Selective Imaging Revisited

Johannes Stüttgen

Lehrstuhl für Informatik 1 (IT-Sicherheitsinfrastrukturen)
Universität Erlangen-Nürnberg

7th International Conference on
IT Security Incident Management & IT Forensics

12.03.2013

Forensic Evidence Acquisition

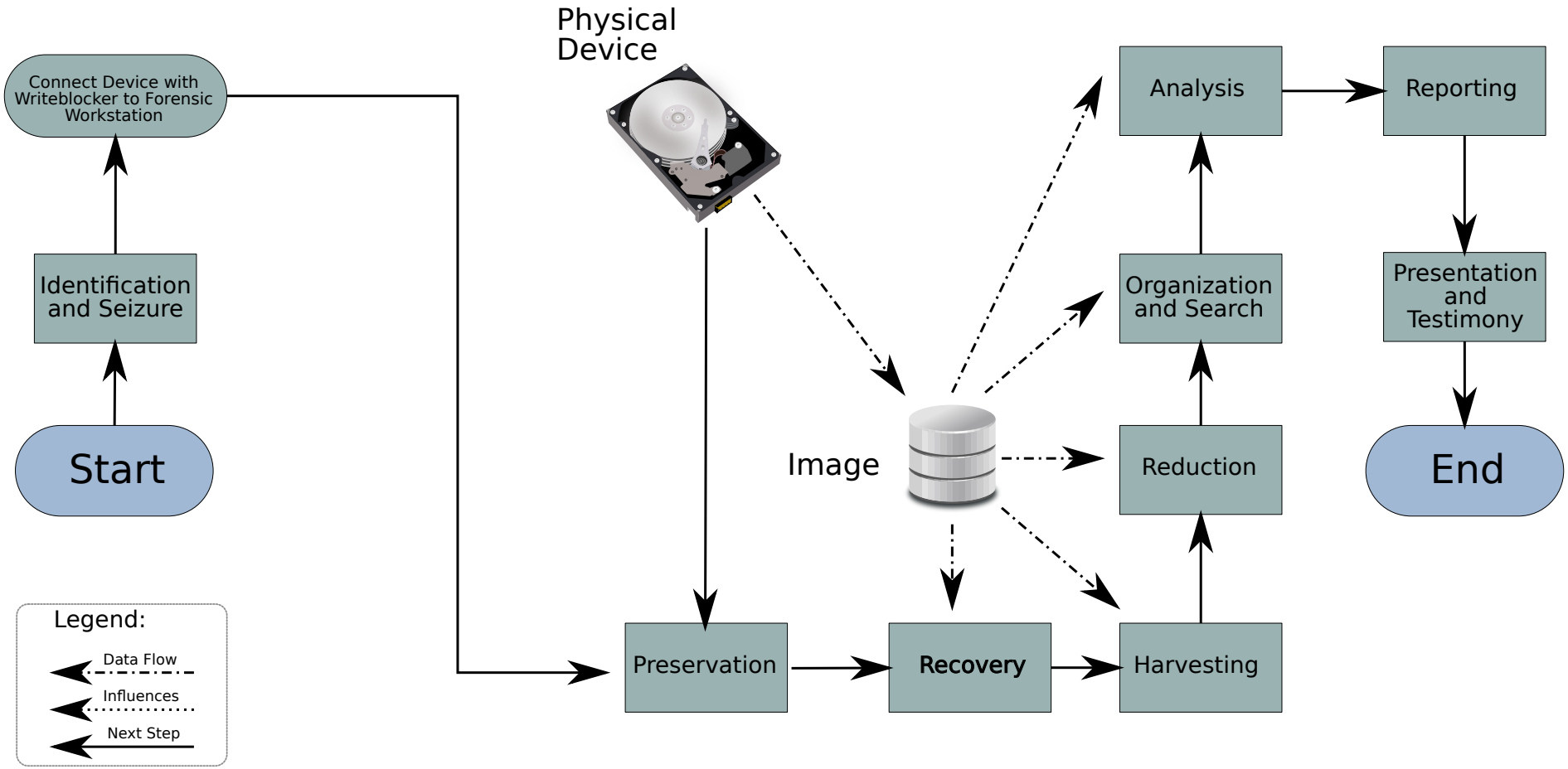
Evidence from hard-disks is usually acquired by creating a sector-wise image:

