10장. 디지털 증거 수집 및 분석 기술

박종혁

서울과학기술대학교 컴퓨터공학과

jhpark1@seoultech.ac.kr

개요

- 학습 목표
  - 디지털 증거 수집을 위한 활성시스템 조사, 디스크 이미징, 임베디드 시스템 조사와 관련한 기술을 살펴본다.
  - 실제 간단한 실습을 통해 디지털 증거 수집에 대한 학습의 이해와 경험을 획득한다.
  - 수집된 디지털 데이터를 분석하여 사건의 실마리 또는 증거를 찾기 위한 다양한 기술들을 살펴본다.
  - 디지털 증거 분석을 데이터 뷰잉, 검색, 통계 분석, 타임라인 분석 기술 등 다양한 기술에 대해 살펴본다.
- 학습 내용
  - 활성 시스템 조사
  - 디스크 이미징 기술
  - 임베디드 시스템 증거 확보 방법
  - 데이터 뷰잉, 디스크 브라우징 기술
  - 다양한 증거 분석 기술
  - 증거 분석 도구를 활용한 증거 분석
  - 안티 포렌식 대응 기술

목 차

1. 디지털 증거 수집 장비 및 SW

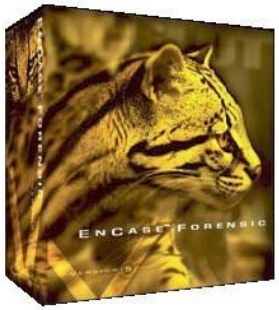
디지털 증거 조사 과정에서 사용되는 장비 / SW

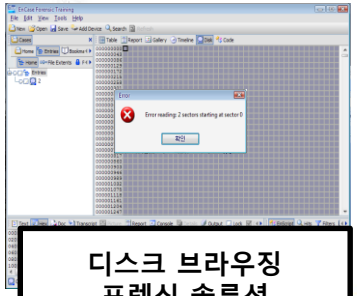
디스크 복제 장치
(ICS ImageMasster Solo4)

디스크 복제 장치
(Logicube Dossier)

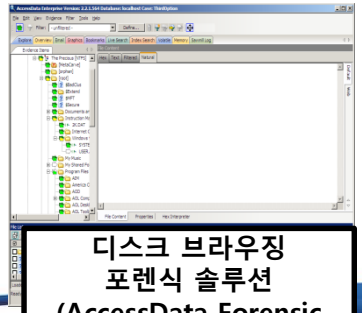
HDD 쓰기 방지 장치
(ICS Super DriveLock)

휴대용 포렌식 도구
(EnCase Portable)





디스크 브라우징
포렌식 솔루션
(Guidance Encase v6)



디스크 브라우징
포렌식 솔루션
(AccessData Forensic
Toolkit v3.0)

라이브 포렌식 도구
(Helix 3)

휴대폰 포렌식 솔루션
(Paraben Device Seizure v3.3)

2. 활성 시스템 조사

활성 시스템 조사

활성 시스템 (Live System)

- 운영 중인 시스템
- **휘발성 데이터(Live Data, Volatile Data)**
  : 시스템의 RAM에 저장되어 있어 전원을 차단하면 수집할 수 없는 데이터
- 활성 시스템 정보: "사건 현장에서 촬영한 즉석 사진" 과 같이 당시 **시스템의 동작 상태를 그대로 나타내는 시스템 사용 정보**

활성 시스템 조사란(Live Forensics)?

- Live Forensics: 활성 상태의 시스템에서 증거 수집 및 분석을 수행하는 일련의 조사 과정
- 디스크 이미지 조사 기반의 일반적 디지털 포렌식 과정과는 다르게 **시스템이 구동 중인 상태에서 데이터를 선별 수집 및 조사를 진행함**
- 시스템의 전원을 차단하면 수집할 수 없는 **휘발성 데이터**와 신속한 조사에 필요한 **비휘발성 데이터**를 선별해서 수집

활성 데이터의 개념

- 활성 데이터
  - 활성 시스템에서 수집할 수 있는 휘발성 데이터와 조사에 필요한 비휘발성 데이터를 포괄하는 개념

활성 시스템 조사에서 수집 데이터의 분류

비휘발성 데이터 수집의 필요성

- 하드디스크 용량이 급격하게 증가함에 따라 디스크 이미지 기반의 증거 조사가 어려워지고, 현장에서 증거의 선별 수집 및 즉각적인 분석이 이루어져 하는 상황 발생
- 서버와 같이 시스템을 압수하기가 용이하지 않은 경우, 시스템을 끄지 않고 조사를 수행할 수 있어야 함
- 법 집행기관이 아닌 단체에서 조사할 경우, 디스크 이미징을 수행할 수 없는 경우가 발생
  - 사건에 연관된 파일만 압수하는 선택적 압수 수색 영장에 대한 필요성이 제기됨

수집 정보 분류

- 휘발성 데이터, 실행 중인 프로세스 덤프, 물리 메모리 이미지
- 데이터 선별을 통한 최소한의 비휘발성 데이터 수집
  - 파일시스템 메타데이터: NTFS의 $MFT 파일, FAT 의 File Allocation Table 등
  - 운영체제 설정 정보: 시스템 이벤트 로그, 레지스트리 파일 등
  - 사용자 정보: 윈도우 계정, 시작 프로그램, 최근 접근 문서 등
  - 응용프로그램 정보: 인터넷 히스토리, 검색어, 웹 계정(ID/Password) 등

활성 시스템 조사 기법 -1

휘발성 데이터 분석

- 수집한 다양한 **휘발성 데이터에 대한 상관 분석을 통해 의미있는 결과 도출**
- 네트워크 데이터 분석 기술
  - 인가되지 않은 프로세스의 네트워크 연결 조사
  - 공유 자원 정보 획득 ➔ 정보 유출 증거 수집
- 물리 메모리 및 가상 메모리 이미지에 대한 분석 기술
  - 덤프한 프로세스 이미지에서 유용한 정보 추출
  - 각 메모리 이미지에 대한 특정 키워드 검색 기술 및 유용한 텍스트 추출 기능

운영체제 사용 흔적 조사

- 운영체제 사용과 관련된 정보 수집 (시스템 기본 설정, 레지스트리 등)
- 운영체제 설정 파일 분석을 통한 사용 패턴 분석
- 포렌식 관점에서 유용한 데이터에 대한 연관 분석

활성 시스템 조사 기법 -2

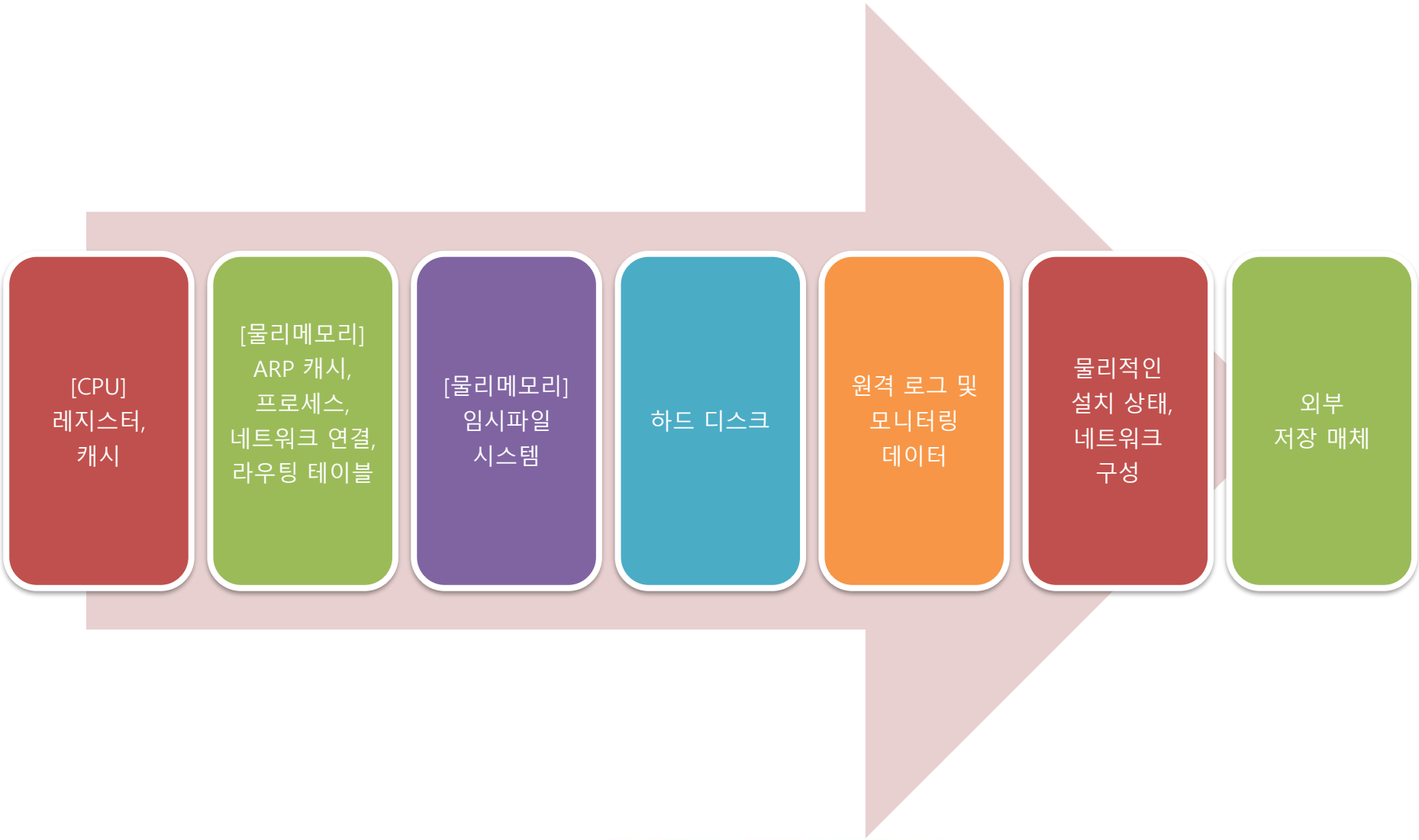
파일시스템 메타데이터

- 키워드 검색 등을 통한 수사 대상 선별
- 수집이 필요한 특정 파일(한글, 오피스 문서 등) 선택적 증거 수집
- 확장자 별 통계 분석을 통한 해당 컴퓨터의 용도 및 사용자 패턴 분석
- 타임라인 분석을 통한 사용자 행위 추적

응용 프로그램 분석

- 웹 브라우저, 전자메일, 메신저 관련 파일 분석을 통한 사용자의 관심사 파악
- 응용프로그램 캐쉬(Prefatch) 분석을 통한 사용 이력 조사

휘발성 정도에 따른 수집 절차 (RFC 3227)



활성 시스템 조사의 한계성

활성 시스템에서 조사를 위한 도구를 실행할 경우, 시스템에 불가피한 데이터 변형을 유발함

- Ex) 파일 접근 시간 변경, 메모리 데이터 무결성 훼손

DLL 후킹 기법 등을 사용하는 악성코드에 의해서 변조된 데이터를 수집할 가능성이 있음

- 메모리 분석을 통한 악성 코드 존재여부 파악이 필요함

삭제 파일에 대한 복구 불가능

- 이미지 분석과정에서 파일 복구를 시도해야 함

활성 데이터 조사 고려사항 (1/2)

조사 대상 시스템의 무결성 보장

- 이론적으로는 수집 데이터의 내용은 물론 메타 데이터까지 변경되지 않아야 함
- **휘발성 데이터 수집 및 분석은 필연적으로 대상 시스템에 영향을 미침**
- 휘발성 저장 장치에서 수집한 증거의 법적 효력에 관한 연구 필요
  - 절차적 방법 : 입회인의 서명
  - 기술적 방법 : 메타데이터 변경의 최소화, 수집 데이터의 위조 불가능성

데이터 수집의 용이성

- 이미지 생성 없이 필요한 데이터만 선별적으로 수집 가능

인증 절차 우회

- 사용자 인증이 불필요하며, 자동 암호화 솔루션 우회 가능

활성 데이터 조사 고려사항 (2/2)

활성 시스템 조사 시 주의사항

- 시스템에 주는 영향을 최소화
- 신뢰성 있는 도구 사용
- 신중한 조사
  - 한번 변경되면 원래 상태로 되돌릴 수 없음
- 조사과정의 기록
  - 조사 시각, 수집 데이터 명시 및 변경 데이터 명시

활성 데이터의 저장

CD를 이용하는 방법

- 활성 포렌식 도구를 CD에 저장하고 수집한 데이터는 USB 저장 장치에 저장하거나 네트워크를 이용하여 증거 수집 서버에 전송함
- 포렌식 도구의 신뢰성을 향상시킬 수 있음

USB Thumb Drive 이용하는 방법

- 대용량의 데이터를 저장할 수 있어 최근 많이 사용됨
- Windows의 Plug and Play 기능 활성화로 어려움 없이 사용 가능
  - Windows 시스템에 로그 정보 남음(setuplog.txt, setupapi.log)
- USB 저장 장치의 일부 영역을 제조사에서 제공하는 고유 도구를 이용하여 CDFS(CD File System)로 설정한 후 도구를 저장하고, 수집한 데이터는 데이터 영역에 저장하는 방식을 이용

