



**Intel[®] NAS
Performance Toolkit
NASPT**

Table of Contents

1	Overview.....	3
1.1	Motivation.....	3
1.2	Benefits.....	4
1.3	Components.....	5
2	Installation.....	6
2.1	Client System Requirements.....	6
2.2	Target Device Requirements.....	6
2.3	Installation Method.....	6
3	Exerciser Usage.....	8
3.1	Test Preparation.....	8
3.2	Test Execution.....	8
3.2.1	Record Keeping.....	8
3.2.2	Running Tests.....	9
3.2.3	Batch Mode.....	9
3.2.4	Viewing Results.....	10
3.3	Tips.....	10
3.3.1	Background File Compression.....	11
3.3.2	File System Buffering.....	11
3.3.3	Disk Fragmentation.....	11
3.3.4	Competition form Other Applications.....	12
4	Analyzer Usage.....	13
4.1	Trace.....	13
4.2	Global Statistics.....	14
4.3	Per File Statistics.....	14
4.4	Service Times.....	14
4.5	Transfer Sizes.....	14
4.6	Throughput.....	15
4.7	Global Access Sequence.....	16
4.8	Per File Access Sequence.....	17
4.9	Tips.....	17

1 Overview

The Intel® NAS Performance Toolkit (NASPT) is a file system exerciser and analysis tool designed to enable direct comparison of network attached storage (NAS) performance. NASPT seeks to discern differences in user level performance when a given client PC uses different remote storage solutions. To that end, NASPT uses a set of real world workload traces gathered from typical digital home applications: HD video playback and record, office productivity applications, video rendering/content creation and more. NASPT reproduces the file system traffic observed in these traces onto whatever storage solution the user provides, records the system response, and reports a rich variety of performance information.

While NASPT runs on a Windows XP* client, the target NAS device may run any operating system.

1.1 Motivation

In trying to express NAS performance in terms non-technical consumers would understand and care about, the NASPT development team found difficulty in applying the performance tools available at the time. Commercial, system-wide benchmarks proved unsuitable. Some focused on isolating particular hardware components like the processor or graphics solution and didn't generate much file system activity at run time. Others behaved more like real applications but for a variety of reasons targeted only local hard drives or more specifically just the C: drive.

A range of disk performance tools, most notably Intel's IPEAK Performance Toolkit as well as IOMeter and others provide extensive insight into the performance of physical, local disks. However, these tools generally work from within the kernel to directly measure disk performance as isolated from the operating system and the rest of the platform. While this is precisely the correct policy for measuring physical disk performance, it doesn't extend easily to the stated goal of measuring the performance of a complete NAS platform as viewed from the end-user's perspective.

For the purpose of measuring NAS performance in a way that is meaningful to developers and consumers alike, NASPT improves over the previously available tools in several important ways:

1. Many storage specific tools work only with locally attached disks or employ a different testing methodology for network disks than for local disks. NASPT delivers one common testing methodology for both local and remote storage devices regardless of the attachment interface.
2. Directly issuing raw commands to the disk results in throughputs at or near the physical limits of the disk drive but precludes transferring any useful data to user level programs. NASPT seeks to measure the performance available to end-user applications.

* Other names and brands may be claimed as the property of others

3. Measuring the disk in isolation from the rest of the platform negates any benefit from file caching and prefetching provided by the operating system, minimizes the impact of other hardware features like processor capabilities and memory bandwidth, and masks the benefit of any innovative performance features provided by the NAS device. NASPT aims to measure the whole NAS platform, including the impact of the operating system and hardware components. In practice, NASPT often reveals significant performance differences between different NAS devices with identical disks.
4. Many tools rely on algorithmically generated traffic, contrary to NASPT's stated objective of measuring system response to real world user level applications. NASPT uses traffic from common digital home applications.
5. Pre-existent tools express results in abstract terms: IOPS, random vs. sequential accesses, transfer size, etc. These details are very difficult to communicate in a meaningful way to the consumer market. NASPT records these details for the technical user, but expresses the results of each test in terms more intuitive and compelling to the consumer.

