How to build a fileserver

Jeff Deifik jeff@deifik.com

About Me

- Using Unix since 1983, Linux since 1996
- Software for first e-commerce system (from 1985-1995)
- Software for the first orbiting radio telescope satellite
- Software for the most advanced pulse oximeter
- Cybersecurity for government satellite ground control, balancing sound cybersecurity with cost and schedule.
- Currently employed at The Aerospace Corp.

Overview

- Why bother
- Hardware considerations
- Networking
- Software

Why Bother

- You can buy a pre-built file server
  - Limited software functionality
    - DIY can use distro of choice and whatever sw you want
  - Limited support – when it is EOL you are SOL
  - Limited expandability
    - DIY can have whatever expandability you want
  - May or may not have ECC memory
  - More expensive

Hardware Considerations

- How much do you want to store
- How fast do you need to access data
- How large is your budget

Disk Options

- Rust drives
  - SATA – common, low power consumption
  - SAS – less expensive, more power consumption
- SSD
  - SATA common, a bit slow, expensive
  - M.2 common, faster than SATA, expensive
- RAID Level
  - RAID 5
  - ZFS

Disk Options

- Rust drives
  - SATA – 12tb $90 used 5 year warranty, $200 new
- SSD
  - SATA – 4tb $200+ new most QLC
  - M.2 – 4tb $200+ some TLC, 8tb - $800+
    - Requires $100 card to hold 4 and pci-e bifrication
    - $800 for card to hold 8 pci-e m.2 from Sonnet
- RAID Level
  - RAID 5 or perhaps RAID 6 when there are lots of drives

Disk Array Speed

- Using a raid-5 array I measured the following with bonnie++

Drives	Read	Write
four 8tb 5400rpm SATA	258 mb/sec	205 mb/sec
seven 4tb 7200rom SATA	443 mb/sec	245 mb/sec
four 10tb 7200rpm SATA	472 mb/sec	265 mb/sec
five 8tb 5400rpm SATA	453 mb/sec	383 mb/sec
four 12tb 7200rpm SAS	656 mb/sec	535 mb/sec

Case Options

- Depends on Hard Drive selection
- Do you want hot swap drives?
  - A bit more expensive, but easy to change drives
  - The Antec 900 has 9 front bays – room for 15 hot swap SATA drives via Supermicro 5-in-3 modules

Motherboard

- How much compute do you want?
- Do you want ECC? (a very good idea)
  - Without it, you can have a memory error and not know about it. Linux tends to use all DRAM as cache, so errors can corrupt data on disk
- Does the mb have enough disk drive ports?

My recommendations

- I have 4 file servers currently
- I want a minimum of 4 SATA/SAS drives
- Using one or two Supermicro 5-in-3 cages for hot swap drives (supports SATA/SAS)
- I am using 7*4tb, 5*8tb, 4*10tb, and 4x12tb drives
- I am using Supermicro X10SLM / X10SLL motherboards

My recommendations

- Supermicro X10 motherboards
  - Xeon E3-12xxv3 80w processors - cheap
  - DDR3 ECC buffered memory – up to 32gb
  - @6 sata II ports
  - Dual gigabit ethernet
  - IPMI
  - Cheap and reliable
  - Under $50 used, cpu $20 used

Power Supply

- 200-300w 80+ gold power supply would be ideal. Idle power typically @55w AC
- These are hard to come by as most power supplies are larger
- Power supply efficiency (except titanium min power 20%, so 400w PS needs @80w to get reasonable efficiency
- If you go with an OEM case, some have reasonably small power supplies

Networking

- Needs to be compatible with your infrastructure
- For rust drives 2.5g to 5g will be fine
  - You will need a 2.5g or 10g switch
  - 5g NIC $17 shipped from aliexpress
  - You will need 2.5/5g NICs in computers
- For SSD drives 10g+ is ideal
  - 10g+ fiber NICs are cheap and low power
  - 10g+ fiber switches are cheap
  - 10g-base-t is expensive or high power
  - 10b-base-t switches are expensive

Software

- Pick the distro of your choice
- If you want to use ZFS, Ubuntu is the distro that includes it, or use FreeBSD
- Mdadm is linux software RAID
- Samba is linux software for talking SMB (microsoft remote disk protocol)
- ssh server to talk to the file server

MDADM

- mdadm --create /dev/md1 --verbose --chunk=128 --level=5 --raid-devices=4 /dev/sd[bcde]
- mdadm --detail –scan
- mdadm --detail /dev/md1
- mkfs -V -t ext4 /dev/md1 -m0
- mkdir /data
- /etc/fstab - /dev/md1 /data ext4 acl,relatime,nofail 1 3
- mdadm --detail --scan --verbose >> /etc/mdadm.conf
- mount -a

Add disk to array

- Grow array
- Erase drive dd if=/dev/zero of=/dev/sdf
  - You can exit after a few seconds. The goal is to remove the partition table & metadata
- mdadm --add /dev/md1 /dev/sdf
- mdadm --grow /dev/md/alpha --raid-devices=5
- This will ‘reshape’ which can take @2 days
- mdadm --detail --scan --verbose >> /etc/mdadm.conf
- Check status with mdadm –detail /dev/md1

More adding disk to array

- umount /data
  - Unmount so you can manage the array
- e2fsck -f /dev/md1
  - Check the file system (-f means force)
- resize2fs -p /dev/md1
  - Resize the file system (-p shows progress)
- tune2fs -m 0 /dev/md/alpha
  - You probably don’t’ need reserve space, so set it to 0%

Samba

- smbpasswd -a <username>
- testparm /etc/samba/smb.conf
- /bin/systemctl restart smb.service

smb.conf

[global]

workgroup = WORKGROUP
max log size = 50
netbios name = Samba_XXX
log file = /var/log/samba/log.%m
Change XXX to One ot Two

hosts allow = localhost 192.168.47.
hosts deny = ALL # same as 0.0.0.0/0
encrypt passwords = yes

use sendfile = true
min receivefile size = 2048
aio write size = 16384
write raw = yes
max xmit = 32768
oplocks = yes
aio read size = 16384
read raw = yes
getwd cache = yes
large readwrite = yes

[data_<username>]
path = /data/j<username>
valid users = <username>
browsable = yes
write list = <username>