Dell HPC NFS Storage Solution - High Availabil

NSS5-HA configurations

By Xin Chen

The latest Dell NSS-HA solution was published on September 2013, of which the version is NSS5-HA. This release leverages Intel Ivy Bridge processors and RHEL 6.4 to offer higher overall system performance than previous NSS-HA solutions (NSS2-HA, NSS3-HA, and NSS4-HA, and, NSS4.5-HA).

Figure 1 shows the design of NSS5-HA configurations. The major differences between NSS4.5-HA and NSS5-HA configurations are:

- Processors:
    - NSS4.5-HA: E5-2680@2.7GHz, 8 cores per processor, (SandyBridge processors)
    - NSS5-HA: E5-2695@2.4GHz, 12 cores per processor, (IvyBridge processors)
- Memory:
    - NSS4.5-HA: 8 x 8GiB, 1600MHz, RDIMMs
    - NSS5-HA: 8 x 8 GiB, 1866MHz, RDIMMs
- OS:
    - NSS4.5-HA: RHEL6.3
    - NSS5-HA: RHEL6.4

Except for those items and necessary software and firmware updates, NSS4.5-HA and NSS5-HA share the same HA cluster configuration and storage configuration. (Refer to NSS4.5-HA white paper for the detailed information about the two configurations.)

Figure 1.   NSS5-HA 360TB architecture
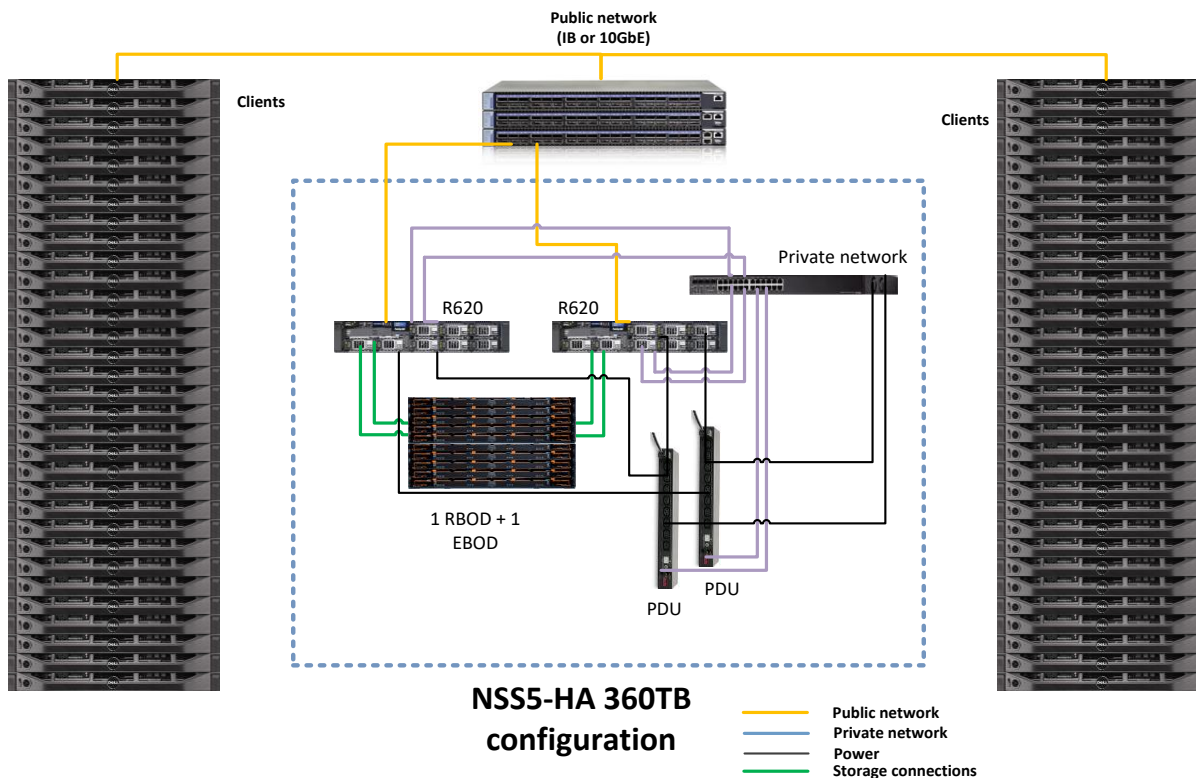
Although Dell NSS-HA solutions have received many hardware and software upgrades to support higher availability, higher performance, and larger storage capacity since the first NSS-HA release, the architectural design and deployment guidelines of the NSS-HA solution family remain unchanged. In the rest of the blog only the I/O performance information of NSS5-HA will be presented, meanwhile, in order to show the performance difference between NSS5-HA and NSS4.5-HA, the corresponding performance numbers of NSS4.5-HA are also presented.

