A Strategy for Testing Graphic File Carving Tools

Jim Lyle & Rick Ayers

National Institute of Standards and Technology

Disclaimer

Certain trade names and company products are mentioned in the text or identified. In no case does such identification imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products are necessarily the best available for the purpose.

CFTT at NIST

- Assurance that the forensics software used in investigations works well enough that the results can be admitted in court.
- Independent testing (or at least an independently designed test methodology)
- NIST develops the test methodology and tests selected tools (CFTT)
- NIST also develops and posts data-sets (CFReDS) for testing forensic tools

Outline

- File Carving Background
- Creating data-sets for file carving
- Measuring results
- Some behaviors observed
- Summary

File Carving

- An investigator may want more than just what is visible within a file system
- Deleted information can be recovered
  - File system meta-data based recovery
  - Data signature based recovery, aka “file carving”
- File carving – reconstructing deleted files from unallocated storage based on file content, file system meta-data can be ignored

Background

- Many file types have recognizable signatures in the file data
    - ➢ Graphic – jpeg, gif, png, bmp & tiff
    - ➢ Video – mp4, wmv, 3gp, ogv, mov, avi
    - ➢ Document – doc, docx, xls, xlsx, pdf, ppt & pptx
    - ➢ Archive – zip, rar, 7z, gz & tar
    - ➢ Others -- ???
- Can’t test all at once

Other Work

- DFRWS file carving challenges
  - Completeness
  - Fragmentation
  - Fragment order

- DFTT data set

Testing Issues

- Dozens of parameters that might affect tool behavior
- Focus on most important parameters
  - Completeness
  - Fragmentation
  - Embedded pictures (thumbnails)
  - Tool option settings (use default values)
- Be aware of other issues like . . .
  - File type specific characteristics
  - Compression level
  - Thumbnails
  - EXIF data
  - Audio track

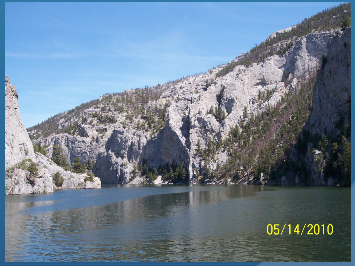
Data Sets for Graphic Files

- Collection of separate graphic files:
  - Barn.gif
  - Winter.tiff
  - River.png
  - Oak.jpg
  - Also bmp
- Eight files of each type
- Can construct “dd disk image file”

