Internet Explorer Cache

Overview

The Internet Explorer disk cache is a storage folder for temporary Internet files that are written to the hard disk when a user views pages from the Internet. Internet Explorer uses a persistent cache and therefore has to download all of the content of a page (such as graphics, sound files or video) before it can be rendered and displayed to the user. Even when the cache is set to zero percent, Internet Explorer requires a persistent cache for the current session. The persistent cache requires 4 MB or 1 percent of the logical drive size, whichever is greater.

Disk Cache Storage Location

The disk cache location varies across operating systems and can usually be found in the following default locations.

Windows 9x & Millennium Edition (Without Profiles)

```
\Windows\Temporary Internet Files\
```

Windows 9x & Millennium Edition (With Profiles)

```
\Windows\Profiles\{Profile Name}\Temporary Internet Files\
```

Windows 2K & XP

```
\Documents and Settings\{User Name}\Local Settings\Temporary Internet Files\
```

Windows Vista, 7

```
\Users\{User Name}\AppData\Local\Microsoft\Windows\Temporary Internet Files\ \Users\{User Name}
\AppData\Local\Microsoft\Windows\Temporary Internet Files\Low\
```

To identify the correct location of the cache for each user, the Registry hive for the particular user must be examined. You cannot rely upon the live cache folder being in the default location. The following figure shows the Registry key containing the cache folder location.

NTUSER.DAT Registry Hive

```
HKCU\Software\Microsoft\Windows\CurrentVersion\Explorer\
```

The "User Shell Folders" subkey stores the paths to Windows Explorer folders for each user of the computer. The entries in this subkey can appear in both the "Shell Folders" subkey and the "User Shell Folders" and in both HKEY_LOCAL_MACHINE and HKEY_CURRENT_USER. The entries that appear in user "User Shell Folders" take precedence over those in "Shell Folders". The entries that appear in HKEY_CURRENT_USER take precedence over those in HKEY_LOCAL_MACHINE.

Temporary Internet Files Folder Structure

Within the Temporary Internet Files folder, you will find a Content.IE5 folder. This folder is the main disk cache for Internet Explorer. Outlook Express also writes data to this location as is explained later in this article. Inside the Content.IE5 folder, you will find a minimum of 4 cache folders. The file names for these cache folders comprise of 8 random characters. When further cache space is required, Internet Explorer will add additional folders in multiples of 4. Figure 1 shows the layout of a typical Internet Explorer disk cache.

