FreeNAS 9.2 vs 9.3

By J.Bronka [Kobetsu]

Table of Contents

Introduction

Reason to create this document and compare more detailed FN9.2 and FN9.3 was great mistake I did when without tests replace FN9.2.1.5 with new version FN9.2.1.8 on production server.

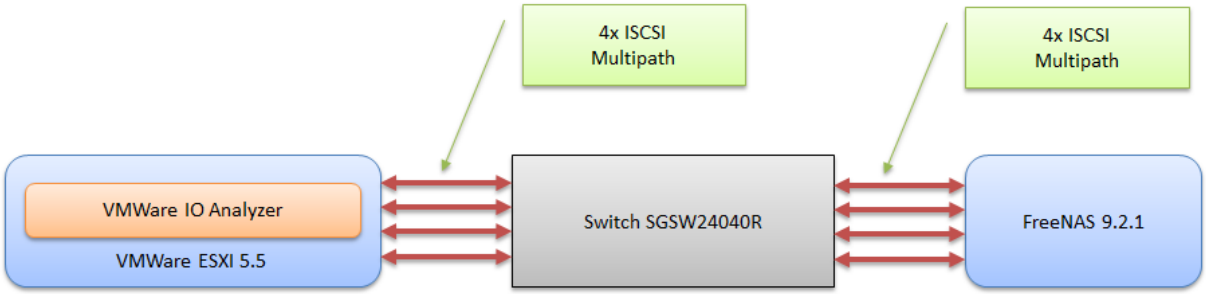
All works but number of issues about performance degradation forced me to go back to FN9.2.1.5.

With FN9.2.1.8 the same problems were with experimental kernel ISCSI target as well as traditional.

Benefits of FN9.3 as full VAAI implementation are interesting to consider migration… but… performance need to be checked.

I. Configurations

I.1 FreeNas 9.2 Test Configuration



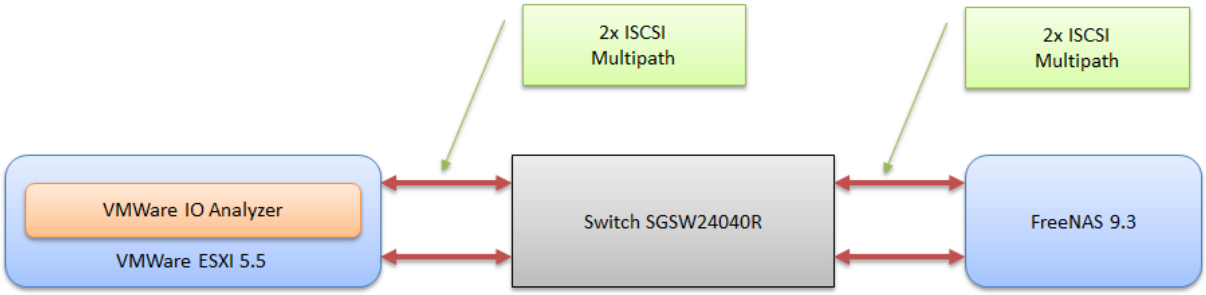
VMWare ESXI:

- Host: VMWare ESXI 5.5
- Test VM: VMWare Appliance IO Analyzer
- 1x XEON E5-2667 v3 / 3.2 GHz
- SuperMicro X10DRi
- 64GB RAM
- 2x NIC I350 / 1GBit (JUMBO 9000)

FreeNas:

- FreeNAS-9.2.1.5-RELEASE-x64
- I7 4820K / 3.7Ghz
- 64GB RAM
- 2x NIC Intel 82574L / 1GBit (JUMBO 9000)
- RAID LSI 9260-16i (There are no support for JBOD but each of disk is working as RAID0)
- RAID 10 (14x Western Digital Red 2TB [WD20EFRX])
- L2ARC (SSD)

I.2 FreeNas 9.3 Test Configuration



VMWare ESXI:

- Host: VMWare ESXI 5.5
- Test VM: VMWare Appliance IO Analyzer
- 1x XEON E5-2667 v3 / 3.2 GHz

- SuperMicro X10DRi
- 64GB RAM
- 2x NIC I350 / 1GBit (JUMBO 9000)

FreeNas:

- FreeNAS-9.3-STABLE-201412240734
- I7 3770K / 3.5Ghz
- 32GB RAM
- 2x NIC Intel 82574L / 1GBit (JUMBO 9000)
- SATA onboard controller

I.3 Disks

The following disks were tested in several configurations:

- Disk Western Digital 1TB Caviar Black 7200 64MB SATAIII (WD1002FAEX)
- Disk Western Digital Red 2TB (WD20EFRX)

II. Test cases

The following cases were tested (regarding disks configuration):

