Beyond CIFS

David Kruse

Microsoft
Agenda

- Previous Version Access
- Access Based Enumeration
- Server Side Data Movement
- Symbolic Link Support in SMB2
- Durable Open Support in SMB2
Previous Version Access

- Offers the user access to previous versions of files or folders.
- Reduces IT costs by enabling end-user to restore content.
- Native support in WinXP SP2+, earlier requires Shadow Copy Client (http://technet.microsoft.com/en-us/windowsserver/bb405951.aspx)
## Previous Version Access

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NumberOfSnapShots</th>
</tr>
</thead>
<tbody>
<tr>
<td>NumberOfSnapShotsReturned</td>
</tr>
<tr>
<td>SnapShotArraySize</td>
</tr>
<tr>
<td>SnapShots (variable)</td>
</tr>
</tbody>
</table>

- Implementation is aligned with volume snapshot features, but true versioning can be supported
- Client must expect that snapshots can be removed at any time
Previous Version Access

- Preprocessing needs to happen for all path-based requests
- SMB2_FLAGS2_REPARSE_PATH hints that @GMT is present
- Downlevel support for previous versions requires scanning regardless of flag
Previous Version Access

- SMB2 Support is simpler:
  - Create Context contains @GMT token converted to LARGE_INTEGER time value
  - Single create path and handle based command set mean that only create logic needs to be Previous Version aware.
Previous Version Access

Client

Server

Open Target File

Return Success

Issue Snapshot FSCTL

Return Snapshot List

Query Attrib of Files at Snapshot Times

Return Attributes

Open Version Requested by User
Previous Version Access

- Downlevel (Win9x/NT4) Support

- <downlevel enumeration example here>
Access Based Enumeration

- Provides a filtered view of the namespace based on user access
- View filters out files and directories which the user does not have READ_DATA/LIST_DIR access to unless the caller is privileged. (backup privilege)
- Enabled/Disabled at share granularity
Access Based Enumeration

- SMB/SMB2 server should filter directory enumeration results as desired
  - Always allow “.” and “..”
  - Handle wildcard and non-wildcard queries
  - Consider error mapping on create requests
- Server MUST set bit in tree connect response
  - Technologies like Offline Files require knowledge that they are seeing a filtered namespace
Resume Key interaction provides a portable identifier for an existing open to allow referencing in new scenarios.

Server Side Copy enables clients to request data movement between source and destination files without pulling the data to the client.
Server Side Data Movement

- Resume Key Format
  - 24 byte opaque value, defined by you
  - Locally unique to the server
    - May be assembled from GUID, FileId, Open Time, Process Id, signature, etc.

- Why not simply use Fid?
Server Side Data Movement

Query Resume Key on SMB Channel

Bind to Open via Alternate Method

<Perf Data Here>
On Success or Failure, returns information on written data.

If Request exceeds configured limits, returns configured limits and INVALID_PARAMETER.
Server Side Data Movement

- Current Client Usage:
  - Vista+ for Remote<->Remote CopyFile
  - Vista+ for Offline Files Bitmap Upload

- <insert CopyFile perf chart>
Symbolic Link Support for SMB2

- Allows SMB2 clients to process absolute and relative symbolic links.
  - (You may have heard of them before)
- Symlink evaluation always done by the client
- Vista client default policy locks down remote symlink evaluation
  - Local Computer Policy \ Computer Configuration \ Administrative Templates \ System \ NTFS Filesystem \ “Selectively allow the evaluation of symbolic links”
  - Enable “Remote->Remote” and potentially “Remote->Local”
Symbolic Link Support for SMB2

<table>
<thead>
<tr>
<th>0 1 2 3 4 5 6 7 8 9</th>
<th>0 1 2 3 4 5 6 7 8 9</th>
<th>0 1 2 3 4 5 6 7 8 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>SymLinkLength</td>
<td>SymLinkErrorTag</td>
<td>ReparseTag</td>
</tr>
<tr>
<td>ReparseDataLength</td>
<td>Reserved</td>
<td></td>
</tr>
<tr>
<td>SubstituteNameOffset</td>
<td>SubstituteNameLength</td>
<td></td>
</tr>
<tr>
<td>PrintNameOffset</td>
<td>PrintNameLength</td>
<td>Flags</td>
</tr>
<tr>
<td>PathBuffer (variable)</td>
<td></td>
<td>...</td>
</tr>
</tbody>
</table>

Storage Developer Conference 2008
© 2008 Microsoft Corporation. All Rights Reserved.

www.storage-developer.org
Symbolic Link Support for SMB2

- <Symlink format examples>
Durable Open Support for SMB2

- Enables a client to re-establish an open held on the server after a disconnect has occurred.
- Client requests this behavior during Create
- Server can grant it as appropriate
- Reconnect is not necessarily guaranteed by the server.

- Supported on Vista+ clients.
Durable Open Support for SMB2

- SMB2_CREATE_DURABLE_HANDLE
  - Sent on create to request durability
  - Acknowledged in successful create response if durability is available
- SMB2_CREATE_DURABLE_HANDLE_RECONNECT
  - Sent on new connection to rebind to durable handle.
  - SessionId, TreeId must be valid
Client is responsible for replaying operations in flight for which response was not received.

Vista Client: Reconnect is retry based, not timeout based (Reconnect attempted 5 times)

Server 2k8 holds durability after disconnect while oplock is present
What else is there?

- SMB1
  - Reauthentication on Kerberos Ticket Expiry
  - Signing Key Protection
  - Extended Responses (with Maximal Access)
  - 32-bit Read/Write lengths
  - Native NT information levels via Pass-thru
Questions?