts of the current data on disk) the following can happen:

regarding your data
1. nothing written, old data is valid
2. new valid data, modifications are correct
3. modificated data is parly written, no chance to detect or repair the problem

regarding the metadata
1. metadate correct updated
2. metadata corrupt=corrupt filesystem, an offline fschk is needed to repair the filesystem structure

No chance to detect metadata problems beside a offline fschk that can last days and this does not even help to detect or repair data corruptions. The result is only valid metadata structures.

ZFS, a new generation filesystem
https://en.wikipedia.org/wiki/ZFS

ZFS stands for a new generation of filesystem that do not try to reduce these problems but to avoid them completely with two basic principles: CopyOnWrite and End to End Checksums on data/OS level.
CopyOnWrite means, that you do not do inline updates of old data but write datablocks always new.
In the above powerloss szenario, ZFS behaves like:

regarding your data and metadata
1. modified data is written completely new, data pointers are updated, new data is valid and verified
2. modified data is not written completely, data pointers are not updated, old data keeps valid and verified
3. If anything goes wrong, this will be detected by checksums on next read and auto repaired (self healing)

That does not mean, that you cannot have a dataloss. If you write/update a large file with a powerloss, this file is corrupt and damaged (no miracle with ZFS) but your filesystem is not damaged ans always valid.

reasons for corrupted data on disk

- Powerloss or a system crash
- I/O errors due to a bad disk, controller, cabling, backplane or PSU
- On-disk data corruption due to cosmic rays, magnetic or electromagnetic fields
- Driver bugs resulting in data being transferred to or from the wrong location
- A user or Virus modifying data by accident or intention

All beside the last problem can be at least detected and mostly autorepaired by checksums and redundancy.
For the last point ZFS offers snapshots. On a multiTerabyte Array you will always find corrupted data over time.

Scrub on every read, on demand or as a planned job

On every read, data is checked and repaired from Raid redundancy when a checksum error is detected. (auto self healing filesystem). If you want to check the whole pool you can start a scrub manually or as a planned napp-it Job. With desktop disks I would do this once a month ex on a low io day like saturday. Unlike a traditional fschk that require an offline filesystem for quite a long time without a real repair option, a scrub is a online process that runs in the background with low priority and verifies/repair all data.

I own many server and systems and see checksum repairs quite often: This feature is mandatory!

16. Data Snapshots and Backup

If you care about your data, you do backups. If you really care about your data, you do multiple backups like tapes that you rotate on a daily or weekly base for a data history. If your data was deleted or modified by accident or intention (virus, staff) you have a chance to regain original data from a former backup.

While you always need a backup for a real disaster like fire or a thief, this concept has three weak points.
The first is that the number of backups is mostly limited due to limited resources.
The second is, that access to a backup means mostly a restore that is at least inconvenient.
The third is, that you often cannot trust the backup, because it usually has no checksum verification and no repair option on problems. Mostly you or the admin discover this when you need the backup – too late..

Data versioning/ secure backup

For the data versioning problem with regular user access to former states, you can save several versions of a file like report-2015.doc, report-2015v2.doc or report-2015-this is the latest.doc.

Another option is a mechanism like Apple Timemachine, where you copy/sync dataversions to a another disk on a regular base like once a day. While this work it is annoying because of the delay when you must copy or restore huge data.

Another option is Windows shadow copies on a Windows server (https://en.wikipedia.org/wiki/Shadow_Copy). This is a block level snapshot mechanism of the whole volume. The result is a versioning filesystem. If you do daily snaps, you can browse/ restore the data with Windows „Previous Versions".

The problem remains, that you should not do too many snaps. I also had a problem with Windows VSS snaps in the past that they were lost after a system crash with a fress OS install. Main problem: you cannot really trust NTFS filesysystems (not always consistent like a CopyOnWrite filesystem) and no checksums (no verified data, no autorepair). This may be different in future with ReFS but currently this is not a comparable option to ZFS.