- **FreeNAS 9.2**
  - **Case FN9.2#1**
    - Virtual Disk 128GB, eager zeroed, filled with random values before test (dd if=/dev/random ….)
    - Lun 500GB over RAID 10 (14x Western Digital Red 2TB [WD20EFRX])
    - ISCSI with the same settings in initiator/target configuration:
      - FirstBurstLength = 256KB
      - MaxBurstLength = 256KB
      - MaxRecvDataSeqLength = 128KB
      - InitR2T = NO
      - ImmedaiteData = YES
      - DelayedACK = YES
    - Additionally initiator has default VMware ESXi value of MaxOutstandingR2T=1
    - No compression
  - **Case FN9.2#2**
    - Virtual Disk 128GB, eager zeroed, filled with random values before test (dd if=/dev/random ….)
    - Lun 500GB over RAID 10 (14x Western Digital Red 2TB [WD20EFRX])
    - ISCSI with the same settings in initiator/target configuration:
      - FirstBurstLength = 1MB
      - MaxBurstLength = 1MB
      - MaxRecvDataSeqLength = 1MB
      - InitR2T = NO
      - ImmedaiteData = YES
      - DelayedACK = YES
    - Additionally initiator has default VMware ESXi value of MaxOutstandingR2T=1
    - No compression
  - **Case FN9.2#3**
    - Virtual Disk 128GB, eager zeroed, filled with random values before test (dd if=/dev/random ….)
    - Lun 500GB over RAID 10 (14x Western Digital Red 2TB [WD20EFRX])
    - ISCSI with the same settings in initiator/target configuration:
      - FirstBurstLength = 64KB (default for FN9.2)
      - MaxBurstLength = 256KB (default for FN9.2)
      - MaxRecvDataSeqLength = 256KB (default for FN9.2)
      - InitR2T = NO
      - ImmedaiteData = YES
      - DelayedACK = YES
    - Additionally initiator has default VMware ESXi value of MaxOutstandingR2T=1

        - No compression
- **FreeNAS 9.3**
    - **Case FN9.3#1**
        - Virtual Disk 128GB, eager zeroed, filled with random values before test (dd if=/dev/random ….)
        - Lun 500GB over single HDD Western Digital 1TB Caviar Black 7200 64MB SATAIII (WD1002FAEX)
        - Default configuration of ISCSI in FreeNas
        - Compression LZ4
        - Initiator ISCSI settings:
            - FirstBurstLength = 256KB
            - MaxBurstLength = 256KB
            - MaxRecvDataSeqLength = 128KB
            - InitR2T = NO
            - ImmedaiteData = YES
            - DelayedACK = YES
            - Additionally initiator has default VMware ESXi value of MaxOutstandingR2T=1
    - **Case FN9.3#2**
        - Virtual Disk 128GB, eager zeroed, filled with random values before test (dd if=/dev/random ….)
        - Lun 500GB over RAID1 (2x HDD Western Digital 1TB Caviar Black 7200 64MB SATAIII (WD1002FAEX))
        - Default configuration of ISCSI in FreeNas
        - Compression LZ4
        - Initiator ISCSI settings:
            - FirstBurstLength = 256KB
            - MaxBurstLength = 256KB
            - MaxRecvDataSeqLength = 128KB
            - InitR2T = NO
            - ImmedaiteData = YES
            - DelayedACK = YES
            - Additionally initiator has default VMware ESXi value of MaxOutstandingR2T=1
    - **Case FN9.3#3**
        - Virtual Disk 128GB, eager zeroed, filled with random values before test (dd if=/dev/random ….)
        - Lun 500GB over RAID10 (4x HDD Western Digital 1TB Caviar Black 7200 64MB SATAIII (WD1002FAEX))
        - Default configuration of ISCSI in FreeNas
        - Compression LZ4
        - Initiator ISCSI settings:
            - FirstBurstLength = 256KB
            - MaxBurstLength = 256KB
            - MaxRecvDataSeqLength = 128KB

        - InitR2T = NO
        - ImmedaiteData = YES
        - DelayedACK = YES
        - Additionally initiator has default VMware ESXi value of MaxOutstandingR2T=1
    - **Case FN9.3#4**
        - Virtual Disk 128GB, eager zeroed, filled with random values before test (dd if=/dev/random ….)
        - Lun 500GB over RAID1 (2x Disk Western Digital Red 2TB (WD20EFRX))
        - Default configuration of ISCSI in FreeNas
        - Compression LZ4
        - Initiator ISCSI settings:
            - FirstBurstLength = 256KB
            - MaxBurstLength = 256KB
            - MaxRecvDataSeqLength = 128KB
            - InitR2T = NO
            - ImmedaiteData = YES
            - DelayedACK = YES
            - Additionally initiator has default VMware ESXi value of MaxOutstandingR2T=1
    - **Case FN9.3#5**
        - Virtual Disk 128GB, eager zeroed, filled with random values before test (dd if=/dev/random ….)
        - Lun 500GB over single Disk Western Digital Red 2TB (WD20EFRX)
        - Default configuration of ISCSI in FreeNas
        - Compression LZ4
        - Initiator ISCSI settings:
            - FirstBurstLength = 256KB
            - MaxBurstLength = 256KB
            - MaxRecvDataSeqLength = 128KB
            - InitR2T = NO
            - ImmedaiteData = YES
            - DelayedACK = YES
            - Additionally initiator has default VMware ESXi value of MaxOutstandingR2T=1
    - **Case FN9.3#6**
        - Virtual Disk 128GB, eager zeroed, filled with random values before test (dd if=/dev/random ….)
        - Lun 500GB over RAID10 (6x HDD Western Digital 1TB Caviar Black 7200 64MB SATAIII (WD1002FAEX))
        - Default configuration of ISCSI in FreeNas
        - Compression LZ4
        - Initiator ISCSI settings:
            - FirstBurstLength = 256KB
            - MaxBurstLength = 256KB