1.2 Benefits

1. NASPT delivers platform-wide measurements of file system performance from user-level applications, incorporating the effect of many hardware and software components of the NAS device, not just the physical disk.
2. Measurements utilize actual traffic patterns recorded from real world applications that are easily related to the interests of non-technical end-users.
3. NASPT measures each device by the same mechanisms regardless of physical attachment or location.
4. Users may easily add other workloads to suit their particular needs.
5. The NASPT tools are intuitive and easy to use so new users can get up and running very quickly.
6. Tests require minimal user interaction, minimizing the labor overhead necessary to run the tests.
7. NASPT reports results so that they are easily communicated to consumers,.

1.3 Components

NASPT Exerciser: The exerciser generates traffic at the file system level using typical user application semantics. The traffic generated emulates actual traces of real applications' file system requests. A selection of application traces based on likely digital home workloads ships with the NASPT distribution. While the exerciser was designed for NAS devices, the user can direct the traffic generation to any of a number of storage resources: network or local drives, USB devices, etc. Any mass storage device with a drive letter can serve as a suitable target for the NASPT Exerciser.

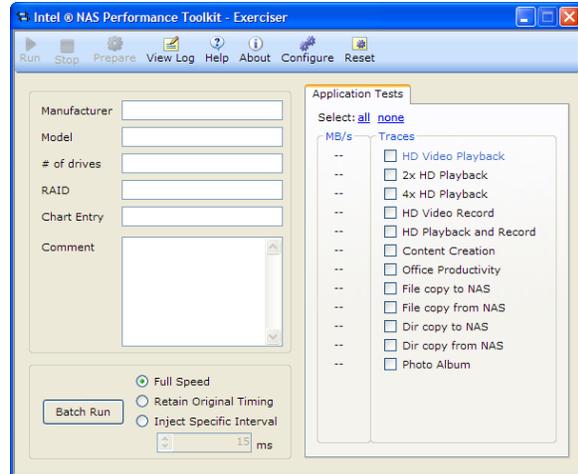


Figure 1 NASPT Exerciser

NASPT Analyzer: The analyzer tool consists of a set of data visualizations designed to assist the user in obtaining a deeper understanding of the results of the exerciser's tests. The tool quotes various statistics across the whole test as well as for individual files. Charts and histograms show throughput over a designated interval, transfer sizes, response times, and access patterns. A sortable, filterable listing of the test traffic shows start, finish, and response times for every transaction. These tools provide powerful mechanisms for investigating the performance characteristics of a given device.

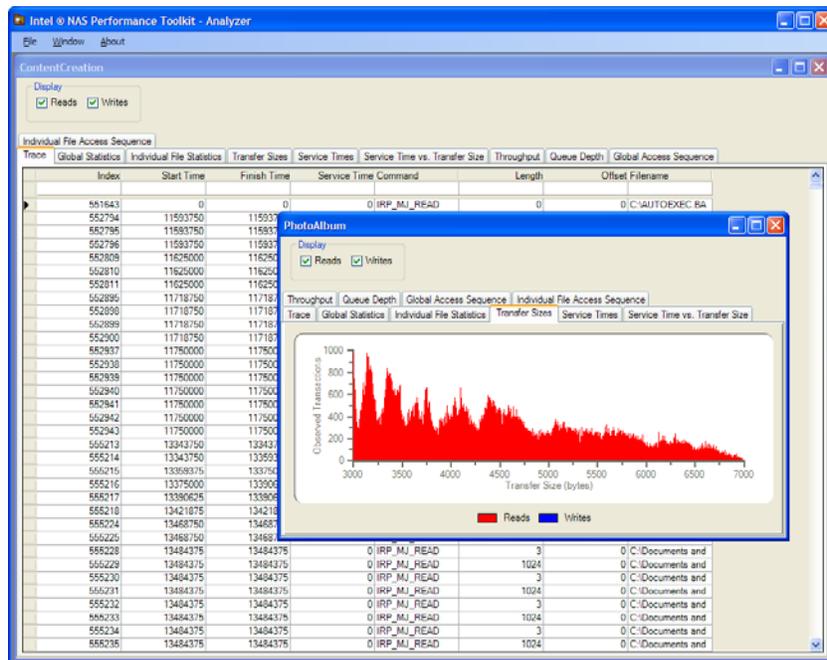


Figure 2 NASPT Analyzer

2 Installation

2.1 Client System Requirements

- The NASPT client is the PC platform that runs the actual software. At runtime, NASPT generates traffic from the client platform to the desired storage device.
- Intel® Architecture 32-bit processor with SSE2 instruction set: Pentium™ 4 processors or later
- Minimum 1GB SDRAM
- Minimum 1GB disk space
- Windows XP* SP2 or Windows Vista* operating system.
- Installation includes the .NET Framework* v2.0