ZFS snapshots

ZFS snapshots are far better than the former solutions. ZFS is a CopyOnWrite filesystem where all modified datablocks are written newly while the former datablocks can be overwritten after a succesfull write.

A ZFS snapshot means that the former datablocks are blocked and cannot be re-used unless the snap is deleted. This requires only to keep some datapointers and can be done without delay and no initial space consumption. Even ten-thousands of snaps can be hold without any problem (Okay as the former state blocks capacity, sometimes the pool is full). As this is done on ZFS storage, checksum verification, scrubbing and autorepair is working - does not matter how old a snap is - ideal for long term storage and archives with regular scrubs.

As this is managed by the ZFS pool itself, you are not in danger too loose them when you move a pool. You can also trust these snaps absolutely. An admin can destoy a snap but cannot modify data as a snap is readonly.
The best is, this is transparent to a user. You can access ZFS snaps on Solaris via „Windows Previous Versions" with all ACL or AD permissions intact even from a Backup. With a snapshot rule like take a snap every 15 Minkeep 4, every hour-keep 24, every week-keep 4, every months- keep 24 you can go back two years on a filer.

ZFS backup

While you can backup data from a ZFS storage to any system, ZFS offers remote and ultrafast incremential replication based on snaps where only modified datablocks are transferred with ZFS security and their own snapshot history. In my own setups, I use two main backup systems in my serverroom where I replicate data based on even or uneven days and a third backup system in another building for important data and a snapshot history that covers at least 90 days (daily snaps)

17. Basic operational settings

Your storage appliance is now up and running. Care about the following settings

Napp-it settings (menu About >> Settings)
all settings are stored in /var/web-gui/_log/napp-it.cfg

- set passwords for admin and operator (encrypted one way hashvalues)
- set email (mailserver, mailuser, mailpw, store unencrypted)
- set push data (alerts to your desktop or smartphone)

System-Settings

- Menu Sytem >> HW and Localization >> Localization
  ex America > New_York, set Language en_US.UTF-8 and your keyboard, you need a reboot

- create bootable snapshots (=BE, bootenvoronments) manually prior serious system modifications
  This is done automatically on OS or napp-it updates and allows a bootup on a former OS state.

Auto-Job Settings

- Enable napp-it auto-job to 5min (Jobs >> autos ervice)
- set other job to sync time via ntpdate > AD server or any other ntpserver

- Set email-alert and status jobs in menu Jobs >> Email >> alert or status
  Per default napp-it sends email unencrypted over port 25
  If your smtp server requires TLS encrypted mail example Googlemail over port 587, you must
    - install TLS modules, see http://napp-it.org/downloads/tls.html
    - switch napp-it to use TLS in menu Jobs >> TLS Email >> enable TLS

- Set push-alert (Pushalot or Pushover) for your desktop or smartphone
    - more see www.pushalot.com (Windows8 and -Phone, free) and www.pushover.net (ios, Android)

- Set a backup job (Jobs >backup >> create backup job) tp backup basic OS and napp-it settings to a pool

- Set autoscrub jobs (see 15.)
  for your pools in menu Jobs >> scrub >> create autoscrub job
    - ex set autoscrub of your system pool (rpool) to every first sat (of a month)
    - set autoscrub of your datapools (with desktop disks I would use once a monty as well)

- Set autosnap jobs (see 16.)
  for your pools in menu Jobs >> snap >> create autosnap job
  As ZFS snaps are readonly and cannot be modified/destroyed from a share, they are virus/user save
  This is your first and most important method to avoid dataloss and undo unwanted modifications ex:
    - set autosnap: snap every 15 min, keep 4
    - set autosnap: snap every hour, keep 24
    - set autosnap: snap every day, keep 31
    - set autosnap: snap every 1st of a month, keep 12

Your primary storage ist where you should care about a highest possible level of raid and data security. Data restoring can be done mostly from your primary storage as ZFS is a versioning filesystem with snaps.