For detailed information about NSS-HA solutions, please refer to our published white papers:

- Dell HPC NFS Storage Solution High Availability Configurations, release version NSS2-HA, published at April 2011.

- Dell HPC NFS Storage Solution High Availability Configurations with Large Capacities, release version NSS3-HA, published at February 2012.

- Dell HPC NFS Storage Solution High Availability (NSS-HA) Configurations with Dell PowerEdge 12th Generation Servers, release version NSS4-HA, published at July 2012.

- Dell HPC NFS Storage Solution - High Availability (NSS-HA) Configuration with Dell PowerVault MD3260/MD3060e Storage Arrays, release version NSS4.5-HA, published at October 2012.

Note: for any customized configuration/deployment, please contact your Dell representative for specific guidelines.

NSS5-HA I/O performance summary

2

Presented here are the results of the I/O performance tests for the current NSS-HA solution. All performance tests were conducted in a failure-free scenario to measure the maximum capability of the solution. The tests focused on three types of I/O patterns: large sequential reads and writes, small random reads and writes, and three metadata operations (file create, stat, and remove).

A 360TB configuration was benchmarked with IPoIB network connectivity. A 64-node compute cluster was used to generate workload for the benchmarking tests. Each test was run over a range of clients to test the scalability of the solution.

The `IOzone` and `mdtest` utilities were used in this study. `IOzone` was used for the sequential and random tests. For sequential tests, a request size of 1024KiB was used. The total amount of data transferred was 256GiB to ensure that the NFS server cache was saturated. Random tests used a 4KiB request size and each client read and wrote a 4GiB file. Metadata tests were performed using the `mdtest` benchmark and included file create, stat, and remove operations. (Refer to Appendix A of the [NSS4.5-HA white paper](#) for the complete commands used in the tests.)

IPoIB sequential writes and reads

Figures 2 and 3 show the sequential write and read performance. For the NSS5-HA, the peak read performance is 4379MB/sec, and the peak write performance is 1327MB/sec. From the two figures, it is obviously that the current NSS-HA solution has higher sequential performance numbers than the previous one.

