napp-it ZFS Storage principles for high performance SMB fileservices

Server:	Dual Xeon, 8/40 GB RAM Intel X540, mtu 9000 (JumboFrames)
Server OS:	Oracle Solaris 11.3 SMB 2.1
	OmniOS 151017 SMB 2.1
Client OS:	MacPro OSX 10.10/10.11 with Sanlink2
Client OS:	Win 8.1/ 10 Pro with Intel X540-T
Questions:	SMB1 vs SMB2 vs NFS mtu 1500 vs MTU 9000 NVMe vs SSD vs Disks Raid-0 vs Mirror vs RaidZ
	Single user vs multiple user Single pool vs multi pools
	Solaris 11.3 vs new OmniOS 8 GB RAM vs 40 GB RAM
	Dedupe and LZ4
	NVMe vs Enterprise SSD vs regular SSD
	Write behaviour of them vs filesize over time
	OS Tuning options
	RAM requirement

published: Mar 2017, Günther Alka

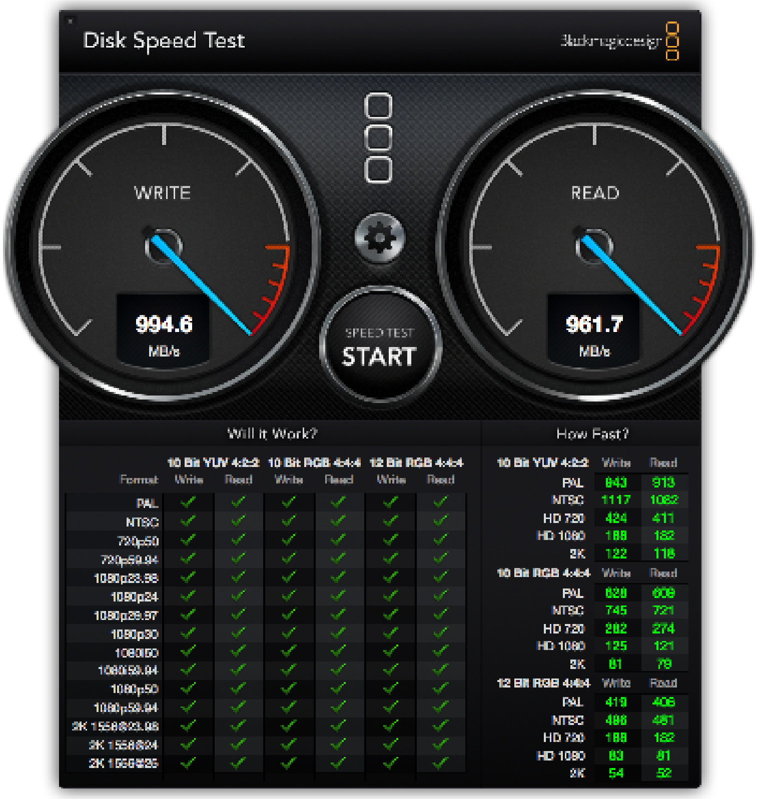
Client

Hardware:	Apple MacPro, OSX 10.10 4 Core 3,7 GhZ, 12 GB RAM Network: 10 Gbase-T Sanlink 2

Local disk performance on Apple NVMe
that is quite similar to an Samsung 950 Pro M2

Performance:

write	994 MB/s
read	961 MB/s



Will it Work?

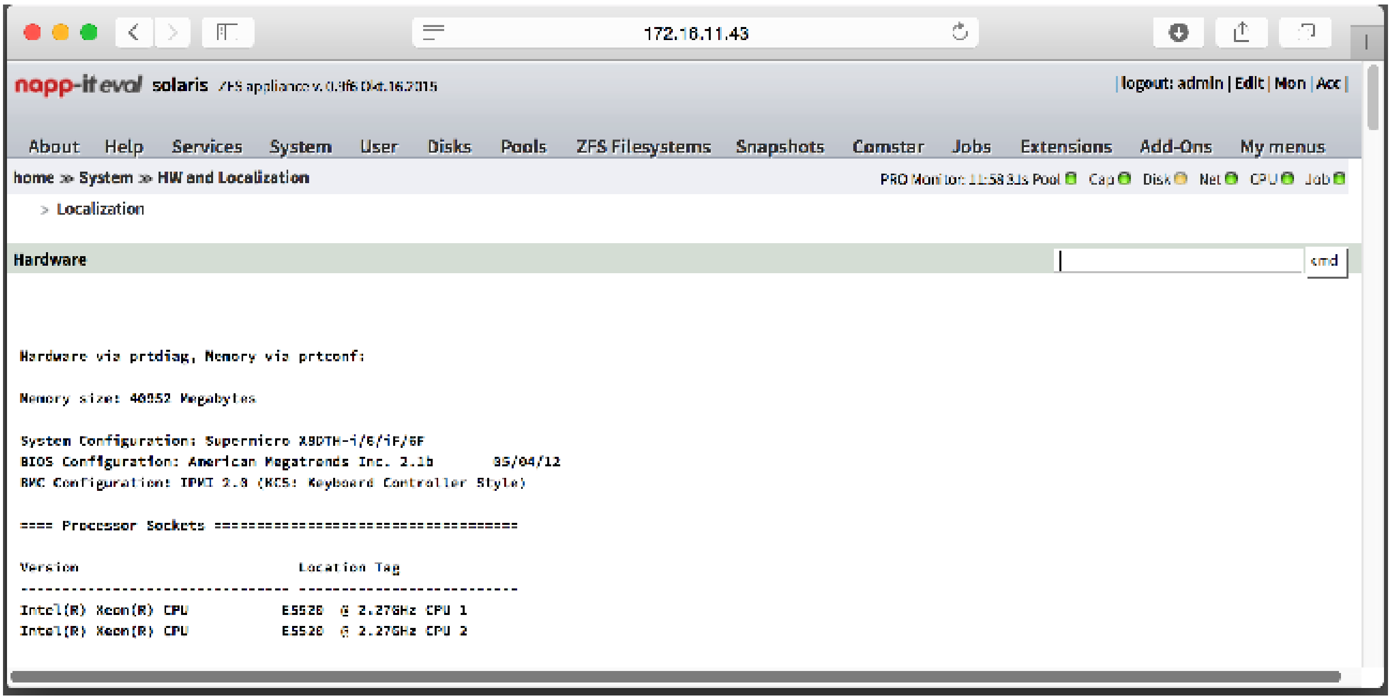
Format	10 Bit YUV 4:2:2 Write	10 Bit YUV 4:2:2 Read	10 Bit RGB 4:4:4 Write	10 Bit RGB 4:4:4 Read	12 Bit RGB 4:4:4 Write	12 Bit RGB 4:4:4 Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✓	✓	✓	✓	✓	✓
720p59.94	✓	✓	✓	✓	✓	✓
1080p23.98	✓	✓	✓	✓	✓	✓
1080p24	✓	✓	✓	✓	✓	✓
1080p29.97	✓	✓	✓	✓	✓	✓
1080p30	✓	✓	✓	✓	✓	✓
1080i50	✓	✓	✓	✓	✓	✓
1080i59.94	✓	✓	✓	✓	✓	✓
1080p50	✓	✓	✓	✓	✓	✓
1080p59.94	✓	✓	✓	✓	✓	✓
2K 1556p23.98	✓	✓	✓	✓	✓	✓
2K 1556p24	✓	✓	✓	✓	✓	✓
2K 1556p25	✓	✓	✓	✓	✓	✓

How Fast?

10 Bit YUV 4:2:2	Write	Read
PAL	843	913
NTSC	1117	1082
HD 720	424	411
HD 1080	188	182
2K	122	118
10 Bit RGB 4:4:4	**Write**	**Read**
PAL	628	609
NTSC	745	721
HD 720	282	274
HD 1080	125	121
2K	81	79
12 Bit RGB 4:4:4	**Write**	**Read**
PAL	419	406
NTSC	488	481
HD 720	188	182
HD 1080	83	81
2K	54	52

Switch
HP 5900, 48 x 10 Gbase-T

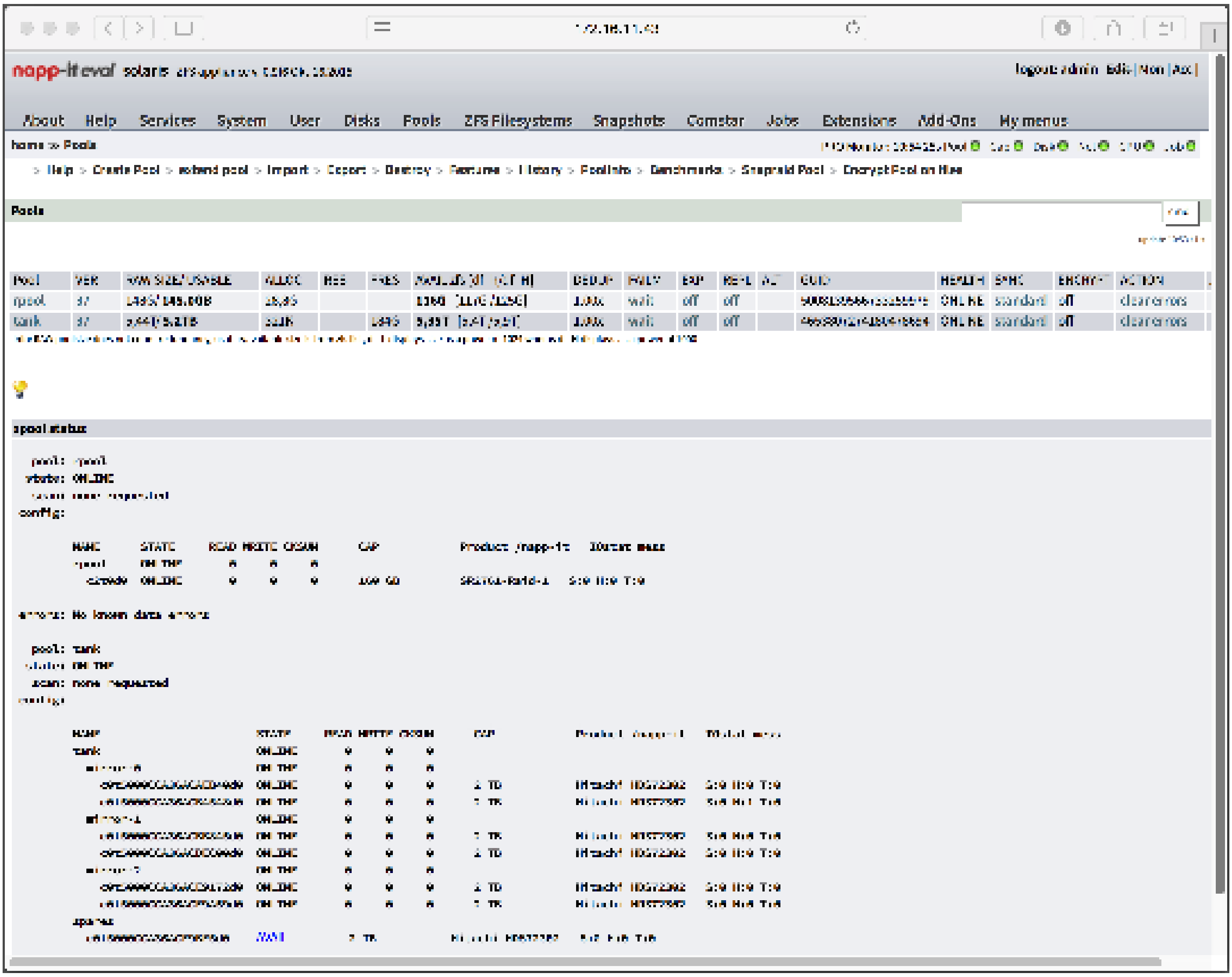
Storage
Dual Xeon, 40 GB RAM. Intel X540 10Gbase-T and Jumboframes enabled on all tests


172.16.11.43
napp-it eval solaris ZFS appliance v. 0.9f6 Oct.16.2015
| logout: admin | Edit | Mon | Acc |
About Help Services System User Disks Pools ZFS Filesystems Snapshots Comstar Jobs Extensions Add-Ons My menus
home >> System >> HW and Localization
PRO Monitor: 11:58 31s Pool Cap Disk Net CPU Job
> Localization
Hardware
cmd

```
Hardware via prtdiag, Memory via prtconf:

Memory size: 40952 Megabytes

System Configuration: Supermicro X8DTH-i/6/iF/6F
BIOS Configuration: American Megatrends Inc. 2.1b       05/04/12
BMC Configuration: IPMI 2.0 (KCS: Keyboard Controller Style)

==== Processor Sockets ====================================

Version                          Location Tag
-------------------------------- --------------------------
Intel(R) Xeon(R) CPU           E5520  @ 2.27GHz CPU 1
Intel(R) Xeon(R) CPU           E5520  @ 2.27GHz CPU 2
```

Server OS (Round 1)
Oracle Solaris 11.3 as it offers SMB 2.1.

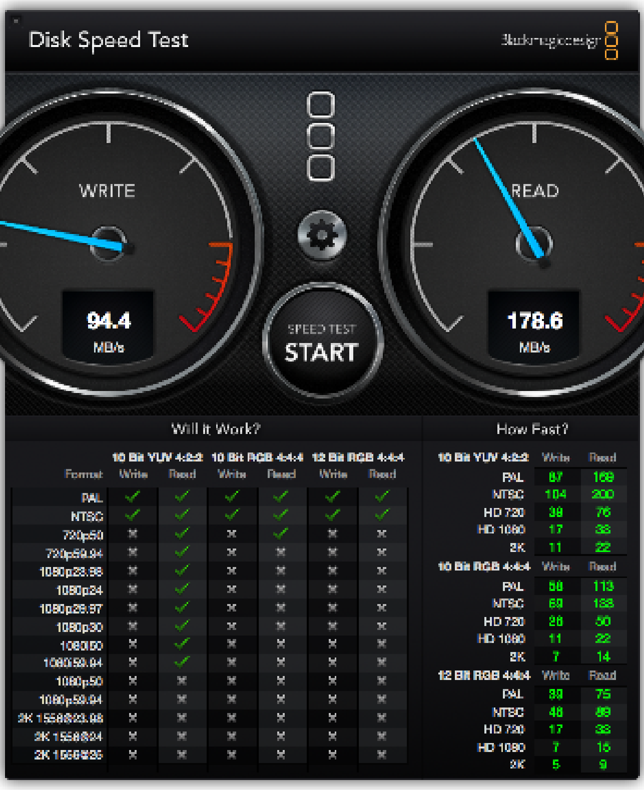
SMB1, 10 GbE (Solaris)
Pool Layout: 3 x Mirror of 2 TB HGST 7k U/m


napp-it eval solaris
logout: admin Edit | Mon | Acc |
About Help Services System User Disks Pools ZFS Filesystems Snapshots Comstar Jobs Extensions Add-Ons My menus
home > Pools
> Help > Create Pool > extend pool > Import > Export > Destroy > Features > History > PoolInfo > Benchmarks > Snapraid Pool > Encrypt Pool on files
Pools

Pool	VER	RAW SIZE/USABLE	ALLOC	RES	FRES	AVAILABLE zfs [df -h]	DEDUP	FAILM	EXP	REPL	ALT	GUID	HEALTH	SYNC	ENCRYPT	ACTION
rpool	37	149G/145.0GB	16.6G			128G [117G/125G]	1.00x	wait	off	off		[illegible]	ONLINE	standard	off	clear errors
tank	37	5.44T/5.2TB	221K		[illegible]	5.35T [5.4T/5.5T]	1.00x	wait	off	off		[illegible]	ONLINE	standard	off	clear errors

zpool status

```
   pool: rpool
  state: ONLINE
   scan: none requested
 config:

        NAME        STATE     READ WRITE CKSUM      CAP            Product /napp-it   IOstat mess
        rpool       ONLINE       0     0     0
          c2t0d0    ONLINE       0     0     0      160 GB         SR2760-Raid-1   S:0 H:0 T:0

errors: No known data errors

   pool: tank
  state: ONLINE
   scan: none requested
 config:

        NAME                      STATE      READ WRITE CKSUM      CAP            Product /napp-it   IOstat mess
        tank                      ONLINE        0     0     0
          mirror-0                ONLINE        0     0     0
            [illegible]           ONLINE        0     0     0      2 TB           Hitachi HDS72202   S:0 H:0 T:0
            [illegible]           ONLINE        0     0     0      2 TB           Hitachi HDS72202   S:0 H:0 T:0
          mirror-1                ONLINE        0     0     0
            [illegible]           ONLINE        0     0     0      2 TB           Hitachi HDS72202   S:0 H:0 T:0
            [illegible]           ONLINE        0     0     0      2 TB           Hitachi HDS72202   S:0 H:0 T:0
          mirror-2                ONLINE        0     0     0
            [illegible]           ONLINE        0     0     0      2 TB           Hitachi HDS72202   S:0 H:0 T:0
            [illegible]           ONLINE        0     0     0      2 TB           Hitachi HDS72202   S:0 H:0 T:0
        spares
          [illegible]             AVAIL           2 TB          Hitachi HDS72202   S:0 H:0 T:0
```



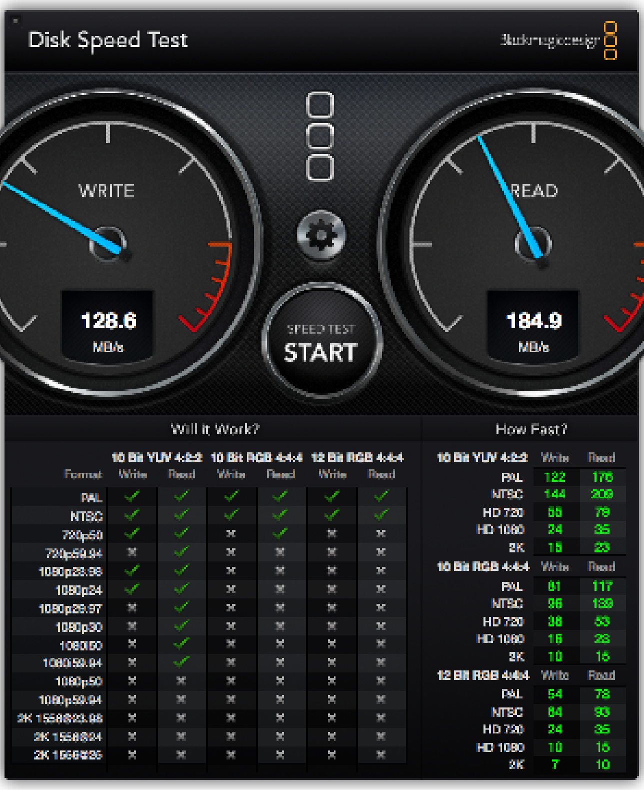
Will it Work?

	10 Bit YUV 4:2:2		10 Bit RGB 4:4:4		12 Bit RGB 4:4:4	
Format	Write	Read	Write	Read	Write	Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✗	✓	✗	✓	✗	✗
720p59.94	✗	✓	✗	✗	✗	✗
1080p23.98	✗	✓	✗	✗	✗	✗
1080p24	✗	✓	✗	✗	✗	✗
1080p29.97	✗	✓	✗	✗	✗	✗
1080p30	✗	✓	✗	✗	✗	✗
1080i50	✗	✓	✗	✗	✗	✗
1080i59.94	✗	✓	✗	✗	✗	✗
1080p50	✗	✗	✗	✗	✗	✗
1080p59.94	✗	✗	✗	✗	✗	✗
2K 1556p23.98	✗	✗	✗	✗	✗	✗
2K 1556p24	✗	✗	✗	✗	✗	✗
2K 1556p25	✗	✗	✗	✗	✗	✗

How Fast?

10 Bit YUV 4:2:2	Write	Read
PAL	87	169
NTSC	104	200
HD 720	39	76
HD 1080	17	33
2K	11	22
10 Bit RGB 4:4:4	**Write**	**Read**
PAL	58	113
NTSC	69	133
HD 720	26	50
HD 1080	11	22
2K	7	14
12 Bit RGB 4:4:4	**Write**	**Read**
PAL	39	75
NTSC	46	89
HD 720	17	33
HD 1080	7	15
2K	5	9

SMB1, MTU 1500, 10 GbE (connect to cifs://)



Will it Work?

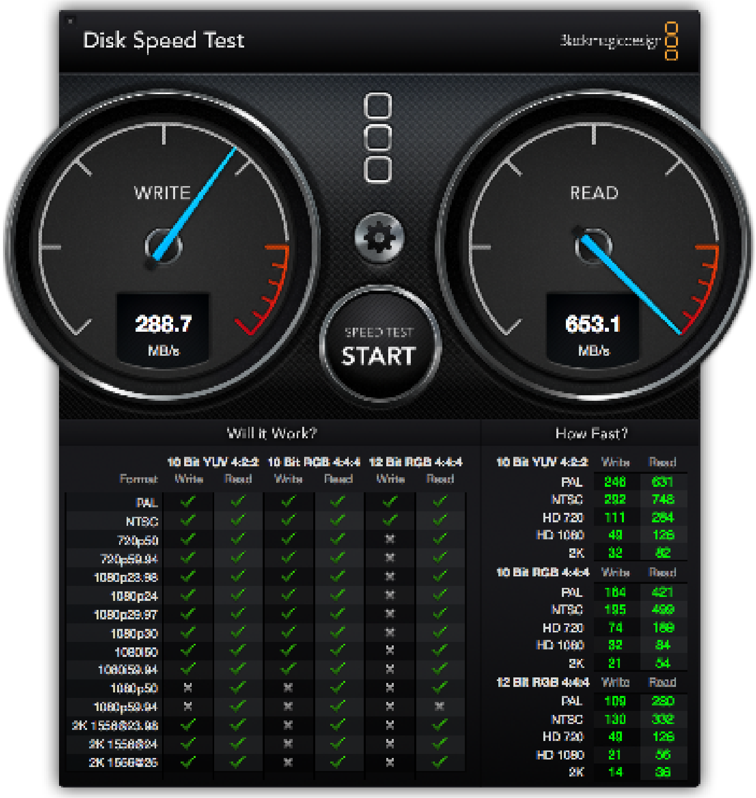
	10 Bit YUV 4:2:2		10 Bit RGB 4:4:4		12 Bit RGB 4:4:4	
Format	Write	Read	Write	Read	Write	Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✓	✓	✗	✓	✗	✗
720p59.94	✗	✓	✗	✗	✗	✗
1080p23.98	✓	✓	✗	✗	✗	✗
1080p24	✓	✓	✗	✗	✗	✗
1080p29.97	✗	✓	✗	✗	✗	✗
1080p30	✗	✓	✗	✗	✗	✗
1080i50	✗	✓	✗	✗	✗	✗
1080i59.94	✗	✓	✗	✗	✗	✗
1080p50	✗	✗	✗	✗	✗	✗
1080p59.94	✗	✗	✗	✗	✗	✗
2K 1556p23.98	✗	✗	✗	✗	✗	✗
2K 1556p24	✗	✗	✗	✗	✗	✗
2K 1556p25	✗	✗	✗	✗	✗	✗

How Fast?