- Reduces risk of accidentally modifying evidence
- Absolute certainty that all possible pieces of evidence have been acquired

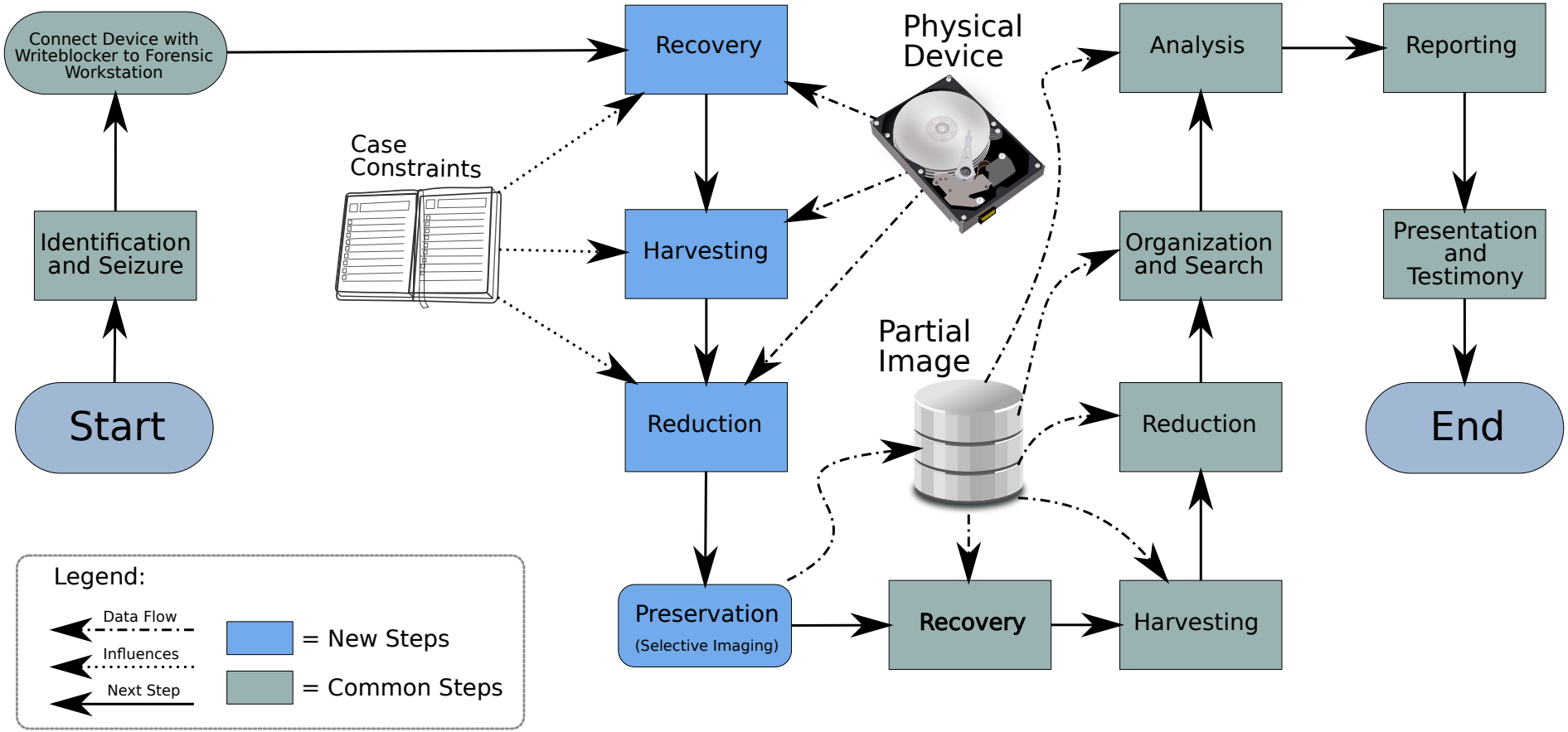
Problems of Sector-Wise Images

- Duration of the process
- Capacity of HDD grows faster than transmission bandwidth
  - Imaging of a HDD with 500 GB capacity with a USB 2.0 Writeblocker (30MB/s) takes about 04:45h
  - Imaging of a HDD with 2 TB capacity using an eSATA Writeblocker (70MB/s) takes about 08:20h
- Data-Protection and privacy concerns don't allow for the acquisition of entire devices in some cases
  - The extend of the acquisition of data should take principle of proportionality into account
  - On a system, used by several people unrelated to a specific case, only the data of the accused is relevant