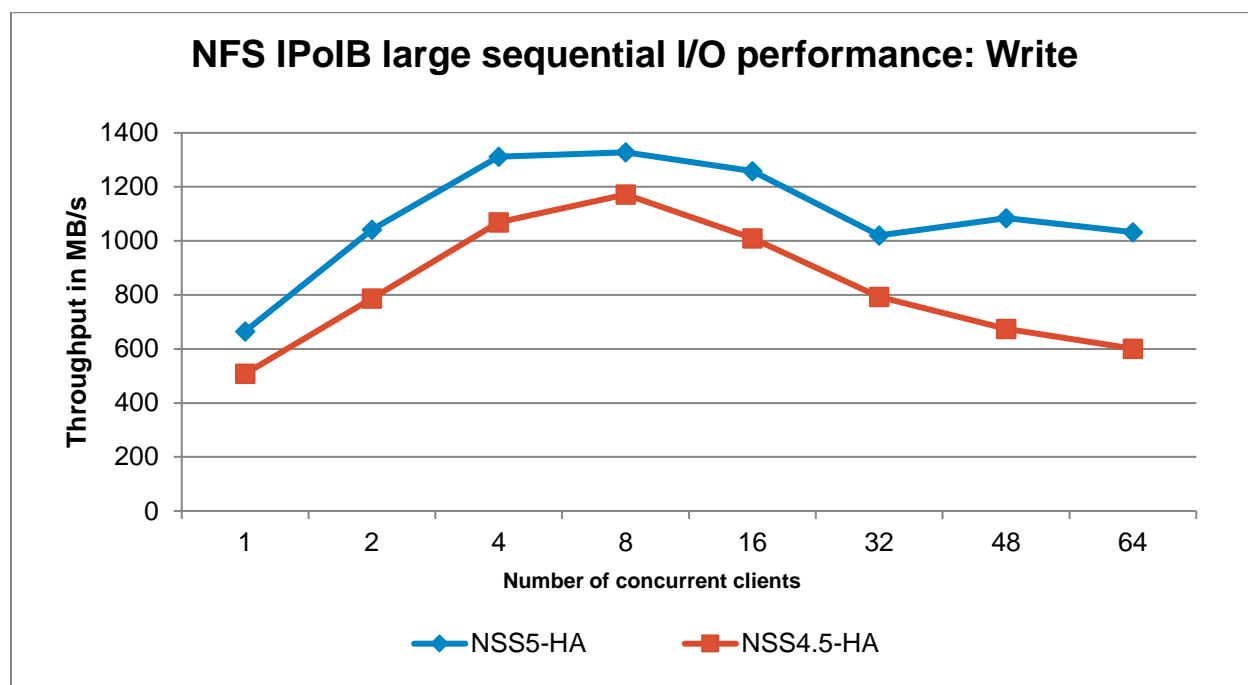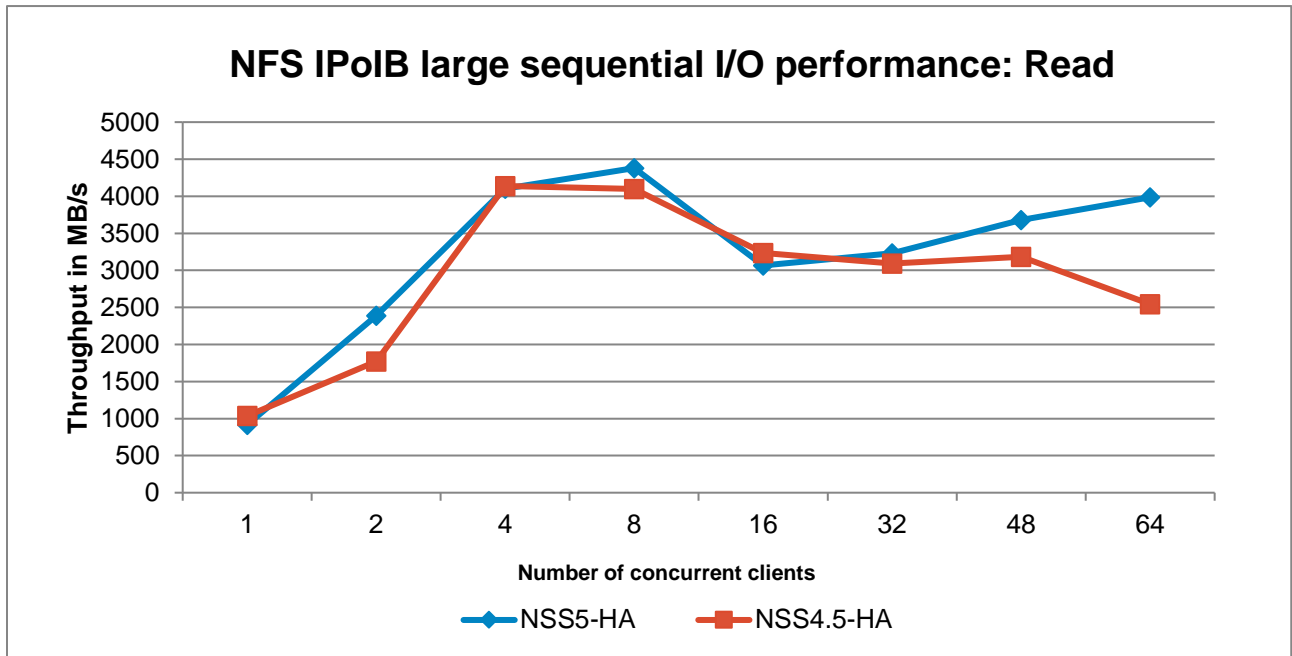Figure 2.   IPoIB large sequential write performance