- MaxRecvDataSeqLength = 128KB
- InitR2T = NO
- ImmedaiteData = YES
- DelayedACK = YES
- Additionally initiator has default VMware ESXi value of MaxOutstandingR2T=1

III. Tests methodology

1. Duration of every test = 120s.
2. Every test executed 5 times to calculate average value
3. Each test case was examined against 12 tests:
   - 512K_RW_RANDOM

     This test will generate 512k random IO with 50% probability of Read IO. Number of Outstanding IO is set to 16.
   - 256K_RW_RANDOM

     This test will generate 256k random IO with 50% probability of Read IO. Number of Outstanding IO is set to 16.
   - 128K_RW_RANDOM

     This test will generate 128k random IO with 50% probability of Read IO. Number of Outstanding IO is set to 16.
   - 64K_RW_RANDOM

     This test will generate 64k random IO with 50% probability of Read IO. Number of Outstanding IO is set to 16.
   - 32K_RW_RANDOM

     This test will generate 32k random IO with 50% probability of Read IO. Number of Outstanding IO is set to 16.
   - 16K_RW_RANDOM

     This test will generate 16k random IO with 50% probability of Read IO. Number of Outstanding IO is set to 16.
   - 8K_RW_RANDOM

     This test will generate 8k random IO with 50% probability of Read IO. Number of Outstanding IO is set to 16.
   - 4K_RW_RANDOM

     This test will generate 4k random IO with 50% probability of Read IO. Number of Outstanding IO is set to 16.
   - MAX_READ_THOUGHPUT

     This test will generate 512k sequential IO with 100% probability of Read IO. Number of Outstanding IO is set to 32. By sequentially reading blocks of large size we intend to mimic workloads similar to streaming video. This configuration reports the maximum throughput for your system.
   - MAX_WRITE_THOUGHPUT

     This test will generate 512k sequential IO with 0% probability of Read IO. Number of Outstanding IO is set to 32. By sequentially writing blocks of large size we intend to mimic workloads similar to video surveillance, video uploading and file backups. This configuration reports the maximum write throughput for your system.
   - MAX_READ_IOPS

     This test will generate 512b sequential IO with 100% probability of Read IO. Number of Outstanding IO is set to 32. By sequentially reading blocks of small size we intend to stress the storage controller cache. This configuration reports the maximum IOPS for your system.
   - MAX_WRITE_IOPS

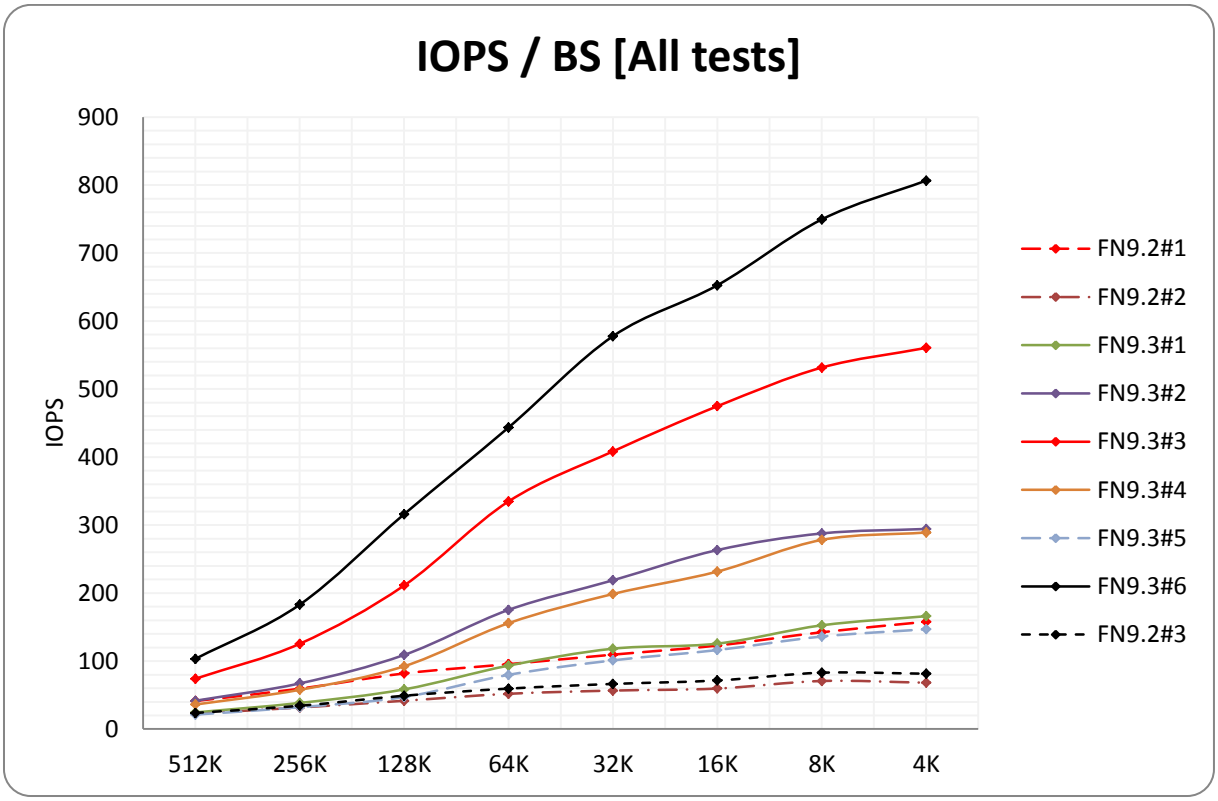
     This test will generate 512b sequential IO with 0% probability of Read IO. Number of Outstanding IO is set to 32. By sequentially writing blocks of small size we intend to stress the backend disks or the write cache in large storage boxes. This configuration reports the maximum write IOPS for your system.