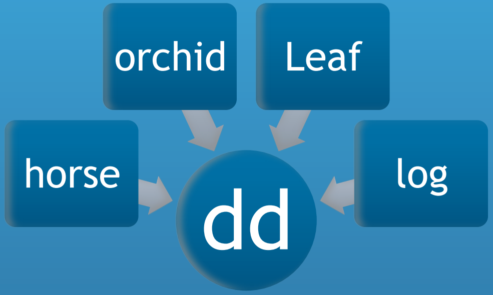
04/12/2009

12/06/2009

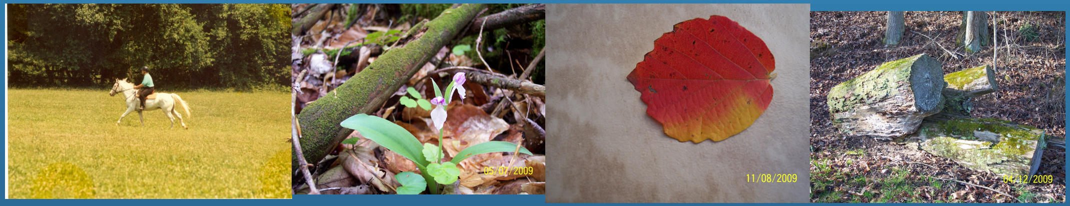
02/07/2010


05/14/2010

Base dd file – Complete & Contiguous Picture Files

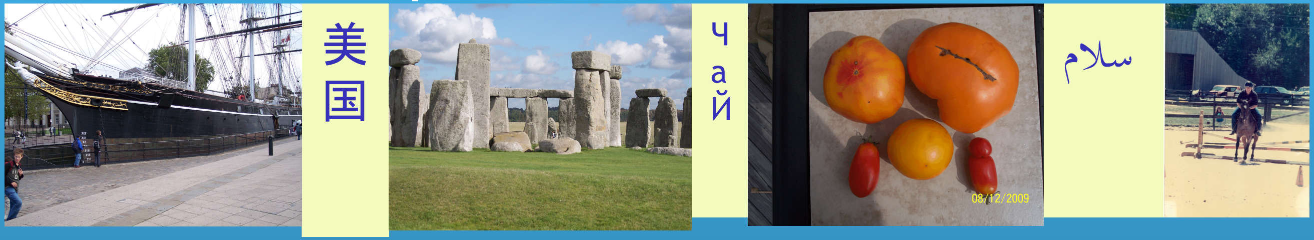


Zero fill to end of last sector



Constructing Other Images

- Padded with cluster sized blocks of text between pictures

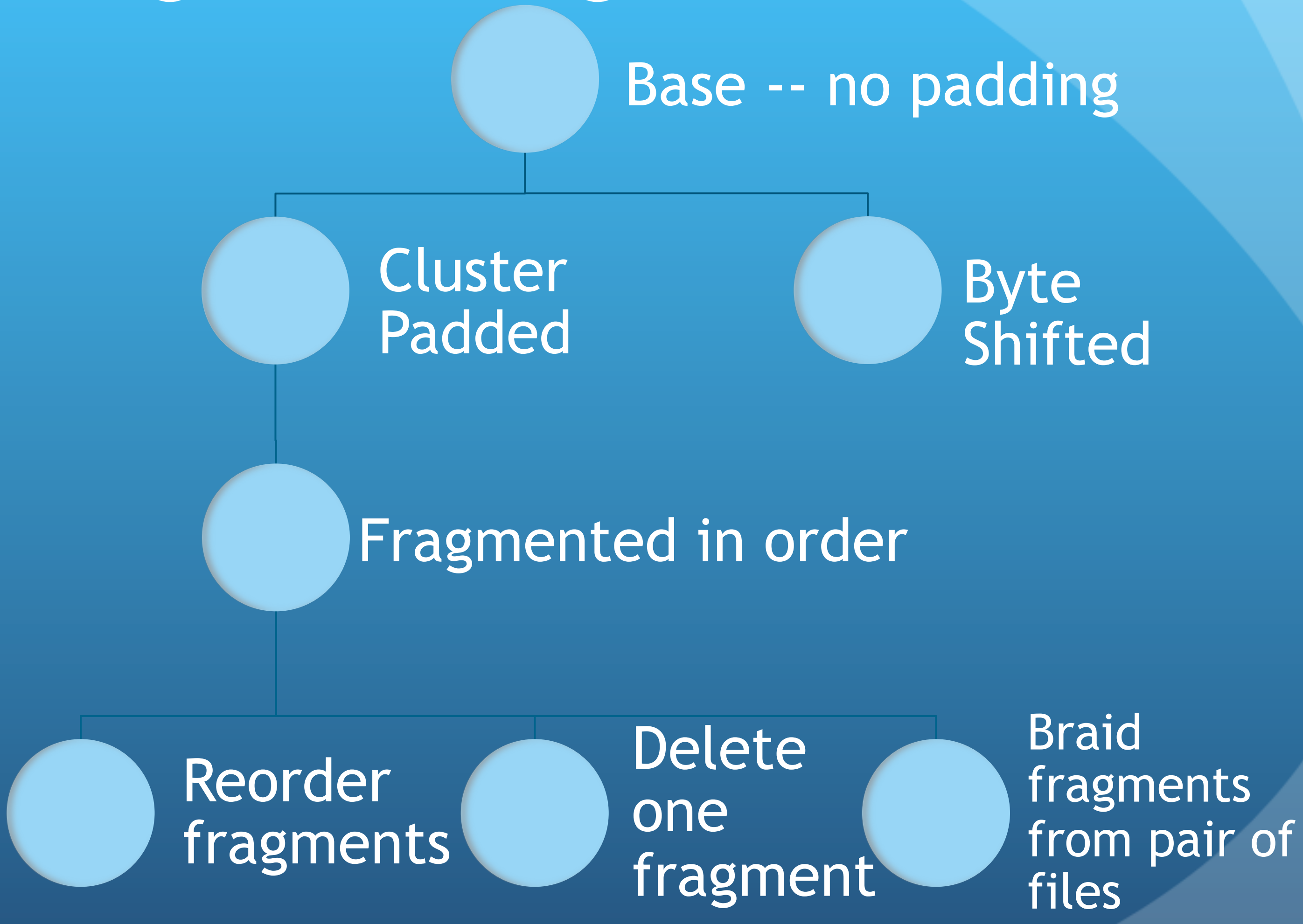


- Fragmented (in order)