To be prepared for a real disaster (sabotage, fire, overvoltage or a thief), you need a disaster backup at least with some snapshots. If data is important, this should be done to two different systems where one must be on a different physical location like another building or offline within a save. You can do this via sync (rsync or robocopy) or via the faster ZFS incremental replication that can be done every few minutes.
- set a replication job to another napp-it appliance (require replication extension and grouping)

18. Security

Restrict access to management functions

- Web management is done via port 81 (http) or port 82 (optional 443) for https
Realtime graphic/ websocket is displayed over port 3000 (http only)
- Remote console via Putty and remote fileaccess via WinSCP is done on SSH port 22
- Replications are done over a random port > 49000

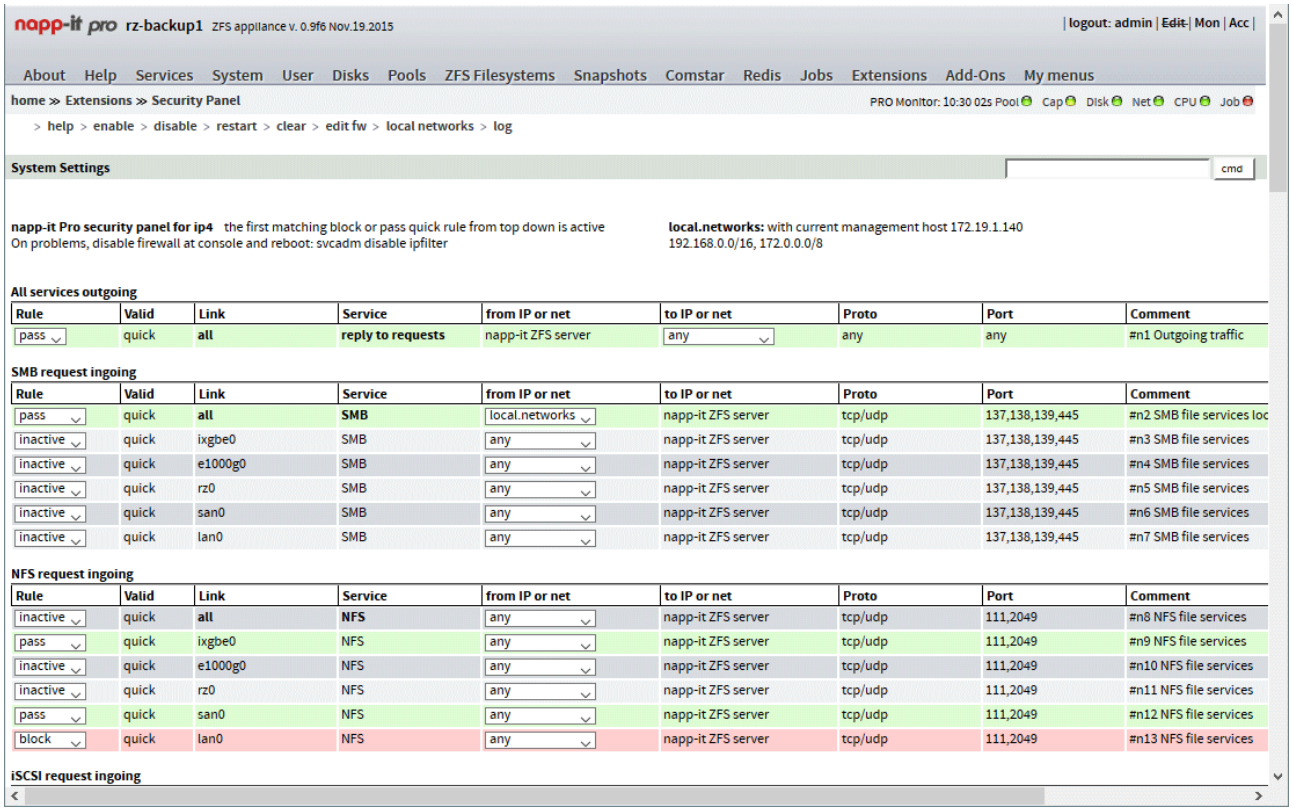
In an unsecure environment, you should restrict the above ports to a secure environment, either based on a network adapter (link) or based on your networks

Restrict access to file services

- Fileserveices like NFS3 do not offer authentication. Access can be only limited to a fakeable source ip. This can be a security problem example when you offer NFS for ESXi where your storage server is accessable over untrusted networks for management or other services.