2.2 Target Device Requirements

- 80GB free storage space. For best results, we suggest that the physical disk be kept empty other than NASPT content and not used for system needs like the paging file or operating system image. This avoids introducing performance impacts due to variations in disk layout and competition for disk accesses.
- Suitable attachment mechanism: User must be able to designate a proper file system path to the target device's storage resource. Most devices mounted with an assigned drive letter will serve as NASPT test targets. However, NASPT does not work with UNC paths (\\server\share\etc).

2.3 Installation Method

1. Run setup.exe from the installation media and follow the on screen prompts.
2. Navigate to the installation directory (default is C:\Program Files\Intel\NASPT) and unzip the file OfficeProductivity.zip to the <installation_directory>\app_traces folder. If you're using Windows' native compressed folder utility, you can simply right click, select extract... and accept the defaults. Otherwise, follow the appropriate method for your chosen compression utility. This particular application trace file is too large to work the .msi file format without being compressed first. The installer puts all of the other traces in the \app_traces folder automatically.
3. Configure the target device by mapping it to a drive letter on the client. NASPT treats each storage resource as a black box so the internal configuration of the target device is up to the user as long as a minimum of 80GB of storage space exists on the target.

4. Launch the NASPT exerciser utility by double-clicking its desktop icon or selecting Start→All Programs→Intel NASPT→NASPT Exerciser from the start menu. From within the exerciser, select Configure from the main menu. This will bring up a dialog box showing the target directory for traffic generation (by default C:\NASPT) and output directory for test results (default: <My Documents>\NASPT_traces). Unless you really want to test your local hard drive, change the target directory to a folder on the target device (e.g.: Z:\NASPT). If this folder doesn't exist, it will be created for you.
5. You are now ready to run tests!

3 Exerciser Usage

3.1 Test Preparation

NASPT is a file system traffic generation tool. As such, it needs destination files to target during the playback test. The preparation phase constructs the files and directories needed to reproduce the traced applications' file system requests on the target storage resource. This step must be completed before NASPT can capture performance data.

To prepare a storage resource to run the test, select the desired set of application based tests and click **Prepare** on the menu bar. This will build the necessary file structure on the target storage resource. This step can take several minutes. Once a given test has been prepared, it will appear highlighted in the exerciser window indicating it is ready to run and need not be prepared again. Further, the exerciser notes the preparation status of each test in the target directory so even if a device is disconnected and later reattached, it retains a record of previously prepared tests.

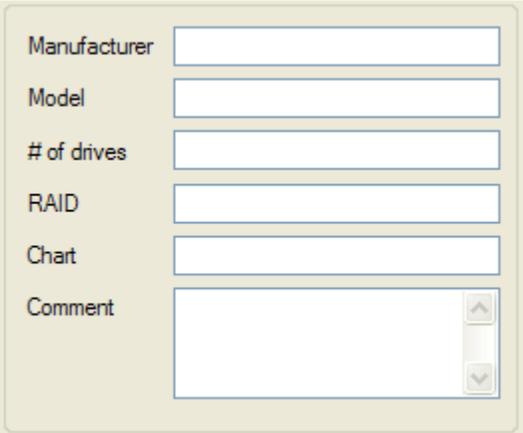
As a convenience, if you're certain you want to run the tests immediately after the preparation phase, you can simply click **Run** on the menu bar and NASPT will do any necessary preparation automatically and then run the tests. Be aware that if you select just a few tests in this case, you could conceivably be accessing the recently cached data from the preparation phase when NASPT runs the performance test and your results won't be indicative of the target device's capabilities.

3.2 Test Execution

3.2.1 Record Keeping

In order to assist with organizing a large number of trace files, NASPT provides a section for the user to enter relevant details about their platform. Any text the user enters into this section is embedded in a header within the output trace file.

Test records are organized into a directory tree by manufacturer name, model name, and date/time stamp.



