BIRMINGHAM CITY
University

Network Forensics

Dr Syed Naqvi

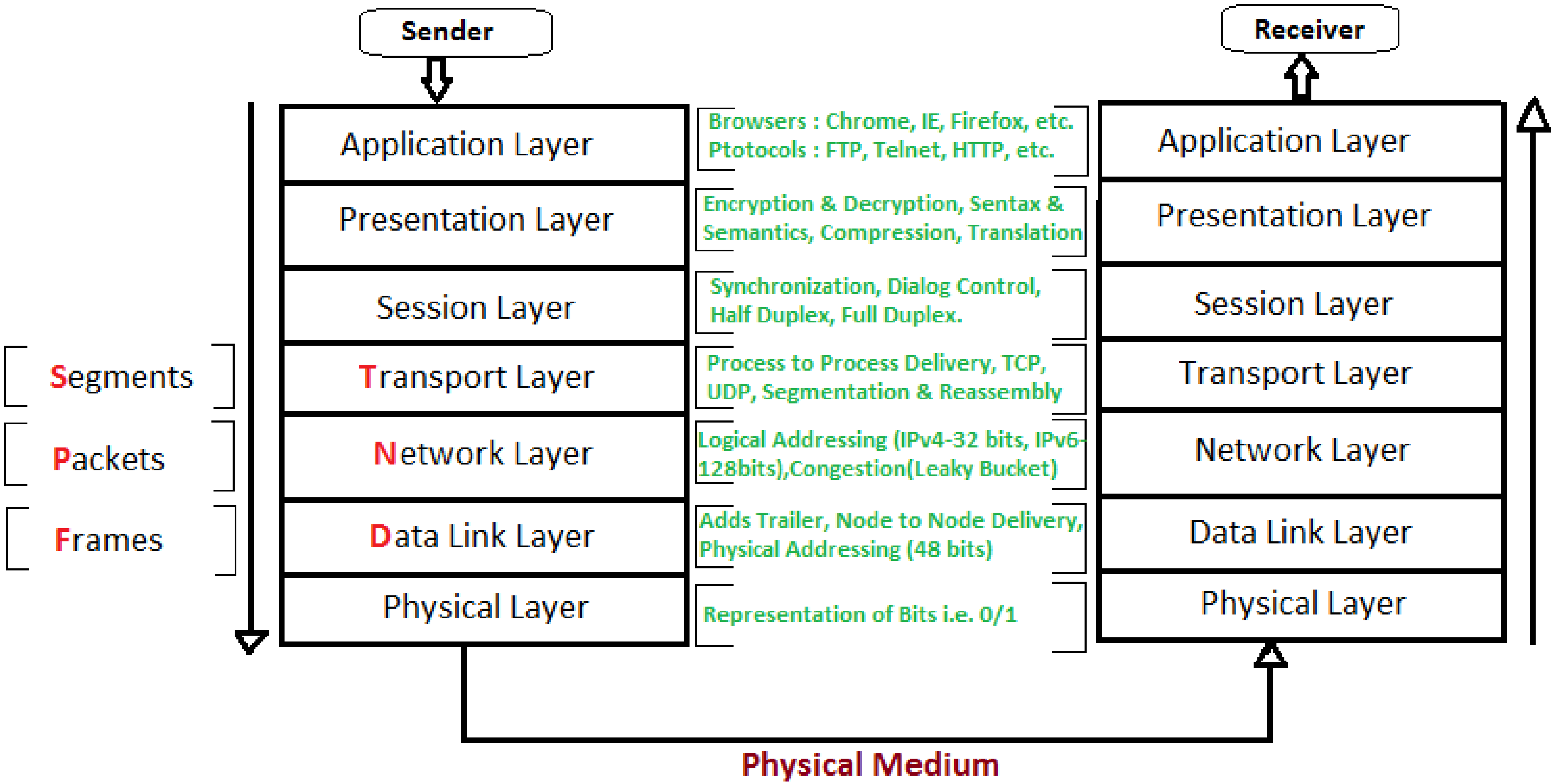
syed.naqvi@bcu.ac.uk

The Higher
Education
Academy

Outline

- Introduction
- Domain Name System (DNS)
- Web browsers and HTML pages
- Forensic Challenges
- Email Forensics

What is a Network?



Sender

Application Layer

Browsers : Chrome, IE, Firefox, etc.
Ptotocols : FTP, Telnet, HTTP, etc.