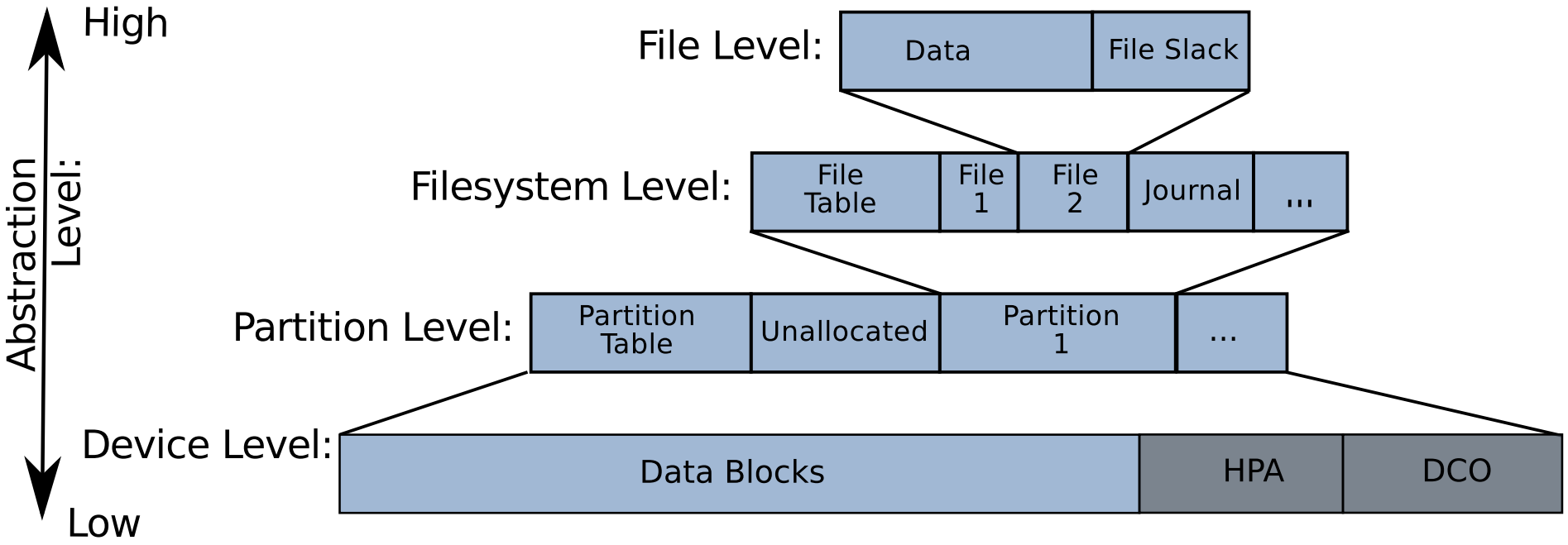
Investigative Process



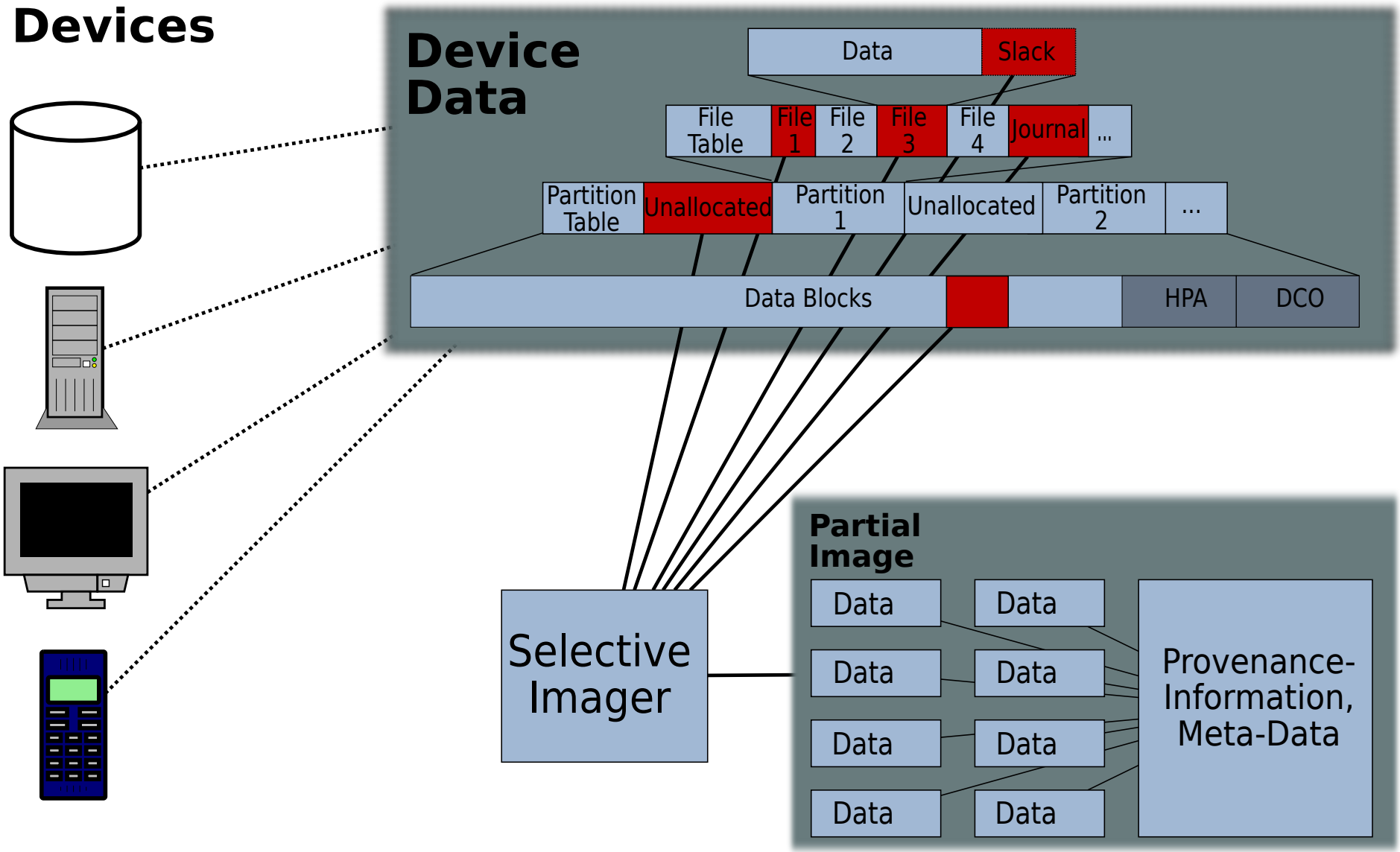
Modified Process



Granularity



Acquisition Procedure



Partial Images

Container for (logical) subsets of a data storage device. Has to fulfil these requirements:

- Storage of arbitrary data objects
- Storage of meta-data of all levels of abstraction (partitions, file-system, ...)
- Storage of results from pre-analysis steps
- Verifiability with the original device

Definition

A Partial Image is a set of data objects from a digital device, together with all relevant metadata, where integrity and provenance is verifiable with the original at all times.

Verifiability I

Two important attributes of a forensic Image have to be verifiable:

- Provenance
- Integrity

With a sector-wise image this can be done by comparing hashes. This guarantees:

- All data in the image at a specific address comes from the exact same address on the original device
- Data in the image has not been tampered with (if the original device is still intact)

Partial Image $\subset$ Data on device $\Rightarrow$ Approach is not feasible

Verifiability II

Hashes can still be used for verification of partial images, as long as:

- Verification is performed separately for every data object
- The partial image stores provenance information for each data object (”Provenential Key”)

Multiple provenential keys are possible:

- Sector-Address of all allocated blocks
- Cluster-Address in the file-system (If object is a file)
- Path in the file-system (If object is a file)

Not every key is applicable in every case, a combination of multiple keys is useful

Metrics of Verification

The selective imager stores multiple provenential keys for the verification of data objects in partial images:

- MD5Sum — Cryptographic Hash
- Byteruns — List of the addresses of each byte on the original device
- Path — The Path in the File-System