Manufacturer	<input type="text"/>
Model	<input type="text"/>
# of drives	<input type="text"/>
RAID	<input type="text"/>
Chart	<input type="text"/>
Comment	<input type="text"/>

Figure 3 -- Platform Configuration

The chart name entry is for spreadsheet users. If desired, one can type their chart legend entry they use in their spreadsheets into this box. NASPT will include this name in the .csv files it generates to facilitate cutting and pasting into an existing spreadsheet database.

3.2.1.1 Options

NASPT includes a few options that allow the user to control the pace at the exerciser generates traffic.

3.2.1.2 Full Speed

When this option is selected, the exerciser will generate a synchronous I/O transaction at every opportunity. That is, as soon as one transaction completes, the application will immediately issue another one. This is the default setting.

3.2.1.3 Retain Original Timings

The “Retain Original Timings” check box directs NASPT to reproduce traffic at the same pace as originally traced. For instance, if a particular transaction appears ten seconds into the original trace, NASPT will not issue that transaction during the replay until at least ten seconds have passed. If the replay starts to fall behind the original trace, NASPT will attempt to catch up. This is useful in isochronous performance studies and whenever the user might be more interested in latencies rather than peak throughput measurements. Note that using the “original timings” switch generally leads to much lower throughputs than otherwise measured with NASPT.

3.2.1.4 Inject Specific Interval

NASPT also allows for the injection of deterministic delays between transactions. The mechanism used to insert these delays provides 15ms granularity.

3.2.2 Running Tests

To run a test, select one or more previously prepared tests from the “Applications Tests” tab and click “Run” on the main menu. Most tests take a few minutes to run depending on the characteristics of the target device. As each test completes, the display updates with the average throughput measured across the duration of that test in the MB/s column. Selecting **Run** for an unprepared workload simply means that NASPT will have to prepare the drive first and then run the test. Should a test fail, any tests that completed successfully are still recorded.

3.2.3 Batch Mode

The Batch Mode button runs five times through the entire list of tests, using a separate disk image for each trial and recording the throughput reported by each run in a BATCH_RESULTS.txt file at the end of the test run. This feature makes it very easy to perform multiple tests with just a single button click and exposes any potential variations in device performance due to disk layout by using five separate disk images.

3.2.4 Viewing Results

In depth analysis of test results is best accomplished with the NASPT Analyzer. The Exerciser provides a few quick ways of viewing the test results. Using the **View Log** button shows a list of recent tests, the dates and times they ran, and the directory in which they reside. Opening one of these entries opens the containing folder with all the files output from the test execution:

- The NASPT-APP.log file within each folder contains a text listing of some statistics calculated across the test's duration. This is the same text as appears in the results dialog box at the end of a test run.
- The NASPT-APP.csv file contains statistics gleaned from each of the individual tests rolled up into one file and presented as lists of comma separated values for easy migration to a spreadsheet or database application.
- The many XXX_log.xml (where XXX is the name of the test) files are direct recordings of the system's response to the traffic generated during the tests. Opening them directly will list their contents in your default web browser. You can open these files with the NASPT Analyzer for more detailed observations of test results.

3.3 Tips

NASPT is subject to the characteristics of the target platform in measuring performance. This is by design: NASPT is a storage platform benchmark, not just a disk benchmark. This means that the operating system of the target device and the choices it makes regarding file-level caching (both prefetching and buffering) will figure into the final results. While this is the desired behavior, it introduces some subtleties that may impact unaware users.

Running a small group of tests repeatedly is an example of the sort of usage that could lead to fetching data from warm file system buffers rather than the target device. This will cause your results to suddenly "improve" markedly, since much of the data is coming from local memory.

Although NASPT attempts to minimize the impact of the client platform, some clients may impact results. NASPT performs only minimal computation during runtime to reduce any impact due to a slow client, thereby allowing the service times of the file system accesses to determine performance. Providing a relatively fast client with plenty of memory, 1GB or more, should be sufficient to ensure consistent results.

Note that tests taken with different clients, while yielding the same overall conclusions, will not be directly comparable. Published measurements should always be taken using identical clients.

Network hubs and switches may skew results. NASPT runs best with just a single, directly connected network cable.

Using different hard drives in the target devices can lead to an unfair comparison. Where there is a choice, use identical hard drives in every NAS device.