10 Bit YUV 4:2:2	Write	Read
PAL	122	176
NTSC	144	209
HD 720	55	79
HD 1080	24	35
2K	15	23
10 Bit RGB 4:4:4	**Write**	**Read**
PAL	81	117
NTSC	96	139
HD 720	36	53
HD 1080	16	23
2K	10	15
12 Bit RGB 4:4:4	**Write**	**Read**
PAL	54	78
NTSC	64	93
HD 720	24	35
HD 1080	10	15
2K	7	10

SMB 1, MTU 9000, 10 GbE (connect to cifs://)

SMB 2.1, 10 GbE (Solaris)
Pool Layout: 3 x Mirror of 2 TB HGST 7k U/m



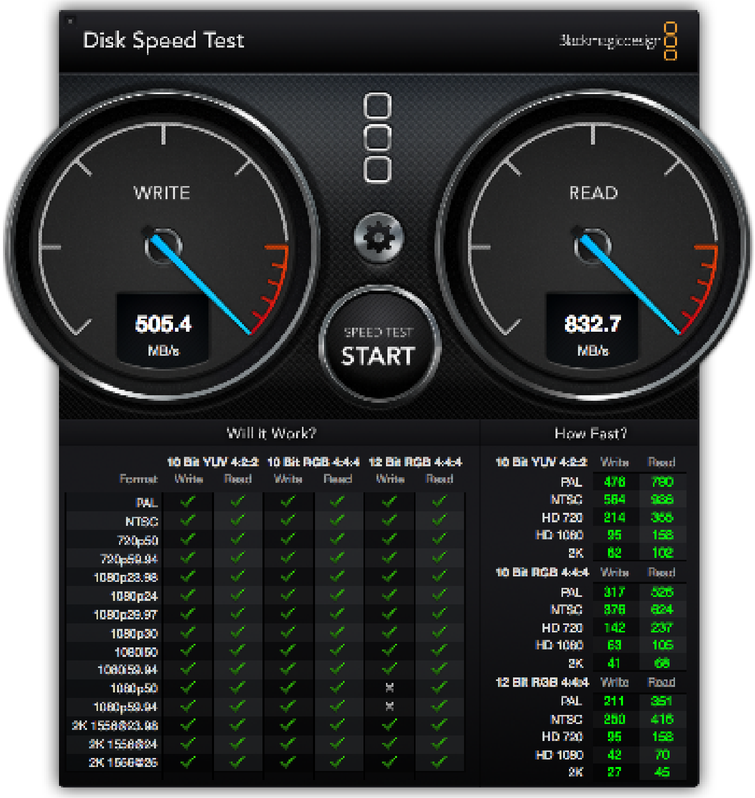
Will it Work?

	10 Bit YUV 4:2:2		10 Bit RGB 4:4:4		12 Bit RGB 4:4:4	
Format	Write	Read	Write	Read	Write	Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✓	✓	✓	✓	✗	✓
720p59.94	✓	✓	✓	✓	✗	✓
1080p23.98	✓	✓	✓	✓	✗	✓
1080p24	✓	✓	✓	✓	✗	✓
1080p29.97	✓	✓	✓	✓	✗	✓
1080p30	✓	✓	✓	✓	✗	✓
1080i50	✓	✓	✓	✓	✗	✓
1080i59.94	✓	✓	✓	✓	✗	✓
1080p50	✗	✓	✗	✓	✗	✓
1080p59.94	✗	✓	✗	✓	✗	✗
2K 1556p23.98	✓	✓	✗	✓	✗	✓
2K 1556p24	✓	✓	✗	✓	✗	✓
2K 1556p25	✓	✓	✗	✓	✗	✓

How Fast?

10 Bit YUV 4:2:2	Write	Read
PAL	248	631
NTSC	292	748
HD 720	111	284
HD 1080	49	126
2K	32	82
10 Bit RGB 4:4:4	**Write**	**Read**
PAL	164	421
NTSC	195	499
HD 720	74	189
HD 1080	32	84
2K	21	54
12 Bit RGB 4:4:4	**Write**	**Read**
PAL	109	280
NTSC	130	332
HD 720	49	126
HD 1080	21	56
2K	14	38

SMB 2.1, MTU 1500, 10 GbE (connect to smb://)



Will it Work?

	10 Bit YUV 4:2:2		10 Bit RGB 4:4:4		12 Bit RGB 4:4:4	
Format	Write	Read	Write	Read	Write	Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✓	✓	✓	✓	✓	✓
720p59.94	✓	✓	✓	✓	✓	✓
1080p23.98	✓	✓	✓	✓	✓	✓
1080p24	✓	✓	✓	✓	✓	✓
1080p29.97	✓	✓	✓	✓	✓	✓
1080p30	✓	✓	✓	✓	✓	✓
1080i50	✓	✓	✓	✓	✓	✓
1080i59.94	✓	✓	✓	✓	✓	✓
1080p50	✓	✓	✓	✓	✗	✓
1080p59.94	✓	✓	✓	✓	✗	✓
2K 1556p23.98	✓	✓	✓	✓	✓	✓
2K 1556p24	✓	✓	✓	✓	✓	✓
2K 1556p25	✓	✓	✓	✓	✓	✓

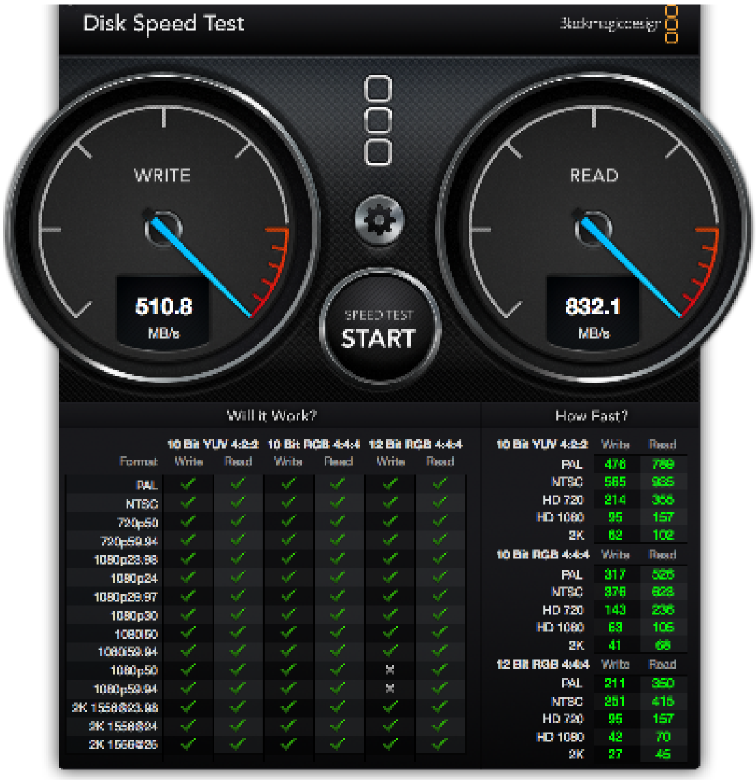
How Fast?

10 Bit YUV 4:2:2	Write	Read
PAL	476	790
NTSC	564	938
HD 720	214	356
HD 1080	95	158
2K	62	102
10 Bit RGB 4:4:4	**Write**	**Read**
PAL	317	526
NTSC	376	624
HD 720	142	237
HD 1080	63	105
2K	41	68
12 Bit RGB 4:4:4	**Write**	**Read**
PAL	211	351
NTSC	250	416
HD 720	95	158
HD 1080	42	70
2K	27	45

SMB 2.1, MTU 9000, 10 GbE (connect to smb://)

SMB 2.1 gives a boost over SMB1 and a second boost when using JumboFrames

SMB 2.1, 10 GbE
Pool Layout: 7 x Raid-0 of 2 TB HGST 7k U/m



Will it Work?

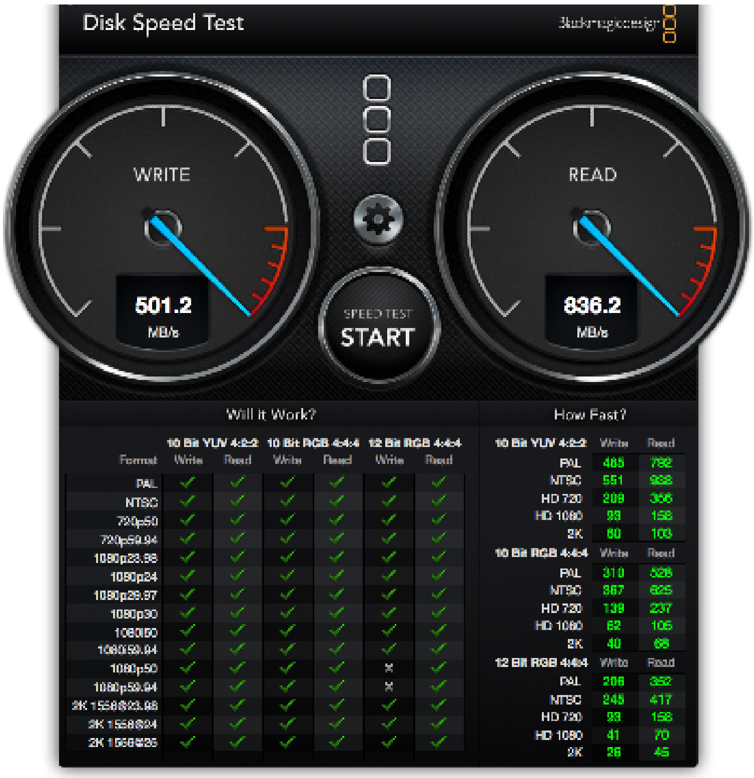
	10 Bit YUV 4:2:2		10 Bit RGB 4:4:4		12 Bit RGB 4:4:4	
Format	Write	Read	Write	Read	Write	Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✓	✓	✓	✓	✓	✓
720p59.94	✓	✓	✓	✓	✓	✓
1080p23.98	✓	✓	✓	✓	✓	✓
1080p24	✓	✓	✓	✓	✓	✓
1080p29.97	✓	✓	✓	✓	✓	✓
1080p30	✓	✓	✓	✓	✓	✓
1080i50	✓	✓	✓	✓	✓	✓
1080i59.94	✓	✓	✓	✓	✓	✓
1080p50	✓	✓	✓	✓	✗	✓
1080p59.94	✓	✓	✓	✓	✗	✓
2K 1556p23.98	✓	✓	✓	✓	✓	✓
2K 1556p24	✓	✓	✓	✓	✓	✓
2K 1556p25	✓	✓	✓	✓	✓	✓

How Fast?

10 Bit YUV 4:2:2	Write	Read
PAL	476	789
NTSC	565	935
HD 720	214	356
HD 1080	95	157
2K	62	102
10 Bit RGB 4:4:4	**Write**	**Read**
PAL	317	526
NTSC	378	628
HD 720	143	236
HD 1080	63	106
2K	41	68
12 Bit RGB 4:4:4	**Write**	**Read**
PAL	211	350
NTSC	251	416
HD 720	95	157
HD 1080	42	70
2K	27	45

SMB 2.1, MTU 9000, 7 x HGST 2TB raid-0

Pool Layout: 10 x Raid-Z2 of 2 TB HGST 7k U/m



Will it Work?

	10 Bit YUV 4:2:2		10 Bit RGB 4:4:4		12 Bit RGB 4:4:4	
Format	Write	Read	Write	Read	Write	Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✓	✓	✓	✓	✓	✓
720p59.94	✓	✓	✓	✓	✓	✓
1080p23.98	✓	✓	✓	✓	✓	✓
1080p24	✓	✓	✓	✓	✓	✓
1080p29.97	✓	✓	✓	✓	✓	✓
1080p30	✓	✓	✓	✓	✓	✓
1080i50	✓	✓	✓	✓	✓	✓
1080i59.94	✓	✓	✓	✓	✓	✓
1080p50	✓	✓	✓	✓	✗	✓
1080p59.94	✓	✓	✓	✓	✗	✓
2K 1556p23.98	✓	✓	✓	✓	✓	✓
2K 1556p24	✓	✓	✓	✓	✓	✓
2K 1556p25	✓	✓	✓	✓	✓	✓

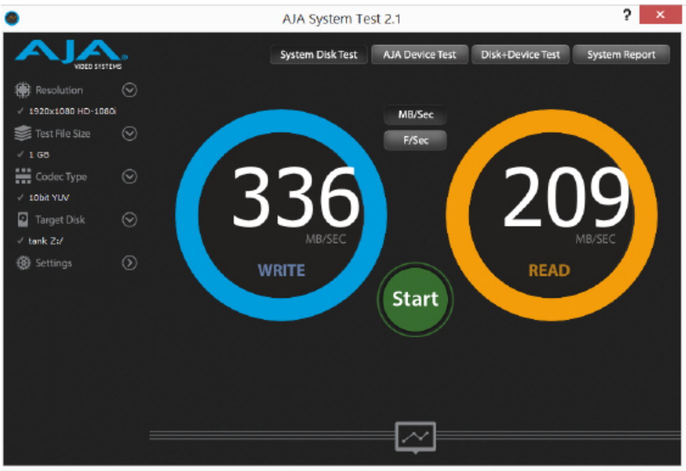
How Fast?

10 Bit YUV 4:2:2	Write	Read
PAL	485	792
NTSC	551	938
HD 720	209	356
HD 1080	93	158
2K	60	103
10 Bit RGB 4:4:4	**Write**	**Read**
PAL	310	526
NTSC	367	625
HD 720	139	237
HD 1080	62	105
2K	40	68
12 Bit RGB 4:4:4	**Write**	**Read**
PAL	206	352
NTSC	245	417
HD 720	93	158
HD 1080	41	70
2K	26	45

SMB 2.1, MTU 9000, 10 x HGST 2TB
Raid Z2, single vdev

There is no advantage of the raid-0 with a much higher iops rate (7 x a single disk vs 1 x a single disk)

Pool Layout: 10 x Raid-Z2 of 2 TB HGST 7k U/m (Solaris)
Windows 8.1 pro, 10 GbE, MTU 1500

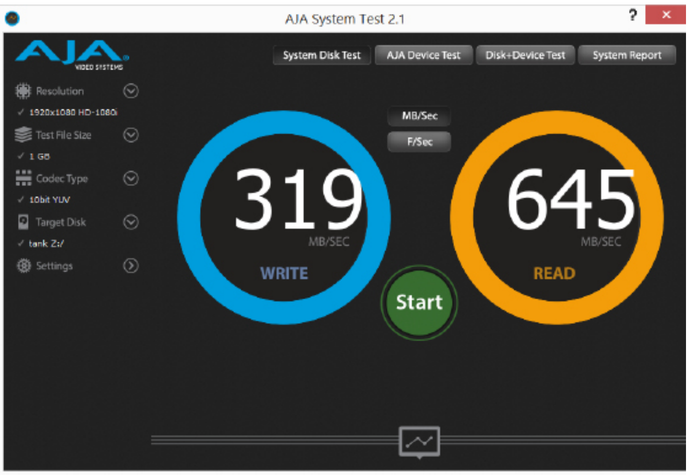


Resolution
✓ 1920x1080 HD-1080i
Test File Size
✓ 1 GB
Codec Type
✓ 10bit YUV
Target Disk
✓ tank Z:/
Settings

336 MB/SEC WRITE

209 MB/SEC READ

Windows 8.1 pro, 10 GbE, MTU 9000



Resolution
✓ 1920x1080 HD-1080i
Test File Size
✓ 1 GB
Codec Type
✓ 10bit YUV
Target Disk
✓ tank Z:/
Settings

319 MB/SEC WRITE

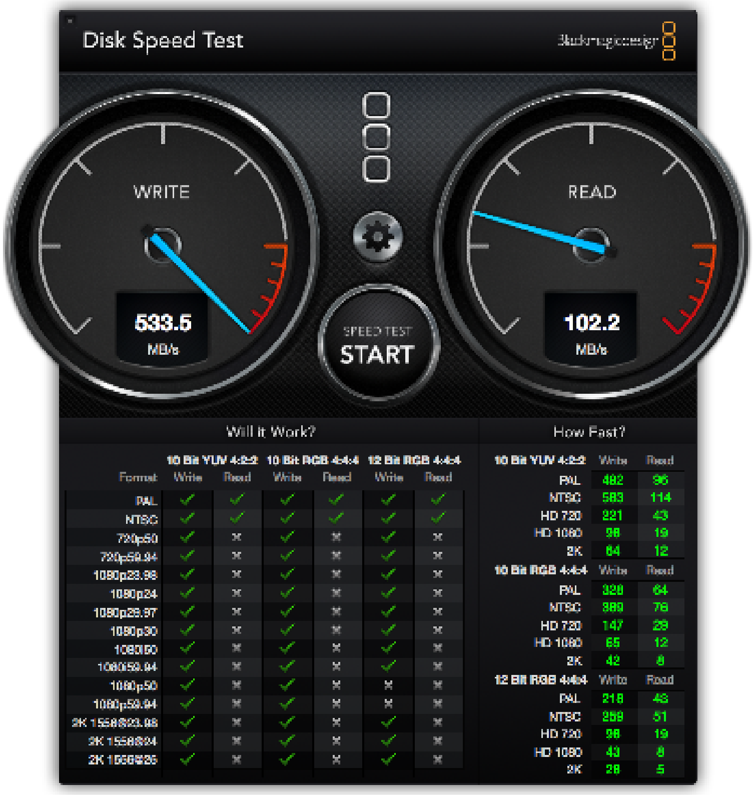
645 MB/SEC READ

On Windows, MTU 1500 is weak on reads

With MTU 9000, read values are much higher

Hardwarde is a 6 core Xeon with 32 GB RAM, an Intel X540-T1 and a local Sandisk Extreme Pro SSD

Pool Layout: 10 x Raid-Z2 of 2 TB HGST 7k U/m
OSX 10.9, MTU 1500



WRITE 533.5 MB/s

READ 102.2 MB/s

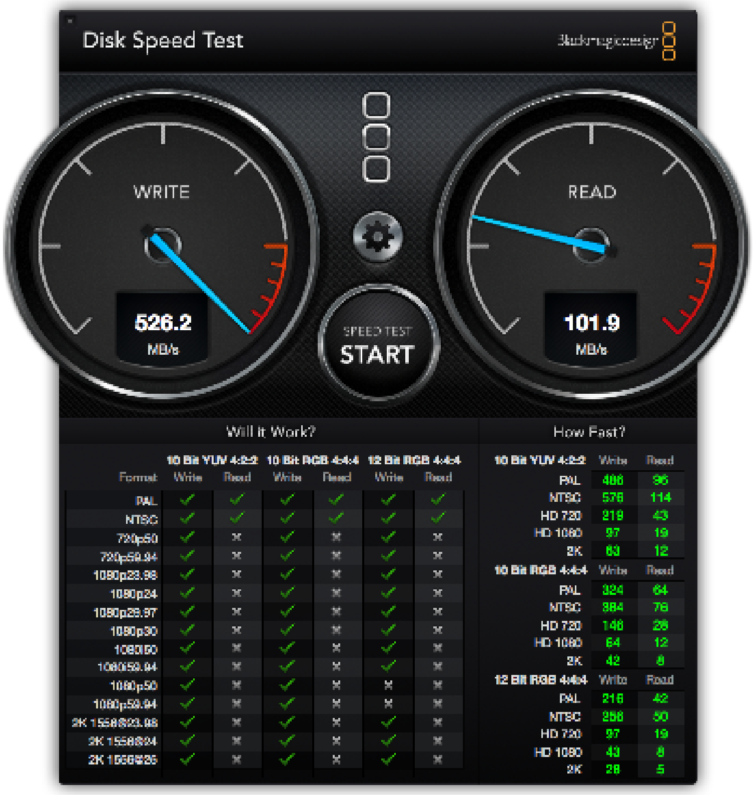
Will it Work?

Format	10 Bit YUV 4:2:2 Write	10 Bit YUV 4:2:2 Read	10 Bit RGB 4:4:4 Write	10 Bit RGB 4:4:4 Read	12 Bit RGB 4:4:4 Write	12 Bit RGB 4:4:4 Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✓	✗	✓	✗	✓	✗
720p59.94	✓	✗	✓	✗	✓	✗
1080p23.98	✓	✗	✓	✗	✓	✗
1080p24	✓	✗	✓	✗	✓	✗
1080p29.97	✓	✗	✓	✗	✓	✗
1080p30	✓	✗	✓	✗	✓	✗
1080i50	✓	✗	✓	✗	✓	✗
1080i59.94	✓	✗	✓	✗	✓	✗
1080p50	✓	✗	✓	✗	✗	✗
1080p59.94	✓	✗	✓	✗	✗	✗
2K 1556@23.98	✓	✗	✓	✗	✓	✗
2K 1556@24	✓	✗	✓	✗	✓	✗
2K 1556@25	✓	✗	✓	✗	✓	✗

How Fast?