Other dd images

- Fragmented (out of order)
- Braided (two files intertwined)
- Incomplete files
- Non-aligned to sectors

Carving Test Images



Base -- no padding

Cluster Padded

Byte Shifted

Fragmented in order

Reorder fragments

Delete one fragment

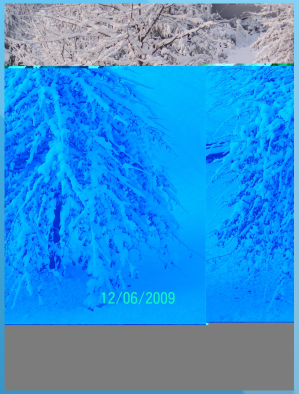
Braid fragments from pair of files

Measuring Results

- Two approaches –
  - Visibility driven – does the tool produce usable (viewable) results
  - Data driven – See what the tool actually does in relation to ground truth
    - Measure fraction of returned data that belongs
    - Measure fraction of possible data returned
- Methods are complementary

Visibility Driven Measurement

- Each file checked for visibility by two independent observers
- Resolve differences if disagreement



Category	Visibility
Viewable Complete	Flaws – minor or none
Viewable Incomplete	Flaws – partial, multiple files
Not viewable	Data matches file type, Flaw prevents display
False Positive	Data doesn’t match file type

Data-driven Measurement

- We know the ground truth
- Based on sectors present in carved files and information retrieval based statistics – evaluate returned data
  - Relevant – sector comes from a source file in dd file
  - Retrieved – sector returned in a carved file
- P = (relevant ∧ retrieved)/retrieved -- fraction of retrieved sectors from a source file -- **how much noise returned**
- R = (relevant ∧ retrieved)/relevant – fraction of relevant sectors retrieved – **how much stuff missed**
- F = 2 x (P x R)/(P + R) – average of P & R

Testing Plan

- Test reports for tools carving . . .
  - Graphic (jpg, gif, etc.) files -- will be published soon
  - Video files – drafting reports now
  - Next class – Documents? Archives? Audio?

General Results

- Most tools find majority of non-fragmented jpg & gif
- Recovered bmp files usually viewable
- Most recovered tif files not viewable
- Tools usually have different behaviors, e.g.,
  - Recover few files, but almost all viewable files
  - Recover many files, but most not viewable
- Occasionally, tool exhibits interesting behavior . . .

A Rabbit-hole of Interesting Behavior

- One tool (A) recovered 8 tiff files from the unpadded dd file
- F score for tiff files was 1.00
- But, only one file was viewable, seven were not viewable
- Examination of the eight files – last sector of tiff file replaced by noise in the carved file
- That last sector is critical to having a displayable file
- Other tools on same data –
    - Tool B Carved 4 with 3 viewable
    - Tool C Carved 10, none viewable
    - Tool D Carved 8, all viewable
- Without both measures we wouldn’t know how close the tool was. Maybe an investigator can repair the file and extract a critical piece of evidence

Summary

- NIST/CFTT is creating downloadable data-sets for testing file carving tools – with ground truth
- Downloadable tools for creating additional test images and analyzing the results
- DHS is publishing test reports for carving tools – graphic files soon, video files later this year
- Tools behaviors can be compared using common data-sets
- NIST/CFTT is publishing raw test data for examination
- The data-sets reveal interesting tool behavior

Sponsors

- NIST OLES
- DHS S&T

Contact

Jim Lyle
JLYLE@NIST.GOV

Barbara Guttman
BARBARA.GUTTMAN@NIST.GOV

Rick Ayers
RICHARD.AYERS@NIST.GOV

Susan Ballou
SUSAN.BALLOU@NIST.GOV

http://www.cfreds.nist.gov	Test Data Sets
http://www.cftt.nist.gov	Test Reports

Thanks, Any Questions?