휘발성 데이터 수집 및 분석

수집 목적

- **프로세스 정보**
  - 시스템에 악영향을 미치는 악성 프로그램, 이상 프로세스 판별
- **네트워크 정보**
  - 허가되지 않은 네트워크 연결 정보 확인
  - 실행 중인 프로세스의 네트워크 연결 정보 비교/분석
- **사용자 정보**
  - 대상 시스템에 대한 사용자의 흔적 정보들을 바탕으로 정황 증거 확보

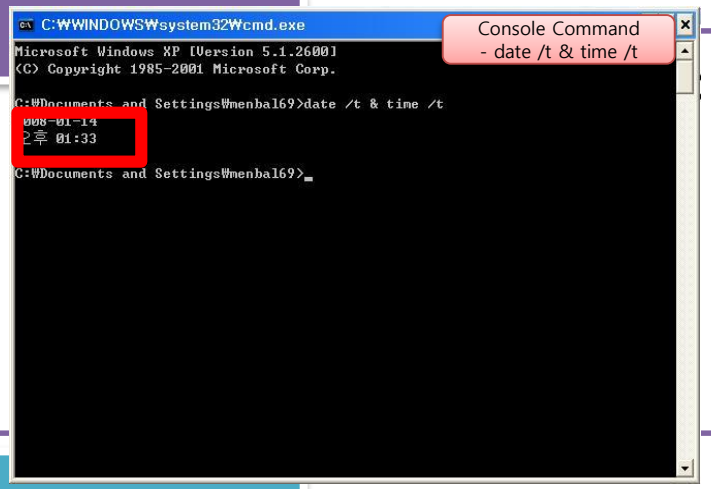
분석 방법

- **상관 관계 분석**
  - 수집 결과에서 서로 관계 있는 정보들을 추출하여 상관 분석을 시행
    - 예) 실행 중인 프로세스 리스트,열린 TCP 포트와 연결된 프로세스 정보, 네트워크 접속 정보 등을 비교 분석하여 비정상적 행위 분석(악성코드 감염 판별)
- **활성 정보를 분석할 수 있는 능력 필요**
  - 각각의 수많은 휘발성 데이터에서 의미 있는 결과를 도출할 수 있는 능력 필요

시스템 기본 정보-시스템 시간

System Time

- 증거 수집 시, 시간 기준이 됨
- 그 외 중요 요소들
  - Real time
  - 시스템의 구동시간(Up time)



수집 방법

- date, API를 이용한 Perl script
- date (/t : 현재 시간 바로 출력)
  - **휴대전화 시각 등 정확한 시간과 비교하여 시간차를 기록함**

시스템 기본 정보- 현재 로그온계정

현재 로그온 계정(Logged-on user)

- 수집 시스템의 현재 계정 정보 확보
- 정보의 주체가 누구인지 알아야 함

수집방법

- net users : netbios 명령어
- psloggedon : Sysinternals 사에서 제공하는 공개프로그램
- net sessions : netbios 명령어

시스템 기본 정보- 디스크 정보

디스크 정보

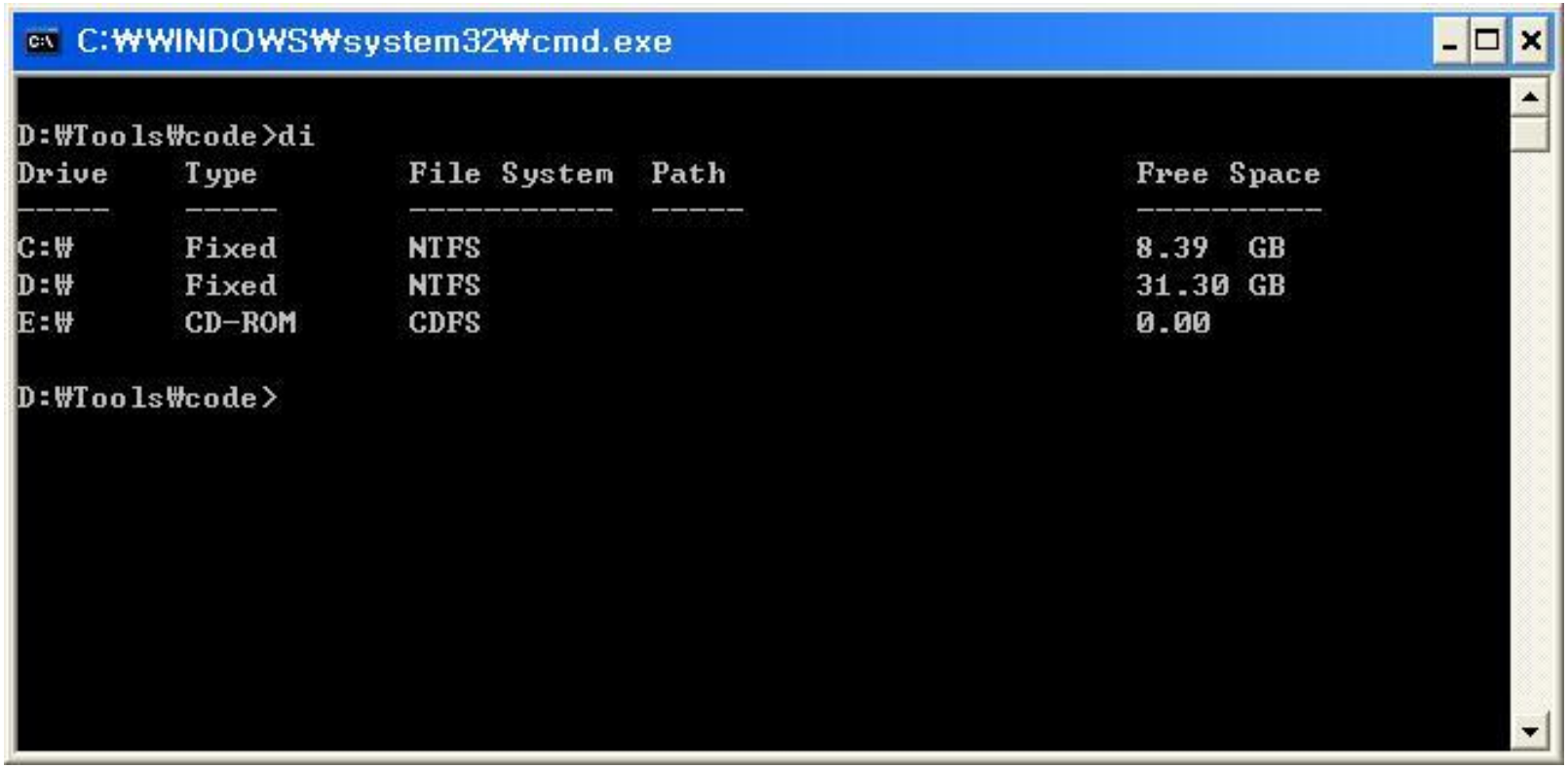
- 수집해야 하는 디스크 목록 확인

수집방법

- Di 명령어(Windows) - 파티션 구성 정보 확인 가능
- Unix
  - ioscan –fC disk
  - 디스크상태 – (cat /proc/diskstats)
  - 파티션정보 – (cat /proc/partitions)

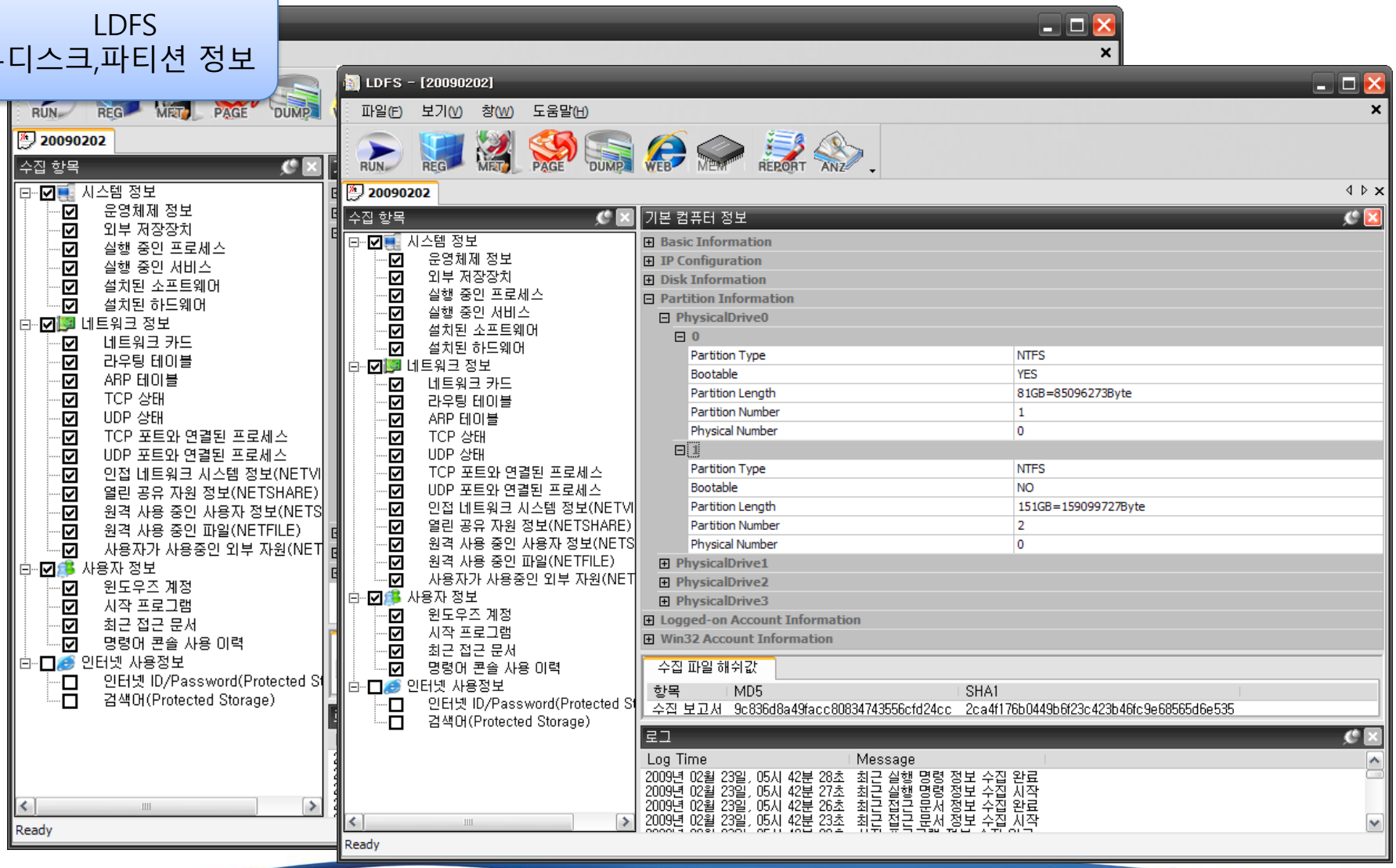
디스크, 파티션 정보(Windows)

- **di (각 Driver의 정보)**



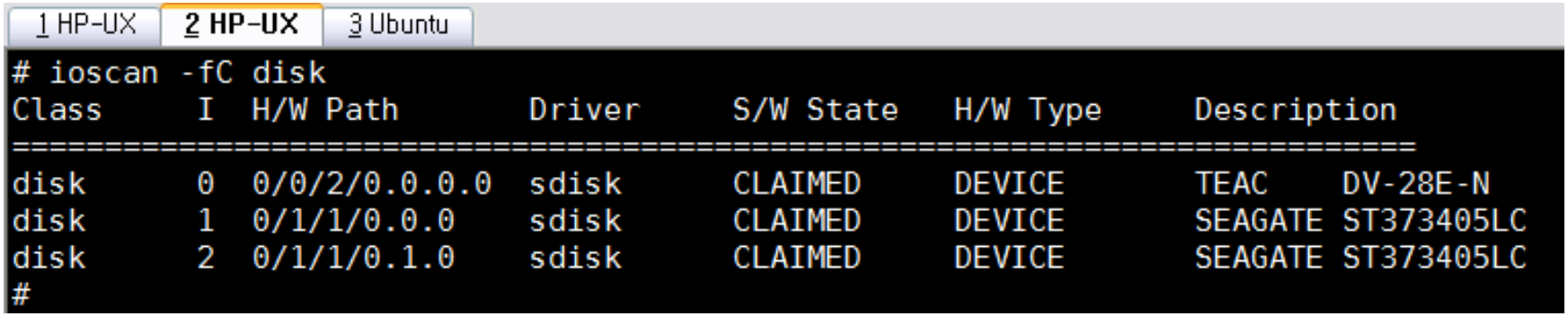
디스크, 파티션 정보(Windows)

LDFS
-디스크,파티션 정보



디스크, 파티션 정보(Unix)