4. Two identical tests (listed in point 3 above) were never started one after the other
5. No other load on FreeNas as well as on ESXI
6. Tested virtual disk size = 120GB
7. Tested virtual disk type eager zeroed
8. Tested virtual disk fully filled with random data

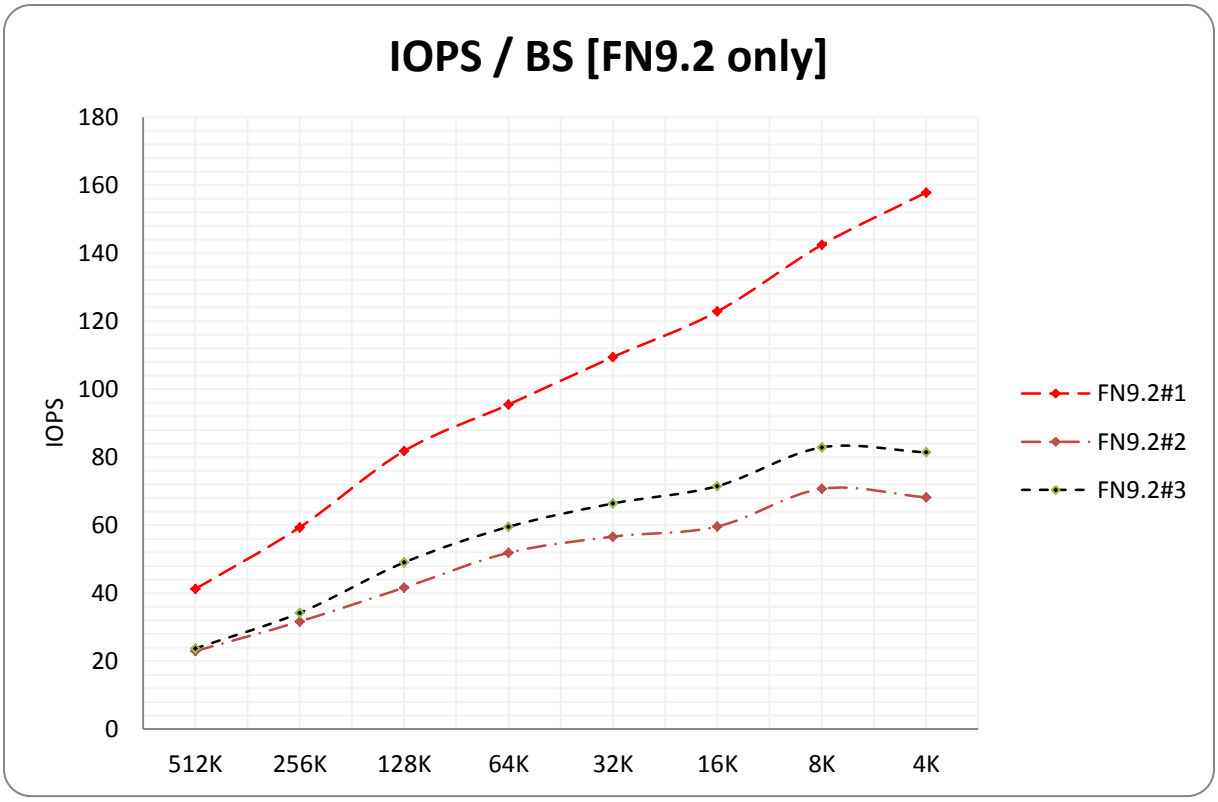
IV. Results

IV.1 IOPS vs Block Size

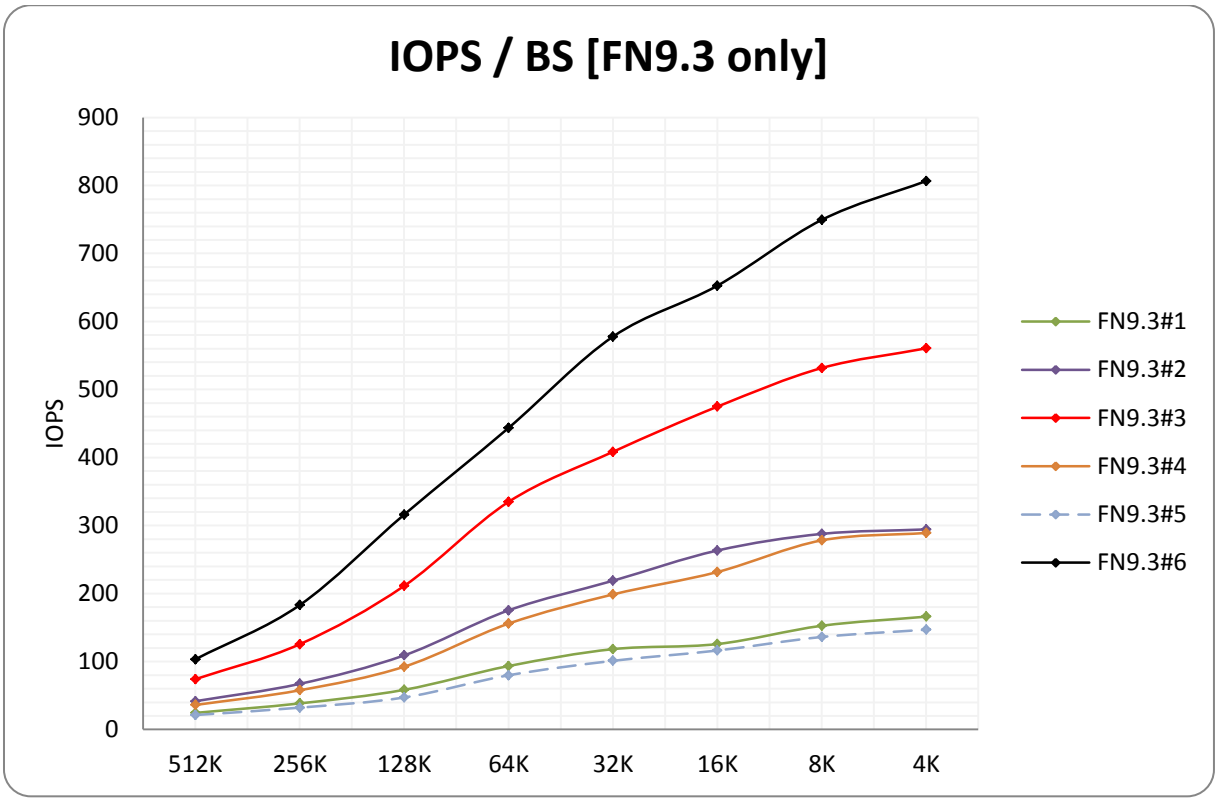
BS	FN9.2#1 IOPS	FN9.2#2 IOPS	FN9.2#3 IOPS	FN9.3#1 IOPS	FN9.3#2 IOPS	FN9.3#3 IOPS	FN9.3#4 IOPS	FN9.3#5 IOPS	FN9.3#6 IOPS
512K	41,208	22,942	23,726	24,452	41,608	73,964	36,072	21,304	103,15