Presentation Layer

Encryption & Decryption, Sentax & Semantics, Compression, Translation

Session Layer

Synchronization, Dialog Control, Half Duplex, Full Duplex.

Segments

Transport Layer

Process to Process Delivery, TCP, UDP, Segmentation & Reassembly

Packets

Network Layer

Logical Addressing (IPv4-32 bits, IPv6-128bits),Congestion(Leaky Bucket)

Frames

Data Link Layer

Adds Trailer, Node to Node Delivery, Physical Addressing (48 bits)

Physical Layer

Representation of Bits i.e. 0/1

Receiver

Application Layer

Presentation Layer

Session Layer

Transport Layer

Network Layer

Data Link Layer

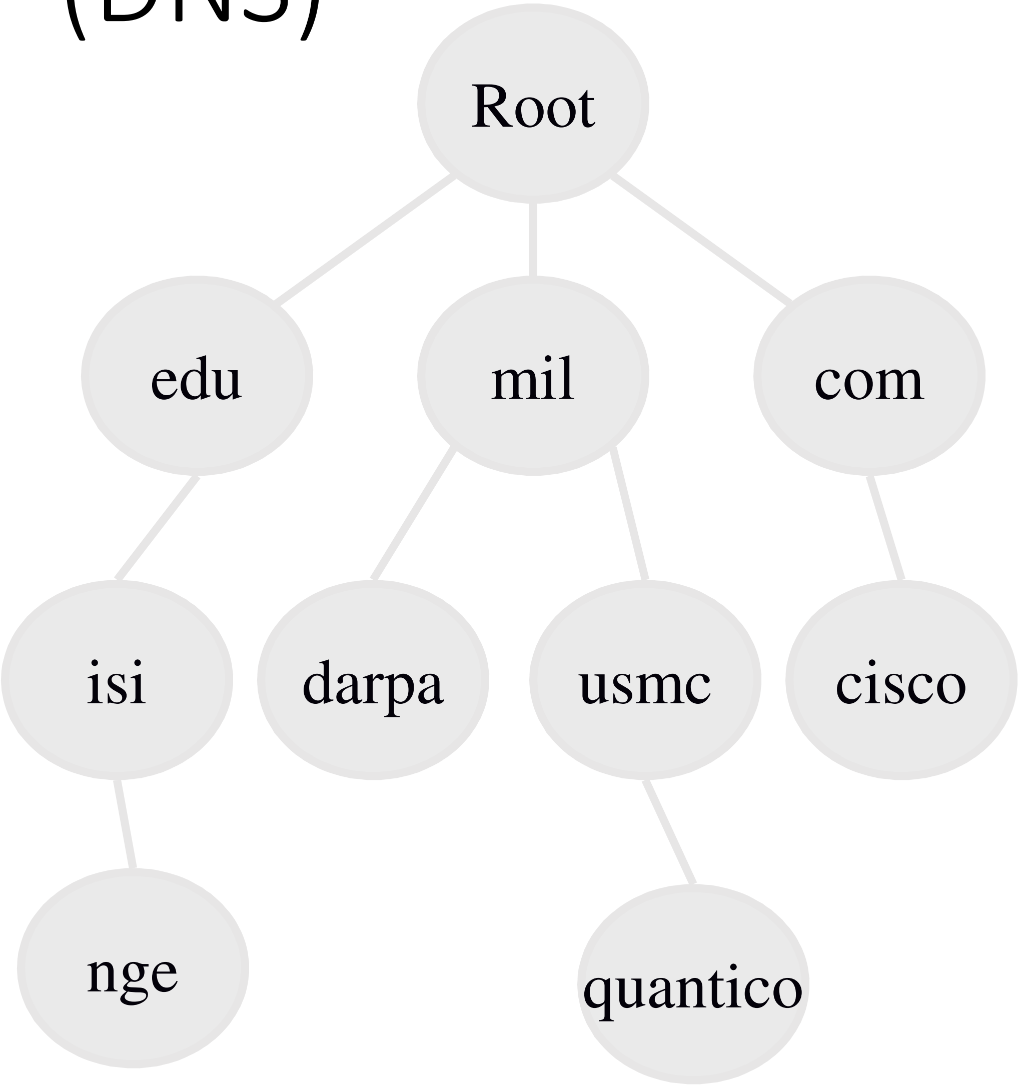
Physical Layer

Physical Medium

Key terms

- **Internet Protocol** (IP) address.
  - IPv4: Identifies a computer system, a 32 bit number formatted into 4 octets e.g. 192.168.1.1
  - IPv6: 128 bit address in eight 16-bit blocks in the format ***global***:***subnet***:***interface***.
    Example: **FE80:0000:0000**:**0000**:**0202:B3FF:FE1E:8329**
  - Address can be static or dynamic, well know systems tend to have a static address.
- **Domain name**.
  - Not easy to remember a system by its IP address, names are easier such as Google, Microsoft, Apple, Facebook, etc.
  - Domain Name System (DNS) is a means of associating a series of text character, the *name*, with an IP address.
  - Domain Name identifies a computer system (hostname) or name space of one domain, e.g. example.com is the domain and www.example.com a specific hostname.

Domain Name system (DNS)



- Virtually every application uses the Domain Name System (DNS).
- DNS database maps:
  - Name to IP address
    www.darpa.mil = 128.9.176.20
  - And many other mappings
    (mail servers, IPv6, reverse…)
- Data organized as tree structure.
  - Each zone is authoritative for its local data.

World Wide Web

- Developed for the transfer of text documents using a protocol called **Hypertext Transfer Protocol** (HTTP).
  - ***Hypertext*** due to the formatting embedded in the text plus the means to refer to other hypertext documents.
- Transfer between client and server involves using software called a **web browser** (**client**) and **HTTP server** (**server**).
- Client is typically a user PC, server is typically a computer system dedicated to the task of handling requests from clients.
  - Server has many CPUs, large amounts of RAM, multiple disk drives to handle multiple simultaneous requests.
- Transfer Protocol defines how a client connects to the server, requests data, manages data transfer and error information.

Data Transfer

- User requests data by typing in a **Uniform Resource Locator** (**URL**) into web browser’s address bar.
- URL describes a resource on the WWW and the means to transfer it.
- URL schema:
  - scheme:protocol specific part.