Figure 3.   IPoIB large sequential read performance

NFS IPoIB large sequential I/O performance: Read

IPoIB random writes and reads

Figure 4 and Figure 5 show the random write and read performance. From the figure, the random write performance peaks at the 32-client test case and then holds steady. In contrast, the random read performance increases steadily beyond going from 32, to 48 to 64 clients indicating that the peak random read performance is likely to be greater than 10244 IOPS (the performance for 64-client random read test case).
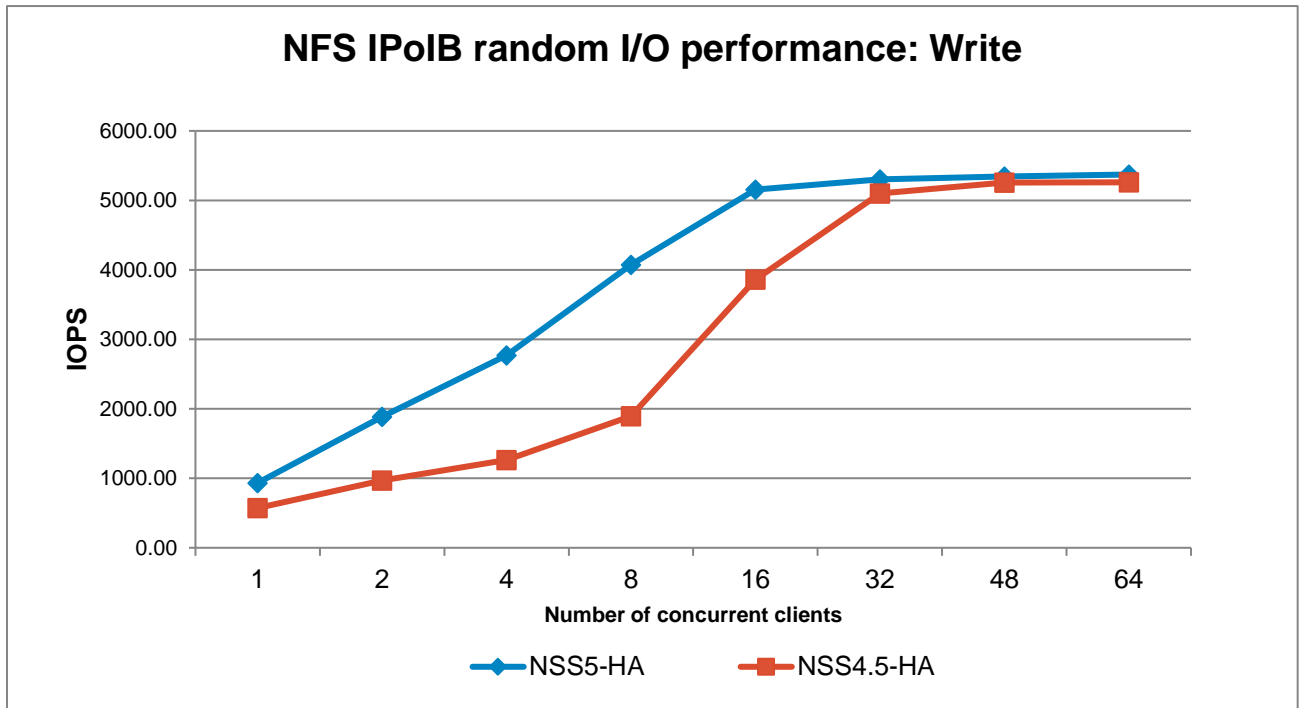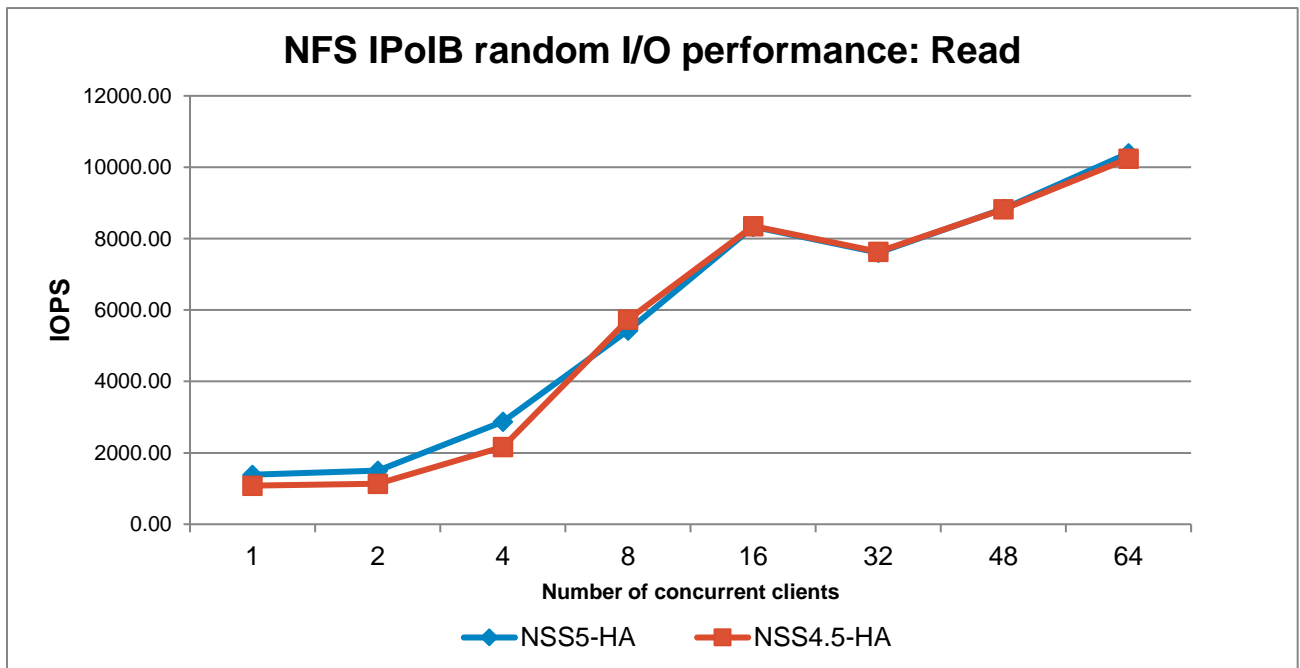
Figure 4. IPoIB random write performance