Finally, use good scientific method and run a few times to make sure you're getting consistent results and understand what is happening during the test. NASPT is designed to make multiple test runs very easy, with minimal labor overhead.

3.3.1 Background File Compression

Some devices compress large files during periods of inactivity. On such systems, if you do not open a file for several days, you may come back to find the device has compressed it. This will cause NASPT to issue "error 38: Reached end of file\n". Simply repeat the prepare step from the NASPT Exerciser to reset the file to its original size. Depending on the device, you may be able to disable file compression.

3.3.2 File System Buffering

All clients and NAS devices will make some attempt to buffer files in local memory to minimize the long latencies incurred when retrieving data from the actual disk drives. How effectively a given device accomplishes this buffering directly impacts the end user experience and is a key constituent of device "goodness". However, it is possible to run afoul of this buffering and obtain erroneous results from NASPT that are too high because the file system buffers already held the needed file data at the beginning of the test. This often occurs when running one of the shorter tests several times in succession without any intervening system activity. It can also occur when running a test immediately after preparing that test, particularly when the test is shorter and consists primarily of read operations. As may be expected, systems with more installed memory are more likely to run into this problem since they often dedicate more local memory to file system buffering.

NASPT measurements should be readily repeatable and consistent. If you find your test results drifting over a wide range or returning unexpectedly high numbers, you may be hitting a warm file system buffer.

3.3.3 Disk Fragmentation

File system addresses represent an abstraction of the storage facilities provided by the disk file system (NTFS, ext3, CIFS, etc..) and the physical disk. As such, addresses within a file which are visible to NASPT do not correspond to addresses on the disk. Thus, a range of accesses that look sequential based on their file offsets may or may not actually be sequential on the disk. While there is no way to guarantee from user level that a file will be laid out sequentially on the disk, most file systems will try to arrange files sequentially where they can. A big, empty disk generally provides a reasonable likelihood that files will be laid out sequentially. Conversely, a relatively full, fragmented disk is much more likely to result in an unpredictable scattering of file fragments. Because rotational media perform significantly better with sequential data, Intel strongly recommends large empty disks for consistent performance measurement between devices.

Note that even with big empty disks, some NAS devices may still exhibit significant fragmentation. Please see [Windows* client CIFS behavior can slow Linux* NAS Performance](#) for an example of how this can happen. NASPT's batch mode option attempts to minimize the effects of varied disk layout by running several tests, each with a separate disk image, and reporting the median result.

3.3.4 Competition form Other Applications

NASPT must share computing resources with other applications running on the PC and often on the NAS device as well. If the virus scanner starts up, a certain application decides to call home to look for updates, or similar activity occurs during the test, NASPT results may be adversely impacted. NASPT includes built-in idle times to allow background processes to run. However, as with all performance measurements, it is best to run at least a few full sets of tests to get a feel for what results are typical for a given platform.

Note that some of the more full-featured NAS products actually run quite a few applications locally, leading to sporadic performance disruptions as the applications compete with I/O traffic for compute resources. NASPT's batch mode option attempts to mitigate this by taking several samples and reporting the median result.

4 Analyzer Usage

The NASPT trace analyzer provides relevant statistics regarding test results and convenient visualization tools for interpreting a system's test performance. The tool is organized into a group of tab pages containing information about a particular test. The exerciser stores test results in XML format using the suffix “_log.xml”. To review a trace, select **file→open** from the menu bar. Navigate to the directory where your traces reside, select the trace file, and click **open**. By default, traces are organized by device manufacturer name, model name, and date/time within the user designated output directory (by default: <My Documents>\NASPT_Tests>). After a few moments, the analyzer displays the trace data in a tabbed format.

4.1 Trace

The “Trace” tab shows every operation recorded by the test along with their chronological order, start and stop times, length, offset, etc. Each column is sortable and filterable. Recorded transactions include reads, writes, and opens. Times are listed in microseconds.

To sort a column, double click a column heading. Clicking again reverses the sort.

To filter, enter the beginning characters of your filter term into the blank space at the top of the column. After a moment, the listing updates to show only those entries whose leading characters match the filter text. To clear the filter, delete the contents of the filter text box.