- The scheme describes the protocol by which the transfer will take place typically http goes here.
- The protocol specific part is the resource defined by:
  - / or // or /// determined by scheme.
  - domain name where resource resides.
  - path to resource on domain name.

Examples

- http://www.w3.org/Addressing/URL/url-spec.txt.
  - http is the scheme.
  - www.w3.org is the domain name.
  - Addressing/URL/url-spec.txt is the path to resource, url-spec.txt is plain a text file.
  - Extensions do have some significance to the web browser and how it displays (renders) the content or prompts to download it.
  - In this case the web browser shows a plain text document.

More examples

- http://moodle.bcu.ac.uk/course/view.php?id=8021
  - This is an example of a **dynamic** web page.
  - **view.php** is a text file in which are programming language (PHP programming language) statements are executed when the user accesses the resource.
  - The ? defines a parameter to pass to the web page so it can configure its behaviour, 8021 is the id of this course (Digital Forensics) Moodle web page.
  - view.php will get data from a database to fill out template content based on the course data input by tutor and additional data about the user.

EnCase File Signatures/Types

- Defines extensions (file types) for htm, html, asp, aspx, mhtml and mht type webpage document files.
- Discounts a lot of webpage extensions such as .php, .asp, .aspx and .jsp used for dynamic webpages.
- Multiple file signatures named HyperText Markup Language N File, where N is a number.
- Cover a number of different possible signatures but not always successful as poorly formatted HTML may not start with <!DOCTYPE or <HTML.

Web browsers

- Source: http://www.w3counter.com/globalstats.php
- Chrome is currently the most popular browser (45.2%).
- Safari next.
- Internet Explorer is supplied with Windows OS.
- Main web browser software in use today.
  - Chrome (Google).
  - Internet Explorer (Microsoft).
  - Firefox (Mozilla).
  - Safari (Apple).
  - Opera (Opera Software).

Web browsing and Computer Forensics

- World Wide Web was originally developed for the dissemination of information (academic).
- Now WWW used for entertainment and recreation (youtube), social networking (facebook, twitter), exchanging messages (webmail) and much more (sadly not always legal).
- On a PC user will interact with the above services using a **web browser**, a program installed on user's PC.
- By default web browser creates a number of artefacts on user's PC due to browsing activity. Can use these to ascertain what the user was browsing and when.
- With some limitations.
- There are a number of web browsers where all offer the same overall behaviour vary in how they store and display webpages.