10 Bit YUV 4:2:2	Write	Read
PAL	482	96
NTSC	583	114
HD 720	221	43
HD 1080	98	19
2K	64	12
10 Bit RGB 4:4:4	**Write**	**Read**
PAL	320	64
NTSC	389	78
HD 720	147	29
HD 1080	65	12
2K	42	8
12 Bit RGB 4:4:4	**Write**	**Read**
PAL	218	43
NTSC	259	51
HD 720	98	19
HD 1080	43	8
2K	28	5

OSX 10.9, MTU 9000



WRITE 526.2 MB/s

READ 101.9 MB/s

Will it Work?

Format	10 Bit YUV 4:2:2 Write	10 Bit YUV 4:2:2 Read	10 Bit RGB 4:4:4 Write	10 Bit RGB 4:4:4 Read	12 Bit RGB 4:4:4 Write	12 Bit RGB 4:4:4 Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✓	✗	✓	✗	✓	✗
720p59.94	✓	✗	✓	✗	✓	✗
1080p23.98	✓	✗	✓	✗	✓	✗
1080p24	✓	✗	✓	✗	✓	✗
1080p29.97	✓	✗	✓	✗	✓	✗
1080p30	✓	✗	✓	✗	✓	✗
1080i50	✓	✗	✓	✗	✓	✗
1080i59.94	✓	✗	✓	✗	✓	✗
1080p50	✓	✗	✓	✗	✗	✗
1080p59.94	✓	✗	✓	✗	✗	✗
2K 1556@23.98	✓	✗	✓	✗	✓	✗
2K 1556@24	✓	✗	✓	✗	✓	✗
2K 1556@25	✓	✗	✓	✗	✓	✗

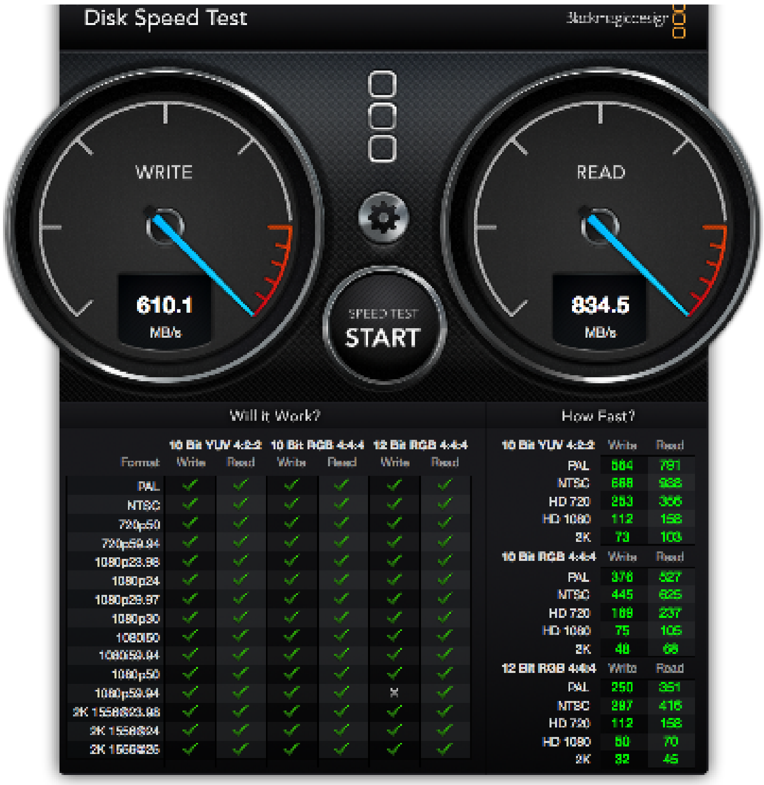
How Fast?

10 Bit YUV 4:2:2	Write	Read
PAL	486	96
NTSC	576	114
HD 720	219	43
HD 1080	97	19
2K	63	12
10 Bit RGB 4:4:4	**Write**	**Read**
PAL	324	64
NTSC	384	78
HD 720	146	28
HD 1080	64	12
2K	42	8
12 Bit RGB 4:4:4	**Write**	**Read**
PAL	216	42
NTSC	256	50
HD 720	97	19
HD 1080	43	8
2K	28	5

There seems to be a problem with OSX 10.9 and MTU 9000
as reads are low even with MTU 9000

I made an update from 10.9 to 10.11 but results were the same.
A clean install of OSX 10.11 fixes this problem.

SMB 2.1, 10 GbE, MTU 9000 (Solaris)
Pool Layout: Single NVMe Intel P750



Will it Work?

	10 Bit YUV 4:2:2		10 Bit RGB 4:4:4		12 Bit RGB 4:4:4	
Format	Write	Read	Write	Read	Write	Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✓	✓	✓	✓	✓	✓
720p59.94	✓	✓	✓	✓	✓	✓
1080p23.98	✓	✓	✓	✓	✓	✓
1080p24	✓	✓	✓	✓	✓	✓
1080p29.97	✓	✓	✓	✓	✓	✓
1080p30	✓	✓	✓	✓	✓	✓
1080i50	✓	✓	✓	✓	✓	✓
1080i59.94	✓	✓	✓	✓	✓	✓
1080p50	✓	✓	✓	✓	✓	✓
1080p59.94	✓	✓	✓	✓	✗	✓
2K 1556p23.98	✓	✓	✓	✓	✓	✓
2K 1556p24	✓	✓	✓	✓	✓	✓
2K 1556p25	✓	✓	✓	✓	✓	✓

How Fast?

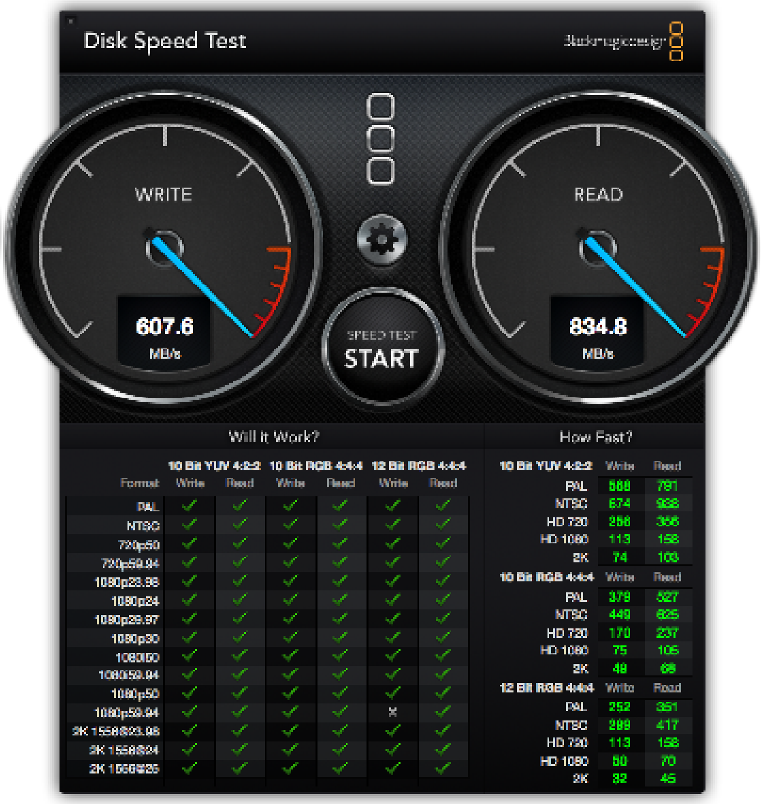
10 Bit YUV 4:2:2	Write	Read
PAL	584	791
NTSC	688	938
HD 720	253	356
HD 1080	112	158
2K	73	103

10 Bit RGB 4:4:4	Write	Read
PAL	376	527
NTSC	445	625
HD 720	189	237
HD 1080	75	105
2K	40	68

12 Bit RGB 4:4:4	Write	Read
PAL	250	351
NTSC	287	416
HD 720	112	158
HD 1080	50	70
2K	32	45

SMB 2.1, MTU 9000,
1 x NVMe Intel p750-400

Pool Layout: 3 x Intel S3610-1,6TB, Raid-0



Will it Work?

	10 Bit YUV 4:2:2		10 Bit RGB 4:4:4		12 Bit RGB 4:4:4	
Format	Write	Read	Write	Read	Write	Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✓	✓	✓	✓	✓	✓
720p59.94	✓	✓	✓	✓	✓	✓
1080p23.98	✓	✓	✓	✓	✓	✓
1080p24	✓	✓	✓	✓	✓	✓
1080p29.97	✓	✓	✓	✓	✓	✓
1080p30	✓	✓	✓	✓	✓	✓
1080i50	✓	✓	✓	✓	✓	✓
1080i59.94	✓	✓	✓	✓	✓	✓
1080p50	✓	✓	✓	✓	✓	✓
1080p59.94	✓	✓	✓	✓	✗	✓
2K 1556p23.98	✓	✓	✓	✓	✓	✓
2K 1556p24	✓	✓	✓	✓	✓	✓
2K 1556p25	✓	✓	✓	✓	✓	✓

How Fast?

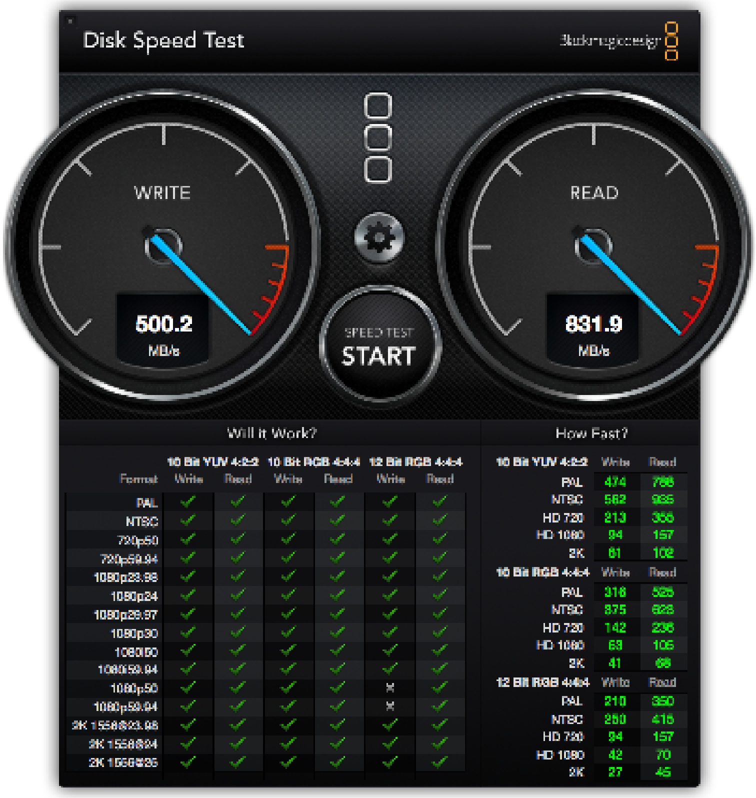
10 Bit YUV 4:2:2	Write	Read
PAL	588	791
NTSC	674	938
HD 720	256	356
HD 1080	113	158
2K	74	103

10 Bit RGB 4:4:4	Write	Read
PAL	379	527
NTSC	449	625
HD 720	170	237
HD 1080	75	106
2K	48	68

12 Bit RGB 4:4:4	Write	Read
PAL	252	351
NTSC	289	417
HD 720	113	158
HD 1080	50	70
2K	32	45

SMB 2.1, MTU 9000,
3 x Intel P3610- 1,6 TB Raid 0

SMB 2.1, 10 GbE, MTU 9000
Single User on 10 disk Raid Z2 (HGST 2TB)



Will it Work?

	10 Bit YUV 4:2:2		10 Bit RGB 4:4:4		12 Bit RGB 4:4:4	
Format	Write	Read	Write	Read	Write	Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✓	✓	✓	✓	✓	✓
720p59.94	✓	✓	✓	✓	✓	✓
1080p23.98	✓	✓	✓	✓	✓	✓
1080p24	✓	✓	✓	✓	✓	✓
1080p29.97	✓	✓	✓	✓	✓	✓
1080p30	✓	✓	✓	✓	✓	✓
1080i50	✓	✓	✓	✓	✓	✓
1080i59.94	✓	✓	✓	✓	✓	✓
1080p50	✓	✓	✓	✓	✗	✓
1080p59.94	✓	✓	✓	✓	✗	✓
2K 1556p23.98	✓	✓	✓	✓	✓	✓
2K 1556p24	✓	✓	✓	✓	✓	✓
2K 1556p25	✓	✓	✓	✓	✓	✓

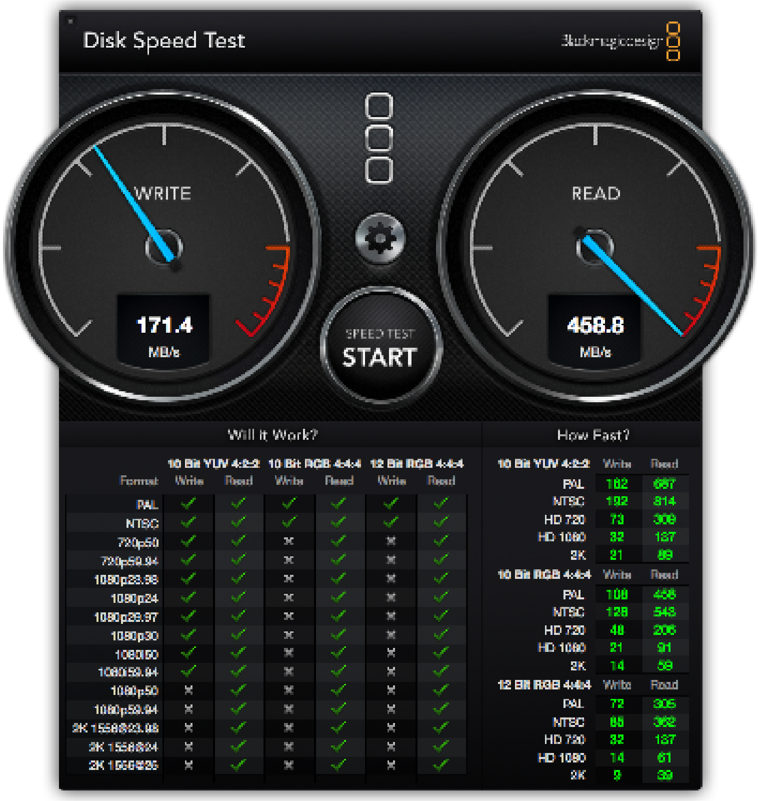
How Fast?

10 Bit YUV 4:2:2	Write	Read
PAL	474	788
NTSC	562	935
HD 720	213	356
HD 1080	94	157
2K	61	102

10 Bit RGB 4:4:4	Write	Read
PAL	316	525
NTSC	375	628
HD 720	142	236
HD 1080	63	105
2K	41	68

12 Bit RGB 4:4:4	Write	Read
PAL	210	350
NTSC	250	416
HD 720	94	157
HD 1080	42	70
2K	27	45

Six concurrent clients on the same Raid Z2



Will it Work?

	10 Bit YUV 4:2:2		10 Bit RGB 4:4:4		12 Bit RGB 4:4:4	
Format	Write	Read	Write	Read	Write	Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✓	✓	✗	✓	✗	✓
720p59.94	✓	✓	✗	✓	✗	✓
1080p23.98	✓	✓	✗	✓	✗	✓
1080p24	✓	✓	✗	✓	✗	✓
1080p29.97	✓	✓	✗	✓	✗	✓
1080p30	✓	✓	✗	✓	✗	✓
1080i50	✓	✓	✗	✓	✗	✓
1080i59.94	✓	✓	✗	✓	✗	✓
1080p50	✗	✓	✗	✓	✗	✓
1080p59.94	✗	✓	✗	✓	✗	✓
2K 1556p23.98	✗	✓	✗	✓	✗	✓
2K 1556p24	✗	✓	✗	✓	✗	✓
2K 1556p25	✗	✓	✗	✓	✗	✓

How Fast?

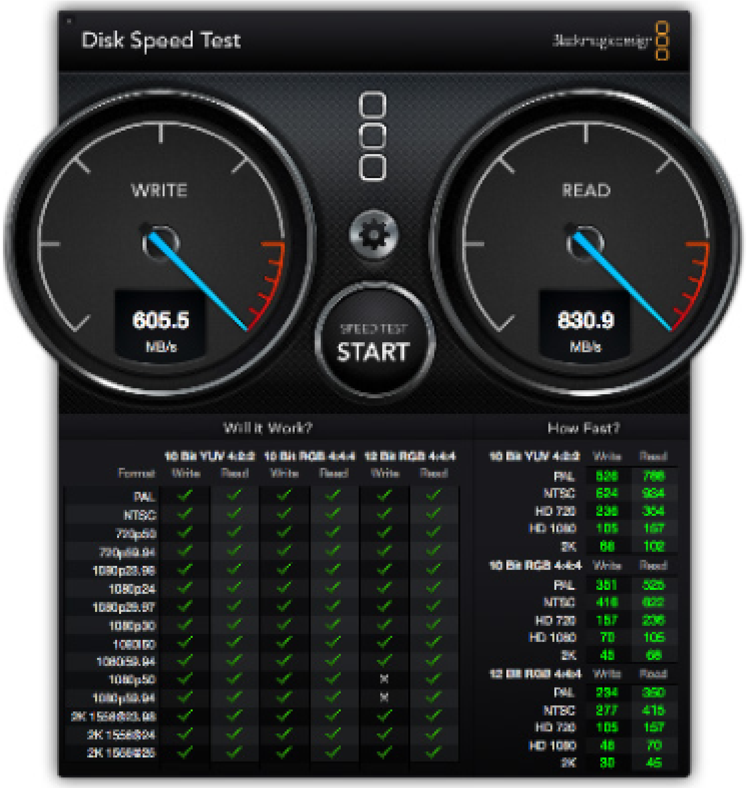
10 Bit YUV 4:2:2	Write	Read
PAL	182	687
NTSC	192	814
HD 720	73	309
HD 1080	32	137
2K	21	89

10 Bit RGB 4:4:4	Write	Read
PAL	100	458
NTSC	128	548
HD 720	40	206
HD 1080	21	91
2K	14	59

12 Bit RGB 4:4:4	Write	Read
PAL	72	305
NTSC	85	362
HD 720	32	137
HD 1080	14	61
2K	9	39

Performance degration is noticable with a Raid-Z2 pool from 10 spindels.
Write values for 6 concurrent users jumps between 130 MB and 200 MB/s with reads always 300 MB/s and better

SMB 2.1, 10 GbE, MTU 9000 (Solaris)
Single User on 1 disk NVMe Intel P750



Disk Speed Test

WRITE 605.5 MB/s

READ 830.9 MB/s

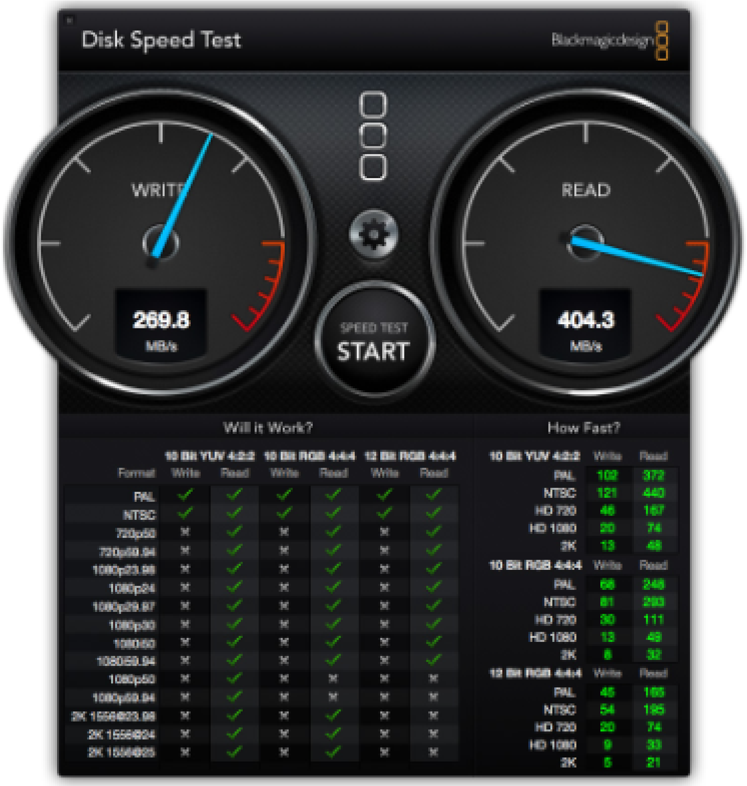
Will it Work?

	10 Bit YUV 4:2:2		10 Bit RGB 4:4:4		12 Bit RGB 4:4:4	
Format	Write	Read	Write	Read	Write	Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✓	✓	✓	✓	✓	✓
720p59.94	✓	✓	✓	✓	✓	✓
1080p23.98	✓	✓	✓	✓	✓	✓
1080p24	✓	✓	✓	✓	✓	✓
1080p29.97	✓	✓	✓	✓	✓	✓
1080p30	✓	✓	✓	✓	✓	✓
1080i50	✓	✓	✓	✓	✓	✓
1080i59.94	✓	✓	✓	✓	✓	✓
1080p50	✓	✓	✓	✓	✗	✓
1080p59.94	✓	✓	✓	✓	✗	✓
2K 1556p23.98	✓	✓	✓	✓	✓	✓
2K 1556p24	✓	✓	✓	✓	✓	✓
2K 1556p25	✓	✓	✓	✓	✓	✓

How Fast?