In an unsecure environment, you should restrict access to services like iSCSI, NFS, SMB or WWW either based on a network adapter (link) or based on your local networks or single ip adresses.

Firewall settings/ Security panel (napp-it Pro)



napp-it pro rz-backup1 ZFS appliance v. 0.9f6 Nov.19.2015 | logout: admin | Edit | Mon | Acc |

About Help Services System User Disks Pools ZFS Filesystems Snapshots Comstar Redis Jobs Extensions Add-Ons My menus

home » Extensions » Security Panel — PRO Monitor: 10:30 025 Pool Cap Disk Net CPU Job

> help > enable > disable > restart > clear > edit fw > local networks > log

System Settings — cmd

napp-it Pro security panel for ip4 the first matching block or pass quick rule from top down is active
On problems, disable firewall at console and reboot: svcadm disable ipfilter

local.networks: with current management host 172.19.1.140
192.168.0.0/16, 172.0.0.0/8

All services outgoing

Rule	Valid	Link	Service	from IP or net	to IP or net	Proto	Port	Comment
pass	quick	all	reply to requests	napp-it ZFS server	any	any	any	#n1 Outgoing traffic

SMB request ingoing

Rule	Valid	Link	Service	from IP or net	to IP or net	Proto	Port	Comment
pass	quick	all	SMB	local.networks	napp-it ZFS server	tcp/udp	137,138,139,445	#n2 SMB file services loc
inactive	quick	ixgbe0	SMB	any	napp-it ZFS server	tcp/udp	137,138,139,445	#n3 SMB file services
inactive	quick	e1000g0	SMB	any	napp-it ZFS server	tcp/udp	137,138,139,445	#n4 SMB file services
inactive	quick	rz0	SMB	any	napp-it ZFS server	tcp/udp	137,138,139,445	#n5 SMB file services
inactive	quick	san0	SMB	any	napp-it ZFS server	tcp/udp	137,138,139,445	#n6 SMB file services
inactive	quick	lan0	SMB	any	napp-it ZFS server	tcp/udp	137,138,139,445	#n7 SMB file services

NFS request ingoing

Rule	Valid	Link	Service	from IP or net	to IP or net	Proto	Port	Comment
inactive	quick	all	NFS	any	napp-it ZFS server	tcp/udp	111,2049	#n8 NFS file services
pass	quick	ixgbe0	NFS	any	napp-it ZFS server	tcp/udp	111,2049	#n9 NFS file services
inactive	quick	e1000g0	NFS	any	napp-it ZFS server	tcp/udp	111,2049	#n10 NFS file services
inactive	quick	rz0	NFS	any	napp-it ZFS server	tcp/udp	111,2049	#n11 NFS file services
pass	quick	san0	NFS	any	napp-it ZFS server	tcp/udp	111,2049	#n12 NFS file services
block	quick	lan0	NFS	any	napp-it ZFS server	tcp/udp	111,2049	#n13 NFS file services

iSCSI request ingoing

You can use the napp-it Pro security panel to restrict access based on a set of ip adresses or local networks or based on a network adapter. With napp-it free, set the according rules manually.