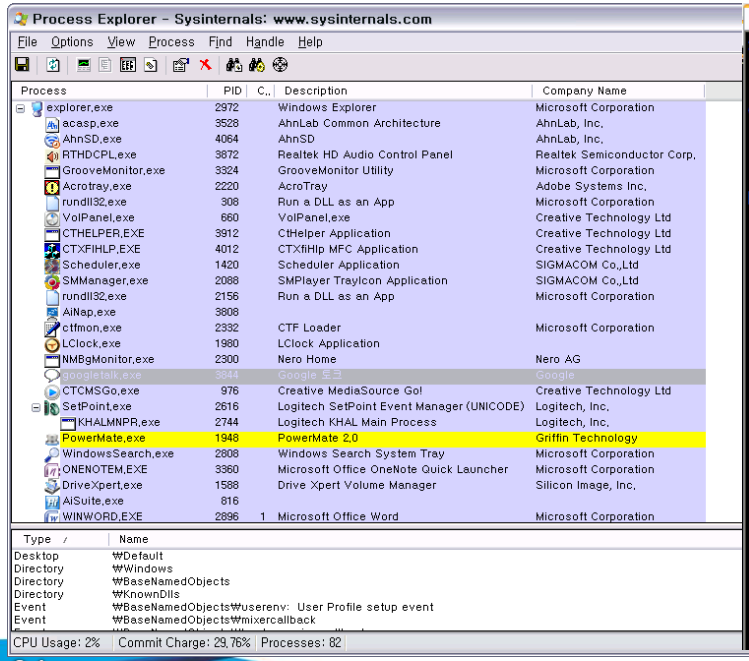
- **SYSTEM INFORMATION – 저장장치 - ioscan –fC disk**

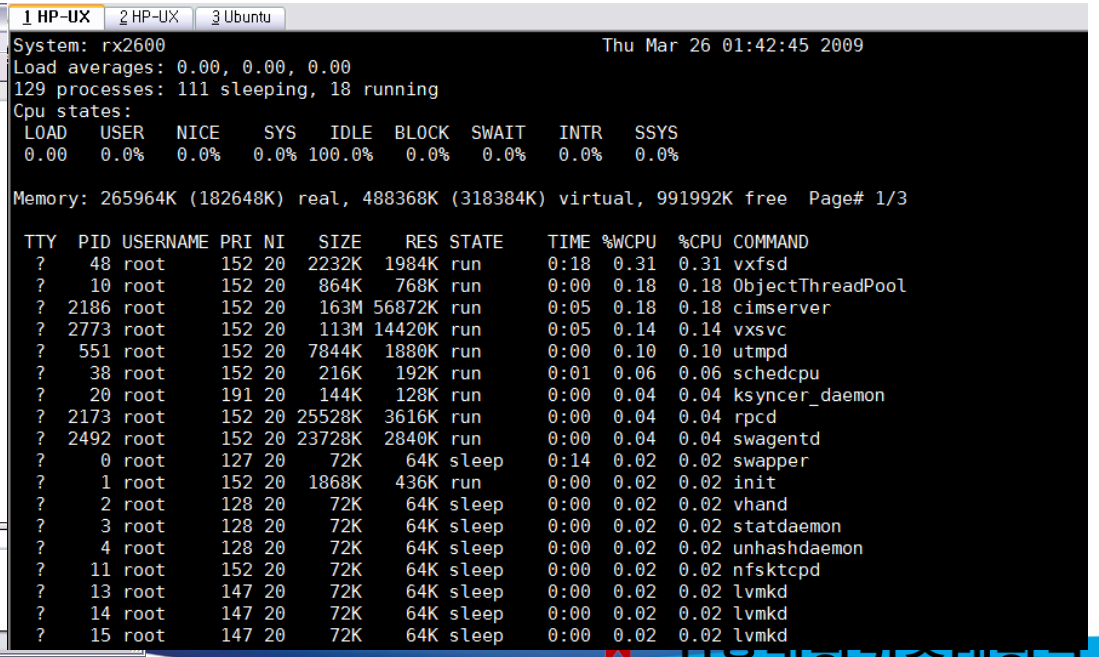


- ✓ **-f : full listing**
- ✓ **-C class : Restrict the output listing to those devices belonging to the specified class**

프로세스 정보 분석

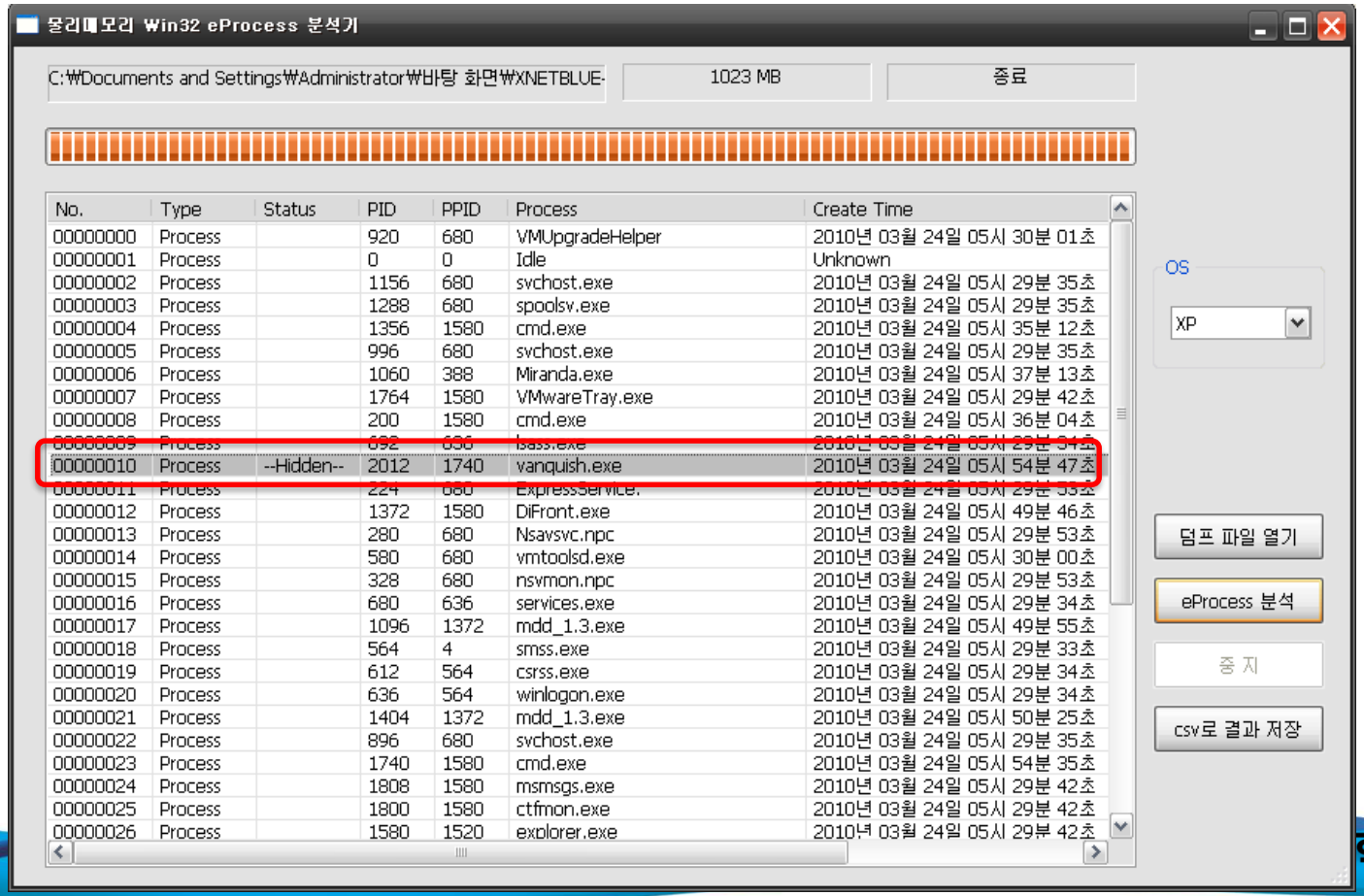
- **목적**: 실행 중인 프로세스 정보 수집 및 분석을 통한 비정상 프로세스 판별 악성코드 탐지
- **방법**:
  - 기존 프로세스 정보들과 비교 분석하여 참조 DLL
  - 실행 경로 등이 이상이 없는지 점검, 또한 각 프로세스의 자식 프로세스의 활동 정보 비교/분석
  - 물리 메모리에서 프로세스 구조체를 추출하여 숨겨진 프로세스 탐지 등
- **도구**: 윈도우 작업 관리자(Windows), Process Explorer(Windows), top(Unix)





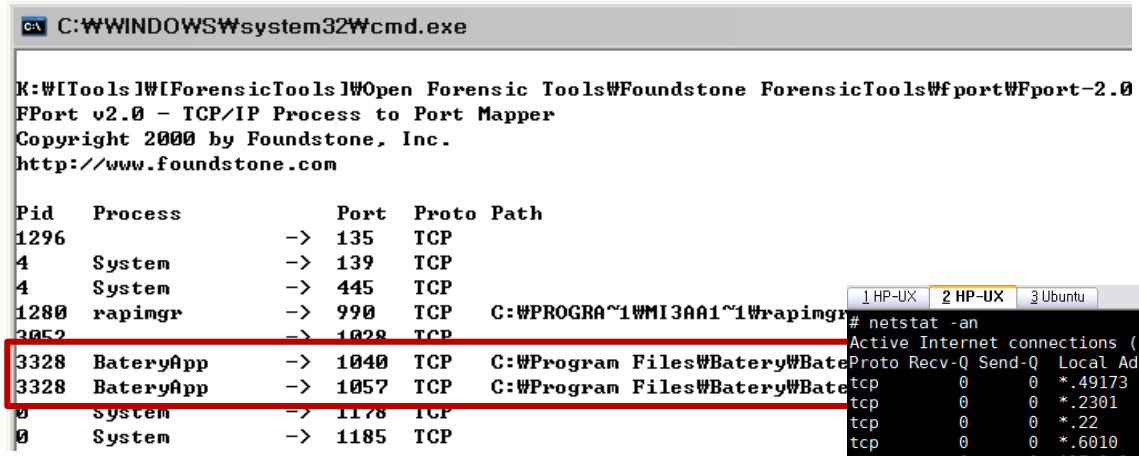
물리메모리에서 프로세스 정보 추출

- 덤프한 물리메모리에서 프로세스 구조체 **(eProcess)** 추출
  - 프로세스 구조체는 포렌식 관점에서 유용한 정보가 다수 존재
- 은닉 프로세스 탐지 및 이전에 실행한 프로세스 목록 추출 가능
  - 전원을 차단하지 않을 경우 메모리에 프로세스 목록이 일주일 이상 잔류



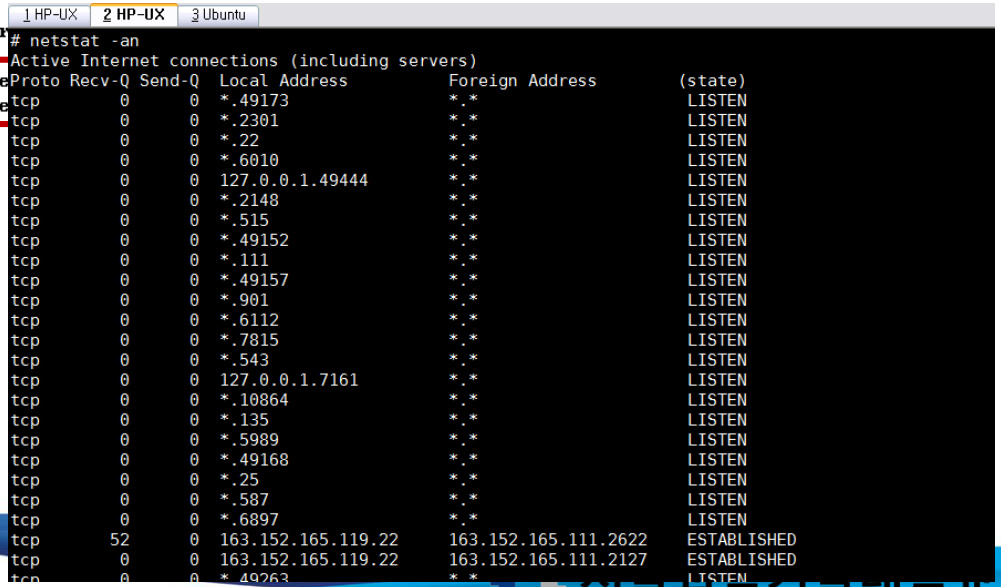
휘발성 데이터 수집 및 분석 -네트워크 정보

- **목적**: 현재 네트워크 연결 및 사용정보를 바탕으로 비인가 접속 판별
- **방법**: 현재 연결된 IP 주소와 포트 등에 대한 점검, 연결 상태와 프로세스 별 연결 포트를 확인하여 허가되지 않은 사용 정보를 판별
- **도구**: netstat, netbios (Windows), Fport (Windows), netstat -an



Fport (Windows)

Netstat (Unix)



네트워크 정보 - 현재 네트워크 연결 정보

현재 네트워크 연결 정보

- 호스트로 들어온 연결
- 호스트에서 나가는 연결
- 침입 흔적 정보 획득

수집방법 (netstat)