10 Bit YUV 4:2:2	Write	Read
PAL	526	788
NTSC	624	934
HD 720	236	354
HD 1080	105	157
2K	68	102
10 Bit RGB 4:4:4	**Write**	**Read**
PAL	351	525
NTSC	416	622
HD 720	157	236
HD 1080	70	105
2K	45	68
12 Bit RGB 4:4:4	**Write**	**Read**
PAL	234	350
NTSC	277	415
HD 720	105	157
HD 1080	48	70
2K	30	45

Six concurrent clients on the same basic pool



Disk Speed Test

WRITE 269.8 MB/s

READ 404.3 MB/s

Will it Work?

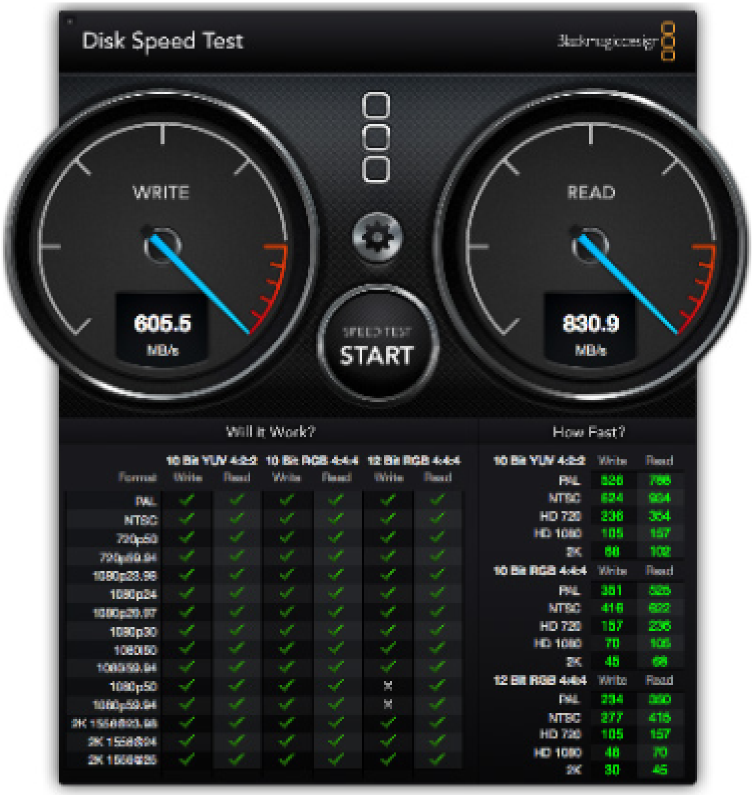
	10 Bit YUV 4:2:2		10 Bit RGB 4:4:4		12 Bit RGB 4:4:4	
Format	Write	Read	Write	Read	Write	Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✗	✓	✗	✓	✗	✓
720p59.94	✗	✓	✗	✓	✗	✓
1080p23.98	✗	✓	✗	✓	✗	✓
1080p24	✗	✓	✗	✓	✗	✓
1080p29.97	✗	✓	✗	✓	✗	✓
1080p30	✗	✓	✗	✓	✗	✓
1080i50	✗	✓	✗	✓	✗	✓
1080i59.94	✗	✓	✗	✓	✗	✓
1080p50	✗	✓	✗	✗	✗	✗
1080p59.94	✗	✓	✗	✗	✗	✗
2K 1556p23.98	✗	✓	✗	✓	✗	✗
2K 1556p24	✗	✓	✗	✓	✗	✗
2K 1556p25	✗	✓	✗	✓	✗	✗

How Fast?

10 Bit YUV 4:2:2	Write	Read
PAL	102	372
NTSC	121	440
HD 720	46	167
HD 1080	20	74
2K	13	48
10 Bit RGB 4:4:4	**Write**	**Read**
PAL	68	248
NTSC	81	293
HD 720	30	111
HD 1080	13	49
2K	8	32
12 Bit RGB 4:4:4	**Write**	**Read**
PAL	45	165
NTSC	54	195
HD 720	20	74
HD 1080	9	33
2K	5	21

The performance degration with 6 users on writes is lower due the higher iops of the NVMe.
Read degration is quite similar to the spindle based Z2 pool. The ZFS Ram cache seems to equal this.

SMB 2.1, 10 GbE, MTU 9000, Multipool-Test vs Z2
Single User on 1 disk NVMe Intel P750



Disk Speed Test

WRITE 605.5 MB/s

READ 830.9 MB/s

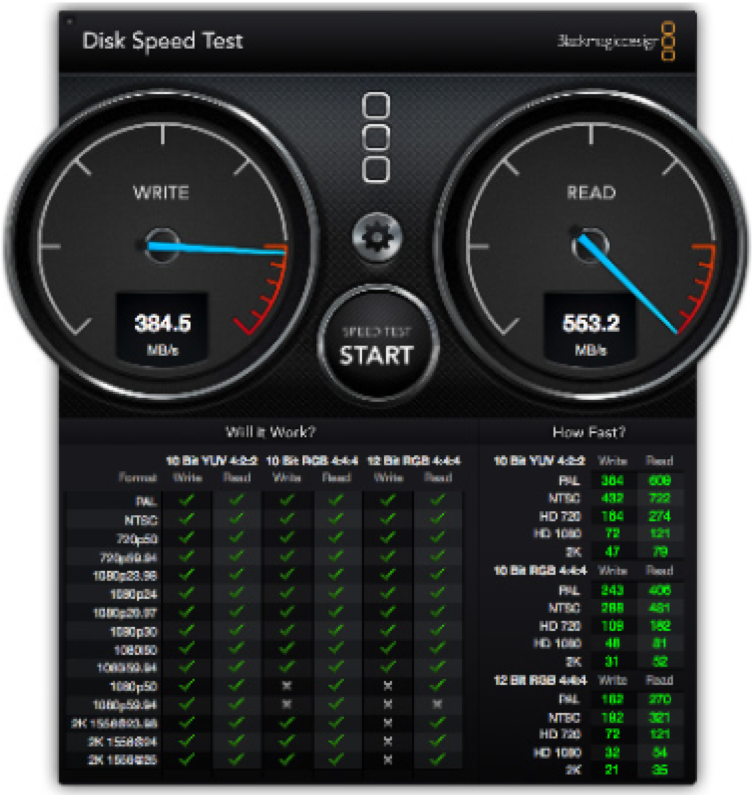
Will it Work?

	10 Bit YUV 4:2:2		10 Bit RGB 4:4:4		12 Bit RGB 4:4:4	
Format	Write	Read	Write	Read	Write	Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✓	✓	✓	✓	✓	✓
720p59.94	✓	✓	✓	✓	✓	✓
1080p23.98	✓	✓	✓	✓	✓	✓
1080p24	✓	✓	✓	✓	✓	✓
1080p29.97	✓	✓	✓	✓	✓	✓
1080p30	✓	✓	✓	✓	✓	✓
1080i50	✓	✓	✓	✓	✓	✓
1080i59.94	✓	✓	✓	✓	✓	✓
1080p50	✓	✓	✓	✓	✗	✓
1080p59.94	✓	✓	✓	✓	✗	✓
2K 1556p23.98	✓	✓	✓	✓	✓	✓
2K 1556p24	✓	✓	✓	✓	✓	✓
2K 1556p25	✓	✓	✓	✓	✓	✓

How Fast?

10 Bit YUV 4:2:2	Write	Read
PAL	526	788
NTSC	624	934
HD 720	236	354
HD 1080	105	157
2K	68	102
10 Bit RGB 4:4:4	**Write**	**Read**
PAL	351	525
NTSC	416	622
HD 720	157	236
HD 1080	70	105
2K	45	68
12 Bit RGB 4:4:4	**Write**	**Read**
PAL	234	350
NTSC	277	415
HD 720	105	157
HD 1080	48	70
2K	30	45

This test is running with 5 user on Raid-Z2.
concurrently to the left single user



Disk Speed Test

WRITE 384.5 MB/s

READ 553.2 MB/s

Will it Work?

	10 Bit YUV 4:2:2		10 Bit RGB 4:4:4		12 Bit RGB 4:4:4	
Format	Write	Read	Write	Read	Write	Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✓	✓	✓	✓	✓	✓
720p59.94	✓	✓	✓	✓	✓	✓
1080p23.98	✓	✓	✓	✓	✓	✓
1080p24	✓	✓	✓	✓	✓	✓
1080p29.97	✓	✓	✓	✓	✓	✓
1080p30	✓	✓	✓	✓	✓	✓
1080i50	✓	✓	✓	✓	✓	✓
1080i59.94	✓	✓	✓	✓	✓	✓
1080p50	✓	✓	✗	✓	✗	✓
1080p59.94	✓	✓	✗	✓	✗	✗
2K 1556p23.98	✓	✓	✓	✓	✗	✓
2K 1556p24	✓	✓	✓	✓	✗	✓
2K 1556p25	✓	✓	✓	✓	✗	✓

How Fast?

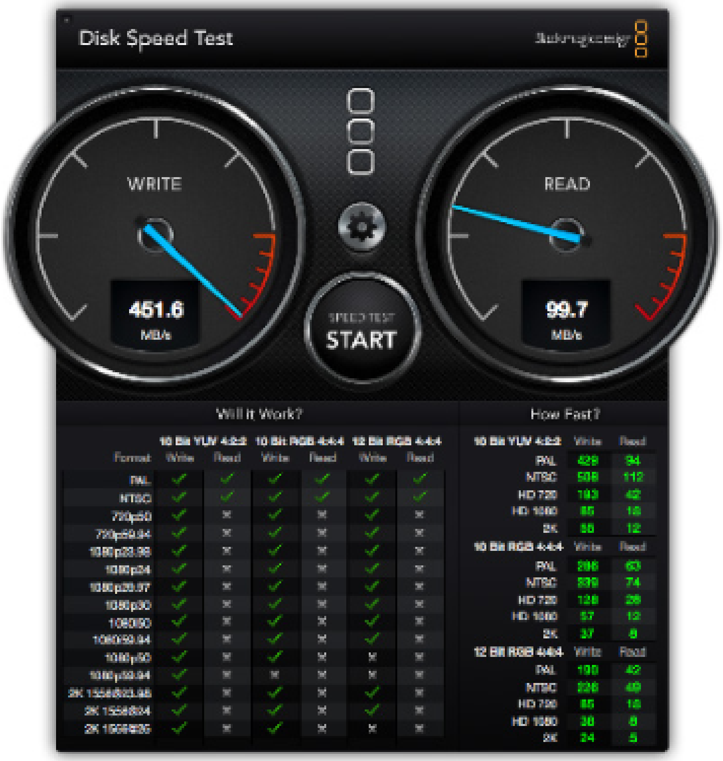
10 Bit YUV 4:2:2	Write	Read
PAL	384	609
NTSC	432	722
HD 720	184	274
HD 1080	72	121
2K	47	79
10 Bit RGB 4:4:4	**Write**	**Read**
PAL	243	406
NTSC	288	481
HD 720	109	182
HD 1080	48	81
2K	31	52
12 Bit RGB 4:4:4	**Write**	**Read**
PAL	162	270
NTSC	192	321
HD 720	72	121
HD 1080	32	54
2K	21	35

Two pools accessed at the same time
The left NVMe pool runs with quite full speed concurrently to the Z2 pool with 5 users

Round 2: SMB2 on OmniOS bloody 151017 beta (Dez 2015)

During this tests, the new OmniOS with SMB2.1 support arrived in a beta state.
I installed the OmniOS 151017 beta to the same hardware with same settings. After the first test it seems that it behaves different to Solaris. While I saw the same behaviour like a good performance on some config and a very bad performance on other. But some configurations that were fine with Solaris are bad with OmniOS and vice versa:

MacPro, OSX 10.5, Raid-Z2, MTU 9000, SMB 2.1
Values on Solaris were 500 MB/s and 800 MB/s



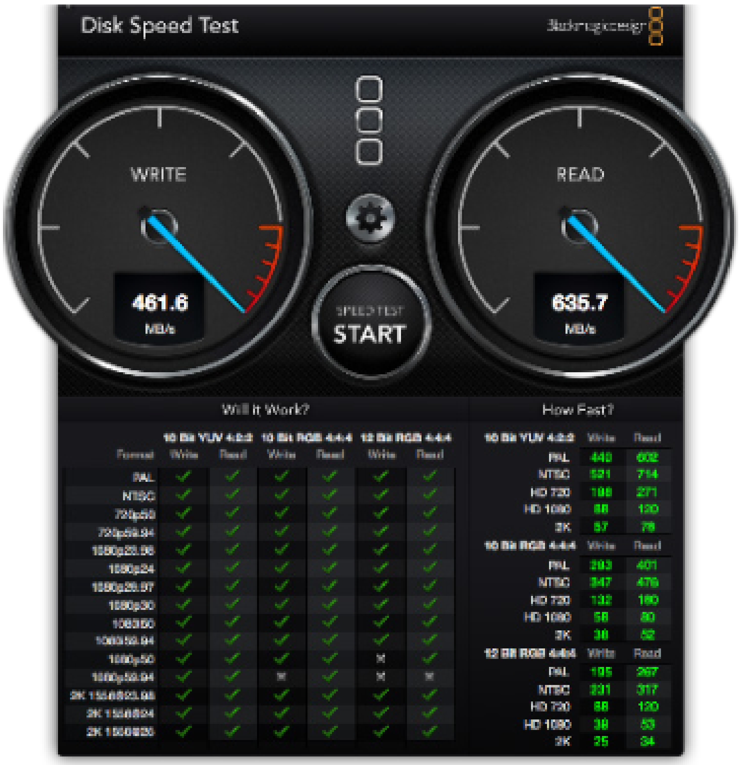
Will it Work?

Format	10 Bit YUV 4:2:2		10 Bit RGB 4:4:4		12 Bit RGB 4:4:4	
	Write	Read	Write	Read	Write	Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✓	✗	✓	✗	✓	✗
720p59.94	✓	✗	✓	✗	✓	✗
1080p23.98	✓	✗	✓	✗	✓	✗
1080p24	✓	✗	✓	✗	✓	✗
1080p29.97	✓	✗	✓	✗	✓	✗
1080p30	✓	✗	✓	✗	✓	✗
1080i50	✓	✗	✓	✗	✓	✗
1080i59.94	✓	✗	✓	✗	✓	✗
1080p50	✓	✗	✓	✗	✗	✗
1080p59.94	✓	✗	✗	✗	✗	✗
2K 1556p23.98	✓	✗	✓	✗	✓	✗
2K 1556p24	✓	✗	✓	✗	✓	✗
2K 1556p25	✓	✗	✓	✗	✗	✗

How Fast?

10 Bit YUV 4:2:2	Write	Read
PAL	429	94
NTSC	508	112
HD 720	193	42
HD 1080	85	18
2K	55	12
10 Bit RGB 4:4:4	Write	Read
PAL	286	63
NTSC	339	74
HD 720	128	28
HD 1080	57	12
2K	37	8
12 Bit RGB 4:4:4	Write	Read
PAL	190	42
NTSC	226	49
HD 720	85	18
HD 1080	38	8
2K	24	5

Values on Solaris were 520 MB/s and 101 MB/s



Will it Work?

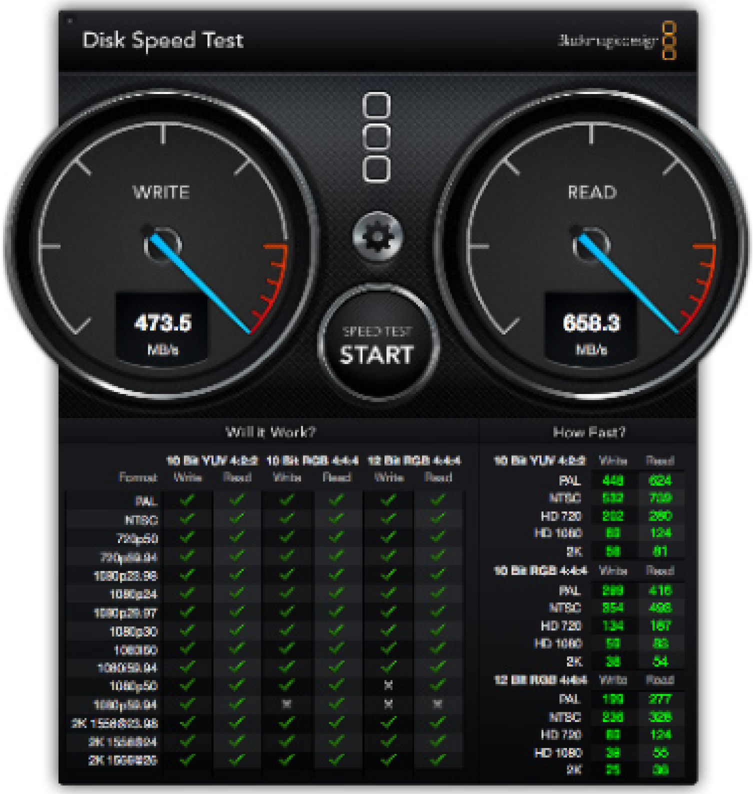
Format	10 Bit YUV 4:2:2		10 Bit RGB 4:4:4		12 Bit RGB 4:4:4	
	Write	Read	Write	Read	Write	Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✓	✓	✓	✓	✓	✓
720p59.94	✓	✓	✓	✓	✓	✓
1080p23.98	✓	✓	✓	✓	✓	✓
1080p24	✓	✓	✓	✓	✓	✓
1080p29.97	✓	✓	✓	✓	✓	✓
1080p30	✓	✓	✓	✓	✓	✓
1080i50	✓	✓	✓	✓	✓	✓
1080i59.94	✓	✓	✓	✓	✓	✓
1080p50	✓	✓	✓	✓	✗	✓
1080p59.94	✓	✓	✗	✓	✗	✗
2K 1556p23.98	✓	✓	✓	✓	✓	✓
2K 1556p24	✓	✓	✓	✓	✓	✓
2K 1556p25	✓	✓	✓	✓	✓	✓

How Fast?

10 Bit YUV 4:2:2	Write	Read
PAL	440	602
NTSC	521	714
HD 720	198	271
HD 1080	88	120
2K	57	78
10 Bit RGB 4:4:4	Write	Read
PAL	293	401
NTSC	347	476
HD 720	132	180
HD 1080	58	80
2K	38	52
12 Bit RGB 4:4:4	Write	Read
PAL	195	267
NTSC	231	317
HD 720	88	120
HD 1080	38	53
2K	25	34

I updated all MacPros to 10.11.2 and Windows on MacPros with newest Sanlink driver for further tests
after this, I saw a consistent high write and read performrmance with OSX and good values on Windows

Pool Layout: 10 x Raid-Z2 of 2 TB HGST 7k U/m



Will it Work?

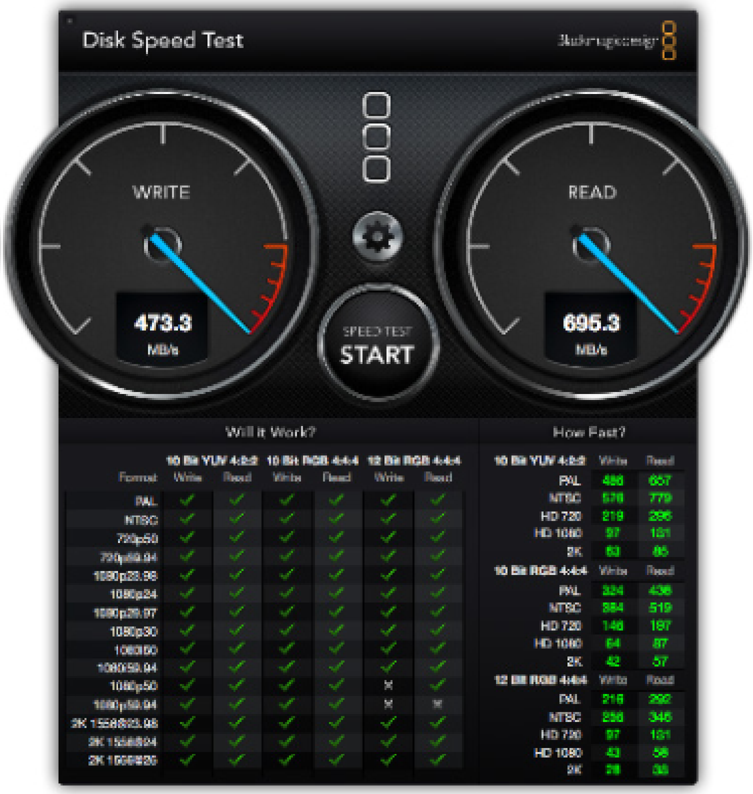
Format	10 Bit YUV 4:2:2		10 Bit RGB 4:4:4		12 Bit RGB 4:4:4	
	Write	Read	Write	Read	Write	Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✓	✓	✓	✓	✓	✓
720p59.94	✓	✓	✓	✓	✓	✓
1080p23.98	✓	✓	✓	✓	✓	✓
1080p24	✓	✓	✓	✓	✓	✓
1080p29.97	✓	✓	✓	✓	✓	✓
1080p30	✓	✓	✓	✓	✓	✓
1080i50	✓	✓	✓	✓	✓	✓
1080i59.94	✓	✓	✓	✓	✓	✓
1080p50	✓	✓	✓	✓	✗	✓
1080p59.94	✓	✓	✗	✓	✗	✗
2K 1556p23.98	✓	✓	✓	✓	✓	✓
2K 1556p24	✓	✓	✓	✓	✓	✓
2K 1556p25	✓	✓	✓	✓	✓	✓

How Fast?

10 Bit YUV 4:2:2	Write	Read
PAL	448	624
NTSC	532	739
HD 720	202	280
HD 1080	88	124
2K	58	81
10 Bit RGB 4:4:4	Write	Read
PAL	299	416
NTSC	354	498
HD 720	134	187
HD 1080	59	83
2K	38	54
12 Bit RGB 4:4:4	Write	Read
PAL	199	277
NTSC	236	328
HD 720	88	124
HD 1080	39	55
2K	25	36

good values, slighty slower than Solaris 11.3 with 501 MB/s write and 836 MB/s read

Pool Layout: 1 x basic Intel P750 NVMe



Will it Work?