Index	Type	Start Time	Finish Time	Service Time	Command	Length	Offset	Filename
1	IRP	0	16384	16384	IRP_MJ_READ	16	0	C:\NASBIO\2HD_vidplay\Documents a
2	IRP	16394	16408	14	IRP_MJ_READ	16	0	C:\NASBIO\2HD_vidplay\Documents a
3	IRP	16419	16444	25	IRP_MJ_READ	8	16	C:\NASBIO\2HD_vidplay\Documents a
4	IRP	16430	18711	2281	IRP_MJ_READ	9636	0	C:\NASBIO\2HD_vidplay\Documents a
5	IRP	18719	18730	11	IRP_MJ_READ	24	0	C:\NASBIO\2HD_vidplay\Documents a
6	IRP	18737	18747	10	IRP_MJ_READ	24	9586	C:\NASBIO\2HD_vidplay\Documents a
7	IRP	18757	18874	117	IRP_MJ_READ	24	9610	C:\NASBIO\2HD_vidplay\Documents a
8	IRP	18767	33899	15132	IRP_MJ_READ	24	3452626...	C:\NASBIO\2HD_vidplay\Documents a
9	IRP	33888	36622	2734	IRP_MJ_READ	24	3452672...	C:\NASBIO\2HD_vidplay\Documents a
10	IRP	36629	36651	22	IRP_MJ_READ	30	0	C:\NASBIO\2HD_vidplay\Documents a
11	IRP	36658	36675	17	IRP_MJ_READ	9556	30	C:\NASBIO\2HD_vidplay\Documents a
12	IRP	36682	36693	11	IRP_MJ_READ	16	9586	C:\NASBIO\2HD_vidplay\Documents a
13	IRP	36699	36709	10	IRP_MJ_READ	8	9602	C:\NASBIO\2HD_vidplay\Documents a

Figure 4 -- Trace Listing

4.2 Global Statistics

This tab shows a few calculations that span the entire trace, such as the total number of read and write operations, the amount of test time spent waiting for read and write fulfillment, how many transactions were to sequential file offsets, etc.

4.3 Per File Statistics

This tab provides performance statistics calculated for each file accessed by the test: bytes read, bytes written, portion of accesses that were sequential, total service time, etc. Like the Trace tab, the Per File Statistics tab is sortable and filterable.

4.4 Service Times

This chart shows a histogram of service times (time spent waiting for a requested transaction to complete) for all read and write operations issued by NASPT during the test.

Initially, the timescale defaults to 0 – 1000ms. Most tests will have a large number of transactions that complete out of the file system cache with service times less than 1ms. All of these transactions will aggregate into the <1ms bucket. You can manually adjust the range of the chart using the chart properties dialog box.

4.5 Transfer Sizes

This chart shows a histogram of transfer sizes organized by transaction type for all read and write operations issued by NASPT during the test. As with the Service Times histogram, this chart provides a properties dialog box for customizing the chart.

4.6 Throughput

NASPT calculates average throughput over discrete time intervals for this chart. By default, this interval is 100ms. As with other charts, reads and writes display separately. An aggregate throughput is shown in green.

As with the other NASPT charts, a chart properties dialog box allows customization of the display. The user can add moving average calculations and specify the calculation interval for both the throughput data and the moving average data. Check boxes allow the user to add and remove individual data lines to suit their needs.