A combination of the Hash and any of the other keys allows reliable verification

Verification

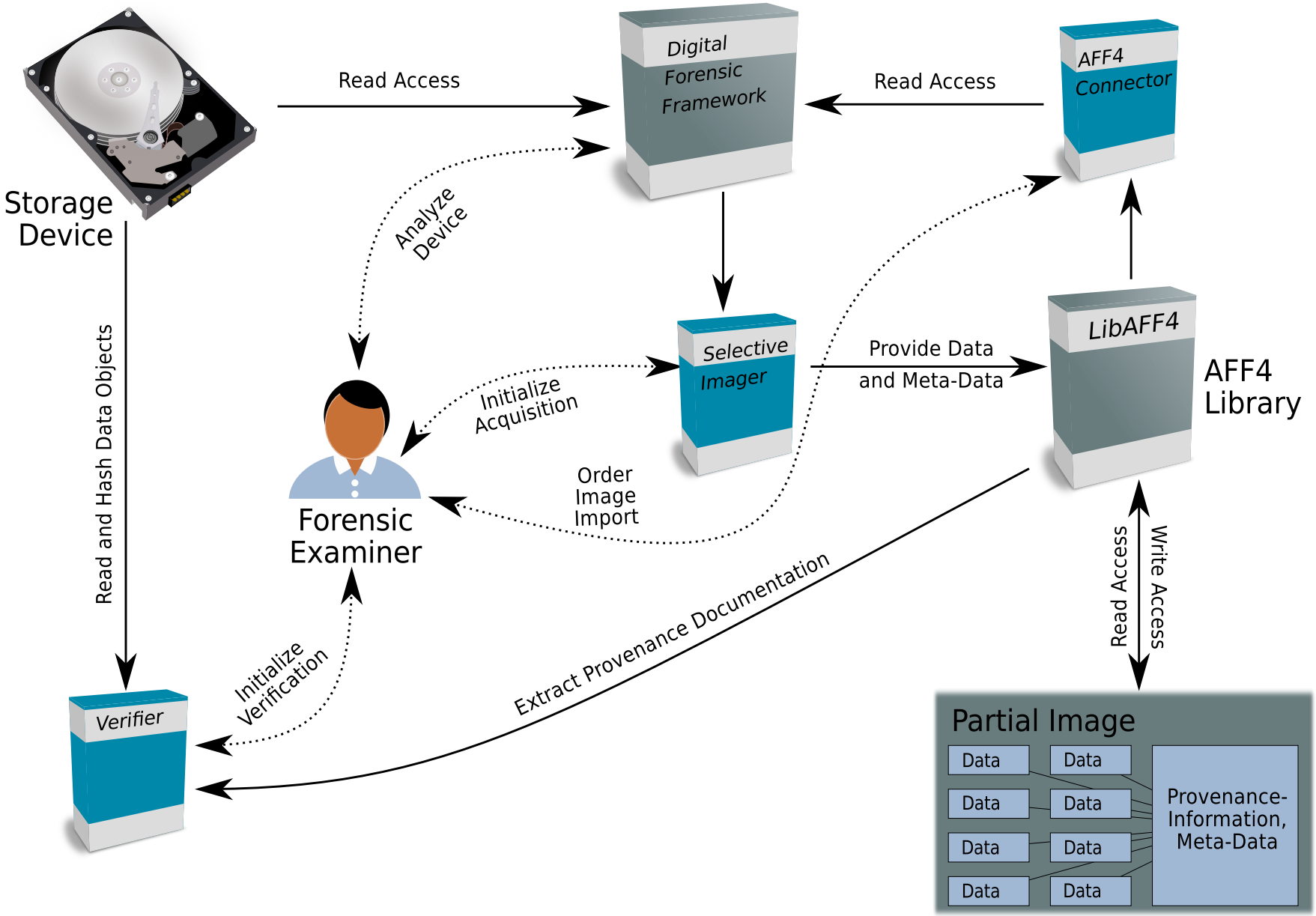
The process is easily automated and also allows verification of heavily fragmented objects:

driest@vogeltop:~/uni/thesis/code

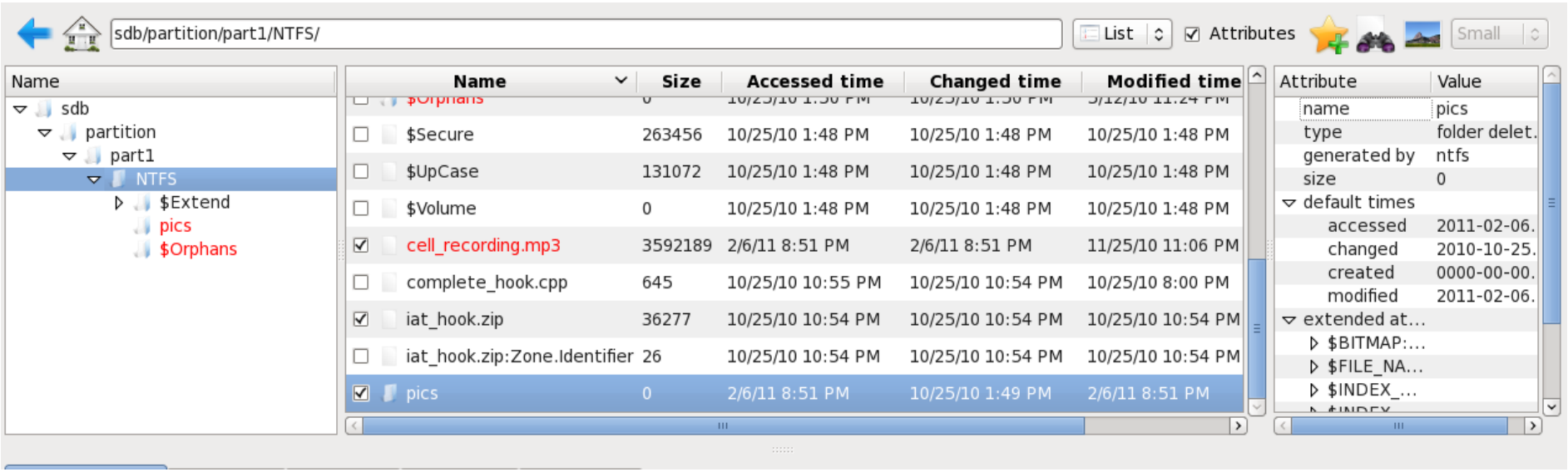
driest@vogeltop:~/uni/thesis/code 112x22