256K	59,348	31,662	34,212	38,402	67,122	125,406	57,684	32,152	182,954
128K	81,796	41,61	49,014	58,322	108,996	211,136	92,366	47,116	315,956
64K	95,454	51,87	59,504	93,318	175,216	334,732	155,692	79,754	443,428
32K	109,434	56,616	66,364	118,182	218,724	408,312	198,624	101,022	577,668
16K	122,83	59,59	71,48	125,69	263,236	474,752	231,604	116,186	652,654
8K	142,414	70,686	82,864	152,54	287,642	531,572	278,174	136,068	749,448
4K	157,824	68,16	81,396	165,964	294,306	560,538	289,062	146,792	806,594



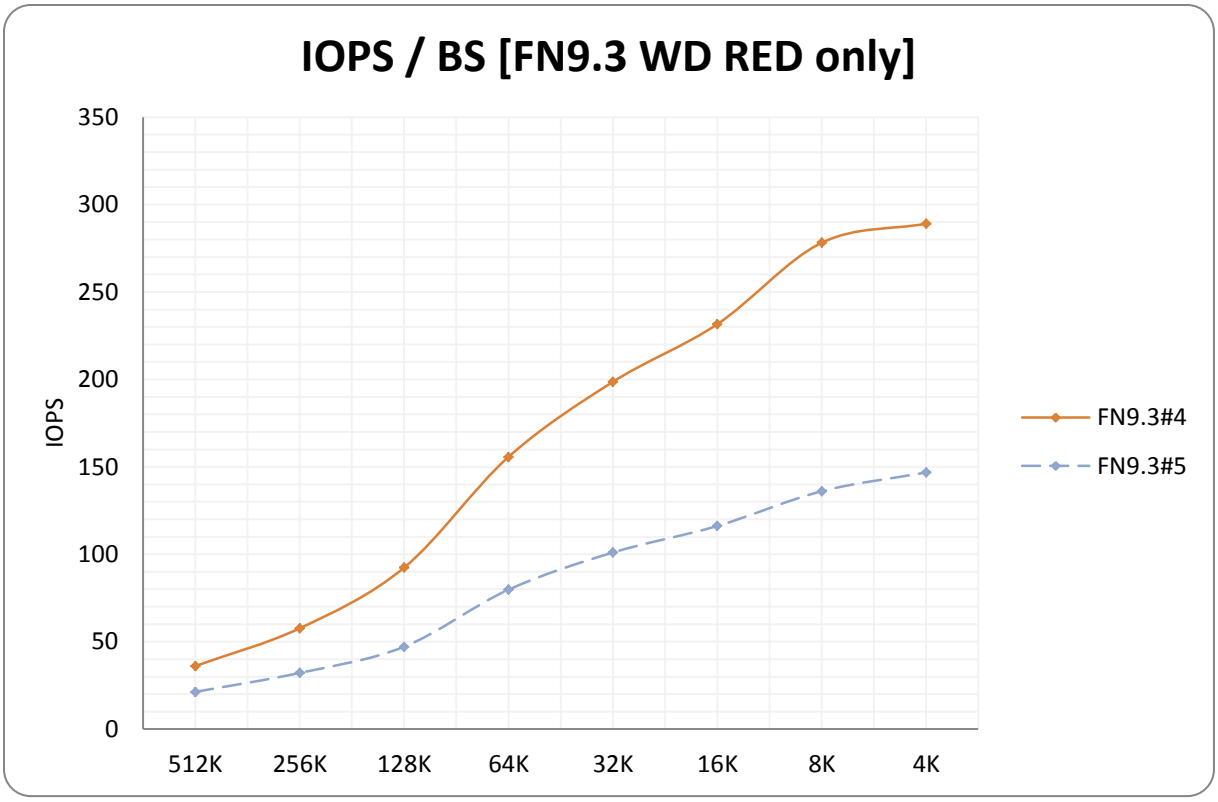
Picture 1: All result from table above on the same chart