19. Tuning

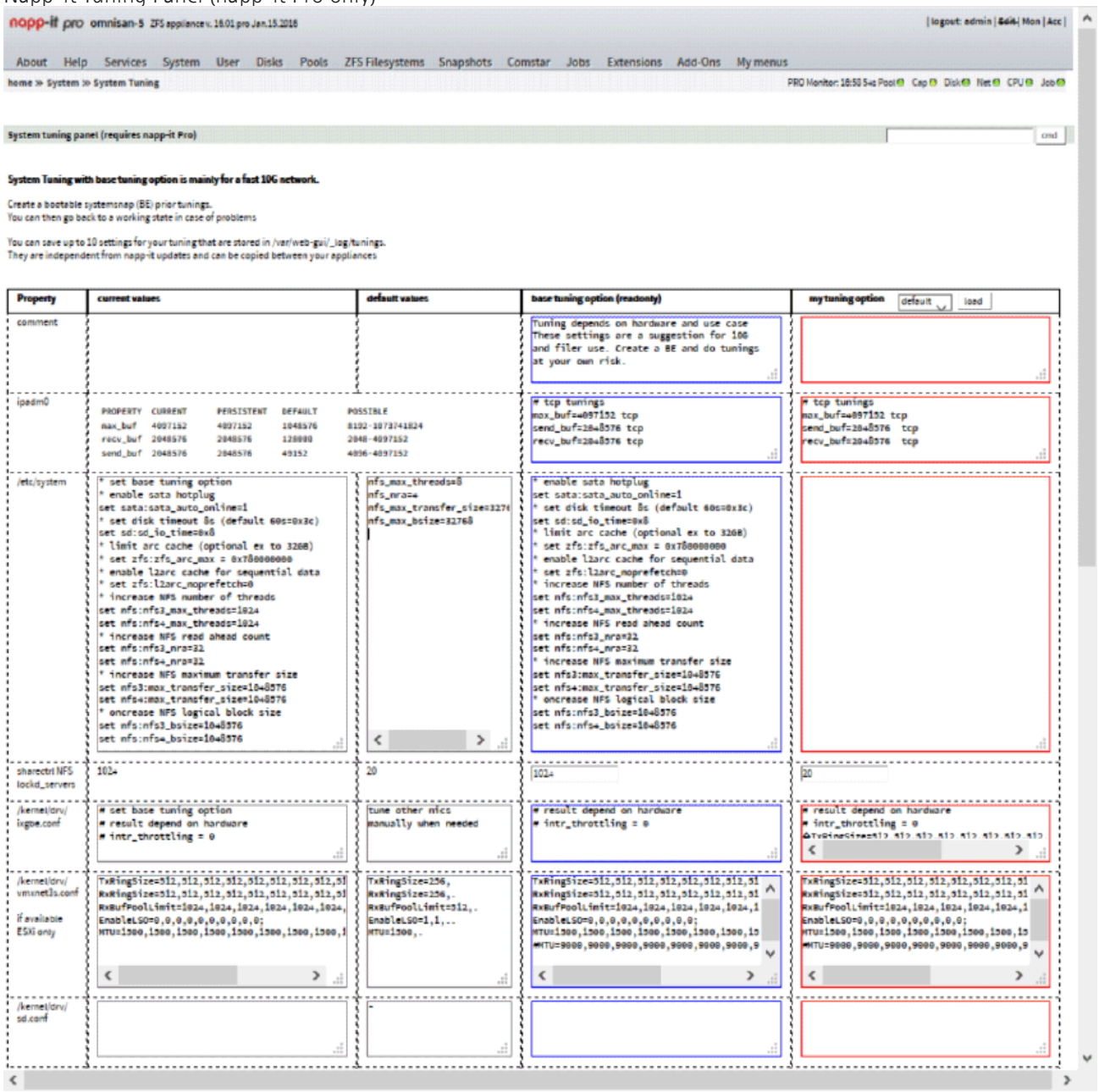
on Nic, System/IP or service level (iSCSI, NFS, SMB)

Storage performance depend on your physical network, your server and storage hardware and your settings that can be optimized to a special hardware environment or a special use case. The defaults are intended for 1Gb/s Ethernet networks and a typical filer use. Above that a Solaris based system offer a lot of tuning options.

When 1Gb/s networks (about 100MB/s) came up many years ago, this was sufficient as storage capacity was quite low and performance of an average internal disk and network performance was quite similar.

Now storage is in the multiple Terabyte area with a view to Petabyte. A simple backup or move of data from one pool to another over network can last days or weeks. Local storage with the new NVMe SSDs can give a performance of 1000MB/s or more. This is why we need faster networks like 10Gb/s Ethernet (about 1000 MB/s) and for this we need some tuning settings to achieve full performance on your hardware or for a workload..

