Intel® NAS Performance Toolkit

Anthony Bock
Frank Hady
Storage Technology Group, Intel Corporation
Agenda

- Home/SMB Performance
- NAS Performance Toolkit Introduction
- Consumer NAS Performance Overview
- Using NASPT to improve NAS Performance
- NASPT v1.7 Announcement
Role of DH/SMB NAS Evolving

- More Files/More Data
- Expanding Usage Models
- New Client Platforms
  - Laptops with SSDs
  - MIDs
  - Nettops/Netbooks
  - Connected TVs

Expanding Role for Consumer NAS Devices
NASPT Designed Specifically for Consumer NAS Requirements

- Existing Tools Repurposed from other Uses
  - Local Storage/Network Tools
  - Commercial Desktop Benchmarks
  - Enterprise Storage Tools
- Need: Model Wide Range of Consumer NAS Workloads
  - Useful for Developers
  - Compelling to Consumers
Performance Tool Requirements

End User Relevant

- Usage derived Test Cases
- Easy to Understand Measurements

Easy to Use

- Minimal Learning Curve
- Easily configured and run
- No special test hardware

Accurate

- Minimal test client impact
- Reproducible measurements
- Comparable across NAS devices
Agenda

- Home/SMB Performance
- NAS Performance Toolkit Introduction
- Consumer NAS Performance Overview
- Using NASPT to improve NAS Performance
- NASPT v1.7 Announcement
Intel NAS Performance Toolkit

- Models real applications
  - Trace based workloads
  - Enables file level buffering
  - Writes to both existing and new files
- Controls for:
  - Drive layout
  - Background processes
- Includes graphical analyzer
NASPT Based on Real Workloads

Serving Two Video Streams
NASPT Based on Real Workloads

Copying a Whole Directory
NASPT Based on Real Workloads

Browsing a Collection of Photos
## Application Based Workloads

<table>
<thead>
<tr>
<th>Test</th>
<th># files</th>
<th>% seq.</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>HD Video Play</td>
<td>1</td>
<td>99.5%</td>
<td>720p HD stream from Windows Media Player® 256kB reads</td>
</tr>
<tr>
<td>2HD Video Play</td>
<td>2</td>
<td>18.1%</td>
<td>2x playback</td>
</tr>
<tr>
<td>4HD Video Play</td>
<td>4</td>
<td>9.6%</td>
<td>4x playback</td>
</tr>
<tr>
<td>HD Video Record</td>
<td>1</td>
<td>99.9%</td>
<td>720p HD stream, 256kB writes</td>
</tr>
<tr>
<td>HD Video Play &amp; Record</td>
<td>2</td>
<td>17.8%</td>
<td>1 playback, 1 record simultaneously</td>
</tr>
<tr>
<td>Directory Copy From NAS</td>
<td>2833</td>
<td>52.5%</td>
<td>64kB reads</td>
</tr>
<tr>
<td>Directory Copy To NAS</td>
<td>2833</td>
<td>52.5%</td>
<td>Predominantly 64kB writes, wide scattering under 16kB</td>
</tr>
<tr>
<td>File Copy From NAS</td>
<td>1</td>
<td>100%</td>
<td>4GB file copy, 64kB reads</td>
</tr>
<tr>
<td>File Copy To NAS</td>
<td>1</td>
<td>100%</td>
<td>64kB writes</td>
</tr>
<tr>
<td>Photo Album</td>
<td>169</td>
<td>80%</td>
<td>All reads – wide distribution of sizes</td>
</tr>
<tr>
<td>Office Productivity</td>
<td>607</td>
<td>81.3%</td>
<td>Reads &amp; writes; small, 1kB &amp; 4kB reads; Mostly 1kB writes</td>
</tr>
<tr>
<td>Content Creation</td>
<td>98</td>
<td>38.6%</td>
<td>95% writes; 1k, 4k &amp; little reads; Writes up to 64kB</td>
</tr>
</tbody>
</table>

More realistic workloads than synthetic tests
Write to New vs. Write to Existing

Overwriting an existing file and writing to a new file differ in performance:
- Applications do both
- Many tools use just one method
- To better model real apps, NASPT can use either approach as specified by the workload