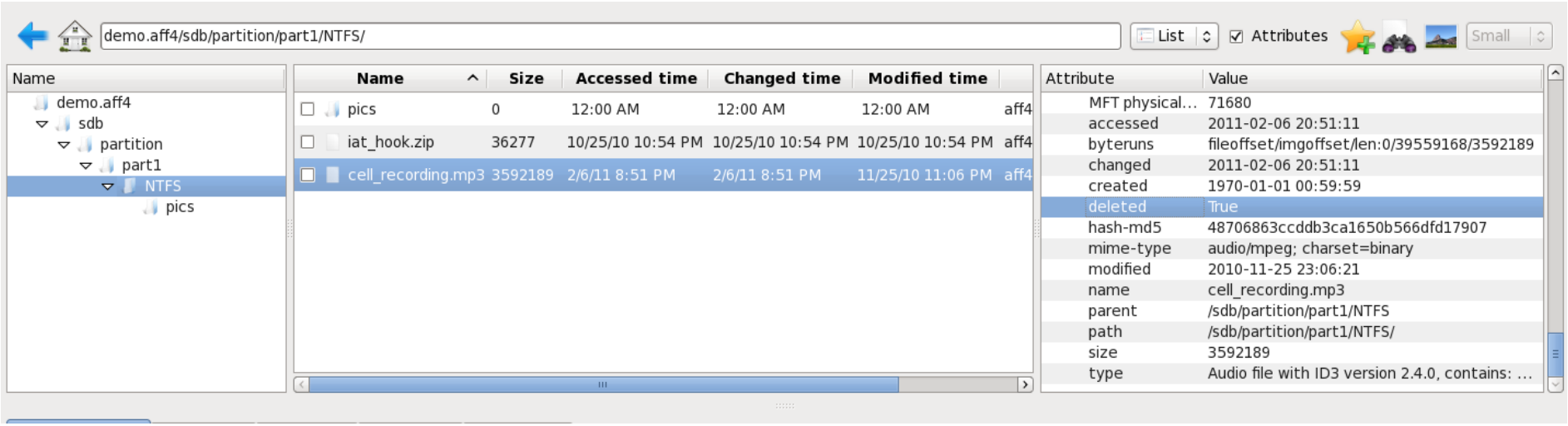
```
Verifying 'GermanWordlist.zip':
        Recorded Provenance:
                Byteruns:'fileoffset/imgoffset/len:0/1183940608/4096 4096/1183944704/4096 8192/1183948800/4096 1
183952896/4096 16384/1183956992/4096 20480/1183961088/4096 24576/1183965184/4096 28672/1183969280/4096 32768/118
6/4096 36864/1183977472/4096 40960/1183981568/4096 45056/1183985664/4096 49152/1183989760/4096 53248/1183993856/
7344/1183997952/4096 61440/1184002048/4096 65536/1184006144/4096 69632/1184010240/4096 73728/1184014336/4096 778
4018432/4096 81920/1184022528/4096 86016/1184026624/4096 90112/1184030720/4096 94208/1184034816/4096 98304/11840
4096 102400/1184043008/4096 106496/1184047104/4096 110592/1184051200/4096 114688/1184055296/4096 118784/11840593
6 122880/1184063488/4096 126976/1184067584/4096 131072/1184071680/4096 135168/1184075776/4096 139264/1184079872/
43360/1184083968/4096 147456/1184088064/4096 151552/1184092160/4096 155648/1184096256/4096 159744/1184100352/409
40/1184104448/4096 167936/1184108544/4096 172032/1184112640/4096 176128/1184116736/1196 '
                MD5:'9f82e852cc53c10138298abf5721a8fe'
                SUCCESS:
                        old hash:'9f82e852cc53c10138298abf5721a8fe'
                        new hash:'9f82e852cc53c10138298abf5721a8fe'


SUCCESS         SUCCESS         SUCCESS
The integrity of all listed streams has been verified!
[driest@vogeltop code]$ ./aff4verify
usage: aff4verify device imagefile [logfile]
[driest@vogeltop code]$
```

Reference Implementation



Acquisition and Import





Efficiency

The procedure has been tested and compared to standard procedure in two very different cases:

- Analysis of a 8GB flash-drive with simple file-system based recovery
- Analysis of a 20GB hard-disc with ”complicated” recovery

Results:

- Results of the flash-drive investigation were obtained 28% faster with selective imaging
- For the HDD without file-carving, results were obtained 40% faster
- For image storage, between 94% and 99.6% of space was saved
- File-Carving destroys the entire time advantage, because it requires all data on the device being read

Speed and Wear

- I/O throughput is substantially lower using the selective approach
- Can probably be increased by sequential disk access
- Amount of data transferred is significantly lower
- Disk Wear is significantly lower

Table : Imaging Speed by Tool and Features

Tool	Compress	Hash	Speed
dd			39.00
aimage			35.00
aimage	•	•	13.30
dff (raw)			32.28
dff (aff4)	•		27.03
dff (aff4)	•	•	26.62
dff (sel.)	•	•	15.56

Table : Device wear by investigative procedures.

Procedure	Sectors	Total
Filesystem Analysis	4,528	0.06%
Selective Imaging	119,624	1.53%
Sector-wise Imaging	7,827,392	100.00%
Carving	7,827,392	100.00%

Interviews

We interviewed forensic experts from industry and government agencies:

- Selective Imaging is already being employed on a file-level
- Often with unfit tools (Windows Explorer, Robocopy, ...)
- Even when Examiners use X-Ways or Encase, they fear overlooking evidence in Unallocated-Space or the File-System Slack Space
- Admission in court is not a problem
- In complicated cases the 100% coverage of a sector-wise image is useful
- Future developments will force investigators to sacrifice this coverage for the ability to operate at all...

Summary

- Forensic acquisition process was modified:
  - Preliminary short analysis directly on the device
  - Selection of relevant data
  - Detailed analysis on partial image
- Partial Images are:
  - Sets of data objects
  - Combined with meta-data
  - Verifiable
- Biggest necessity in cases involving:
  - Servers
  - Networks
  - Cases with previously known, strong constraints

The End

Any Questions?