Napp-it Tuning Panel (napp-it Pro only)



- TCP/IP tuning of Server side

This includes send and receive buffers that can be modified on a running system and mtu settings (Jumboframes, mtu=9000) that require a reboot.

- TCP/IP tuning on switches
you must enable mto=9000 or Jumboframes or the switch blocks Jumboframes

- TCP/IP on client side
This depend on your hardware. For examle with Windows and the mainstream 10G adapter Intel X540, you should disable interrupt throtteling.

- System and service tuning in /etc/system that require a reboot
like NFS buffers, hotplug behaviour of AHCI Sata, timeout of disks on problems and many other aspects.

- Nic (ex in ixgbe.conf) and vnic (vmxnet3s.conf) settings
These are mainly mtu and buffer settings that should be increased for 10G

- Disk settings in sd.conf
Mainly settings for advanced sector disks, power management and other aspects of disks

- SAS controller settings ex in mpt_sas.conf
Settings like mpio and other controller dependent SAS aspects

- napp-in-one tuning
If you use napp-in one to connect ESXi over the internal vswitch with the ZFS storage VM, all transfers are in software so most netwok and Ethernet centric tunings aspects are not needed, You should use vmxnet3 as vnic as it is much faster than the e1000 vnic with the base vmxnet3 tuning. The other tuning aspects are mainly relevant for external access or if you use a 10G nic in pass-through mode for fastest external access.

Remarks about napp-in-one tuning (ESXi related aspects)
You mainly use NFS to offer ZFS storage for ESXi. If you modify OmniOS settings like the vmxnet3 settings, it can happen that you need to reboot OmniOS twice (check console on boot as vmxnet3 settings are displayed on bootup). Some settings result in an „All Path Down=NFS not availabe„ error in ESXi. You can fix this with an ESXi reboot. Some NFS service modifications hinder ESXi to reconnect at all. You can fix this when you delete are VMs from inventory (NOT from disk), delete the NFS mount, readd the NFS mount and reimport the VMs to inventory (use ESXi filebrowser and a right mouse click on the .vmx file in the VM folder)

Tuning for napp-it Free
You can manually modify system settings via console, Putty or remotly edit files via WinSCP
You can try the settings from the napp-it Pro tuning console