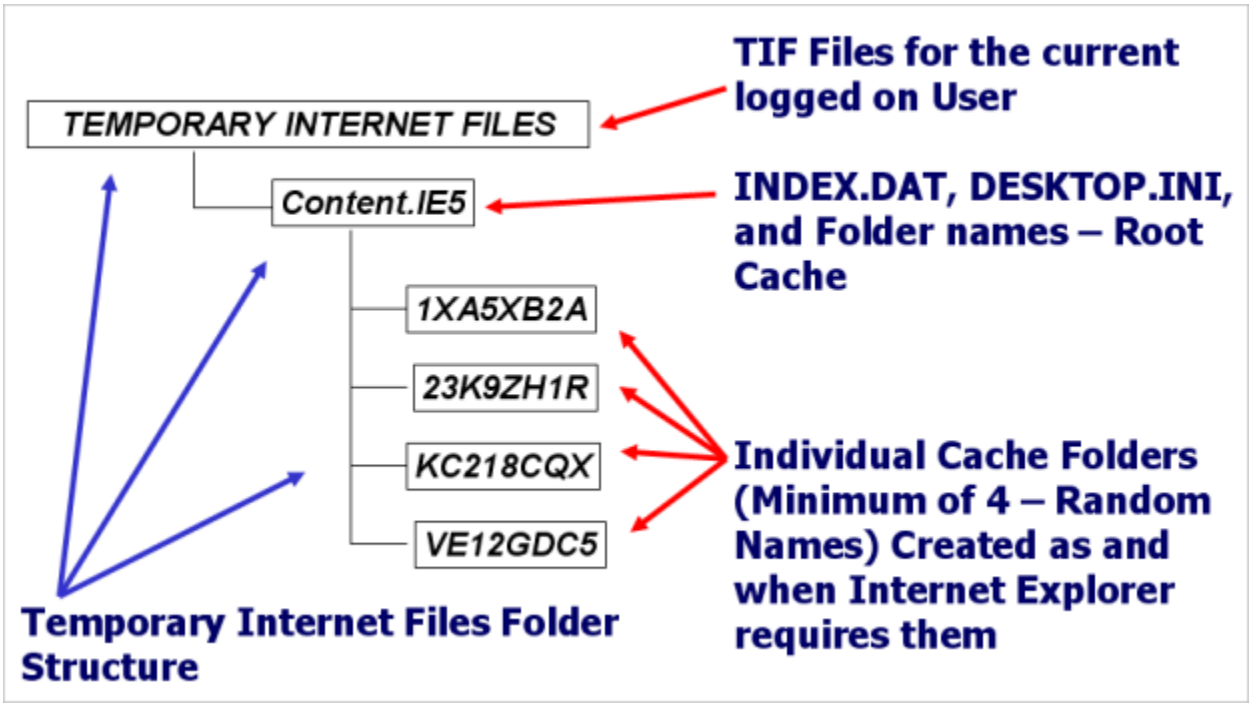

Figure 1

Content.ie5 INDEX.DAT File

The cache INDEX.DAT file is a database of cache entries. It holds information relating to individual cached items so that the browser can check whether the resource needs to be updated (eTag) and information relating to the location of the cached item. It also stores the HTTP (Hypertext Transfer Protocol) response header for the resource. At the start of the INDEX.DAT file, you will find a zero based array holding the names of the cache folders. Each individual URL record contains an index value which refers to a specific folder in the array. This allows Internet Explorer to correctly identify the location of a cached item. This can be seen in Figure 2 below.

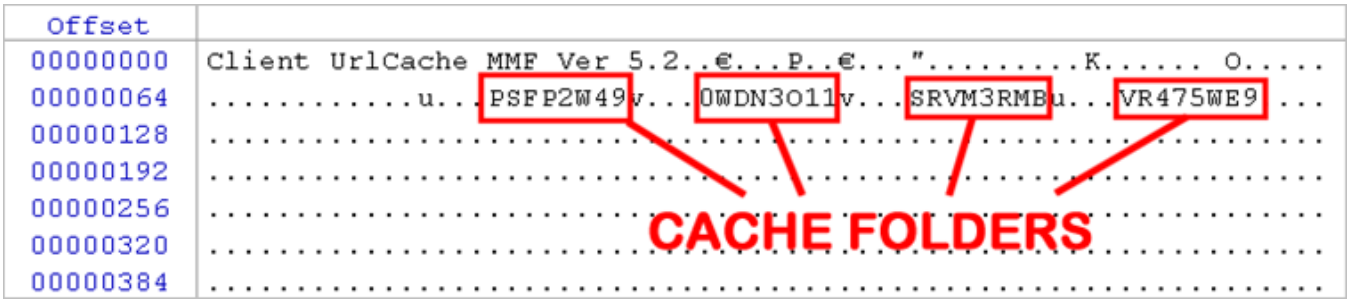

Figure 2

Cache File Naming Convention

As files are saved to the cache, Internet Explorer uses a specific naming convention as shown below. The only exception to this rule is when data is written to this location by Outlook Express or Outlook.

Internet Explorer Cache File Naming Convention
`Filename[Counter].Extension`

As pages and resources are cached, they can easily end up in different folders. Internet Explorer attempts to keep the volume of data and number of cached items across each cache folder as level as possible. As Internet Explorer writes files to the cache folders, it checks to see if a file with the same name already exists. This is frequently the case when web developers do not use imaginative or descriptive names for their files. If the file already exists within the folder, Internet Explorer will increment the counter. If no file exists, the counter portion of the file name is set to 1.

Files with the same naming structure within a cache folder do not necessarily belong to the same web site. Also, multiple visits to the same web site can easily result in files with similar naming conventions being spread across all the cache folders. Cached resources can also become orphaned when the INDEX.DAT entry is not written to disk (such as when Internet Explorer crashes before the entries are written). Figure 3 shows a typical cache folder. There are two files within this folder which would have had the original name "images.jpg". Internet Explorer has renamed them "images[1].jpg" and "images[2].jpg".