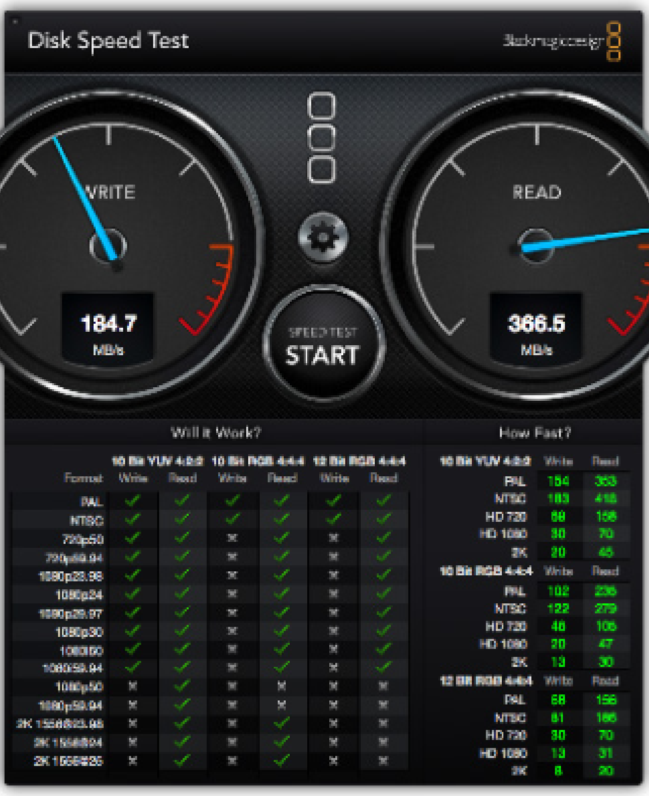
Format	10 Bit YUV 4:2:2		10 Bit RGB 4:4:4		12 Bit RGB 4:4:4	
	Write	Read	Write	Read	Write	Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✓	✓	✓	✓	✓	✓
720p59.94	✓	✓	✓	✓	✓	✓
1080p23.98	✓	✓	✓	✓	✓	✓
1080p24	✓	✓	✓	✓	✓	✓
1080p29.97	✓	✓	✓	✓	✓	✓
1080p30	✓	✓	✓	✓	✓	✓
1080i50	✓	✓	✓	✓	✓	✓
1080i59.94	✓	✓	✓	✓	✓	✓
1080p50	✓	✓	✓	✓	✗	✓
1080p59.94	✓	✓	✓	✓	✗	✗
2K 1556p23.98	✓	✓	✓	✓	✓	✓
2K 1556p24	✓	✓	✓	✓	✓	✓
2K 1556p25	✓	✓	✓	✓	✓	✓

How Fast?

10 Bit YUV 4:2:2	Write	Read
PAL	486	657
NTSC	576	779
HD 720	219	296
HD 1080	97	131
2K	63	85
10 Bit RGB 4:4:4	Write	Read
PAL	324	438
NTSC	384	519
HD 720	146	197
HD 1080	64	87
2K	42	57
12 Bit RGB 4:4:4	Write	Read
PAL	216	292
NTSC	256	346
HD 720	97	131
HD 1080	43	58
2K	28	39

similar on NVMe compared to Solaris with 605 MB/s write and 830 MB/s read

MacPro, OSX 10.5, Raid-Z2, MTU 9000, SMB 2.1 (OmniOS)
5 concurrent user/ MacPros



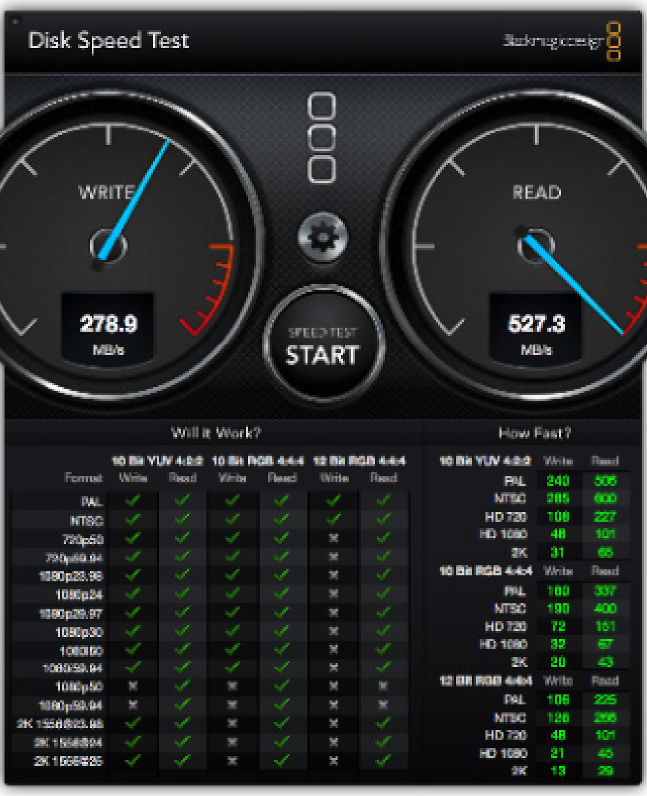
Will it Work?

	10 Bit YUV 4:2:2		10 Bit RGB 4:4:4		12 Bit RGB 4:4:4	
Format	Write	Read	Write	Read	Write	Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✓	✓	✗	✓	✗	✓
720p59.94	✓	✓	✗	✓	✗	✓
1080p23.98	✓	✓	✗	✓	✗	✓
1080p24	✓	✓	✗	✓	✗	✓
1080p29.97	✓	✓	✗	✓	✗	✓
1080p30	✓	✓	✗	✓	✗	✓
1080i50	✓	✓	✗	✓	✗	✓
1080i59.94	✓	✓	✗	✓	✗	✓
1080p50	✗	✓	✗	✗	✗	✗
1080p59.94	✗	✓	✗	✗	✗	✗
2K 1556p23.98	✗	✓	✗	✓	✗	✗
2K 1556p24	✗	✓	✗	✓	✗	✗
2K 1556p25	✗	✓	✗	✓	✗	✗

How Fast?

10 Bit YUV 4:2:2	Write	Read
PAL	154	363
NTSC	183	418
HD 720	68	158
HD 1080	30	70
2K	20	45
10 Bit RGB 4:4:4	Write	Read
PAL	102	236
NTSC	122	279
HD 720	46	106
HD 1080	20	47
2K	13	30
12 Bit RGB 4:4:4	Write	Read
PAL	68	156
NTSC	81	186
HD 720	30	70
HD 1080	13	31
2K	8	20

Concurrent access: 1 user to another NVMe pool



Will it Work?

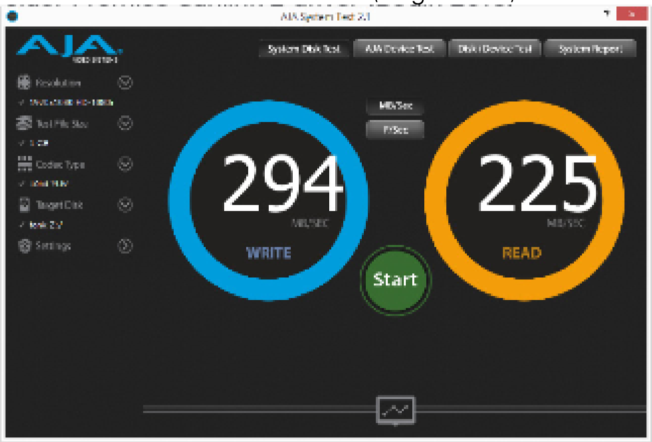
	10 Bit YUV 4:2:2		10 Bit RGB 4:4:4		12 Bit RGB 4:4:4	
Format	Write	Read	Write	Read	Write	Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✓	✓	✓	✓	✗	✓
720p59.94	✓	✓	✓	✓	✗	✓
1080p23.98	✓	✓	✓	✓	✗	✓
1080p24	✓	✓	✓	✓	✗	✓
1080p29.97	✓	✓	✓	✓	✗	✓
1080p30	✓	✓	✓	✓	✗	✓
1080i50	✓	✓	✓	✓	✗	✓
1080i59.94	✓	✓	✓	✓	✗	✓
1080p50	✗	✓	✗	✓	✗	✗
1080p59.94	✗	✓	✗	✓	✗	✗
2K 1556p23.98	✓	✓	✗	✓	✗	✓
2K 1556p24	✓	✓	✗	✓	✗	✓
2K 1556p25	✓	✓	✗	✓	✗	✓

How Fast?

10 Bit YUV 4:2:2	Write	Read
PAL	240	506
NTSC	285	600
HD 720	108	227
HD 1080	48	101
2K	31	65
10 Bit RGB 4:4:4	Write	Read
PAL	160	337
NTSC	190	400
HD 720	72	151
HD 1080	32	67
2K	20	43
12 Bit RGB 4:4:4	Write	Read
PAL	106	225
NTSC	126	266
HD 720	48	101
HD 1080	21	45
2K	13	29

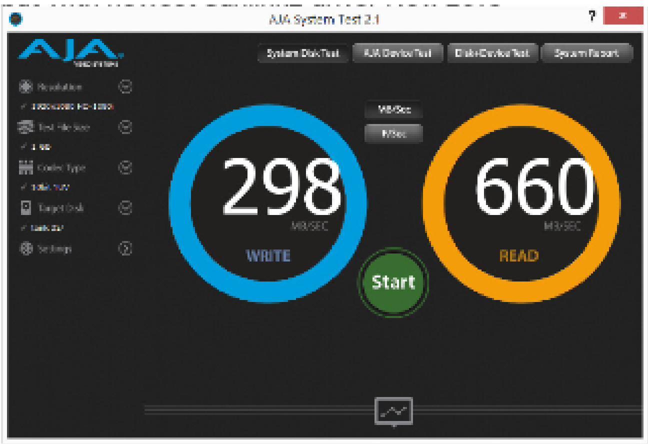
Same like on Solaris
With many user working on the same pool, especially writes values are quite low
Another user that is working on a different pool das better values so pool is more a limiting factor than network

MacPro with Win 8.1, Raid-Z2, MTU 9000, 1 user
older Promise Sanlink 2 driver (Begin 2015)



Resolution: 1920x1080 HD-1080
Test File Size: 1 GB
Codec Type: 10bit YUV
Target Disk: tank Z:/
Settings

Same MacPro with Windows 8.1
but with newest Sanlink2 driver Nov. 2015



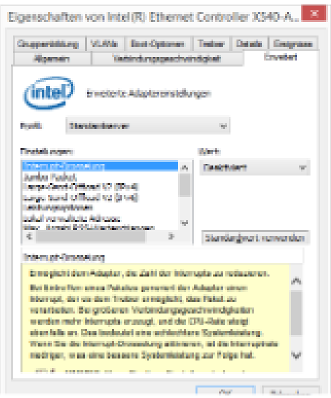
Resolution: 1920x1080 HD-1080
Test File Size: 1 GB
Codec Type: 10bit YUV
Target Disk: tank Z:/
Settings

A newer driver can double read values. This is similar to OSX where another OS release can have the same effect.
In general Windows is lightly slower than OSX with default settings. Solaris and OmniOS perform similar.

Optimal settings X540

Interrupt throttling: deactivated
Jumbo Packet: 9014

use newest Intel driver:
(Sanlink2: Nov 2015)


Eigenschaften von Intel (R) Ethernet Controller X540-A...
Gruppenbildung
VLANs
Boot-Optionen
Treiber
Details
Ereignisse
Allgemein
Verbindungsgeschwindigkeit
Erweitert
intel
Erweiterte Adaptereinstellungen
OK
Abbrechen

Profil: Standardserver

Einstellungen:
Interrupt-Drosselung
Jumbo Paket
Large-Send-Offload V2 (IPv4)
Large-Send-Offload V2 (IPv6)
Leistungsoptionen
Lokal verwaltete Adresse
Max. Anzahl RSS-Warteschlangen

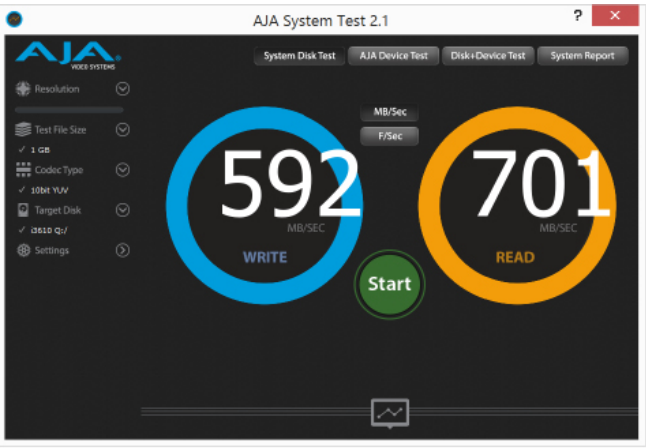
Wert: Deaktiviert

Standardwert verwenden

Interrupt-Drosselung

Ermöglicht dem Adapter, die Zahl der Interrupts zu reduzieren.

Bei Eintreffen eines Paketes generiert der Adapter einen Interrupt, der es dem Treiber ermöglicht, das Paket zu verarbeiten. Bei größeren Verbindungsgeschwindigkeiten werden mehr Interrupts erzeugt, und die CPU-Rate steigt ebenfalls an. Das bedeutet eine schlechtere Systemleistung. Wenn Sie die Interrupt-Drosselung aktivieren, ist die Interruptrate niedriger, was eine bessere Systemleistung zur Folge hat.



Resolution
Test File Size: 1 GB
Codec Type: 10bit YUV
Target Disk: S610 Q:/
Settings

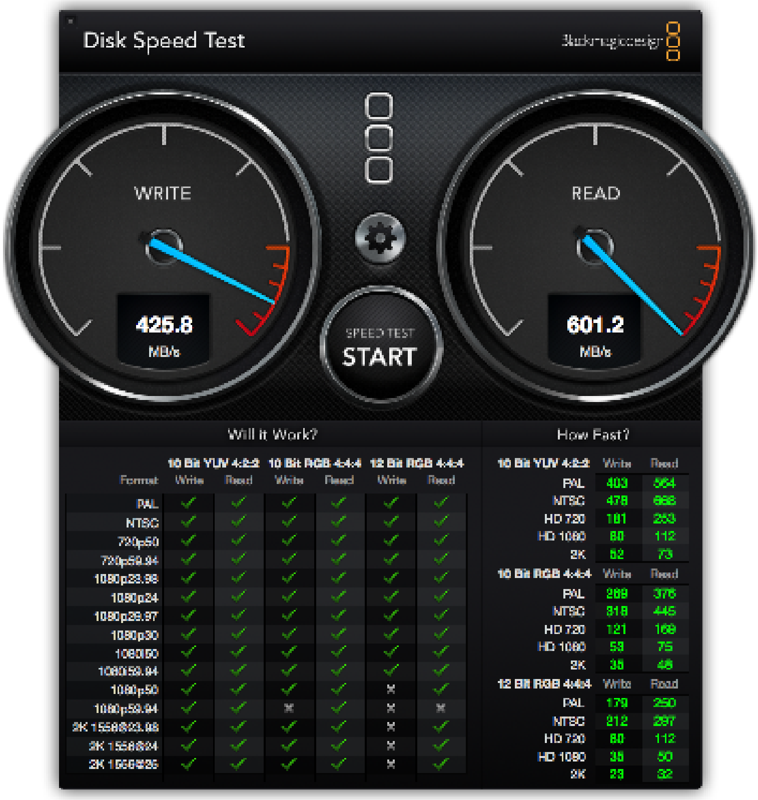
Intel S3610, 3 x raid-0, jumboframes, int-throtteling off

Round 3 - What happens when I reduce Server RAM from 40 GB to 8 GB

on same hardware and sequential tests

Single user access to the Raid-Z2

10 x HGST 2 TB with 8GB RAM



WRITE 425.8 MB/s

READ 601.2 MB/s

Will it Work?

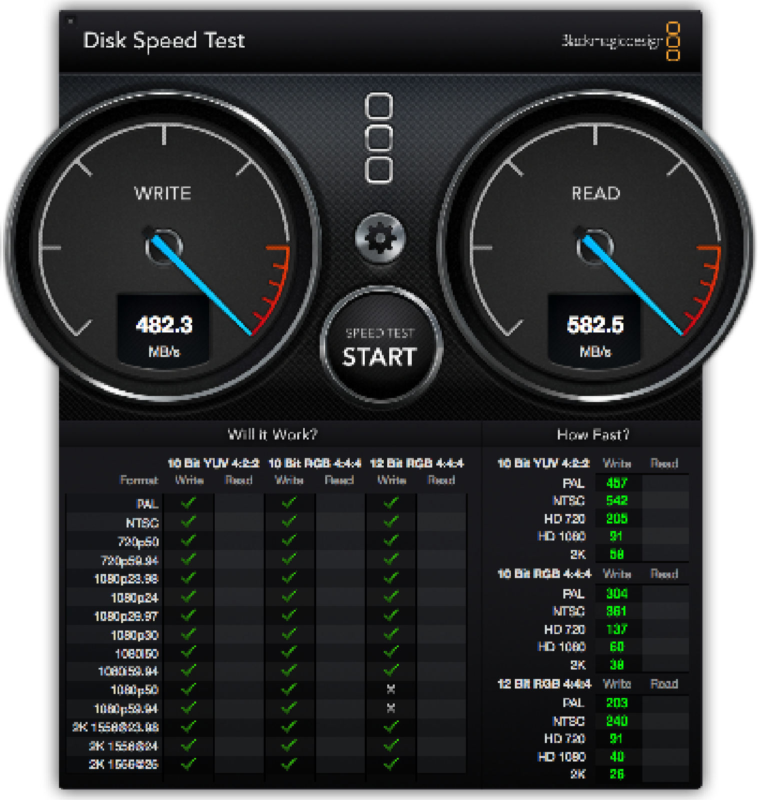
	10 Bit YUV 4:2:2		10 Bit RGB 4:4:4		12 Bit RGB 4:4:4	
Format	Write	Read	Write	Read	Write	Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✓	✓	✓	✓	✓	✓
720p59.94	✓	✓	✓	✓	✓	✓
1080p23.98	✓	✓	✓	✓	✓	✓
1080p24	✓	✓	✓	✓	✓	✓
1080p29.97	✓	✓	✓	✓	✓	✓
1080p30	✓	✓	✓	✓	✓	✓
1080i50	✓	✓	✓	✓	✓	✓
1080i59.94	✓	✓	✓	✓	✓	✓
1080p50	✓	✓	✓	✓	✗	✓
1080p59.94	✓	✓	✗	✓	✗	✗
2K 1556p23.98	✓	✓	✓	✓	✗	✓
2K 1556p24	✓	✓	✓	✓	✗	✓
2K 1556p25	✓	✓	✓	✓	✗	✓

How Fast?

10 Bit YUV 4:2:2	Write	Read
PAL	403	564
NTSC	478	668
HD 720	181	253
HD 1080	80	112
2K	52	73
10 Bit RGB 4:4:4	**Write**	**Read**
PAL	269	376
NTSC	318	445
HD 720	121	169
HD 1080	53	75
2K	35	48
12 Bit RGB 4:4:4	**Write**	**Read**
PAL	179	250
NTSC	212	297
HD 720	80	112
HD 1080	35	50
2K	23	32

Single user access to the raid-0

3 x Intel s 3610-1,6 TB



WRITE 482.3 MB/s

READ 582.5 MB/s

Will it Work?

	10 Bit YUV 4:2:2		10 Bit RGB 4:4:4		12 Bit RGB 4:4:4	
Format	Write	Read	Write	Read	Write	Read
PAL	✓		✓		✓	
NTSC	✓		✓		✓	
720p50	✓		✓		✓	
720p59.94	✓		✓		✓	
1080p23.98	✓		✓		✓	
1080p24	✓		✓		✓	
1080p29.97	✓		✓		✓	
1080p30	✓		✓		✓	
1080i50	✓		✓		✓	
1080i59.94	✓		✓		✓	
1080p50	✓		✓		✗	
1080p59.94	✓		✓		✗	
2K 1556p23.98	✓		✓		✓	
2K 1556p24	✓		✓		✓	
2K 1556p25	✓		✓		✓	

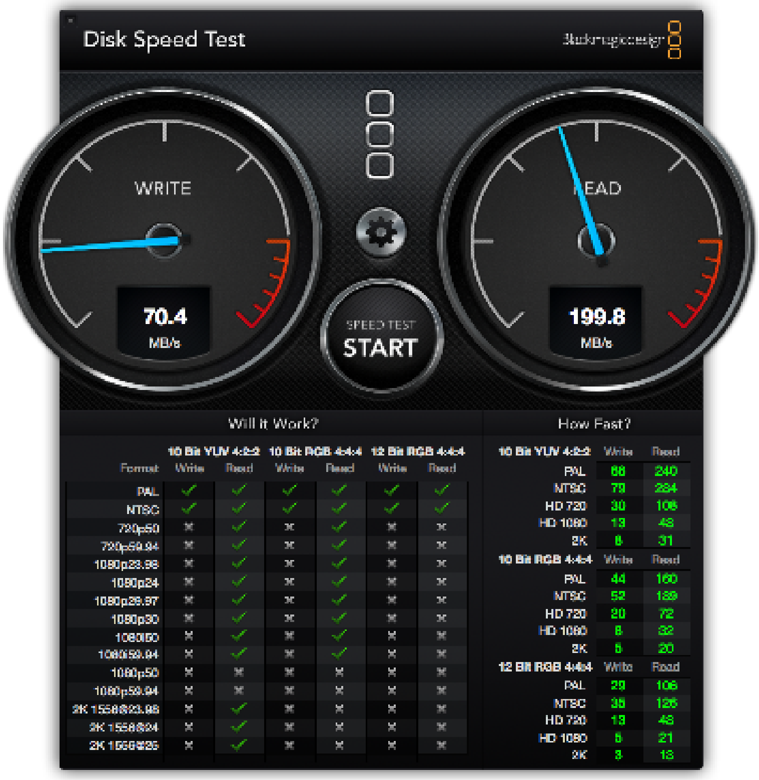
How Fast?