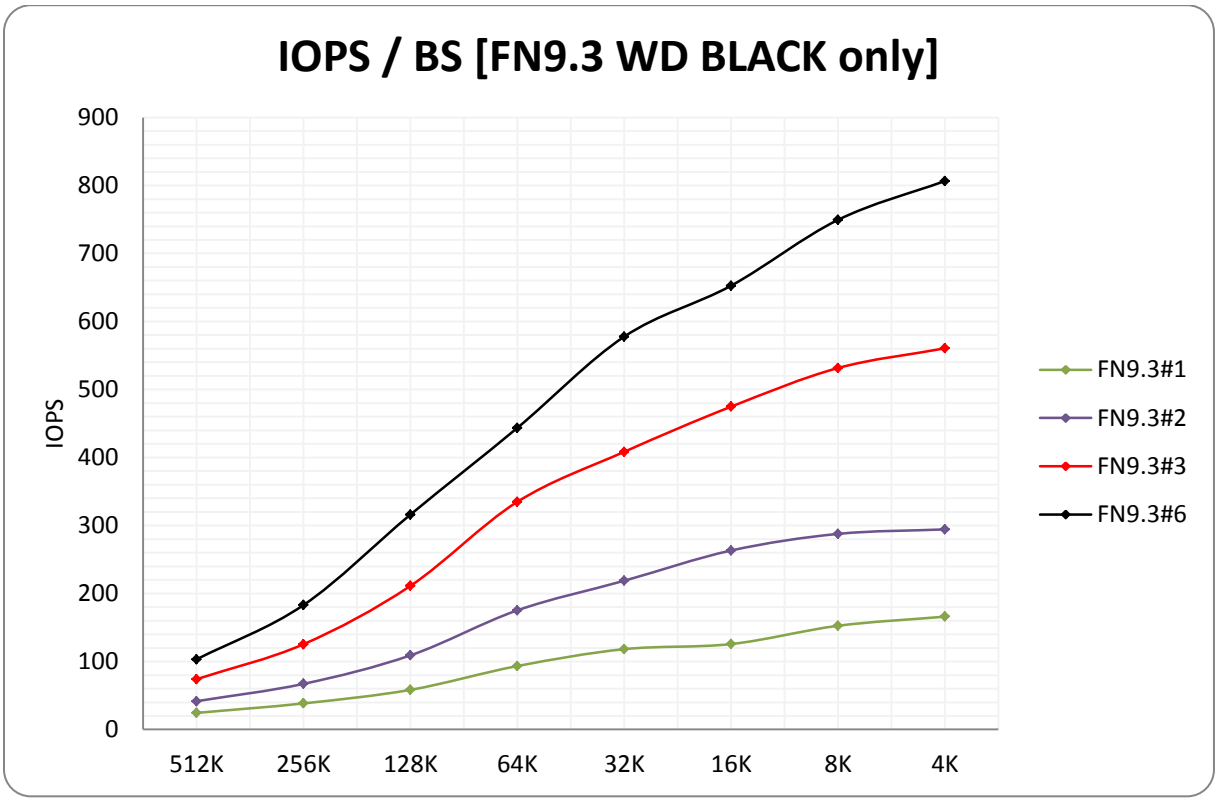
Picture 2: Only data refer to FreeNas 9.2