Browsing history

- Computer forensic examiner needs to be familiar with the differences between major web browsing software.
- Browsers store a history of user browsing activity generally known as the **Internet history cache**.
- A throwback to when individuals used dial-up connections that were slow (typical download speeds of between 2 and 5KB/s).
- Downloaded webpage plus any associated images were stored in cache so that if the user returned to the webpage, web browser would get webpage from cache not download again (within a certain amount of time). Results in faster operation as quicker to load from disk.
- Location and format of data in cache differs for each browser.

Web browsing artefacts

- HTML documents (webpages).
- Image files.
- Cookies – small text file exchanged between web browser and http server. Used to record a small amount of data such as to record the act of logging on. Can be useful in corroborating individual browsed website but not absolute as artefacts on webpage may link to other http servers and create a cookie.
- CSS files (Cascading Style Sheets). A text file format that helps with making a webpage aesthetically pleasing.
- Downloaded files (via web browser).
- JavaScript 'program' files, make the page dynamic including implementing simple games.
- Audio/video data in the form of Macromedia Flash files.

Examining webpages

- Advantage for the examiner is that webpages are text data.
- Easy to view in EnCase and search for using keywords.
- However problem is useful text is interspersed with tags.
- May have to construct keywords based on tags used in data.
- Email addresses when marked up often use a tag where the email address is an hyperlink.
- Click on link results in starting email client software so user can send an email to that address.

Challenges for the computer examiner

- Viewing HTML document as it would be when viewed in browser as web page can difficult as some web browsers modify web page including name for quick retrieval for cache.
    - Need to find what the original web page was called and source URL.
- Makes viewing in EnCase difficult.
- Some web browsers don’t change the sources to images, for example, in document used in tags so viewing the document in web browser is dangerous as web browser will try to download the image from source.
- Image may be stored in cache so need some way to make browser use cache image not download image. Could be dangerous for examiner PC to download images/files.

Temporary Internet Files folder

- Contains a number of sub-folders where there will be at least four sub-folders containing files downloaded as a result of web browsing.
- Folders have random names, contain files downloaded by Internet Explorer.
- Files are in their original format, e.g. HTML files are text, image files will be in the format they were downloaded in such as JPEG or PNG.
- However no specific order to the files other than in time order.

Index.dat

- Stores browsing activity that Internet Explorer displays in order of days for the last week, then weeks for the last two weeks.
- Flat file containing a number of different types of records.
- URL are 384 byte structure detailing:
    - URL of the resource accessed.
    - Timestamps at locations 08 and 10 (hex) from start of record.
- Location 105 is start of URL where resource originated from, stored in ASCII.
- Separate index.dat files for cookies (in Cookies folder) and History folder, one for daily and one for each of the two previous weeks.

HTTP Cookies

- Some webservers initiate the creation of a cookie on user’s PC when they access the website.
- Cookie contains information the webserver can use when the user revisits website.
- Webserver retrieves cookie (web browser returns it) and analyses contents.
- Webpage returned is configured based on this content, e.g. cookie may store a friendly name of the user so web page says “Hello “ followed by name.
- Potential privacy issues where today websites have to state use of cookies (EU law passed May 2011).
- Cookies are text file but the contents is not meant to be human readable.
- Source of browsing history.