- netstat –ano
  - a:모든 연결, n:IP로 표시
    o:Process ID
  - 모든 네트워크 연결 정보 획득

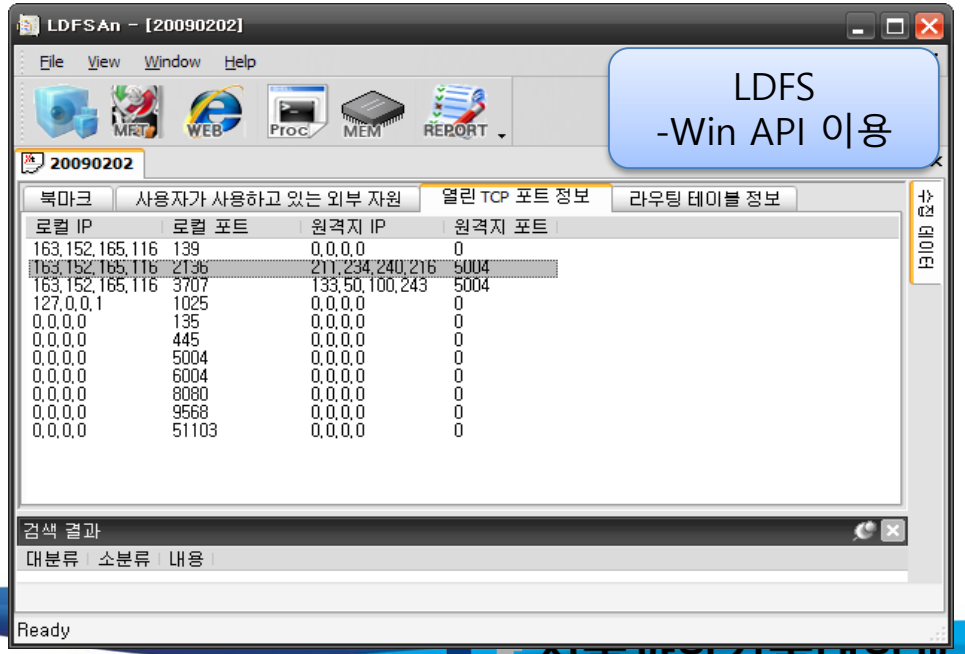
netstat –ano

```
D:₩Tools₩code>netstat -ano

Active Connections

  Proto  Local Address          Foreign Address        State           PID
  TCP    0.0.0.0:135            0.0.0.0:0              LISTENING       1568
  TCP    0.0.0.0:445            0.0.0.0:0              LISTENING       4
  TCP    0.0.0.0:5004           0.0.0.0:0              LISTENING       508
  TCP    0.0.0.0:6004           0.0.0.0:0              LISTENING       508
  TCP    127.0.0.1:1035         0.0.0.0:0              LISTENING       2764
  TCP    163.152.146.233:139    0.0.0.0:0              LISTENING       4
  TCP    163.152.146.233:1041   207.46.111.65:1863     ESTABLISHED     1080
  TCP    163.152.146.233:1050   211.234.239.138:5004   ESTABLISHED     508
  TCP    163.152.146.233:1842   220.69.247.1:80        CLOSE_WAIT      2772
  UDP    0.0.0.0:445            *:*                                    4
  UDP    0.0.0.0:500            *:*                                    1284
  UDP    0.0.0.0:1025           *:*                                    1744
  UDP    0.0.0.0:1031           *:*                                    472
  UDP    0.0.0.0:1063           *:*                                    1744
  UDP    0.0.0.0:4500           *:*                                    1284
```

LDFS
-Win API 이용



네트워크 정보 -열린 네트워크 포트와 프로세스 정보

열린 네트워크 포트와 연결된 프로세스 정보

- 특정 활성 포트와 매핑된 프로세스 정보
- 의심 가는 프로세스가 포트를 열고 데이터를 주고 받는다면 악성 코드 의심

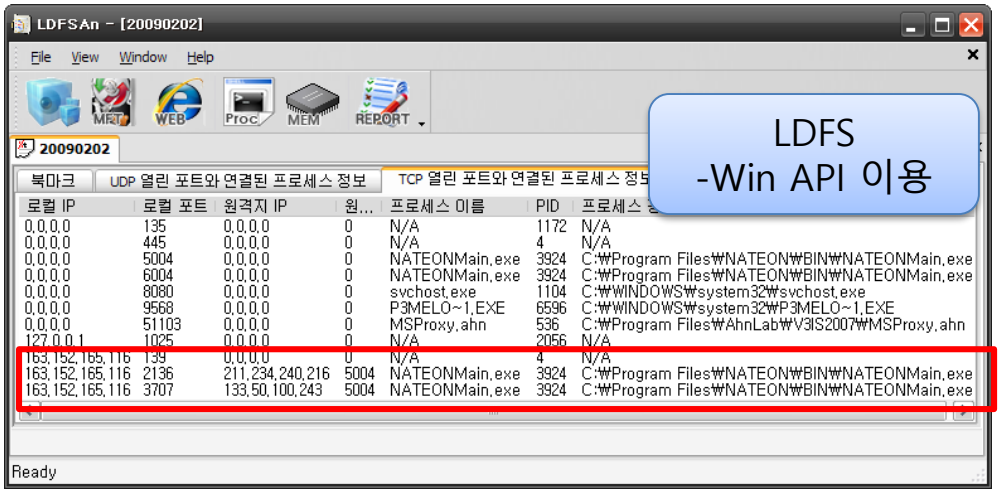
```
C:₩Documents and Settings₩Luke369x2>"H:₩[Tools]₩[ForensicTools]₩Live Forensics₩F
port-2.0₩Fport.exe"
FPort v2.0 - TCP/IP Process to Port Mapper
Copyright 2000 by Foundstone, Inc.
http://www.foundstone.com

Pid   Process            Port  Proto Path
1172                 ->  135   TCP
4     System         ->  139   TCP
4     System         ->  445   TCP
2056                 ->  1025  TCP
3924  NATEONMain     ->  2136  TCP   C:₩Program Files₩NATEON₩BIN₩NATEONMain.exe
5232  Melon          ->  3330  TCP   C:₩Program Files₩MelOn Player₩Melon.exe
3924  NATEONMain     ->  3707  TCP   C:₩Program Files₩NATEON₩BIN₩NATEONMain.exe
3924  NATEONMain     ->  5004  TCP   C:₩Program Files₩NATEON₩BIN₩NATEONMain.exe
3924  NATEONMain     ->  6004  TCP   C:₩Program Files₩NATEON₩BIN₩NATEONMain.exe
1104  svchost        ->  8080  TCP   C:₩WINDOWS₩system32₩svchost.exe
6596  P3MELO~1       ->  9568  TCP   C:₩WINDOWS₩system32₩P3MELO~1.EXE
536   MSProxy        ->  51103 TCP   C:₩Program Files₩AhnLab₩V3IS2007₩MSProxy.ah
```

fport

수집방법

- Fport
  - Foundstone 사에서 제공하는 공개 버전의 콘솔 명령어



LDFS
-Win API 이용

네트워크 정보 - 공유 파일

Open Files

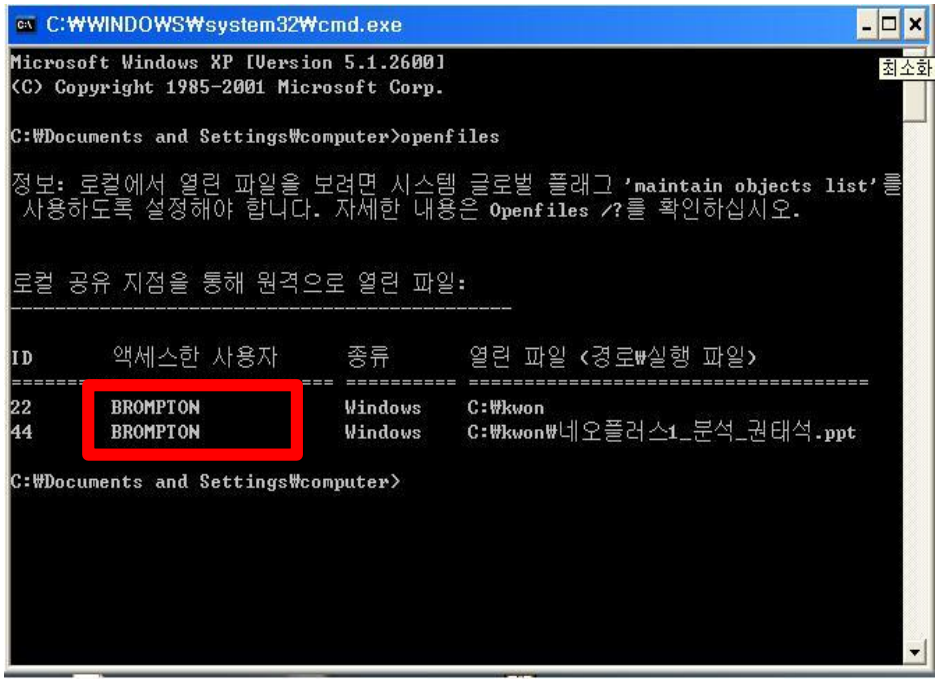
- 외부에서 원격으로 열려있는 파일 획득
- 허가 받지 않은 사용자가 접근해서 자원을 사용하는지 확인 필요

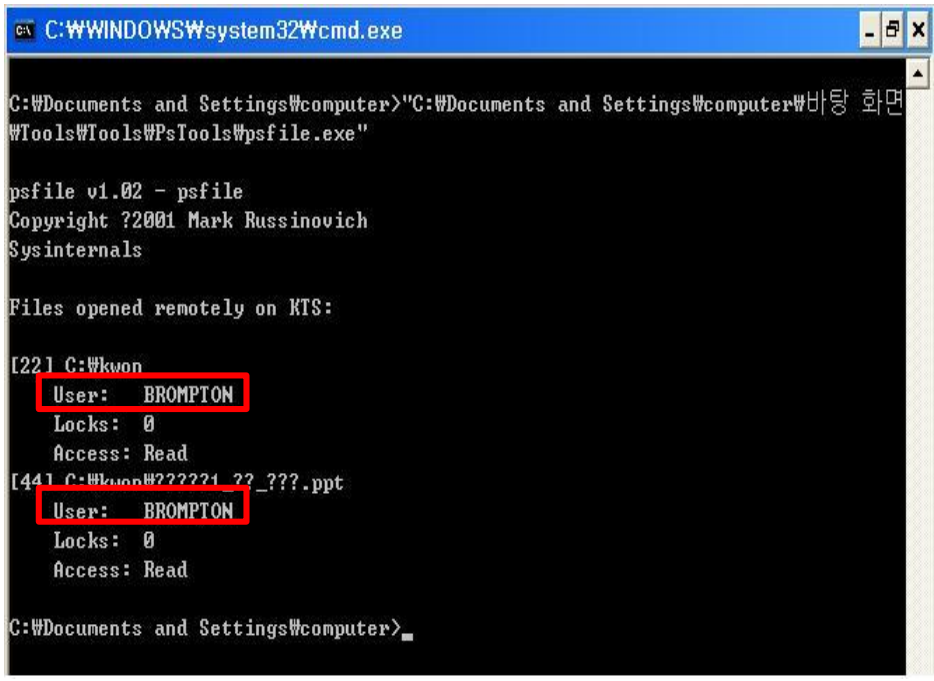
수집방법

- net file : 공유 파일 정보 출력 명령어
- psfile : sysinternal 사에서 제공하는 콘솔 명령어
- openfiles : 콘솔 명령어

네트워크 정보 - 공유 파일

- **openfiles**
  - 열린 파일 조사
- **psfile**
  - 원격에서 접근한 파일 정보 획득





네트워크 정보 - 외부 시스템 연결 정보

외부 시스템 연결 정보

- 사용자가 원격으로 사용하고 있는 공유 자원 정보 획득
- 허가 받지 않은 외부 자원 ➔ 정보 유출

수집방법

- **nbtstat –c (c : Cached NetBIOS Name Table)**
  - **IP, 네트워크에 연결된 후 현재까지의 시간 등의 정보 조사**

```
C:₩Documents and Settings₩computer>nbtstat -c

로컬 영역 연결:
Node IpAddress: [163.152.146.203] Scope Id: []

                NetBIOS Remote Cache Name Table

       Name               Type         Host Address    Life [sec]

   163.152.146.233<20>  UNIQUE          163.152.146.233     295
   DONCOM         <20>  UNIQUE          163.152.146.194     422

C:₩Documents and Settings₩computer>_
```

비휘발성 데이터

- 전원을 차단해도 사라지지 않는 데이터
- 쓰기방지 장치를 부착하여 수집하는 것이 원칙
- 시스템을 종료시키지 못하는 환경에서는 활성 시스템에서 직접 데이터 수집 및 분석이 필요
- 데이터 선별을 통해 사건과 관련된 비휘발성 데이터만을 수집하는 경우도 있음