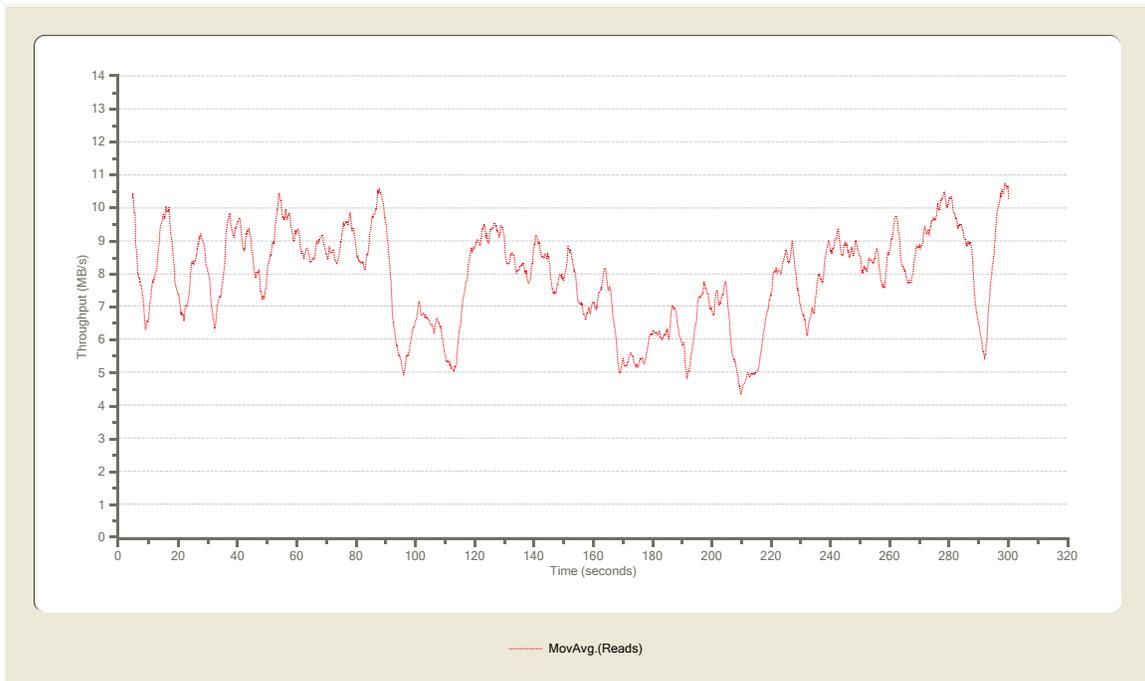


Figure 5 -- Throughput Chart Showing Moving Average Calculations

4.7 Global Access Sequence

This chart shows a conceptual mapping of the file offsets accessed during the test. Each file is represented by a different color. (To map colors to filenames, consult the Individual File Access Sequence chart.) The x-axis is time, initially adjusted to fit the whole test in the window. The y-axis represents the offset within each file expressed in bytes with the upper limit sized so the graph completely contains the largest file accessed during the test.

All files will begin at offset zero. Depending on how large the files in a test are, smaller files may be confined to a relatively small region along the bottom of the graph.

The more time a particular access takes, the wider it will appear in the graph. Similarly, transactions that transfer a lot of data will appear taller.

Zoom in on any section of the chart by click-dragging a selection box around the desired range. When zoomed in, right-clicking anywhere on the chart area will reverse the last zoom operation. Double right-clicking resets the chart.

If NASPT accesses a file sequentially, it will appear to be sequential in this chart: a generally upward sloping diagonal line. Bear in mind, however, that file addresses do not map directly to disk logical block addresses (LBAs). Files that appear to be accessed sequentially may or may not be physically sequential on the hard disk.