In-private browsing

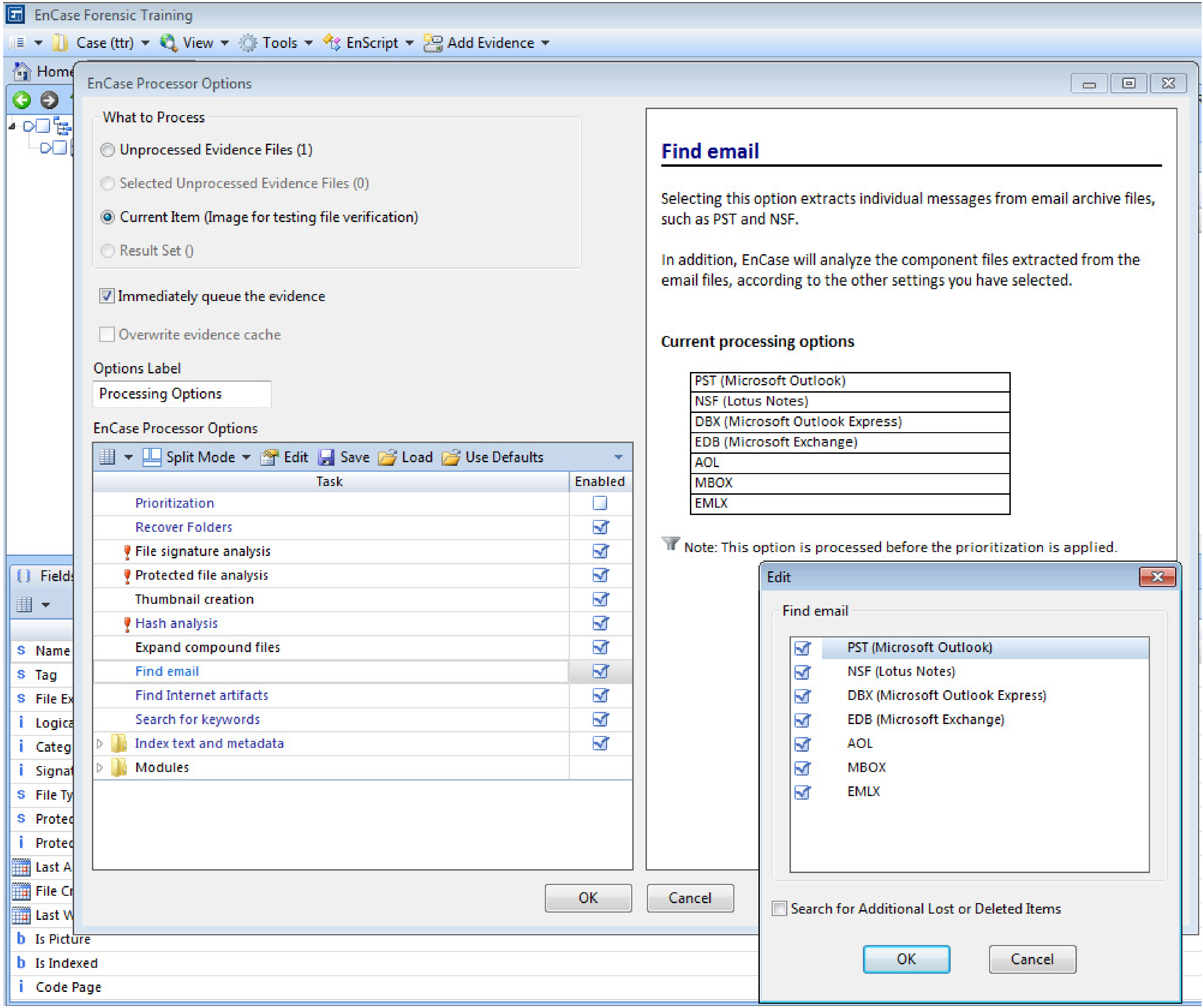
- Most modern browsers have this mode (also known as in-cognito mode).
- Allows browsing without storing any artefacts of browsing activity on PC or deletes browsing history (temporary files)

Downloaded files

- OS tracks downloaded files using Alternate Data Stream to record source of file.
- ADS identified by the Zone.Identifier name on the file name.
- Contains text typically:
  - [ZoneTransfer] ZoneId=3
- ZoneId = 3 indicates source was Internet.
- OS inspects this number and takes appropriate action, e.g. displays a warning prompt for program files as program is deemed unsafe by default.

Email forensics

- Exercise



EnCase Processor Options

What to Process

○ Unprocessed Evidence Files (1)

○ Selected Unprocessed Evidence Files (0)

● Current Item (Image for testing file verification)

○ Result Set ()

☑ Immediately queue the evidence

□ Overwrite evidence cache

Options Label

Processing Options

EnCase Processor Options

Task	Enabled
Prioritization	□
Recover Folders	☑
File signature analysis	☑
Protected file analysis	☑
Thumbnail creation	☑
Hash analysis	☑
Expand compound files	☑
Find email	☑
Find Internet artifacts	☑
Search for keywords	☑
Index text and metadata	☑
Modules	

Find email

Selecting this option extracts individual messages from email archive files, such as PST and NSF.

In addition, EnCase will analyze the component files extracted from the email files, according to the other settings you have selected.

Current processing options

PST (Microsoft Outlook)
NSF (Lotus Notes)
DBX (Microsoft Outlook Express)
EDB (Microsoft Exchange)
AOL
MBOX
EMLX

Note: This option is processed before the prioritization is applied.

Edit

Find email

☑ PST (Microsoft Outlook)
☑ NSF (Lotus Notes)
☑ DBX (Microsoft Outlook Express)
☑ EDB (Microsoft Exchange)
☑ AOL
☑ MBOX
☑ EMLX

□ Search for Additional Lost or Deleted Items