활성 시스템 데이터 수집

- 운영체제 종류별로 수집에 사용할 수 있는 도구 및 시스템 **API**가 다름

  → 활성 시스템 데이터 수집도구는 이러한 특성을 고려하여 개발

- 시스템 상태의 변경을 최소화 해야 하므로 **CLI(Command Line Interface)**의 사용을 권장

- 시스템 명령어, 라이브러리를 정적으로 포함하고 있는 도구를 사용, 읽기만 가능한 매체에서 실행할 것을 권장

3. 디스크 이미징

- 디스크 이미징 장비 **(Disk Imaging Hardware)**
  - 단독으로 복제 디스크를 생성할 수 있는 포렌식 장비
  - 디스크를 컴퓨터에 연결하지 않고 다른 하드디스크에 사본을 생성
  - LogiCube Talon, Dossier
  - ICS ImageMasster Solo 3 & 4 등

Logicube Talon

Logicube Dossier

ICS Image Masster Solo3

ICS Image Masster Solo4

디스크 이미징 장비를 이용한 사본 생성 절차

[조사 대상 시스템]

[조사 대상 시스템]
하드디스크 분해

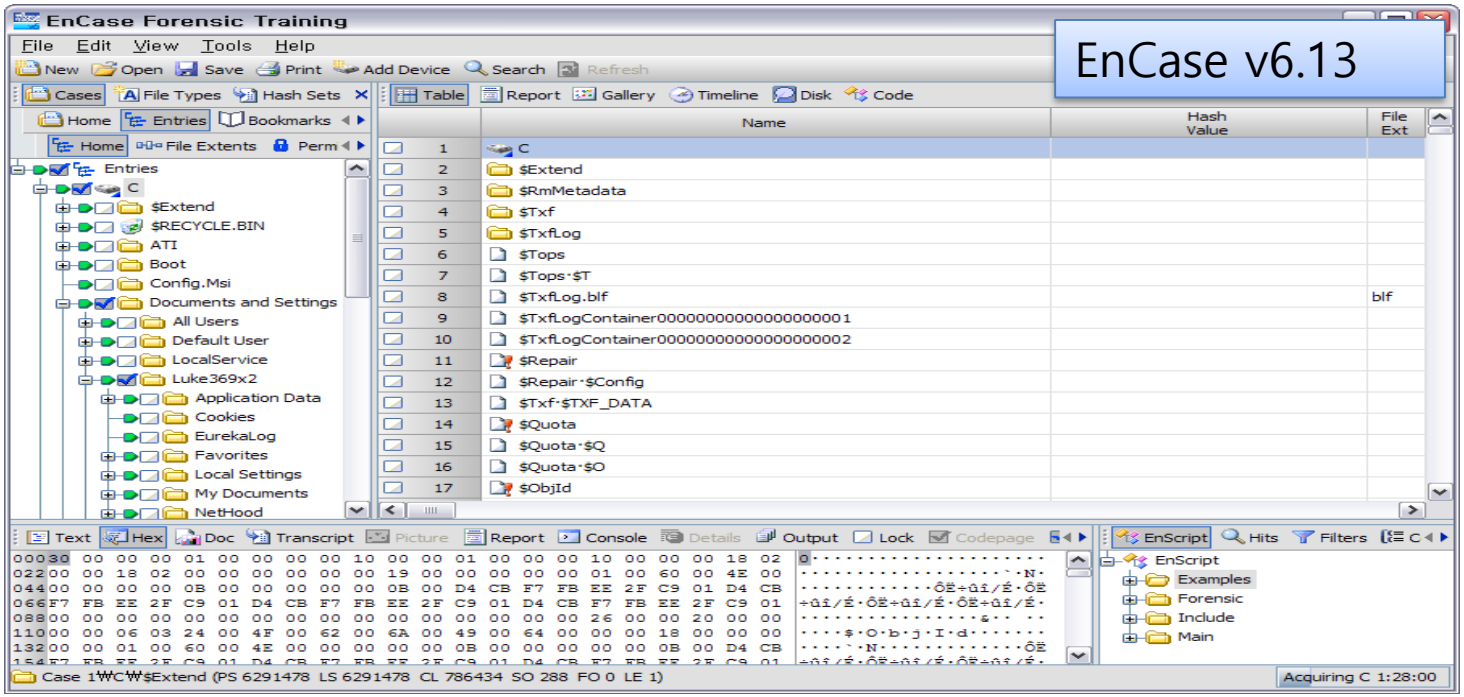
[수집 대상 디스크]

원본 디스크에 대하여
사본 디스크를 복제
(사본 2개까지 생성)

[이미징 장치]

디스크 이미징 및 분석

- **EnCase**
  - Guidance Software 에서 개발한 세계에서 가장 널리 쓰이는 포렌식 S/W
  - 그래픽 인터페이스 바탕으로 디스크 이미징, 디스크 브라우징 기능 제공
  - 증거 미리보기 및 데이터 검색/분석 기능 제공
  - 윈도우, Palm OS등의 플랫폼과 RAID 방식 지원



서울과학기술대학교 SEOULTECH

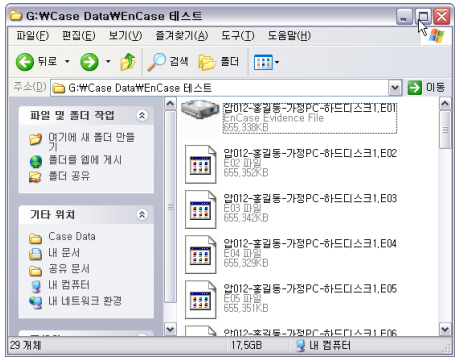
디스크 이미징 S/W를 이용한 사본 생성 절차

[조사 대상 시스템]

분해

[수집 대상 디스크]
쓰기 방지 장치에 연결

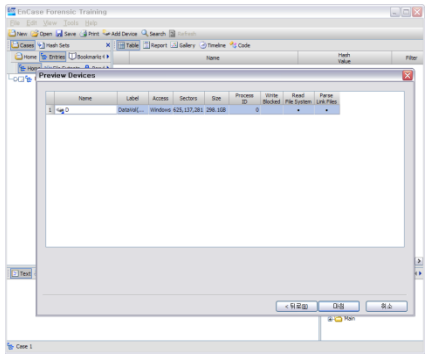
[이미지 파일]



획득

[쓰기방지 장치]
분석 시스템과 쓰기 방지 장치 연결

[디스크 이미징 도구]



수집

[분석 시스템]

4. 임베디드 시스템의 증거 확보

스마트폰 증거자료 추출 방법

- 스마트 폰 주요 데이터
  - 휴대폰 데이터 + **이메일, 아웃룩(MS), 인터넷 사용 기록, GPS 정보** 등

- 데이터 획득 방법

논리적 수집 방법

MS ActiveSync

or

Remote API

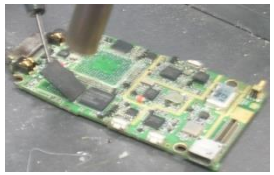
Request Files

Sending Files

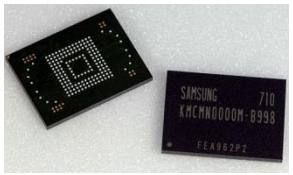
물리적 수집 방법

using JTAG Interface

or



using Flash Memory Reader

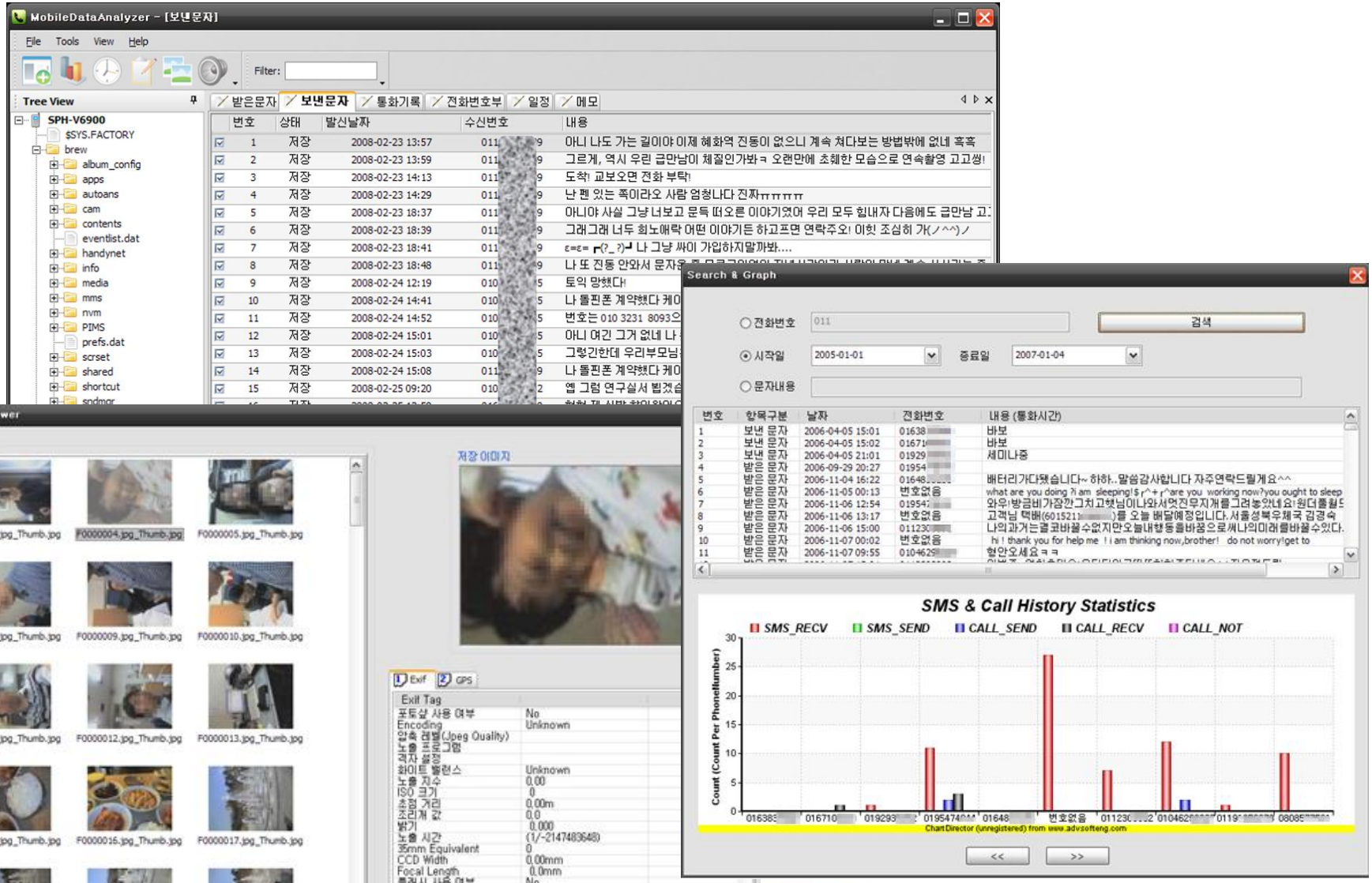


모바일 데이터의 분석

- 모바일 포렌식
  - 휴대폰 데이터 분석 항목

데이터	항목
수신 문자메시지(SMS)	수신 일시, 송신자 전화번호, 문자내용
발신 문자메시지(SMS)	발신 일시, 수신자 전화번호, 문자내용
임시저장 문자메시지(SMS)	저장 일시, 수신자 전화번호, 문자내용
최근 통화 기록(Call History)	통화종류, 송수신자 전화번호, 통화시간
전화번호부(PhoneBook)	저장된 이름, 전화번호, 단축번호, 그룹
일정(Schedule)	일정 일시, 일정내용
메모(Memo)	메모 일시, 메모내용
사진(Photo)	사진 콘텐츠, 사진 촬영 정보
멀티미디어(Multimedia)	동영상, 음성 메모, 음악 등
휴대폰 전자일련번호(ESN)	제조사 식별 코드, 기기 일련 번호
휴대폰 비밀번호	※ 휴대폰 잠금 해제