10 Bit YUV 4:2:2	Write	Read
PAL	457	
NTSC	542	
HD 720	205	
HD 1080	91	
2K	58	
10 Bit RGB 4:4:4	**Write**	**Read**
PAL	304	
NTSC	361	
HD 720	137	
HD 1080	60	
2K	38	
12 Bit RGB 4:4:4	**Write**	**Read**
PAL	203	
NTSC	240	
HD 720	91	
HD 1080	40	
2K	26	

With a single user/ single thread sequential performance with only 8 GB server-RAM is about 30% lower than with 40 GB RAM but more than enough for most workloads and 4x to 6x to whats possible with 1G networks.

Six concurrent MacPro/user access to the Raid-Z2

10 x HGST 2 TB with 8GB RAM



WRITE 70.4 MB/s

READ 199.8 MB/s

Will it Work?

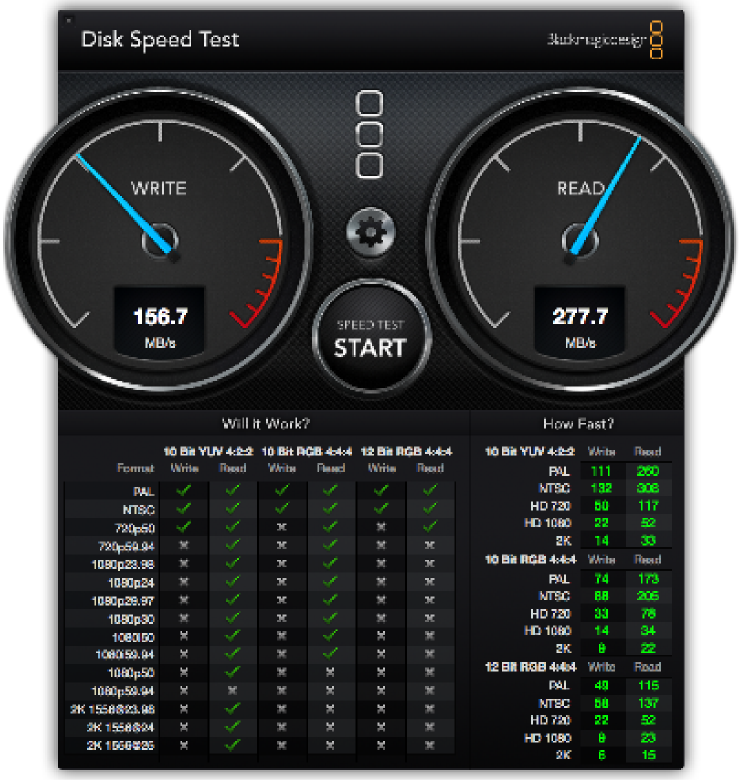
	10 Bit YUV 4:2:2		10 Bit RGB 4:4:4		12 Bit RGB 4:4:4	
Format	Write	Read	Write	Read	Write	Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✗	✓	✗	✓	✗	✗
720p59.94	✗	✓	✗	✓	✗	✗
1080p23.98	✗	✓	✗	✓	✗	✗
1080p24	✗	✓	✗	✓	✗	✗
1080p29.97	✗	✓	✗	✓	✗	✗
1080p30	✗	✓	✗	✓	✗	✗
1080i50	✗	✓	✗	✓	✗	✗
1080i59.94	✗	✓	✗	✓	✗	✗
1080p50	✗	✗	✗	✗	✗	✗
1080p59.94	✗	✗	✗	✗	✗	✗
2K 1556p23.98	✗	✓	✗	✗	✗	✗
2K 1556p24	✗	✓	✗	✗	✗	✗
2K 1556p25	✗	✓	✗	✗	✗	✗

How Fast?

10 Bit YUV 4:2:2	Write	Read
PAL	66	240
NTSC	79	284
HD 720	30	106
HD 1080	13	48
2K	8	31
10 Bit RGB 4:4:4	**Write**	**Read**
PAL	44	160
NTSC	52	189
HD 720	20	72
HD 1080	8	32
2K	5	20
12 Bit RGB 4:4:4	**Write**	**Read**
PAL	29	106
NTSC	35	126
HD 720	13	48
HD 1080	5	21
2K	3	13

Six concurrent MacPro/user access to the raid-0

3 x Intel SSD S3610-1,6 TB



WRITE 156.7 MB/s

READ 277.7 MB/s

Will it Work?

	10 Bit YUV 4:2:2		10 Bit RGB 4:4:4		12 Bit RGB 4:4:4	
Format	Write	Read	Write	Read	Write	Read
PAL	✓	✓	✓	✓	✓	✓
NTSC	✓	✓	✓	✓	✓	✓
720p50	✓	✓	✗	✓	✗	✓
720p59.94	✗	✓	✗	✓	✗	✗
1080p23.98	✗	✓	✗	✓	✗	✗
1080p24	✗	✓	✗	✓	✗	✗
1080p29.97	✗	✓	✗	✓	✗	✗
1080p30	✗	✓	✗	✓	✗	✗
1080i50	✗	✓	✗	✓	✗	✗
1080i59.94	✗	✓	✗	✓	✗	✗
1080p50	✗	✓	✗	✗	✗	✗
1080p59.94	✗	✗	✗	✗	✗	✗
2K 1556p23.98	✗	✓	✗	✗	✗	✗
2K 1556p24	✗	✓	✗	✗	✗	✗
2K 1556p25	✗	✓	✗	✗	✗	✗

How Fast?

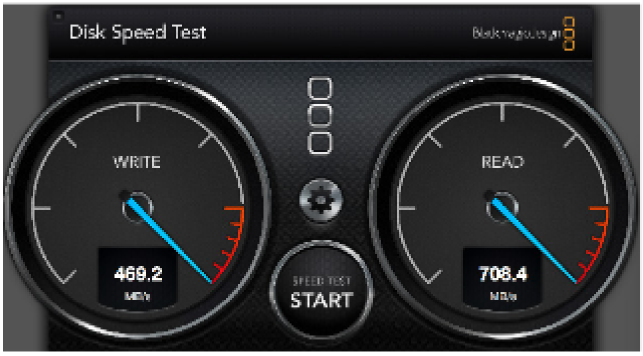
10 Bit YUV 4:2:2	Write	Read
PAL	111	260
NTSC	132	308
HD 720	50	117
HD 1080	22	52
2K	14	33
10 Bit RGB 4:4:4	**Write**	**Read**
PAL	74	173
NTSC	88	205
HD 720	33	78
HD 1080	14	34
2K	9	22
12 Bit RGB 4:4:4	**Write**	**Read**
PAL	49	115
NTSC	58	137
HD 720	22	52
HD 1080	9	23
2K	6	15

With only 8 GB RAM, many reads must come from disk. With the limited iops of the spindle Z2 pool performance degration especially on writes is quite massive. The SSD pool with a much higher iops rate is more than twice as fast but only half as fast as on samed tests with 40 GB RAM

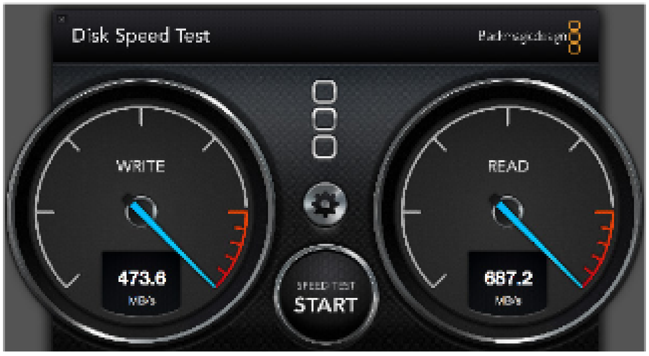
Round 5: Some tests

Dedupe and Compress (Tests run against an Intel P750-400)

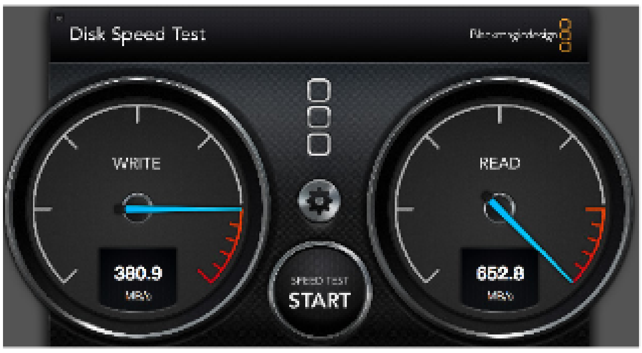
Referenz: dedup and Compress off (469/708 MB/s)



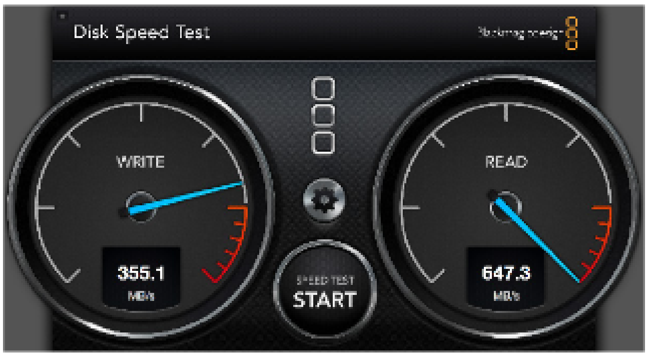
Compress LZ4 enabled (473/687 MB/s)



Dedupe enabled (380 / 652 MB/s)



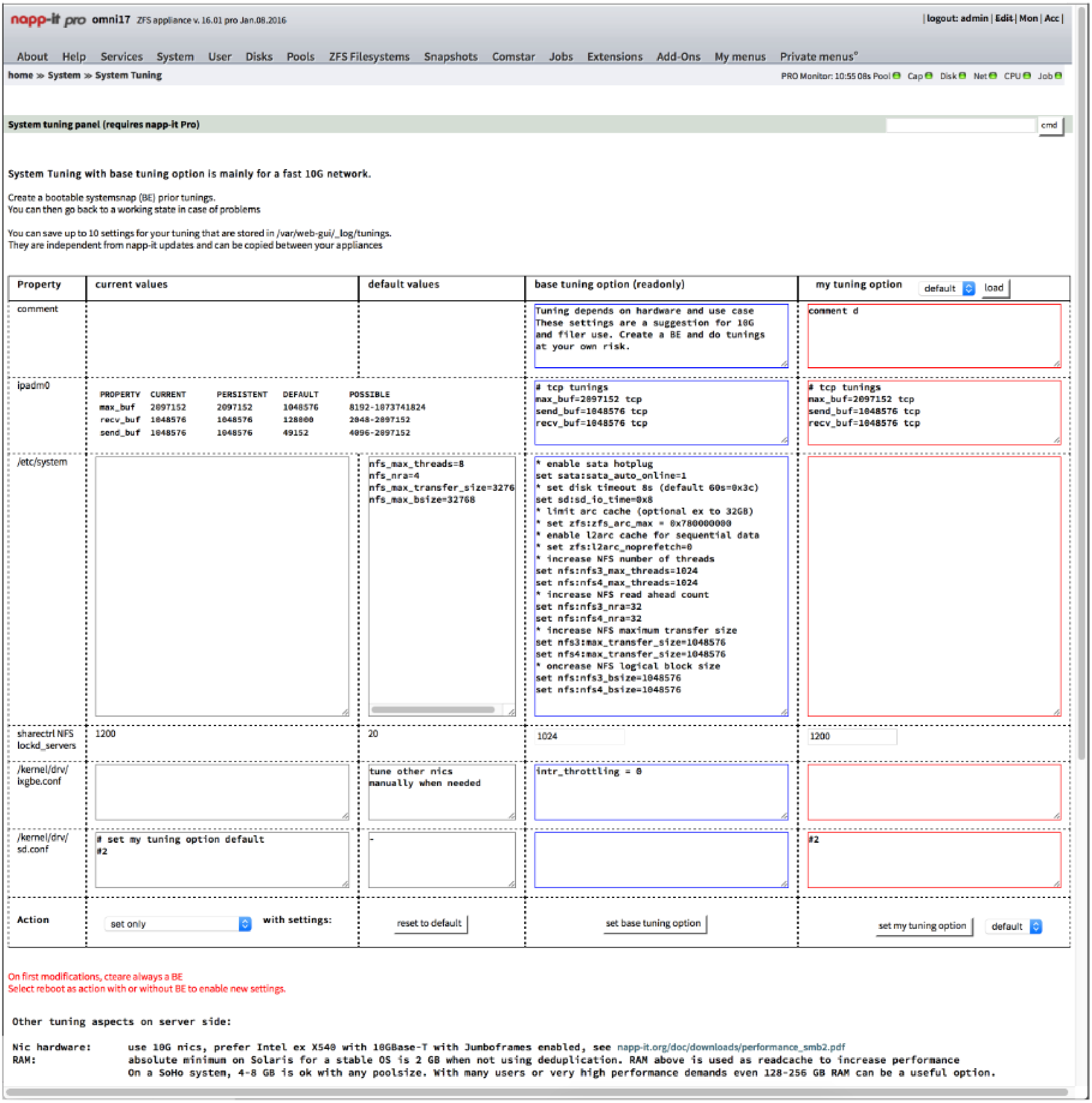
Dedupe and LZ4 enabled (355 / 647 MB/s)



There is a slight performance degration with compress and dedupe. But while dedupe has a massive impact on RAM needs (count up to 5 GB RAM per TB data + RAM needs for ARC cache), LZ4 is a good idea even on data without a good compressability.

Round 6: OS tuning options

Solaris and OmniOS are very fast but defaults are more optimized for 1G networks.
Lets try some tunings (Napp-it 16.01pro, menu System > System Tuning) to a single Intel P750-400



napp-it pro omni17 ZFS appliance v. 16.01 pro Jan.08.2016
| logout: admin | Edit | Mon | Acc |
About Help Services System User Disks Pools ZFS Filesystems Snapshots Comstar Jobs Extensions Add-Ons My menus Private menus°
home » System » System Tuning
PRO Monitor: 10:55 08s Pool Cap Disk Net CPU Job
System tuning panel (requires napp-it Pro)
cmd

System Tuning with base tuning option is mainly for a fast 10G network.

Create a bootable systemsnap (BE) prior tunings.
You can then go back to a working state in case of problems

You can save up to 10 settings for your tuning that are stored in /var/web-gui/_log/tunings.
They are independent from napp-it updates and can be copied between your appliances

Property	current values	default values	base tuning option (readonly)	my tuning option default load
comment			Tuning depends on hardware and use case These settings are a suggestion for 10G and filer use. Create a BE and do tunings at your own risk.	comment d

ipadm0	PROPERTY CURRENT PERSISTENT DEFAULT POSSIBLE
	max_buf 2097152 2097152 1048576 8192-1073741824
	recv_buf 1048576 1048576 128000 2048-2097152
	send_buf 1048576 1048576 49152 4096-2097152
			# tcp tunings
			max_buf=2097152 tcp
			send_buf=1048576 tcp
			recv_buf=1048576 tcp
				# tcp tunings
				max_buf=2097152 tcp
				send_buf=1048576 tcp
				recv_buf=1048576 tcp

/etc/system		nfs_max_threads=8
		nfs_nra=4
		nfs_max_transfer_size=3276
		nfs_max_bsize=32768
			* enable sata hotplug
			set sata:sata_auto_online=1
			* set disk timeout 8s (default 60s=0x3c)
			set sd:sd_io_time=0x8
			* limit arc cache (optional ex to 32GB)
			* set zfs:zfs_arc_max = 0x780000000
			* enable l2arc cache for sequential data
			* set zfs:l2arc_noprefetch=0
			* increase NFS number of threads
			set nfs:nfs3_max_threads=1024
			set nfs:nfs4_max_threads=1024
			* increase NFS read ahead count
			set nfs:nfs3_nra=32
			set nfs:nfs4_nra=32
			* increase NFS maximum transfer size
			set nfs3:max_transfer_size=1048576
			set nfs4:max_transfer_size=1048576
			* oncrease NFS logical block size
			set nfs:nfs3_bsize=1048576
			set nfs:nfs4_bsize=1048576	

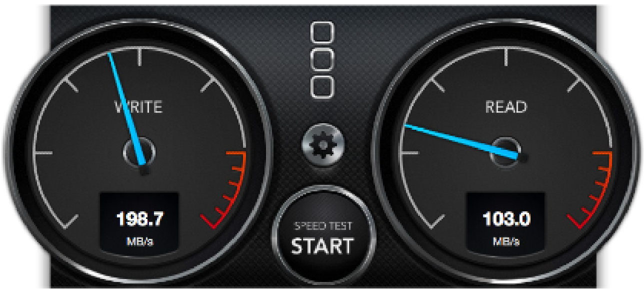
sharectrl NFS lockd_servers	1200	20	1024	1200
/kernel/drv/ ixgbe.conf		tune other nics manually when needed	intr_throttling = 0	
/kernel/drv/ sd.conf	# set my tuning option default #2	-		#2
Action	set only with settings:	reset to default	set base tuning option	set my tuning option default

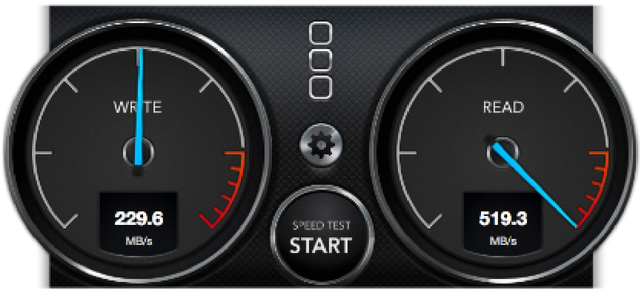
On first modifications, cteare always a BE
Select reboot as action with or without BE to enable new settings.

```
Other tuning aspects on server side:

Nic hardware:     use 10G nics, prefer Intel ex X540 with 10GBase-T with Jumboframes enabled, see napp-it.org/doc/downloads/performance_smb2.pdf
RAM:              absolute minimum on Solaris for a stable OS is 2 GB when not using deduplication. RAM above is used as readcache to increase performance
                  On a SoHo system, 4-8 GB is ok with any poolsize. With many users or very high performance demands even 128-256 GB RAM can be a useful option.
```

NFS results on OSX with „napp-it default tuning“ enabled

For whatever reason, default NFS performance on OSX is weak

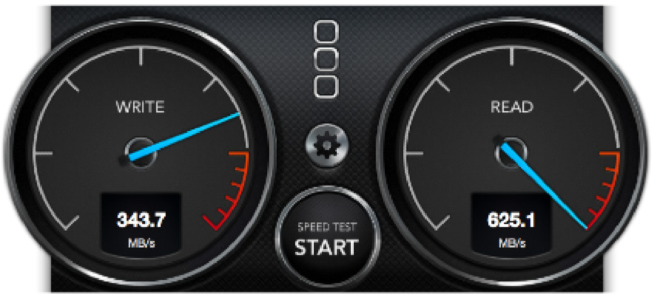




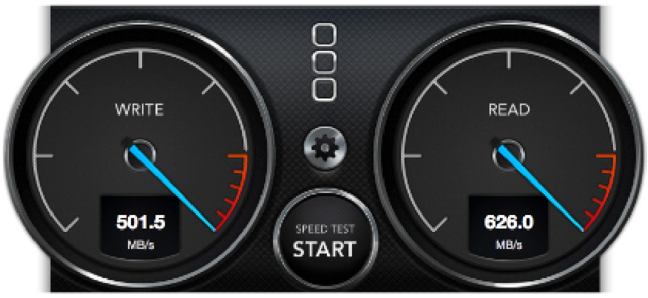
Result:

With default OmniOS settings, NFS performance is bad but with the default tuning options in napp-it: much better

SMB2 results on OSX

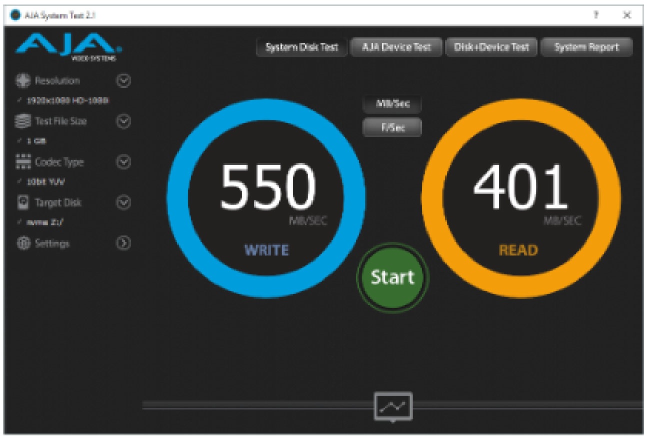


default OmniOS settings
Result: similar read values but much better write values



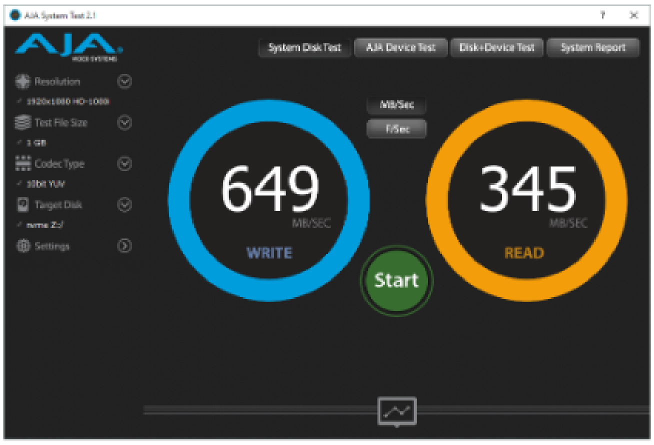
OmniOS with „napp-it basic tuning“ enable

SMB2 results on Windows 10 (MacPro via Bootcamp)



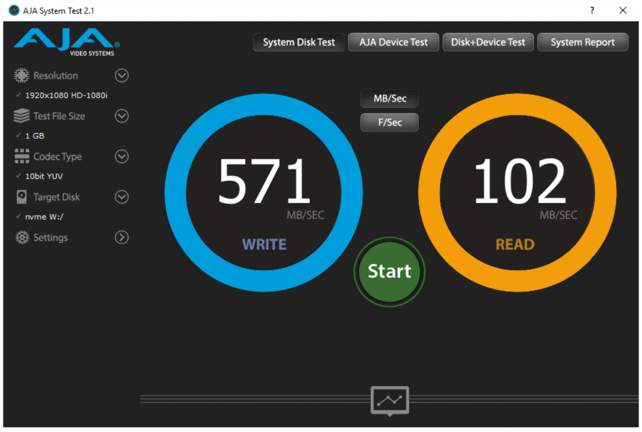
Resolution: 1920x1080 HD-1080i
Test File Size: 1 GB
Codec Type: 10bit YUV
Target Disk: nvme Z:/
Settings

default OmniOS settings,
Windows MTU 1500 and int_throtteling disabled



Resolution: 1920x1080 HD-1080i
Test File Size: 1 GB
Codec Type: 10bit YUV
Target Disk: nvme Z:/
Settings

default OmniOS settings
MTU 9000 and int_throtteling disabled



Resolution: 1920x1080 HD-1080i
Test File Size: 1 GB
Codec Type: 10bit YUV
Target Disk: nvme W:/
Settings

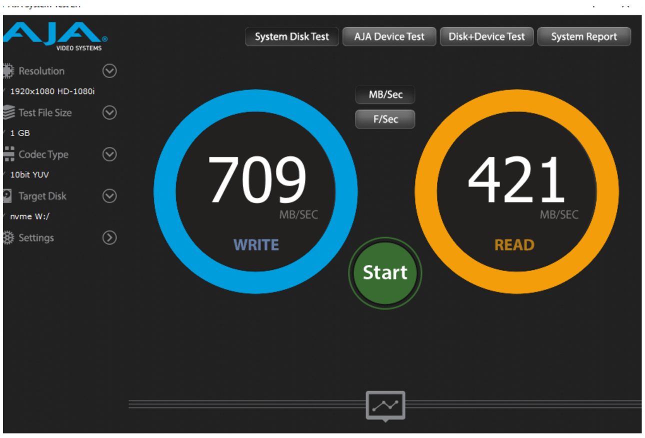
second check with OS defaults

I made a second check with OS-Default.