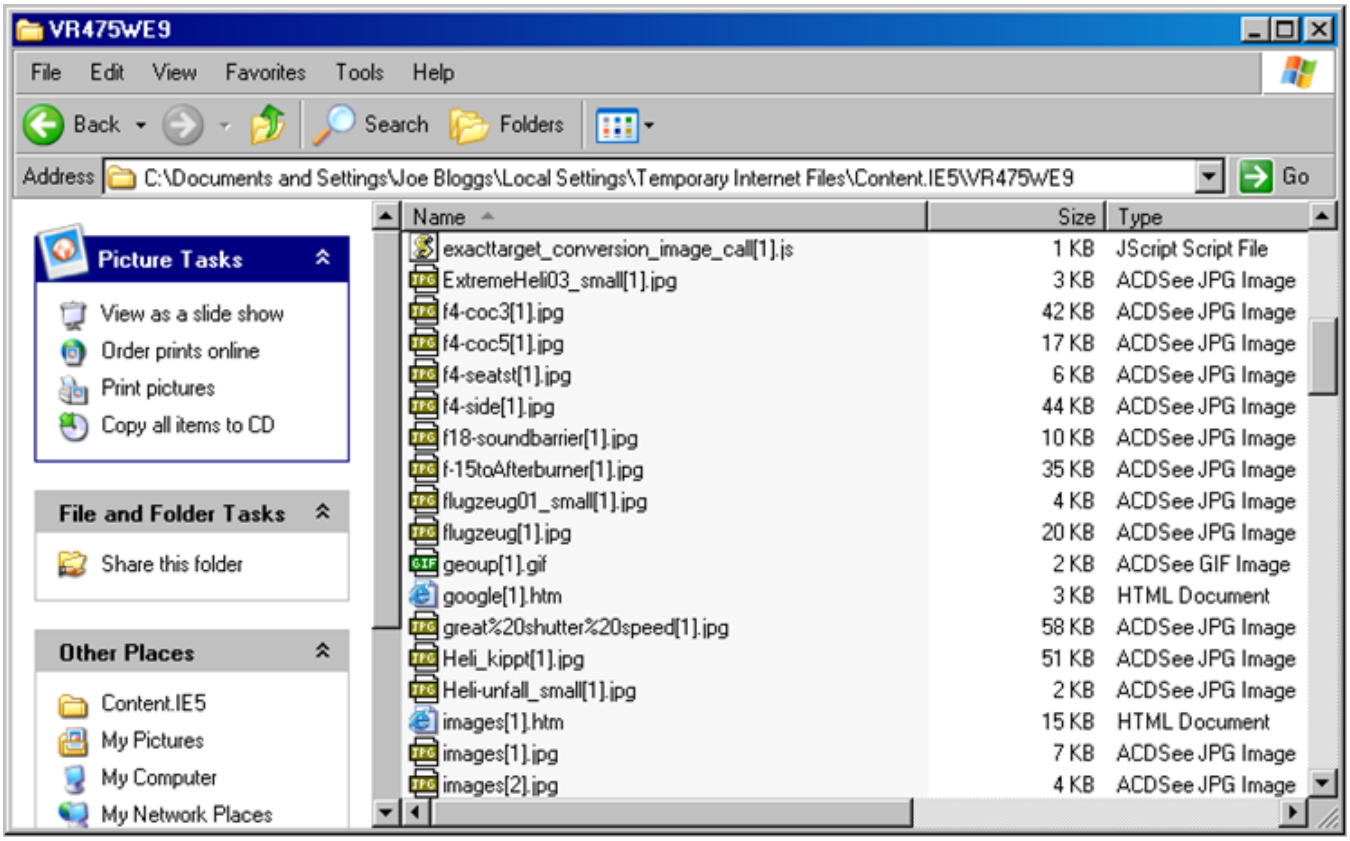

Figure 3

Outlook Express Email Client

Microsoft Outlook Express also uses the Content.IE5 folder as a temporary cache. When a user selects an email message within the client, Outlook Express reads the data from the corresponding DBX file and caches the components of the email so that it can be rendered on screen for the user as an email message. The structure of the DBX files is such that the email message is broken down into blocks and has to be rebuilt before it can be rendered. If the message contains any attachments, they are stored in Base64 format and stored within the file. The attachments also have to be extracted, decoded and cached prior to rendering on screen. As the message is rebuilt, Outlook Express saves the different elements of the message to the disk cache as a temporary file. The naming convention is different to Internet Explorer. In the past, some forensic examiners have not been aware of this and have incorrectly attributed data in the cache to a visit to a web page when it in fact was there as the result of viewing an email message. The file structure and naming convention is shown below.

Naming Convention
wbk***.dat

Examples
\Temporary Internet Files\Content.IE5\B6KY005Q\wbk39D.tmp \Temporary Internet Files\Content.IE5\Q2XY001Q\wbkA1F.tmp \Temporary Internet Files\Content.IE5\AB32QX81\wbkF10.tmp