모바일 데이터 분석 도구 실행화면



디스크 브라우징 기술

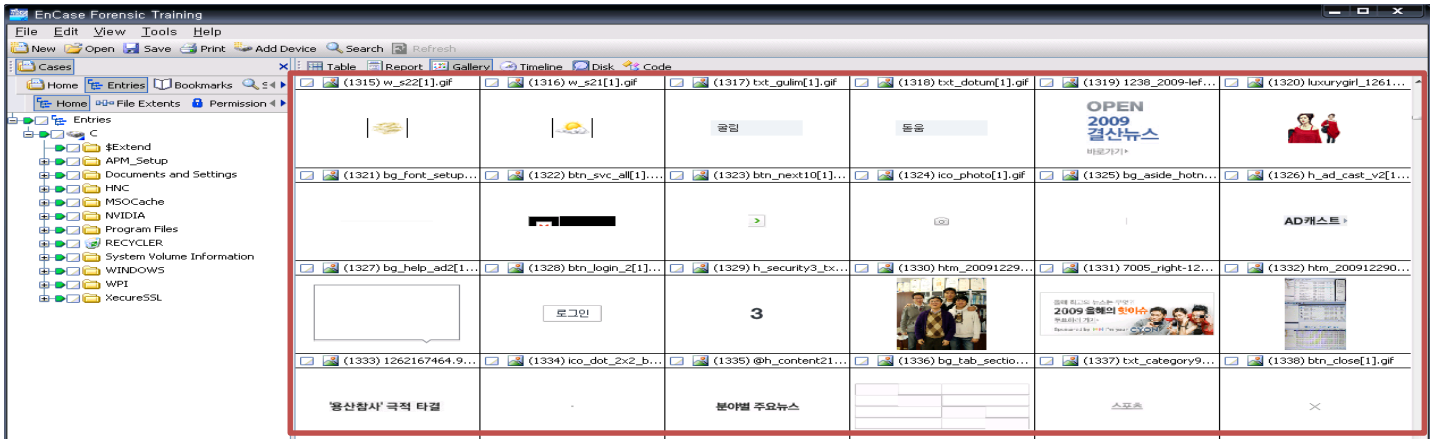
- 기본적인 증거 분석 대상은 확보한 저장 매체
  - 복제한 디스크 사본, 디스크 이미지 파일
  - USB 드라이브, CD 등의 저장 매체에서 생성한 이미지 파일

- 디스크 브라우징(Disk Browsing)
  - 저장매체 또는 하드디스크 이미지의 내부 구조와 파일 시스템을 확인하고, 파일시스템 내부에 존재하는 파일에 대응되는 응용 프로그램의 구동 없이 쉽고 빠르게 분석할 수 있도록 하는 기법
  - 복제한 이미지를 사용자가 수동으로 마운팅해서 열람할 필요가 없어 분석 시간을 줄일 수 있음
  - 디스크 브라우징 도구
    - EnCase, FTK, Final Forensic 등

EnCase의 디스크 브라우징

- 기본적인 디스크 브라우징 기능
  - 파일 시스템의 구조를 확인하고 메타데이터를 출력
  - 각 파일과 관련된 정보들 (생성 • 수정 • 접근 시간, 해쉬값, 시그니처, 저장 위치 등)을 파악
  - 검색, 타임라인 분석, 미리보기 기능 등

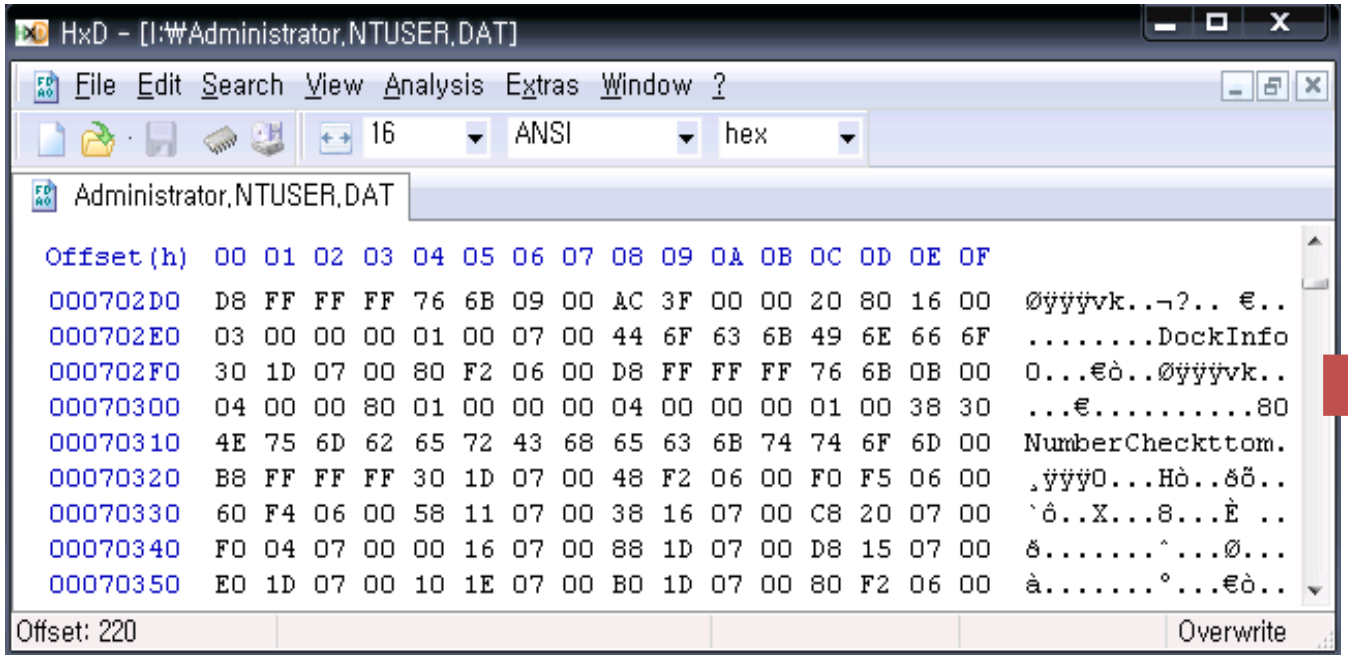
Name	Description	Is Deleted	Last Accessed	File Created	Last Written	Entry Modified
$Extend	Folder, Internal, Hidden, System		12/21/09 11:31:22오후	12/21/09 11:31:22오후	12/21/09 11:31:22오후	12/21/09 11:31:22오후
APM_Setup	Folder		12/30/09 07:51:40오후	12/24/09 04:42:11오후	12/24/09 04:43:02오후	12/24/09 04:43:02오후
Documents and Settings	Folder		12/30/09 08:00:20오후	12/21/09 02:33:45오후	12/21/09 02:58:49오후	12/21/09 02:58:49오후



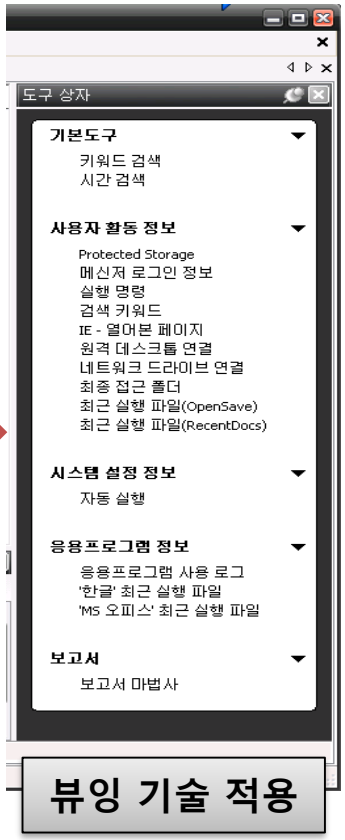
  - 파일 확장자 변경 여부, 암호 파일 등을 확인할 수 있으며, 복구 가능한 삭제 파일과 비할당 영역에 있는 파일 파편들을 검토

데이터 뷰잉 기술

- 파일 포맷이 있는 데이터를 가시적으로 확인할 수 있도록, 디지털 데이터의 구조를 파악해서 시각적으로 정보를 출력하는 기술
- 헥스 에디터를 사용하기도 하지만, 좀더 효과적인 분석을 위해서는 개선된 데이터 뷰잉 기술이 필요



헥스 에디터(hex-editor)로 확인한 윈도우 레지스트리 파일



뷰잉 기술 적용

검색 기술

- 검색 기술의 필요성
  - 저장매체가 대용량화 됨에 따라 수집되는 디지털 데이터의 양도 매우 많아지고 있어, 사건의 단서나 증거를 찾는 것은 점차 어려워짐
  - 사건과 관련된 자료들을 선별하기 위한 검색 기술 개발이 필요함

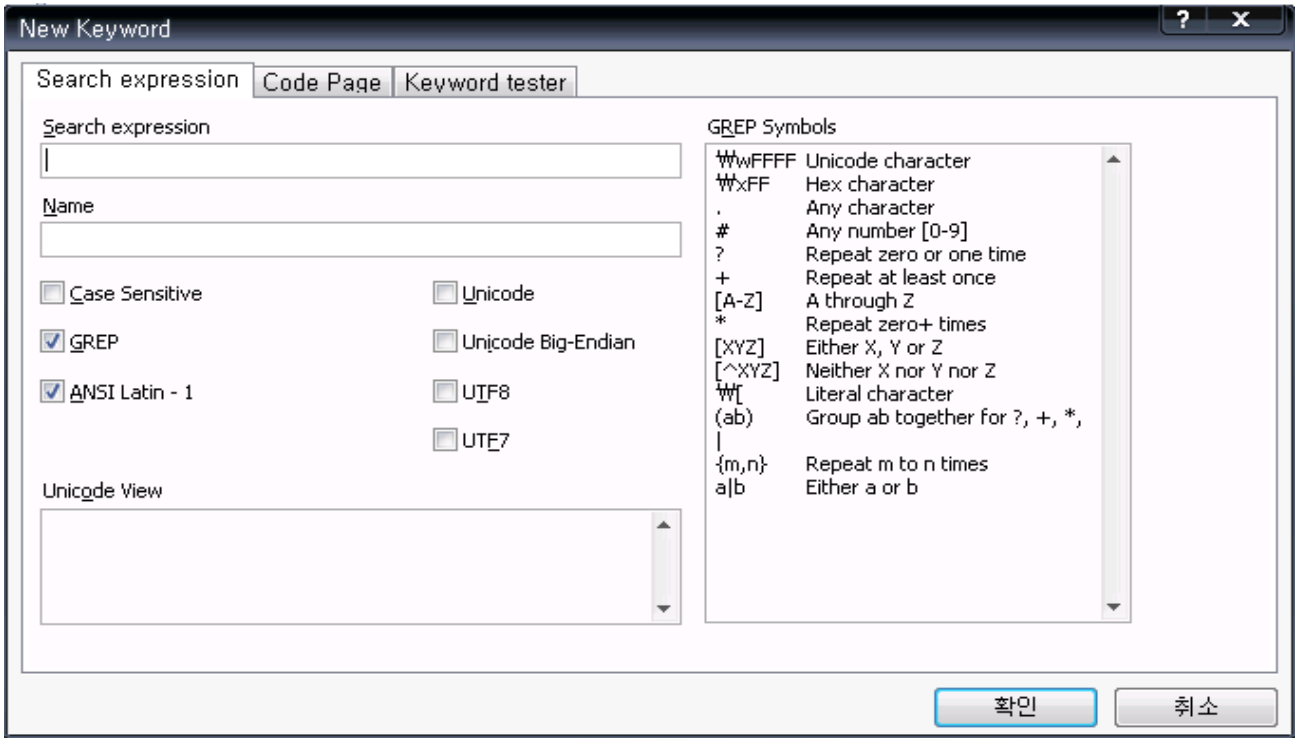
- 포렌식 조사/분석은 연속되는 검색의 반복
  - 모든 파일들의 키워드, Signature에 대해 검색을 반복해야 함
  - 잘 알려진 파일은 검색 대상에서 제외하고, 주목해서 검색할 대상을 선정하여, 검색 범위를 축소하는 것이 중요함.
  - 발전된 형태의 검색기술을 사용하여 조사/분석 단계에 투입되는 시간 비용을 줄일 수 있어야 함

검색 기술의 종류

- **일반 검색 (키워드 검색)**
  - 파일 또는 저장매체 전체를 대상으로 특정 키워드를 입력하여 검색
  - 키워드 검색을 통해 필요한 증거를 찾기 위해서는 텍스트 인코딩, 대·소문자 등의 사항을 고려해야 함
  - 같은 키워드라도 인코딩 방식에 따라 전혀 다른 값이 되기 때문에 찾고자 하는 키워드의 형태를 결정해서 검색을 수행
  - 파일이름, 속성, 내부 문자열/코드값, 시그니쳐 등을 선정하여 목적 파일을 쉽고 빠르게 찾는 기술 (String, Index Search)
- **해쉬 검색**
  - 기존에 구축된 알려진 파일의 해쉬 셋(Reference Data Set)을 사용하여, 조사 분석 대상을 식별하고 검색 수준을 선정할 수 있는 기술
- **슬랙 검색**
  - 파일 시스템의 잉여 공간에 남아있는 기존 파일들의 조각 정보를 찾아내는 기술

검색 기술 –키워드 검색

- **GREP(Globally Find Regular-Expression and Print)**
  - Unix 계열의 운영체제에서 검색을 위해 사용자가 정의할 수 있는 표현 명령어
  - 정규 표현식을 사용해서 다양한 형태의 키워드를 하나의 식으로 설정 가능