With MTU=1500 and MTU=900 I got bad readvalues.
I was not able to get the same values than above.

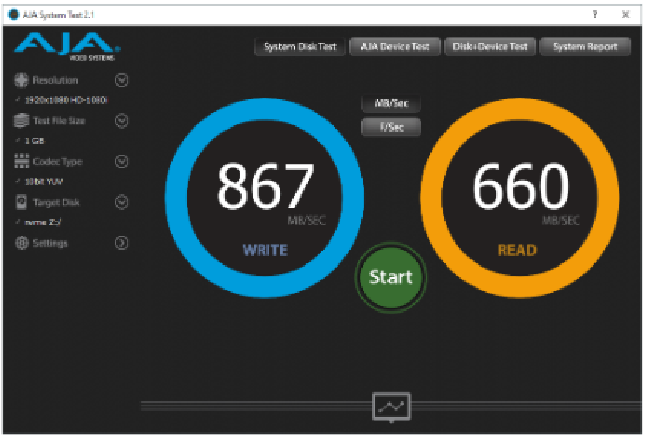
This seems to be related to a weak cable !!
Another cable fixes the problem

The values with 867 MB write/s and 660 MB/ s with
- OmniOS base tunings
- Jumboframes enabled
- int_throttling disabled on Windows are always perfect, does not matter if I reboot or switch settings



Resolution: 1920x1080 HD-1080i
Test File Size: 1 GB
Codec Type: 10bit YUV
Target Disk: nvme W:/
Settings

OmniOS basic tuning
Windows MTU 1500, and int _throtteling disabled



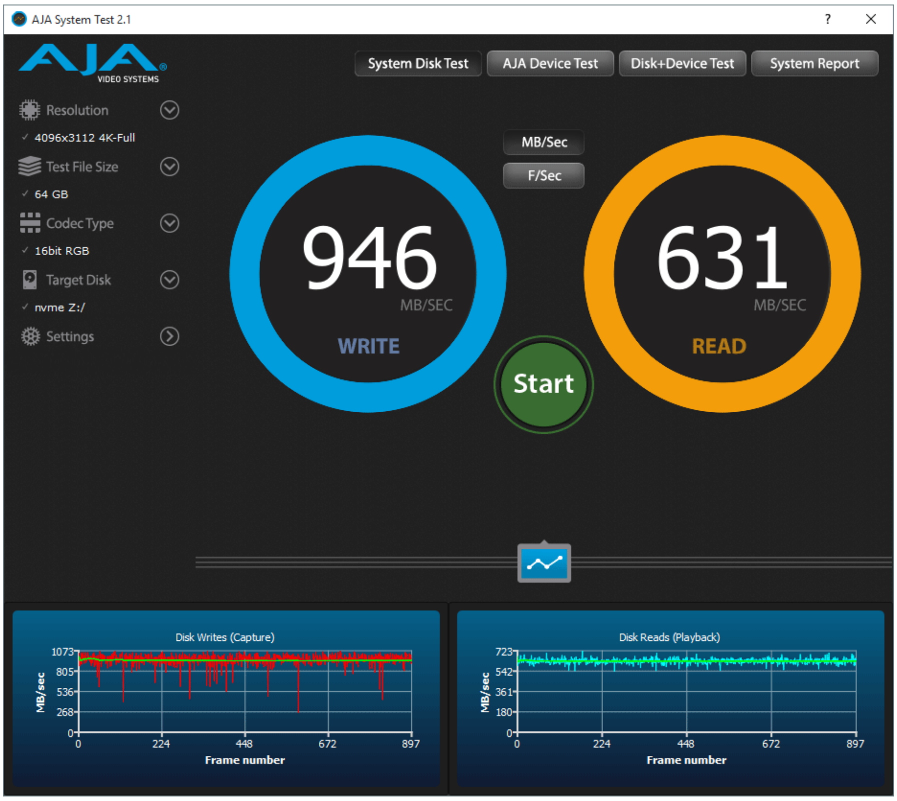
Resolution: 1920x1080 HD-1080i
Test File Size: 1 GB
Codec Type: 10bit YUV
Target Disk: nvme Z:/
Settings

OmniOS basic tuning
Windows MTU 9000, and int _throtteling disabled

Single NVMe vs Enterprise 6G SSD vs Consumer 6G SSD vs 7k rpm 6G disk with large filesize

With the AJA testtool, you can select different videoresolutions, compressors and filesizes where Aja writes a file per frame with a size of about 75 MB per frame with this resolution. The total filesize determins the number of frames. I selected large 4k frames with uncompressed RGB16 and 64 GB filesize to check behaviour over time.

Intel P750 Nvme with large 4k framesize (75MB per frame)



Resolution ✓ 4096x3112 4K-Full

Test File Size ✓ 64 GB

Codec Type ✓ 16bit RGB

Target Disk ✓ nvme Z:/

Settings

946 MB/SEC WRITE

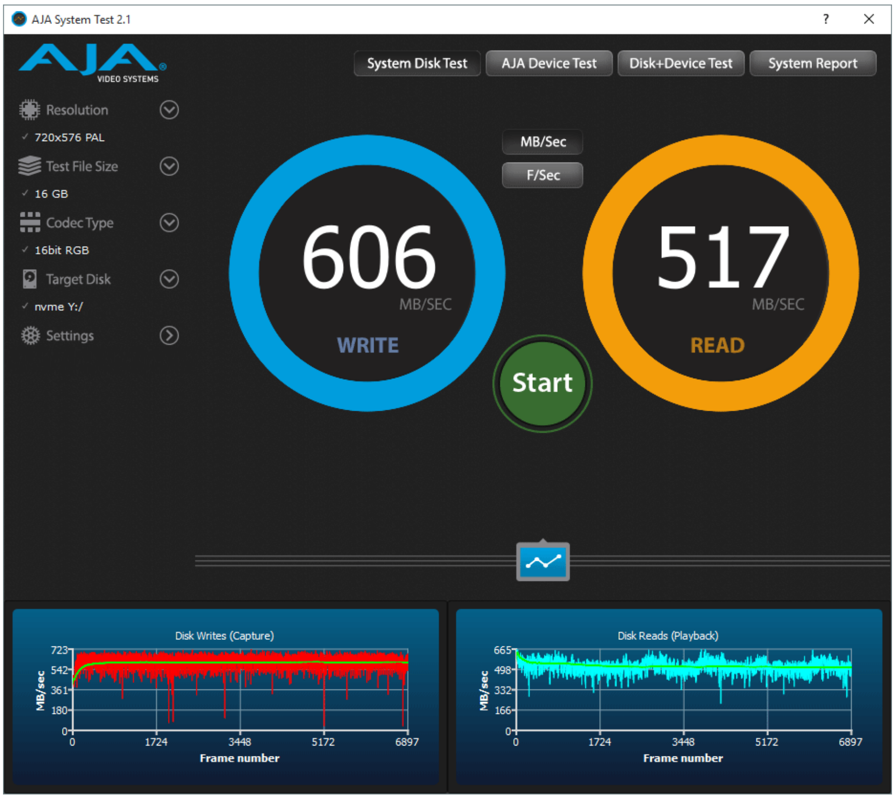
631 MB/SEC READ

Result:

Write transfer starts with full network performance of about 1000 MB/s and remains at that level with some minor drops.

Read performance remains constantly over 600 MB/s

Intel P750 Nvme with small PAL SD framesize (2,5 MB per frame)



Resolution ✓ 720x576 PAL

Test File Size ✓ 16 GB

Codec Type ✓ 16bit RGB

Target Disk ✓ nvme Y:/

Settings

606 MB/SEC WRITE

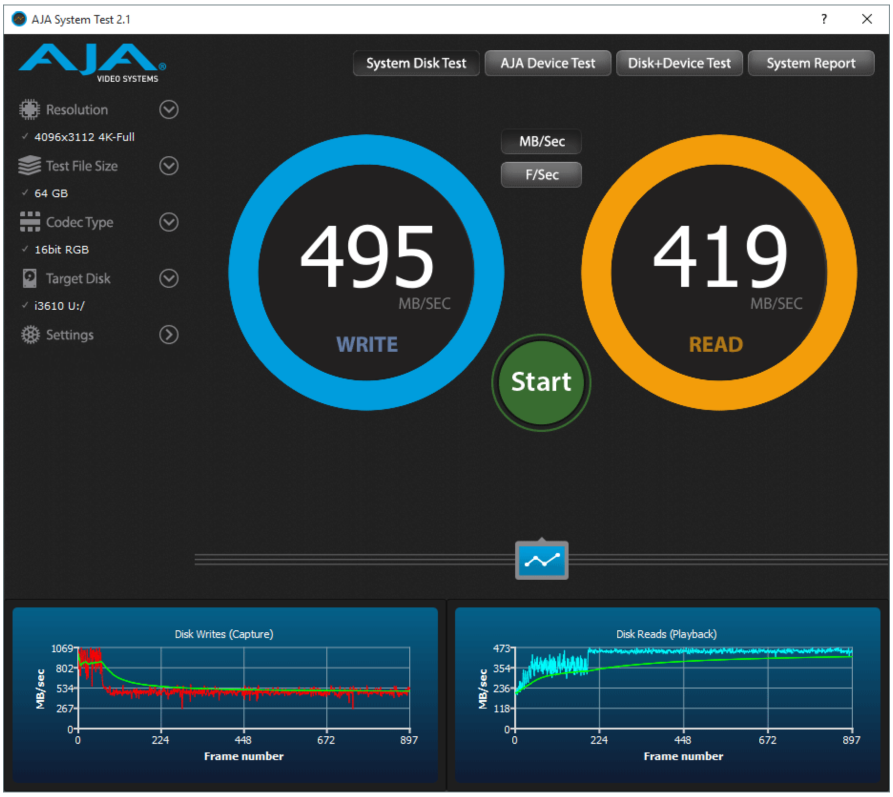
517 MB/SEC READ

Result

During the write of nearly 7000 frames, write performance was constantly at an average of about 600 MB/s

Read is constantly over 500 MB/s what shows the quality of the Intel NVMe disks regarding transfer rate and iops.

Intel S3610 SSD with large 4k framesize (75MB per frame)



AJA System Test 2.1

Resolution ✓ 4096x3112 4K-Full

Test File Size ✓ 64 GB

Codec Type ✓ 16bit RGB

Target Disk ✓ i3610 U:/

495 MB/SEC WRITE

419 MB/SEC READ

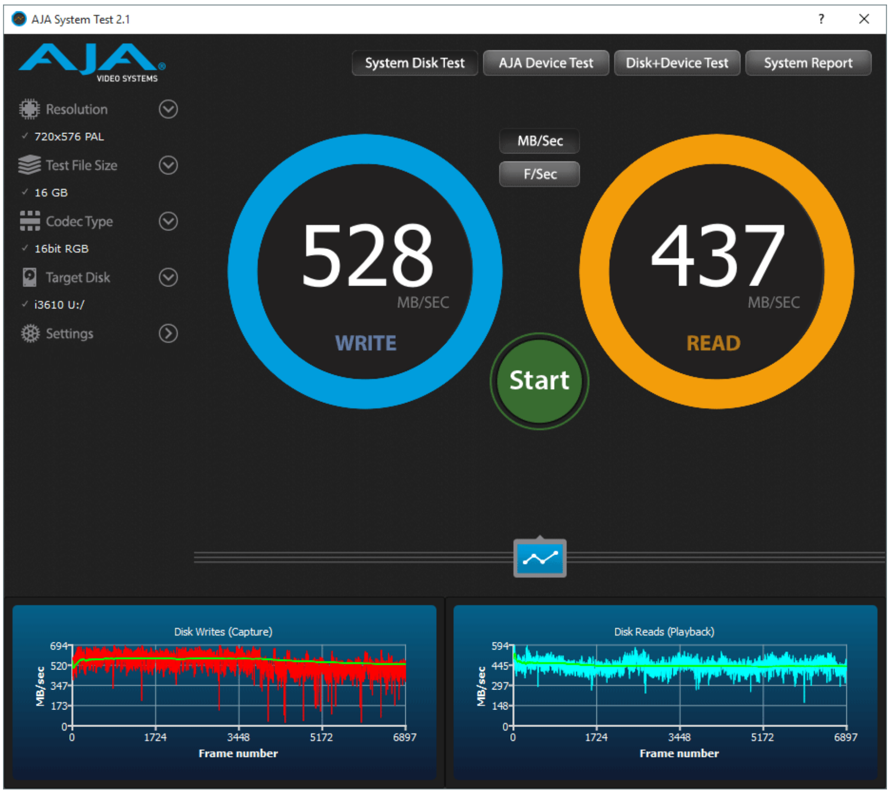
Result:

Write transfer starts with full network performance of about 1000 MB/s and drops after some frames to about 500 MB/s.

The initial peak seams the effect of the ZFS write buffer.

Read performance remains constantly at 470MB/s after an initial lower value.

Intel S3610 SSD with small PAL SD framesize (2,5 MB per frame)



AJA System Test 2.1

Resolution ✓ 720x576 PAL

Test File Size ✓ 16 GB

Codec Type ✓ 16bit RGB

Target Disk ✓ i3610 U:/

528 MB/SEC WRITE

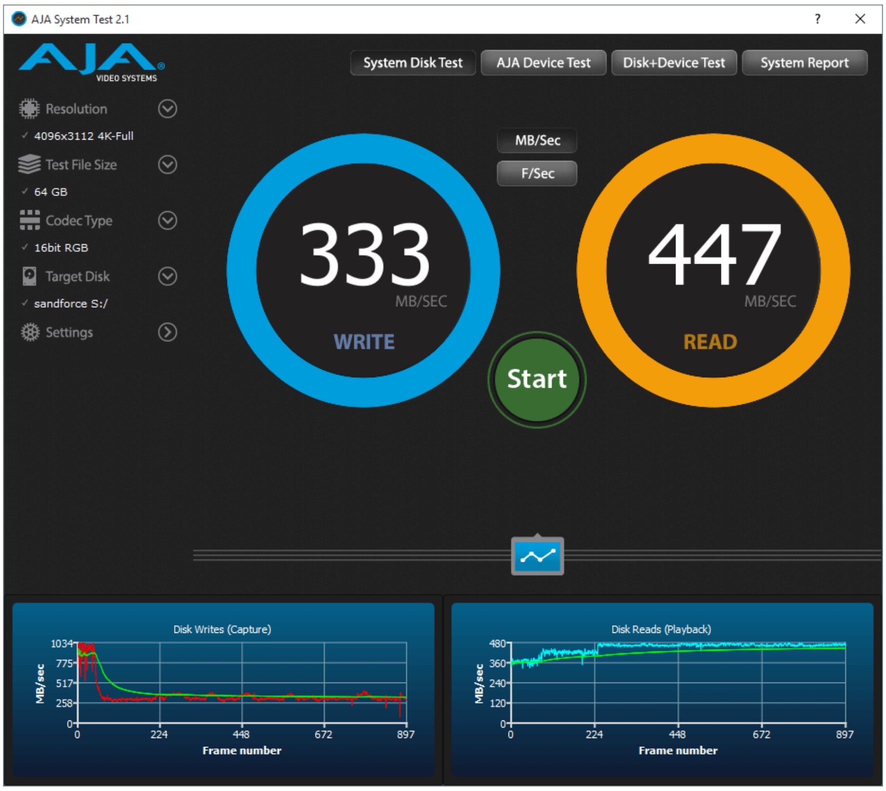
437 MB/SEC READ

Result

During the write of nearly 7000 frames, write performance was constantly at an average of about 500 MB/s (near limit of 6G Sata)

Read remains constantly at about 450 MB/s

Sandforce based SSD with large 4k framesize (75MB per frame)



Resolution
✓ 4096x3112 4K-Full
Test File Size
✓ 64 GB
Codec Type
✓ 16bit RGB
Target Disk
✓ sandforce S:/
Settings

333 MB/SEC WRITE

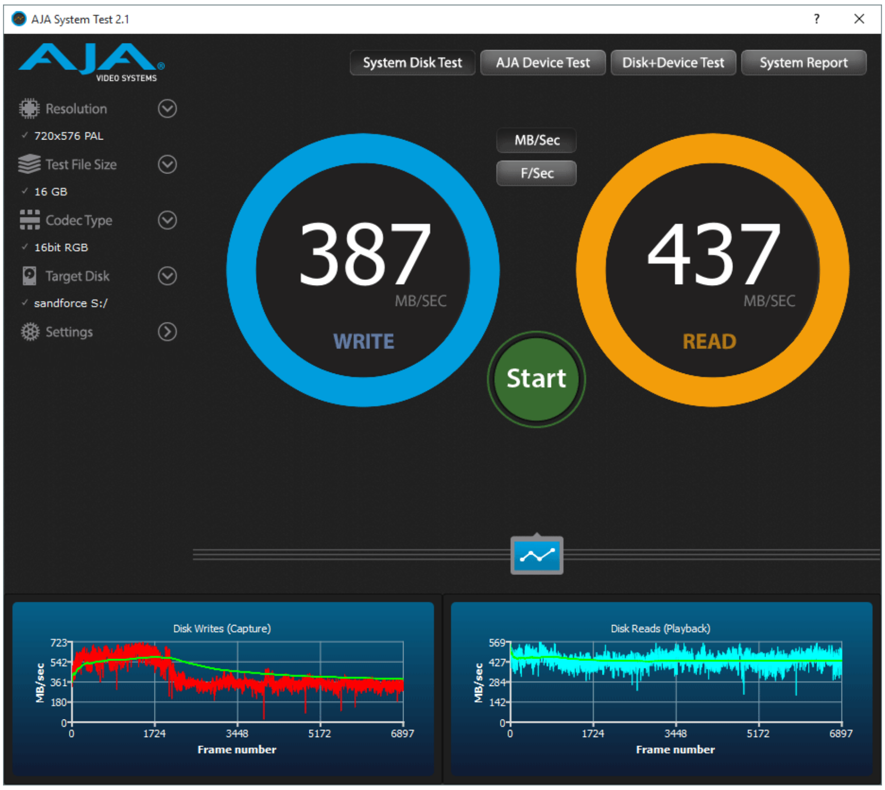
447 MB/SEC READ

Result:

Write transfer starts with full network performance of about 1000 MB/s and drops after some frames to about 250 MB/s. The initial peak seams the effect of the ZFS write buffer.