Picture 3: Only data refer to FreeNas 9.3



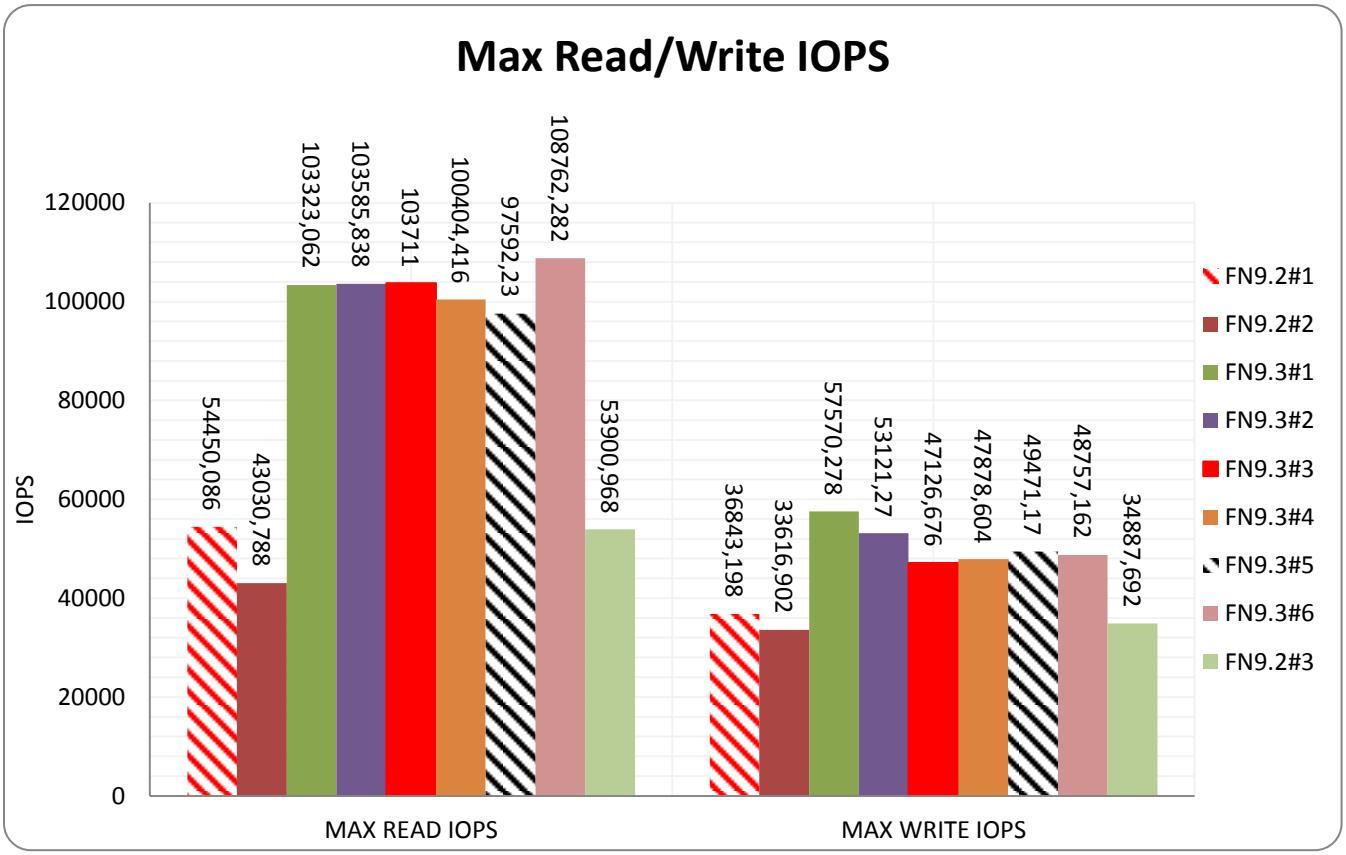
Picture 4: Only data refer to disk WD RED 2TB