Source: NASPT Measurements
Disk Layout Matters to NAS Performance

- Disk layout has significant impact on performance
- Applications have no control over layout
- Some tools use direct/unbuffered I/O to force idealized layout
- NASPT’s Batch Mode performs 5 trials with independent disk images
  - Result is median of all five trials
  - Retains output traces from all trials

Source: NASPT Measurements
Agenda

- Home/SMB Performance
- NAS Performance Toolkit Introduction
- Consumer NAS Performance Overview
- Using NASPT to improve NAS Performance
- NASPT v1.7 Announcement
NAS Usage: Still & Streaming Media

Digital Photo Browsing
(~10 minute fetch of 100+ images)

- Local HD
- Fastest NAS
- Slowest NAS

Storage Wait Time (Minutes)

- 5.6x slower than local HD
- 2.7x slower than fastest NAS

Video Distribution

- Maximum Throughput (MB/s)
- Source: NASPT Measurements

Interactive Usage Highlights Value of Performance
Consumer NAS Performance Varies Widely

Source: NASPT measurements

Be a NAS performance leader – Users will notice
Users Will Notice...

- Many users will already notice superior NAS performance
- As users upgrade networks, NAS performance becomes more and more visible

Source: NASPT Measurements
Agenda

- Home/SMB Performance
- NAS Performance Toolkit Introduction
- Consumer NAS Performance Overview
- Using NASPT to improve NAS Performance
- NASPT v1.7 Announcement
Realistic File Usage Enables NASPT to find Issues

NTFS and EXT3 Handle “Data Holes” Differently

**Application**
Two writes to file “foo” with never-written “data hole” between

**NTFS**
Gap is preserved as allocated but invalid disk space

**EXT3**
Gap is eliminated by abutting discontiguous writes

And with “file buffering” on, Windows* generates 1-byte hints

Combination Leads to Unintended Consequences for Linux* NAS

**Application**
Contiguous write stream

**Network**
Windows OS adds one byte writes to high offsets

**EXT3 result**
Severe fragmentation - ~18/MB
Low performance – 50% drop
Strict Allocation Doesn’t Quite Work

Samba’s included “strict allocate” feature fills gaps with zeroes

Disk layout improves dramatically

Files with many discontinuous writes and large gaps experience severe delays

**Strict Allocation Doesn’t Quite Work**

Application

Two writes to file “foo” with never-written “data hole” between

Strict Allocate

Gap is preserved by filling intervening space with zeroes

000000000

1

2

unwritten
Solution: Modify Samba Zero Fills

Two small changes improve strict allocate behavior with Windows* clients

- Skip zero fills when copying files
  - 1-byte hints don’t occur when file size is known ahead of time
- Only fill to 2MB past current end of file

```c
samba/source/smbd/vfs.c
vfs_fill_sparse(...)
{
  ...
  if (len <= st.st_size)
    return 0;
  //Impose limit on how much to write ahead of current position
  #define ALLOCATION_LIMIT 0x200000
  if (len - st.st_size > ALLOCATION_LIMIT)
    return 0;
  ...
}
```

```c
samba/source/modules/vfs_default.c:
vfswrap_ftruncate(...)
{
  int result = -1;
  SMB_STRUCT_STAT st;
  char c = 0;
  SMB_OFF_T currpos;
  START_PROFILE(syscall_ftruncate);
  /* ignore file fill when presented with new file of known size.
   * if (lp_strict_allocate(SNUM(fsp->conn)))
   *       result = strict_allocate_ftruncate(handle, fsp, fd, len);
   * END_PROFILE(syscall_ftruncate);
   * return result;
   */
  ...
}
```
Agenda

- Home/SMB Performance
- NAS Performance Toolkit Introduction
- Consumer NAS Performance Overview
- Using NASPT to improve NAS Performance
- NASPT v1.7 Announcement
NASPT Version 1.7 Available

- General usability enhancements
- Reduced run time
- Better controls for sources of variance
  - Variations in disk layout
  - Competition from background processes
- Users may add custom workloads

http://www.intel.com/software/naspt
http://www.intel.com/design/servers/storage/NAS_Perf_Toolkit.htm
Would this tool be valuable to you as an open source project?
How many would anticipate contributing code?