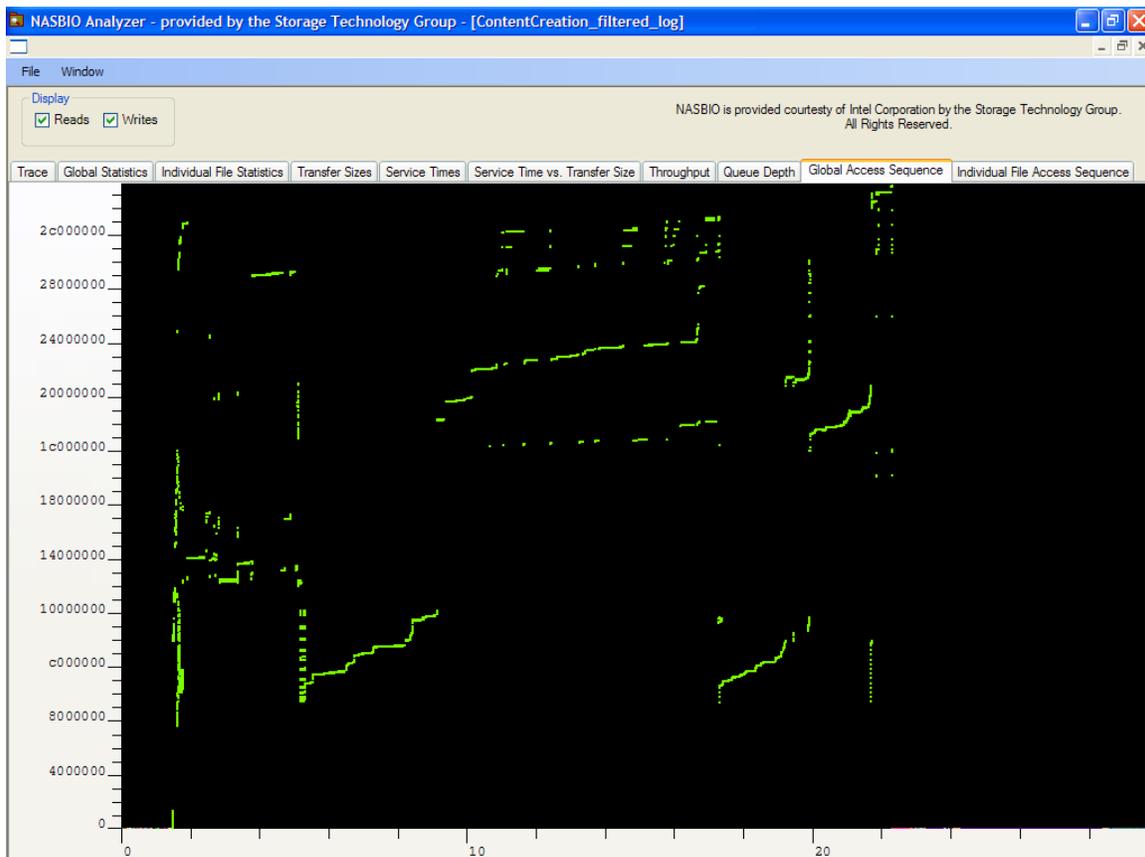


Figure 6 -- Access Chart (All of these accesses are to the same file)



4.8 Per File Access Sequence

Similar to the Global Access Sequence chart, this page shows a conceptualized map of file offsets and transfer sizes. The difference here is that the user can select which files appear in the chart using the “Select Files” button in the upper left corner. Initially, the graph only shows the file with the largest amount of traffic during the test as measured by total bytes transferred.

Click **Select Files** to open the file selection dialog box. This dialog box shows a listing of all file names sorted by total bytes transferred. Those files with the most bytes transferred appear at the bottom of the list.

The colors displayed on this chart match those on the Global Access Sequence chart. By selecting individual files and pushing “Apply”, one can discern which files are displayed using which colors.

4.9 Tips

The NASPT Analyzer is a valuable tool for investigating the behavior of NASPT playback tests and workload traces. It is important to remember that NASPT is a file system test addressing performance available to user level applications. Platform-wide file access performance is the metric of interest. Disk behavior is but one component of the test outcome.

- The throughputs reported by disk utilities are generally higher than those reported by NASPT. Many disk performance tests use raw commands issued directly to the disk. This method maximizes data transfer rates to and from the disk but precludes making the transferred data useful to user-level programs. NASPT attempts to discern the throughput available to user level programs using the facilities provided by the operating system and traces of actual traffic generated by real workloads.
- File offsets do not map predictably to disk blocks. Sequential accesses within a file will often be more or less sequential on the disk, but there is no guarantee. The relationship between one file and another with regards to the physical disk is unpredictable. Using large, empty disks will improve the likelihood that files will actually be laid out sequentially on the physical storage media.

<http://www.intel.com/software/naspt>

http://www.intel.com/design/servers/storage/NAS_Perf_Toolkit.htm

The *Intel® NAS Performance Toolkit* ("Software") is furnished under license and may only be used or copied in accordance with the terms of that license. No license, express or implied, by estoppel or otherwise, to any intellectual property rights is granted by this document. The Software is subject to change without notice, and should not be construed as a commitment by Intel Corporation to market, license, sell or support any product or technology. Unless otherwise provided for in the license under which this Software is provided, the Software is provided AS IS, with no warranties of any kind, express or implied. Except as expressly permitted by the Software license, neither Intel Corporation nor its suppliers assumes any responsibility or liability for any errors or inaccuracies that may appear herein. Except as expressly permitted by the Software license, no part of the Software may be reproduced, stored in a retrieval system, transmitted in any form, or distributed by any means without the express written consent of Intel Corporation.

Copyright © 2008, Intel Corporation