검색 기술 –키워드 검색

- **GREP(Globally Find Regular-Expression and Print)**

정규표현식	의 미	사용 예제
^	문자열의 처음	^aaa : 문자열의 처음에 aaa를 포함하면 매칭
$	문자열의 끝	aaa$: 문자열의 끝에 aaa를 포함하면 매칭
.	하나의 문자 매칭	.lue : alue, blue, clue ...
?	바로 이전 문자의 빈도가 0 또는 1인 경우 매칭	forensics? : forensic 또는 forensics
*	바로 이전 문자의 빈도가 0 이상인 경우 매칭	digital_*for : digitalfor, digital_for, digital__for ...
+	바로 이전 문자의 빈도가 1이상인 경우 매칭	digital_ + for : digital_for, digital__for ...
[ABC]	A, B, C 중의 하나의 문자 매칭	dig[ei]tal : digetal, digetal
[^ABC]	A, B, C 이외의 문자 매칭	dig[^i]tal : digetal, digftal ... (digital제외) [^a-z] : 소문자 이외의 문자 매칭
[A-C]	A부터 C중에 하나의 문자 매칭	[a-d] : a, b, c, d [0-2] : 0, 1, 2
₩	특수문자를 일반문자로 사용하는 경우	[?[₩]] : ?, [,]
X{M}	문자 X를 M번 반복	h{3} : h가 3번 이상 반복 (hhh, hhhh ...)
X{M,N}	문자 X를 M회 이상 N회 이하 반복	h{3,5} :hhh, hhhh, hhhhh
a|b	a, b 둘 중에 하나인 경우 매칭	web₩.(com)|(net) : web.com, web.net

검색 기술 –키워드 검색

- **GREP(Globally Find Regular-Expression and Print)**

Name	Kim	Number	870101-1234567
Name	Lee	Number	802110-2894561
Name	Park	Number	841218-1236874
Name	Han	Number	835577-1654789
Name	Hong	Number	880202-2345678

- 유닉스에서 **GREP** 명령어 사용하여 검색

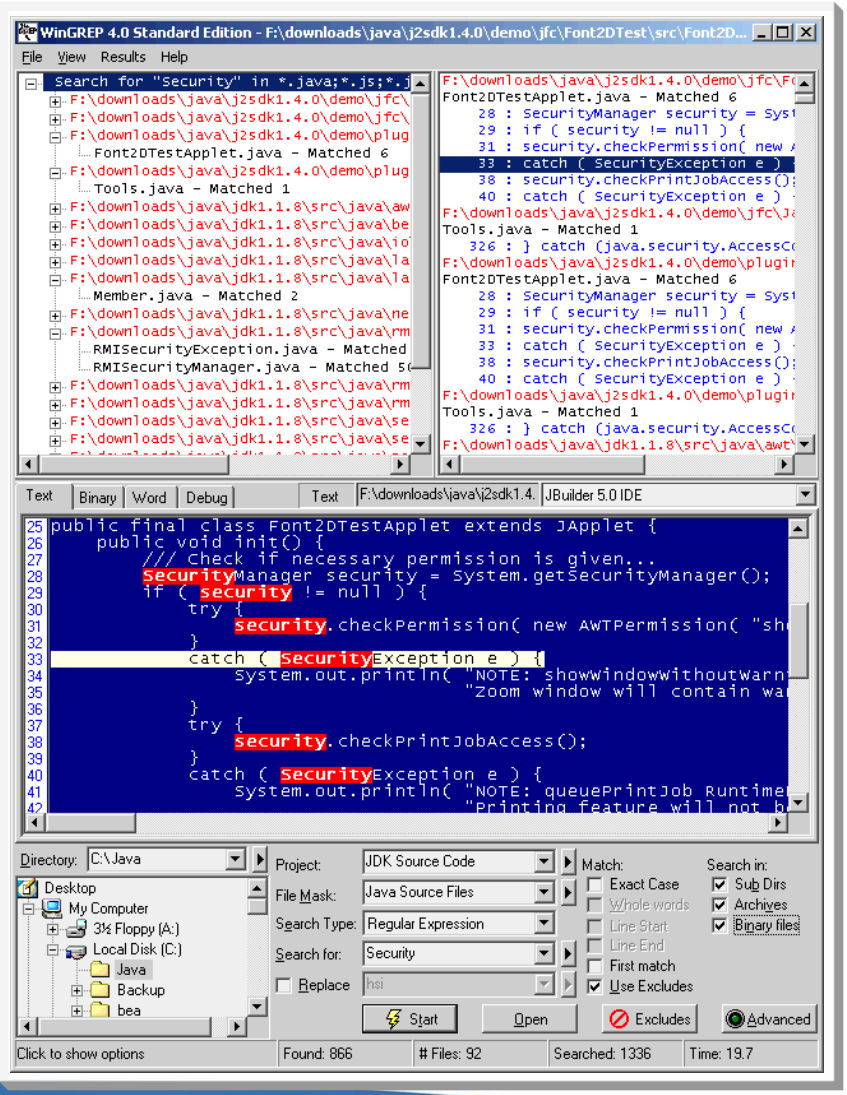
 # grep '.*[0-9]{2}[01]{1}[0-9]{1}[0-3]{1}[0-9]{1}-[0-9]{7}' DataFile

명령어 출력 결과

Name	Kim	Number	870101-1234567
Name	Park	Number	841218-1236874
Name	Hong	Number	880202-2345678

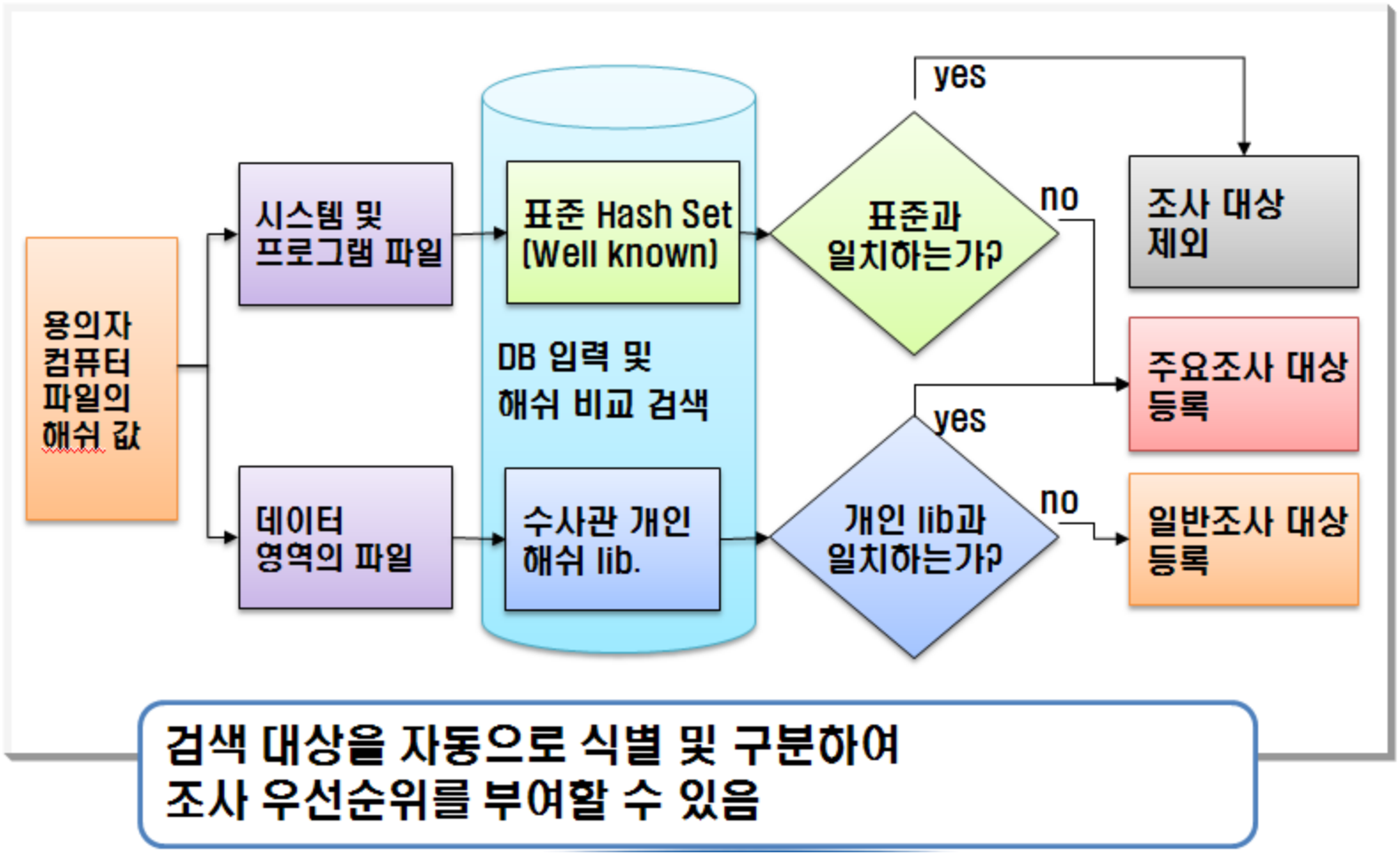
검색 기술-문자열 검색

- 파일 및 문자열 검색
  - 검색 작업은 지속적으로 이루어지므로 포렌식 전문 검색 도구를 쓰는 것이 유용함
  - WinGREP Search Tool (Hurricane Software)
    - 다중 디렉토리를 포함, 제외 검색
    - 키입력을 최소화한 사용자 인터페이스
    - 미리보기 기능
    - 압축/바이너리 파일에 대한 검색 기능
    - PDF/MS-Office 파일 검색 기능



검색 기술-파일 검색

- **Hashed Search 원리**



검색 기술-파일 검색

National Software Reference Library (NSRL)

- 美 NIST 산하 CFTT에서 제공하는 국가 표준 참조 데이터
- Justice's National Institute of Justice (NIJ)의 지원
- NSRL의 목적
  - 범죄에 사용되는 컴퓨터 파일의 식별 자동화
  - 증거에 포함된 파일 조사를 효율적으로 지원
- NSRL의 세부 내용
  - 다년간 각종 S/W 및 알려진 파일을 수집, 이에 대한 정보와 hash 값을 DB 목록화 (TOTAL 50,121,818 files, 15,722,076 unique hash values)
  - 전세계 8천여 개 S/W, 35개국 언어 OS의 참조 데이터 셋(RDS:Reference Data Set) 구축

```
"SHA-1",                                    "MD5",                                  "CRC32",   "FileName",       "FileSize","ProductCode"
"00000F6ED90D946C057B55545597C3125lDC24E4","F4129AC77F806601BDD44620C17675E7","38CC50B7","004i200r.gif"   ,  1551,228,    "WIN"
"00000FF9D0ED9A6B53BC6A9364C07074DE1565F3","A5D49D6DA9D78FD1E7C32D58BC7A46FB","2D729A1E","cmnres.pdb.dll",76800,2471,    "WIN"
"00000FF9D0ED9A6B53BC6A9364C07074DE1565F3","A5D49D6DA9D78FD1E7C32D58BC7A46FB","2D729A1E","cmnres.pdb.dll",76800,2704,    "WIN"
"00000FF9D0ED9A6B53BC6A9364C07074DE1565F3","A5D49D6DA9D78FD1E7C32D58BC7A46FB","2D729A1E","cmnres.pdb.dll",76800,2741,    "WIN"
"00000FF9D0ED9A6B53BC6A9364C07074DE1565F3","A5D49D6DA9D78FD1E7C32D58BC7A46FB","2D729A1E","cmnres.pdb.dll",76800,2797,    "WIN"
"00000FF9D0ED9A6B53BC6A9364C07074DE1565F3","A5D49D6DA9D78FD1E7C32D58BC7A46FB","2D729A1E","cmnres.pdb.dll",76800,2912,    "WIN"
```

검색 기술-파일 검색

NSRL 활용방안

- 용의자 컴퓨터내의 파일내용과 NSRL 목록을 비교 분석, 알려진 파일을 쉽게 식별하여 조사 범위 집중 가능
- 수사관은 평소 표준 참조 데이터를 입수하거나 제작하여 분석·조사 과정을 효율적으로 체계화하여야 함