NFS IPoIB random I/O performance: Write

Figure 5. IPoIB random read performance

NFS IPoIB random I/O performance: Read

IPoIB metadata operations

Figure 6, Figure 7, and Figure 8 show the results of file create, stat, and remove operations, respectively. As the HPC compute cluster has 64 compute nodes, in the graphs below, each client executed a maximum of one thread for client counts up to 64. For client counts of 128, 256, and 512, each client executed 2, 3, or 4 simultaneous operations.

From the three figures, both NSS5-HA and NSS4.5-HA have very similar performance behaviors, as the two lines for NSS5-HA and NSS4.5-HA in each figure are almost identical; it indicates that the changes we have with NSS5-HA do not have obvious impact on the performance of metadata operations.

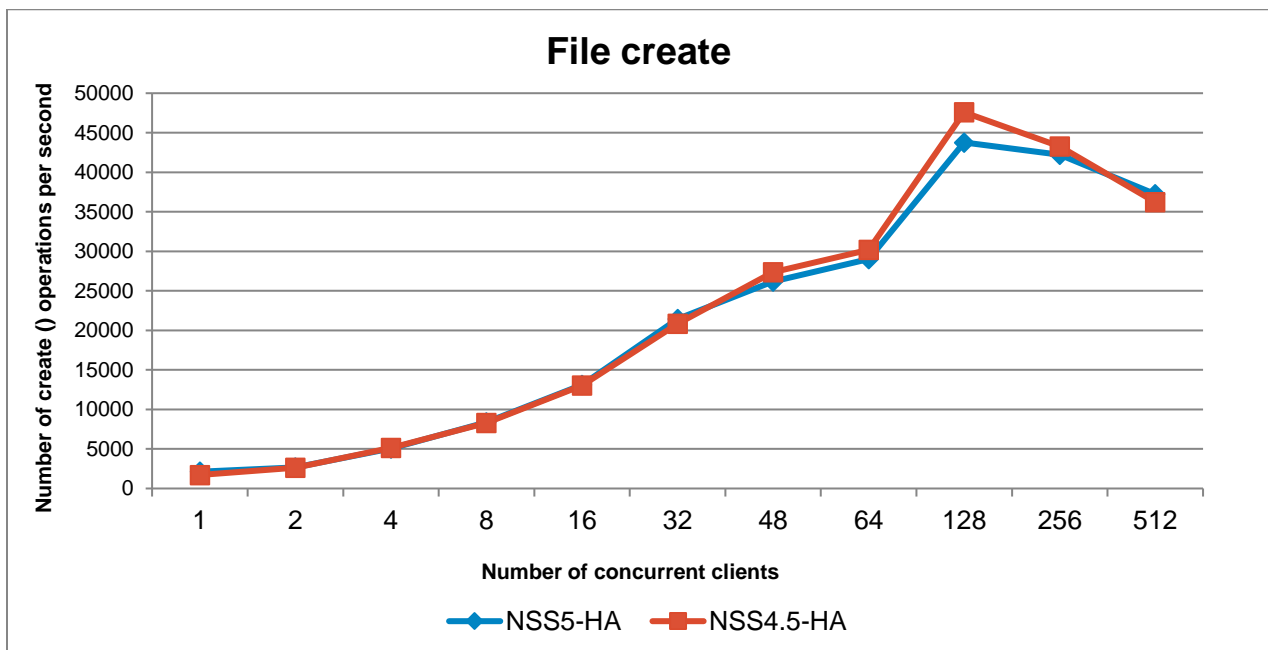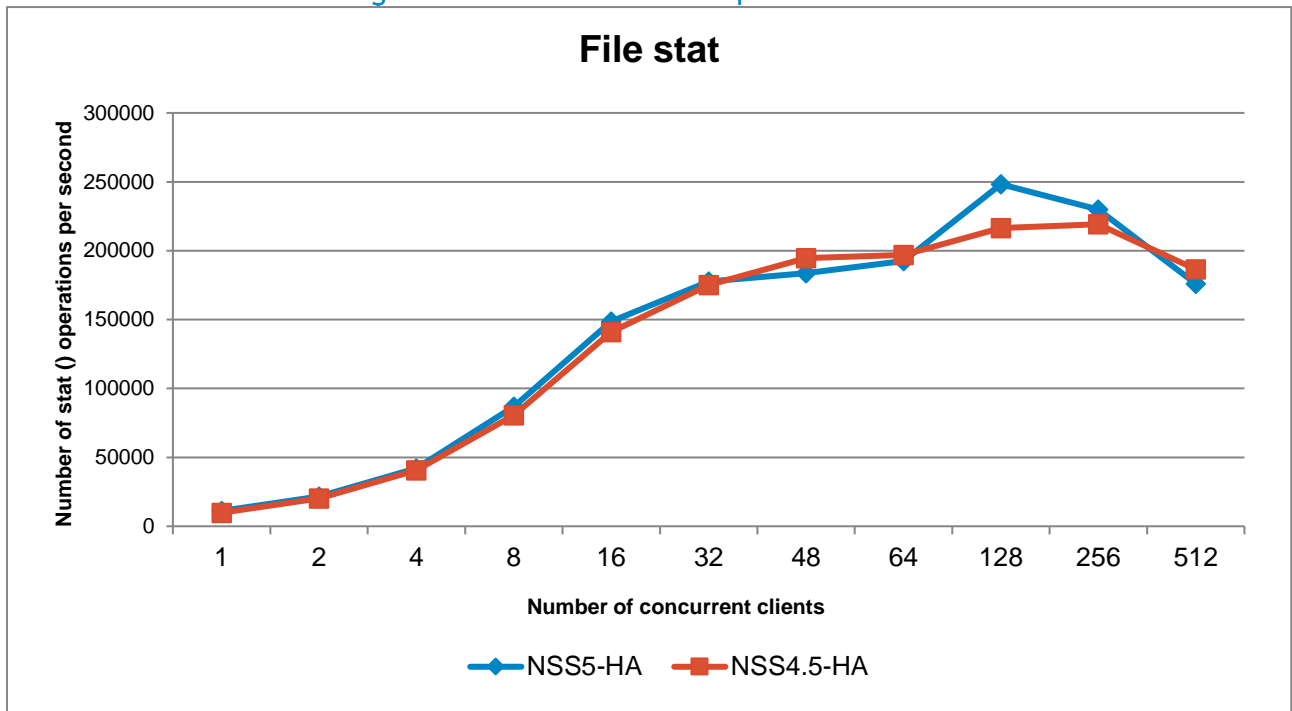**Figure 6.   IPoIB file create performance**

Figure 7.   IPoIB file stat performance

Figure 8.   IPoIB file remove performance