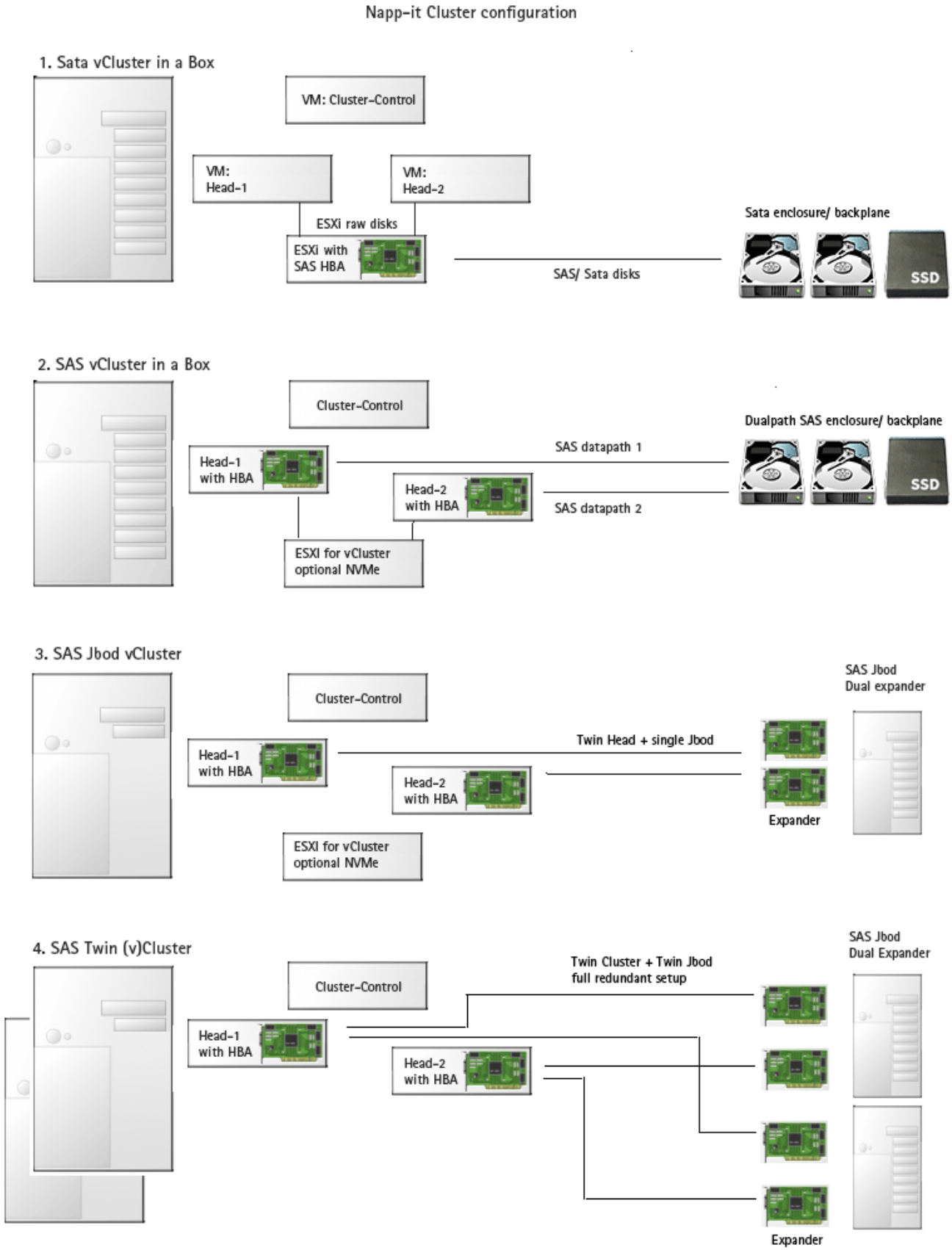
Tuning for napp-it Pro
You can use the tuning console that allows
- enable a basic tuning with the options for an immediate reboot and a BE creation
- use up to ten private editable tuning configurations that can be easily moved to other appliances

Hints about other aspects like SSD, iSCSI settings or Pool considerations see napp-it menu System > System tuning

20. ZFS vCluster in a Box

A napp-in-one setup can be extended to a vCluster with two nodes to provide services in a redundant HA setup with shared storage, either Sata via ESXi shared disk access or dualpath SAS.

A third VM Cluster-Control allows a web-managed Cluster with manual or auto failover



Good to know

Pass-through with Intel Optane
There are problems with Intel Optane NVMe pass-through
A workaround can be to add Optane 900P device ID (2700) to passthru.map

enable SSH on ESXi and login via WinSCP as user root

- edit /etc/vmware/passthru.map
- add following lines at the end of the file:

Intel Optane 900P
8086 2700 d3d0 false

- restart hypervisor

How to update ESXi ex ESXi 6.x7.x to 6.7 u3
https://www.thomas-krenn.com/de/wiki/VMware_ESXi_updaten

ESXi 6.7 u3 is available for download (search via Google, VMware site makes me crazy to search even as a paying user), https://docs.vmware.com/en/VMware-vSphere/6.7/rn/vsphere-esxi-67u3-release-notes.html

Main advantage for All-in- One:
ESXi 6.7u2 comes with a bug that hinders the deployment of my ova template. This is fixed now

download file from VMware: „update-from-esxi6.7-6.7_update03.zip"

create datastore1/updates (ESXi filebrowser)
upload zip

stop all VMs
enable esxi shell und ssh

switch to maintenance mode, connect via putty
per putty: esxcli software vib update -d /vmfs/volumes/datastore1/updates/ESXi670-201806001.zip

reboot
end maintenance mode

see
VMware ESXi updaten – Thomas-Krenn-Wiki