Picture 5: Only data refer to disk WD Black 1TB

IV.2 Max IOPS

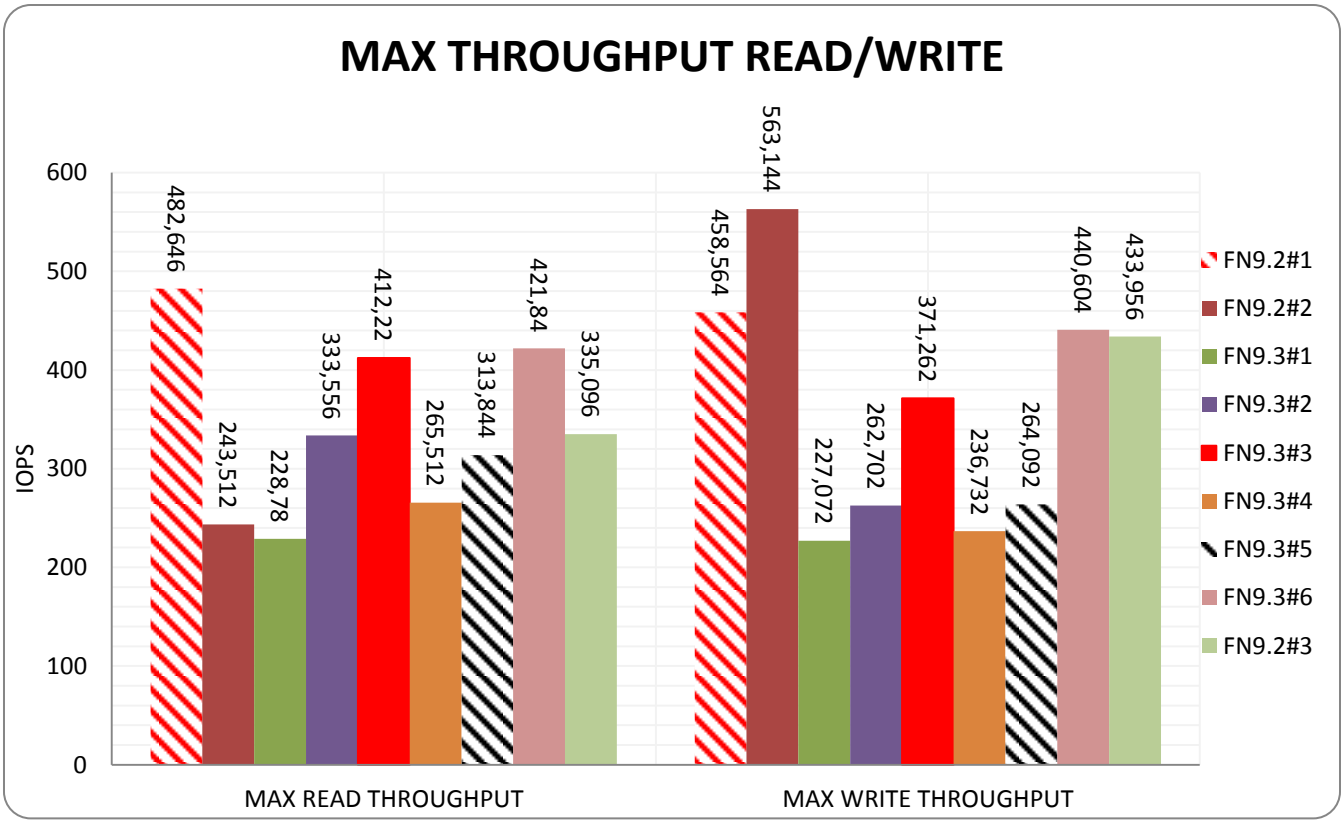
TEST	FN9.2#1 IOPS	FN9.2#2 IOPS	FN9.2#3 IOPS	FN9.3#1 IOPS	FN9.3#2 IOPS	FN9.3#3 IOPS	FN9.3#4 IOPS	FN9.3#5 IOPS	FN9.3#6 IOPS
MAX READ IOPS	54450,086	43030,79	53900,97	103323,062	103585,838	103711	100404,416	97592,23	108762,3
MAX WRITE IOPS	36843,198	33616,9	34887,69	57570,278	53121,27	47126,676	47878,604	49471,17	48757,16



Picture 6: Max Read IOPS and Max Write IOPS

IV.3 Max Throughput

TEST	FN9.2#1 IOPS	FN9.2#2 IOPS	FN9.2#3 IOPS	FN9.3#1 IOPS	FN9.3#2 IOPS	FN9.3#3 IOPS	FN9.3#4 IOPS	FN9.3#5 IOPS	FN9.3#6 IOPS
MAX READ THROUGHPUT	482,646	243,512	335,096	228,78	333,556	412,22	265,512	313,844	421,84
MAX WRITE THROUGHPUT	458,564	563,144	433,956	227,072	262,702	371,262	236,732	264,092	440,604



Picture 7: Max throughput for read/write

V. Conclusions & problems

1. At the beginning great surprise. Picture 1 shows that performance of FN9.3 with RAID 10 composed with 6x WD BLACK has better performance than FN9.2 with RAID 10 composed with 14x WD RED. That's really strange. Strange is also registered performance of FN9.2 with RAID 10 of 14 disks, it's extremely too low.

   So, or methodology of measurement in case of FN9.2 was wrong or something is really not ok with this FN installation. Differences between FN9.2 and FN9.3 test installations were only on FN server and number of network connections (2 in case of FN9.3 and 4 in case of FN9.2). The same test machine (VM on ESXI), the same switches.

   FN9.2 is in fact not test installation. It is production installation where normally 71 VMs (over 5 ESXi) are working and by experience there are no reported any issues about performance.

   So, really looks like test methodology was not correct in case of this installation of FN9.2

2. Performance of FN9.3 presented on picture 2 looks as expected. WD Blacks brigs better performance compare to WD Red and more disks in RAID 10 we have better performance we have.

3. Max throughput in case of FN9.3 is not exactly correct because of only 2 network paths (saturation was observed). FN9.2 as having 4 network paths perform better, but saturation of network only for short period of test time (test time = 120s) rises till almost 4Gbit/s.

Questions:

1. How to get on FreeNas (9.2 and 9.3) statistic about iSCSI about data block sizes, workload ?
2. How to know iSCSI parameters on FN9.3 (like it was available on FN9.2) ? Especially about FirstBurstLength, MaxBurstLength etc. ?
3. What is relation of matrix occupation vs performance ? – in case of FN9.2 there was occupation level about 70% (so no issue for FN where warnings are over 80%) but in case of FN9.2 there were only test virtual disk, nothing more.
4. What is influence of LSI RAID controller to FN9.2 performance ?