Read performance remains constantly at about 470MB/s

Sandforce based SSD with PAL SD framesize (2,5 MB per frame)



Resolution
✓ 720x576 PAL
Test File Size
✓ 16 GB
Codec Type
✓ 16bit RGB
Target Disk
✓ sandforce S:/
Settings

387 MB/SEC WRITE

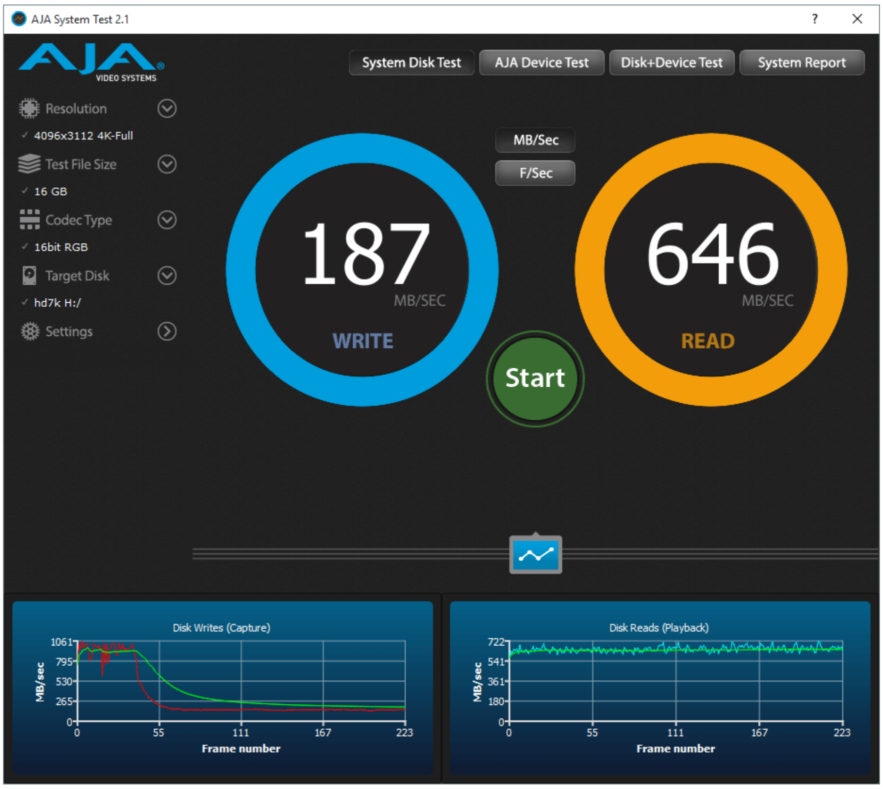
437 MB/SEC READ

Result

During the write of nearly 7000 frames, write performance starts with over 500 MB/s and drops after some frames to about 360 MB/s

Read is constantly at about 450 MB/s .

HGST 2TB disk with 7k rpm and large 4k framesize (75MB per frame)



Resolution
✓ 4096x3112 4K-Full
Test File Size
✓ 16 GB
Codec Type
✓ 16bit RGB
Target Disk
✓ hd7k H:/
Settings

187 MB/SEC WRITE

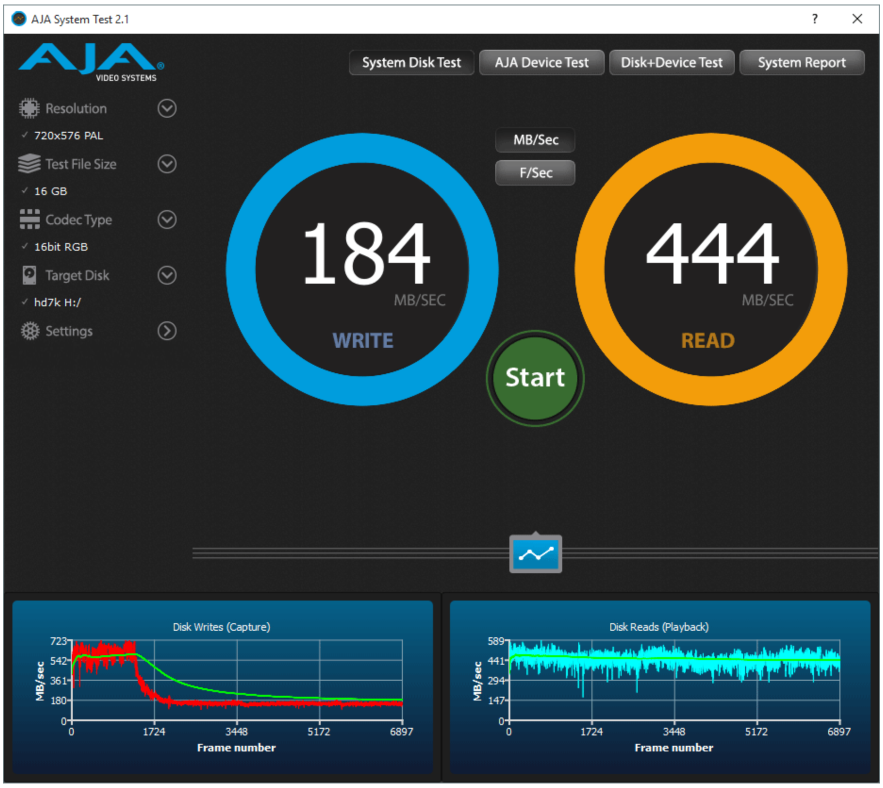
646 MB/SEC READ

Result:

Write transfer starts with full network performance of about 1000 MB/s and drops after some frames to about 150 MB/s. The initial peak seams the effect of the ZFS write buffer.

Read performance remains constantly at the limit of 6G

HGST 2TB disk with 7k rpm with PAL SD framesize (2,5MB per frame)



Resolution
✓ 720x576 PAL
Test File Size
✓ 16 GB
Codec Type
✓ 16bit RGB
Target Disk
✓ hd7k H:/
Settings

184 MB/SEC WRITE

444 MB/SEC READ

Result

During the write of nearly 7000 frames, write performance starts with over 500 MB/s and drops after some frames to about 170 MB/s

Read is constantly at about 450 MB/s .

Fazif of disk types:

Only NVMe and Enterprise SSD can offer a constant high write performance behaviour.
Older SSDs like the Sandforce based Winkom Pro are not as good but offer twice the write performance of disks.

While sequential performance with a single large file on a Raid array is quite similar between SSD and disk pools, the difference is huge when you check with many small files over time. Iops is the difference as an NVMe offers 100000 iops, Enterprise SSDs up to 40000, average SSDs up to 10000 and spindle disks around 100 iops under load.

Fazit

SMB1 vs SMB 2.1

SMB1 on OSX is very slow. SMB 2.1 is a must. With more than 1 GbE use SMB2+ on Windows as well.

MTU 1500 vs MTU 9000 (JumboFrames)

You should use Jumboframes for performance. Performance improvement is up to 20-50%

SMB2 vs NFS

I have had stability problems with NFS on OSX as disconnects happens as well as a wery bad performance of NFS.
Even with Jumboframes the best what I got without any tweaks was 70 MB/s write and 150 MB/s read.
Together with the missing authentication or authorisation of NFS, this is not an option.

Raid-0 vs Mirror vas Raid-Z

For a single user test, differences are minimal as long as your pool is capable of a sequential performance > 1GB/s
In a multi-user/ multi-thread environment, there is a difference as iops performance of raid-Z is lower

NVMe vs SSD vs Spindle based pools

For a single user test, differences are minimal as long as your pool is capable of a sequential performance > 1GB/s
In a multi-user/ multi-thread environment, there is a difference as iops of SSDs is much higher than with spindels.
A single NVMe has nearly no advantage over 3 x 6G SSDs

Solaris 11.3 vs OmniOS 151017

Solaris is on OSX slightly faster than the OmniOS beta. Results with Windows as client are similar.

Windows 8.1/10 vs OSX 10.11

I have not expected the result as OSX is slightly faster on 10G out of the box than Windows (same MacPro hardware)
Other Windows results on a 6core Xeon PC (X540-T1) workstation are similar to the results on the MacPro with Windows. Deactivate Interrupt throttling on Windows and enable Jumboframes (X540 nic) is required.

Server-RAM 40 GB vs 8 GB

RAM is an important factor as reads from RAM-cache are 1000x faster than reads from disk. The reduced load on the disks also improves writes. The more concurrent user or processes the more important is RAM. For a single user 8GB is perfect even with 10GbE. For multiuser or multi process access, use much more RAM in a high performance setup.

OS Tuning

with a base tuning, see napp-it menu System > System tuning: can improve performance up to 20-30%

In General

In the last 10 years disk capacity grows from multi Gigabyte to multi Terabyte with a factor 1000 while network stay at 1Gb/s for years. If you need to copy or backup a disk or pool this can last days or over a week. Modern SSDs can give 10x the performance of 1 GB/s so you really need a 10 GB/s network.

10G Performance depend on settings like Jumboframes and SMB2+. Driver quality and general OS settings are essential. On some configs 10G performance is only minimal better than on an 1G network.

While 1G networks are at about 100 MB/s without special settings, the 10G is in the best case 6-8x as fast but only with settings like SMB2, Jumboframes or disabling interrupt throttling ex in X540 settings on Windows.
I am under the impression that the OS defaults are more optimized for 1G than 10G performance.

Conclusion:
Use 10 GbE, SMB2+ and Jumboframes + some OS tunings if you need 10G performance
Care about OS version and drivers and about some driver settings like the irq throttling in X540 on Windows

Oracle Solaris is out of the box slightly faster than current OmniOS beta.
Both are a huge step forward compared to their predecessor especially on OSX or with 10 GbE.

For single large file and single user the sequential performance is ok with a pool from any sort of disks.
If you need a really high continuous write performace with many small files, you must use enterprise class SSDs

Cabling with 10G Base-T is very critical. Different cables can give 20-30% difference sometimes down to 1 Gb/s - ex 600MB/s write and 105 MB/s read, a new or shorter cable can fix the problem.

about using AJA or other benchmark tools

What to expect
Prior any test, calculate a thumb rule about expected values.
For example, a single harddisk cannot deliver more than 100-200 iops and 200-300 MByte/s
With small files in the outer tracks and a higher fragmentation expect 50-80 MByte/s
If values are outside of the expected range, you have a problem or you test cache qualities not disk qualities.

Cache/RAM effects
If you want to test overall storage quality, use all the RAM as readcache. This is where ZFS performance comes from. as ZFS needs to calculate and prozess more data than other filesystems due the additional checksums. This is the price of the extra security. Optionally enable sequential caching for L2ARC devices, If you want to compare your disks only, not overall storage performance use only 2-4 GB RAM to reduce cache effects.
Usually disable sync-write and enable write back cache for iSCSI devices for best performance.
Enable sync or disable write back cache if you want to test performance on secure powerloss safe writes or a Slog.

Multiuser
Mostly you do not have only a single user on your storage that is viewing a video. For such a use case on 1G networks, any cheap NAS can give up to 100 MByte/s and you do not need high performance storage with striped Raid and a lot of RAM as cache and CPU power. In such a case, data security is the only ZFS advantage.

With AJA, you can test effects of Multi-User use cases quite easily.
Create for example four shares on your pool, start AJA four times and connect each to a different share.
Then start a test on all four instances of AJA simultaniously, best with a large testfile size (16 or 64GB) and compare results with single user values. Mostly you will find, that disk performance fall far below your network performance. To reduce network effects you can use four client computers with a dedicated server nic.

File size
Some tests ex dd create a filestream to test a device without any parallel actions (queuedepth=1)
This is not a realworld scenario and can give the result that a harddisk is faster than a modern SSD..
Use also filesizes that are typical for your workload. Example if you are using the storage for maildate, you have many small files. If you use the storage to edit videodata, you are more interested in the behaviour with large files.
In AJA you can simulate this with the videosize. If you enter a videosize of 720x576 PAL you create many small files while a setting of 4k creates less but larger files. Use 4k videosize if you want to test sequential performance, use PAL if you are interested in io behaviour with small files.

CPU effects
With AJA storage tests you should set 16 bit RGB as video codec
If you use a videocompressor, CPU performance of your **client PC/desktop** is part of the result.
If you enable compress or encryption ex on Solaris, CPU performance of the server is relevant

Performance over time
Especially SSDs are offering best iops performance. On desktop SSDs you find values up to 100000 on the advertisements while enterprise SSDs offer more realistic values of 20000 - 80000 that they can keep under constant load while cheap desktop SSDs can break down to 5000 iops after some time of usage.
If you want to test this effect, you need a test that runs for 1-2 hours and produces a constant write load. You must check AJA graphs of performance over time

Fillrate and usage
Perfomance is quite often a function of fillrate and on SSDs of already used flash cells or overprovisioning.
If you want to compare results use empty disks and do a secure erase on SSDs prior tests

System settings of OmniOS/ Solaris
System settings are always an optimasation for a workload or a system/RAM setup. For an iSCSI, SMB or NFS storage system test with RAM > 1-2GB, use the napp-it default tuning option in menu System to increase tcp buffers as the defaults are optimized for very low RAM and use a current OmniOS with SMB 2.1 enabled (151017up).
LZ4 compress enabled gives mostly better performance values than compress disabled as read/writes are reduced.

Conclusion
If you want to compare two storage systems regarding these questions:
- use more that one AJA instance or use more clients simultaniously
- select a videosize according your usecase, ex 720 x 576 PAL for behaviour with small files
- select a large testfile ex 64GB for a long running test, check AJA read/write graph over time
- select a videocompressor to add some CPU load

Now you can compare the results of two storage systems or two storage layouts/ configurations.

Setup configuration ideas

For multi-user video-editing or other high performance lab setups you need high performance and/or high capacity. If you need both with the same storage, this can be quite expensive as you not only need the storage but also high performance disks and networking and can be very load (requires a server room). Some typical configurations:

1. High Performance (4k) lab or desktop storage, capacity up to several Terabyte for current working data.

You can use a a silent desktop case with a 10 x Sata mainboard with up to seven PCI-e slots, example a SuperMicro X10SRL-F where you can insert for example 3 - 5 Intel P750 NVMe disks with up to 1,2 TB each. With a Raid-Z1 this gives you up to 4,8 TB of ultrafast storage, capable to serve several 4k video editing clients simulatiously with the same data.

As an additional option add up to 9 SSDs like 3,8TB Samsung PM 863 to onboard Sata what can give you up to 26 TB in a Raid-Z2 setup. Sata nr 10 is needed for your bootdisk.

Such a storage can be built ultra silent so you can place it beside the clients. If you use the remaining two slots for 2 x Dual 10G nics, you can connect 4 x 10G clients directly plus 2 x 1G clients or use one as uplink without the need of an extra expensive and loud 10G switch as you can enable bridgung between the nics. This will allow any client to see the server, each other and everything that is connected on the 1G uplink. Up from napp-it Pro 2016.05 dev you can enable the bridging function (OmniOS acts like a normal network switch) in menu System > Network Eth.

For more clients, you can use a 10G switch. As a future option there are the upcoming Intel X710 QSFP+ network adapters. They offer up to 2 x 40G QSFP+ or 8 x 10G SFP+ per nic. Drivers for OmniOS are on the way. With such a card, you can connect up to 8 x 10G clients per nic without the need of a switch. At the moment, QSFP+ optic tranceivers are quite expensive but cheap copper QSFP+ to 4 x SFP+ cables with up to 5m should be available soon.

2. Affordable very high capacity storage

You can use for example a 24 x 3,5" case like SuperMicro SC846BA-R1K28B with a a SuperMicro X10SRL-F and 3 x LSI/AVAGO/Broadcom HBAs with IT mode firmware like the model 9207-8i (expanderless, best for Sata disks). If you insert 24 x 10 TB disks in a 4 x 6 Raid Z2 or 20 disks in a 2 x 10 disk raid z2 (half the ipops) you achieve 160 TB usable. Add enough RAM and at least one NVMe as L2ARV and you have affordable but quite fast high capacity storage. This case cannot be placed near a desktop (too loud).

If capacity should go beyond, you can use cases like the SuperMicro https://www.supermicro.nl/products/chassis/4U/?chs=946 with expander and up to 90 SAS disks per case.

3. Affordable high performance/ high capacity storage

You can use a SuperMicro barebone like the http://www.supermicro.com/products/system/2U/2028/SSG-2028R-ACR24L.cfm with up to 24 x 2,5" disks. If you insert 20 x 3,8 TB SSDs like the Samsung PM 863 in a 2 x 10 SSD Raid-Z2 you have up to 60 TB ultrafast SSD only storage. This case cannot be placed near a desktop (too loud).

OSX 10.12 Sierra

I have repeated the benchmarks with newer OSX 10.11.5 or newest OSX 10.12 releases and found that 10G performance was not as good as before. I have tried some tuning recomendations like signing_required=no that are suggested on the internet but was not able to get an SMB performance better than around 200-300 MB/s - far below the former 600-800 MB/s.

Effect of Readcache (Cache for Metadate and/or Data)

ZFS is known to be fast with its advanced readcaches
But how fast or slow is ZFS without caches?

For these tests I use **filebench/ singlestreamread** (napp-it menu Pools > Benchmark > Filebench):
Caches can be enabled/disabled in napp-it menu Pools >

Filer with a 60% fillrate, 23% fragmentation, 2 x 6 disk raid-z2 vdev of HGST HUS724040AL,
filer under decent load from students during test. The term mb/s in filebench means MegaByte/s.

all caches on
IO Summary: 157739 ops, 5257.764 ops/s, (5258/0 r/w), **5256.7mb/s**, 211us cpu/op, 0.2ms latency
19437: 34.518: Shutting down processes

all caches off
IO Summary: 13757 ops, 458.554 ops/s, (459/0 r/w), **458.5mb/s**, 2168us cpu/op, 2.2ms latency
27837: 35.348: Shutting down processes read drops to 458 MB/s

Now values of single disks
A singlestreamread on a single disk HGSH HE8 Ultrastar

IO Summary: 2554 ops, 85.130 ops/s, (85/0 r/w), **85.1mb/s**, 3578us cpu/op, 11.7ms latency
13519: 41.172: Shutting down processes

the same but with metadata caching on (data caching off)

IO Summary: 2807 ops, 93.565 ops/s, (94/0 r/w), **93.5mb/s**, 3317us cpu/op, 10.7ms latency
4776: 41.229: Shutting down processes
helps a little. With higher fragmentation, the difference may become bigger

and a single **Intel S3610-480** (all chaches off)

IO Summary: 8874 ops, 295.792 ops/s, (296/0 r/w), 295.7mb/s, 2216us cpu/op, 3.4ms latency
16725: 37.100: Shutting down processes

and finally a **very old and slow WD green 1TB** (all chaches off)

IO Summary: 94 ops, 3.133 ops/s, (3/0 r/w), **3.1mb/s,** 67645us cpu/op, 304.0ms latency
17184: 48.344: Shutting down processes

same WD Green with a **singlestreamwrite and sync disabled** (write cache on)

IO Summary: 6318 ops, 210.594 ops/s, (0/211 r/w), **210.6mb/s,** 3791us cpu/op, 4.7ms latency
21691: 31.116: Shutting down processes

same WD Green with a **singlestreamwrite and sync set to always** (write cache off for sync write)

IO Summary: 673 ops, 22.432 ops/s, (0/22 r/w), **22.4mb/s**, 14577us cpu/op, 44.5ms latency
27214: 31.261: Shutting down processes ok.

now a low rpm WD RE4 (low power 2TB enterprise disks)

IO Summary: 463 ops, 15.433 ops/s, (15/0 r/w), 15.4mb/s, 205954us cpu/op, 64.5ms latency
24202: 47.712: Shutting down processes

ok.

Result with low RAM

The result clearly shows. that ZFS performance especially on read highly depend on its caches, propably more than on other filesystems. So especially with very low RAM and disks with very low iops especially read performance of ZFS seems really worse. The faster the disks the better the values. With readcache performance can increase dramatically. High iops is a key value for performance. With the WD green, results with 3,1 MB/s are dramatically bad while results with the newer HGST 8TB disk and 85MB/s and S3610 and 295 MB/s are ok. I am as well astonished about this very bad values on a slow disk with low iops values and cache disabled as mostly you do tests with all caches on to check overall system performance.
On writes, you always use the write cache that optimizes writes so default write values are much better. With sync=on you see similar weak write values as well with low iops disks. On a regular load, ZFS will hide these bad behaviour mostly with the readcache. So the result: avoid low iops disks and use as much RAM as possible for readcache.

Performance considerations regarding RAM: Summary

Aspect 1: sequential raw disk performance

This scales:

- on mirrors sequential read performance	= n x number of disks
- on mirrors sequential write performance	= n x number of vdevs
- on Raid-Z sequential read/write performance	= n x number of datadisks
ex a raid-z2 of 6 disks has 4 datadisks (150MB/s each)	= 4 x 150 MB/s = 600 MB/s

Aspect 2: iops raw disk performance

This scales:

- on mirrors iops read performance	= n x number of disks
- on mirrors iops write performance	= n x number of vdevs
- on Raid-Z iops read/write performance	= n x number of datadisks
ex a dual raid-z2 pool (100 iops/ each)	= 2 x 100 iops = 200 iops/s

Aspect 3: RAM effects

A Copy On Write (CoW) filesystem like ZFS does not place data sequentially on disk but spread them over the disk with a higher fragmentation than on non-CoW filesystems. This is the price of the much higher data security of ZFS. To compensate or overcompensate this, ZFS comes with one of the most advanced RAM based or SSD based cache strategies. With enough RAM most reads are delivered not from disk but from cache to avoid limited disk performances. On writes all small and slow random writes are collected in a rambased write cache and written after a few seconds as a large and fast sequential write.

RAM requirements:
- A Solarish OS requires around 2 GB RAM for stable operation with any poolsize
- 10% of RAM up to 4 GB RAM is used for the write cache

- RAM above is dynamically used as readcache (when not required otherwise) for metadata and small random reads.

total RAM suggestions:

A SoHo filer with a few users or a media server:	4-8 GB RAM
A Filer for VM use or more users	8-32 GB RAM

Special use cases:

If you want to cache nearly all data in RAM	1% of active data required to cache metadata + all random data. This can mean 32-256 GB RAM A rule of thumb: 1-2 GB RAM/TB data
If you want to use realtime dedup	add 2-5 GB RAM per TB dedup data
If you want to cache sequential video data	add a fast NVMe as L2Arc and enable sequential caching size of L2Arc should not exceed 5-10 x RAM

Attention: If you use an Slog device:
An Slog is not a write cache, but a logdevice for a secure write behaviour like to a cache+battery on hardwareraid