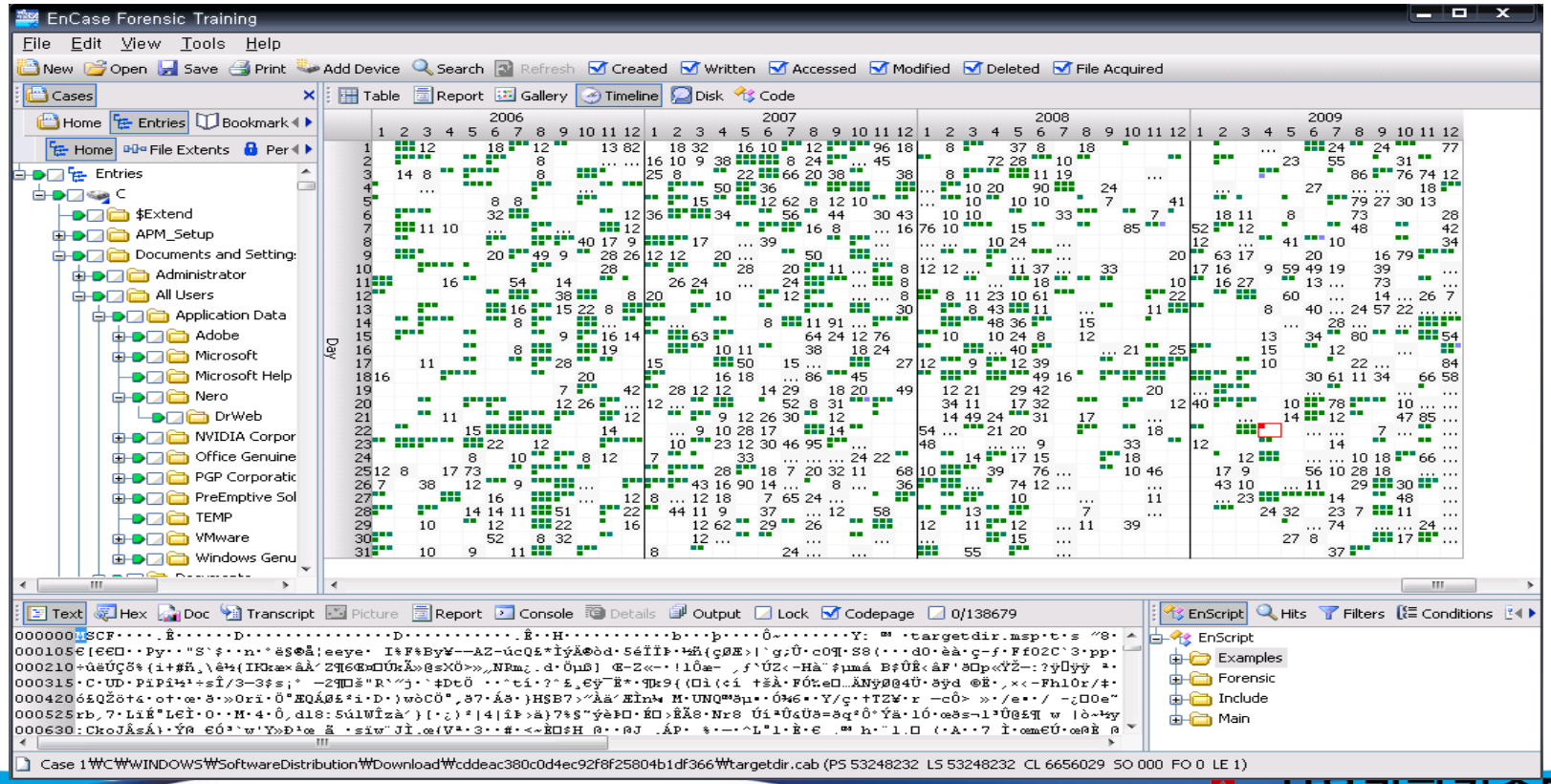
인덱스 기반 검색 기술

- 새로운 검색 기술의 필요성
  - 데이터의 대용량화
    - 하드디스크 기술 발달로 S-ATA 1TB까지 시판
    - 문서, 그림, 동영상의 다양한 데이터를 인터넷으로 공유함으로써 조사에 필요한 데이터 량이 크게 증가
    - OS와 응용 프로그램의 크기 증가 등

  - 소요 시간의 증가
    - 대용량 데이터를 수집하고 분석하는데 많은 시간이 소요됨
    - 컴퓨터 처리 속도 증가에 비해 처리할 데이터는 엄청나게 증가함
- 해결 방안
  - 디스크 이미지에 대한 인덱스를 생성하여 데이터를 빠르게 검색

타임라인 분석

- 타임라인 분석
  - 디지털 데이터의 시간 정보는 범죄 사실을 규명하기 위해 매우 중요한 정보
  - 파일 시스템 상에 저장되는 파일의 시간 정보, 파일 내부의 메타데이터에 저장되는 시간 정보 등 다양한 곳에 저장되어 있는 시간 정보를 이용,타임라인(Timeline)을 구성함으로써 시스템 사용자의 행위를 추적할 수 있음



로그 분석

- 로그(LOG)란?
  - 시스템에 접속한 사용자의 행위 및 시스템의 상태를 주기적으로 저장해 놓은 기록
  - 로그를 이용하여 외부 침입의 흔적과 사용자가 어떠한 명령어를 사용했는지, 그리고 시스템이 처리한 업무와 에러 등의 정보 등을 파악
  - 서버 시스템의 침해사고조사와 같은 경우 가장 기본적으로 행해지는 분석 중의 하나

- 로그의 종류
  - Unix 시스템 계열 로그, Windows 계열 로그, 웹(Web) 로그 등
  - 시스템의 종류에 따라 특별한 설정 없이 기본적으로 생성되는 로그가 있는 반면, 사용자의 설정이 있어야만 생성되는 로그도 존재
  - 조사자는 각 시스템의 기본 로그와 그렇지 않은 로그의 분석 방법을 숙지해야 함

로그분석 - Unix 시스템 로그 분석

- 시스템 별 로그 디렉터리

Unix 시스템	디렉터리
HP-UX	/usr/adm
Solaris, AIX	/var/adm
Linux, BSD	/var/log

- 로그 파일의 종류 및 기본적인 기능

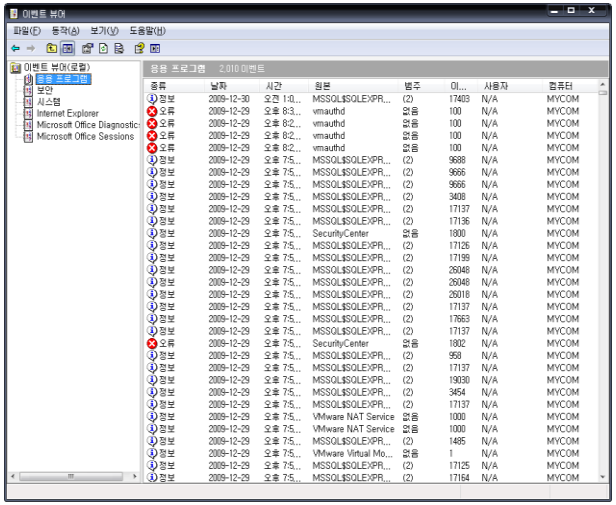
파 일 명	기 능
acct 또는 pacct	사용자별로 실행되는 모든 명령어를 기록
aculog	다이얼-아웃 모뎀 관련 기록(자동 호출 장치)
lastlog	각 사용자의 가장 최근 로그인 시간을 기록
loginlog	실패한 로그인 시도를 기록
messages	부트 메시지 등 시스템의 콘솔에서 출력된 결과를 기록하고 syslog에 의하여 생성된 메시지도 기록
sulog	su 명령 사용 내역 기록
utmp	현재 로그인한 각 사용자의 기록
utmpx	utmp 기능을 확장(extended utmp), 원격 호스트 관련 정보 등 자료 구조 확장
wtmp	사용자의 로그인, 로그아웃 시간과 시스템의 종료 시간, 시스템 시작 시간 등을 기록
wtmpx	wtmp 기능 확장 (extended wtmp)
vold.log	플로피 디스크나 CD-ROM과 같은 외부 매체의 사용에서 발생하는 에러를 기록
xferlog	FTP 접근 기록

로그분석 - Windows 시스템 로그 분석

- 이벤트 로그
  - Windows는 기본적으로 이벤트(event) 로그를 시스템 운영 전반에 걸쳐서 저장
  - 조사자는 이벤트 로그의 분석을 통해 해당 시스템의 전반적인 동작을 알 수 있으며, 증거 자료를 획득할 수도 있음

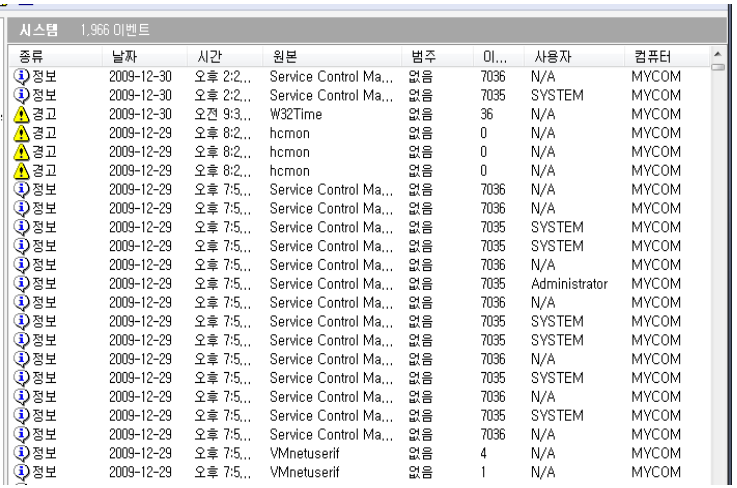
- 이벤트 로그의 종류
  - 응용프로그램 로그
    - 응용프로그램이나 기타 프로그램의 동작에 대한 이벤트가 저장되며, 기록되는 이벤트는 소프트웨어 개발자에 의해 결정
  - 보안 로그
    - 유효하거나 유효하지 않은 로그온 시도 및 파일 생성, 열람, 삭제 등에 관련된 이벤트를 기록
  - 시스템 로그
    - Windows 시스템 구성요소가 기록하는 이벤트로 시스템 부팅 시 드라이버가 로드되지 않는 경우와 같이 구성요소의 오류를 기록

로그분석 - Windows 시스템 로그 분석



이벤트 헤더 정보

정 보	의 미
날 짜	이벤트가 발생한 날짜
시 간	이벤트가 발생한 시간
사 용 자	이벤트를 발생시킨 사용자의 이름
컴 퓨 터	이벤트가 발생한 컴퓨터의 이름
원 본	이벤트를 기록한 소프트웨어 (이벤트가 일어난 프로세스)
이벤트ID	해당 원본의 특정 이벤트 유형을 식별하는 번호
범 주	이벤트의 원본에 의한 이벤트 분류로 주로 보안 로그에서 사용됨
종 류	이벤트 심각도의 분류로 오류, 정보, 경고, 성공, 감사, 실패 감사로 분류



이벤트 정보의 종류

이벤트 유형	설 명
오 류	데이터 손실이나 기능 상실 같은 중대한 문제로 시스템을 시작하는 동안 서비스가 로드되지 못했을 경우와 같은 이벤트 기록
경 고	시스템에 문제가 발생할 수 있는 문제를 미리 알려 주는 이벤트로 디스크 공간이 부족할 때와 같은 이벤트 기록
정 보	응용 프로그램, 드라이버 또는 서비스가 성공적으로 수행되었음을 설명하는 이벤트
성공감사	사용자가 시스템에 성공적으로 로그온 했을 경우와 같이 보안 이벤트가 성공했음을 나타냄
실패감사	사용자가 시스템에 로그온 실패했을 경우와 같이 보안 이벤트가 실패했음을 나타냄

로그분석 - Web 로그 분석

- 웹 서비스를 제공하는 웹 서버의 종류에 따라 다른 형태로 저장

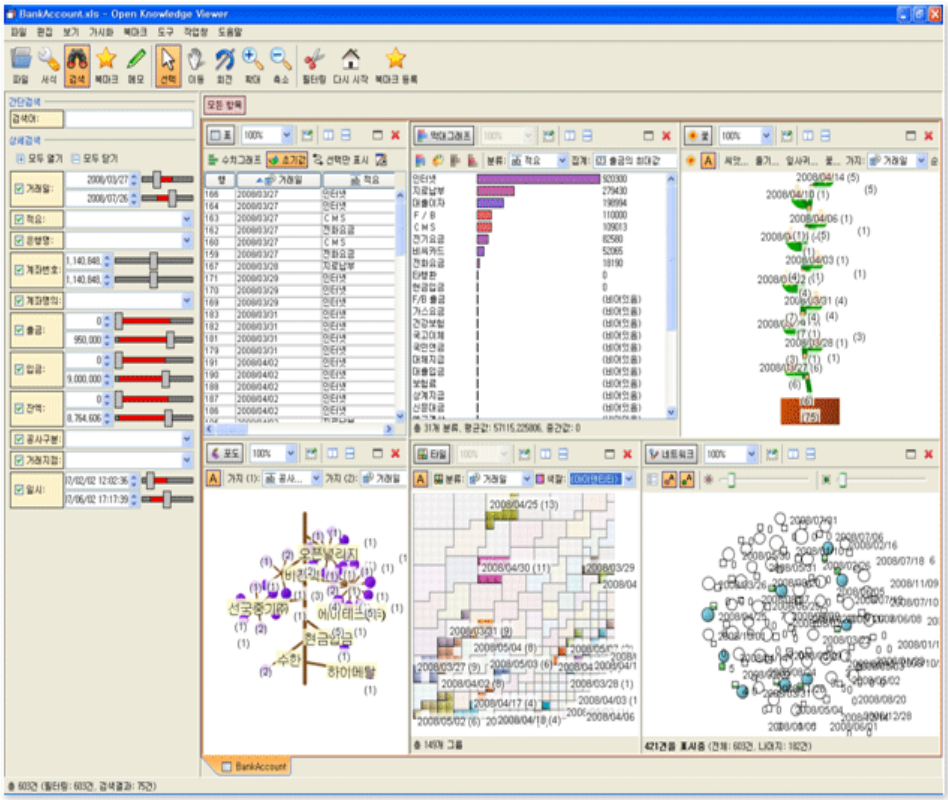
- **CLF(Common Log Format)**로 파일을 생성, 웹 서버의 종류와 설정에 따라 조금씩 차이

CLF 저장 정보 및 설명

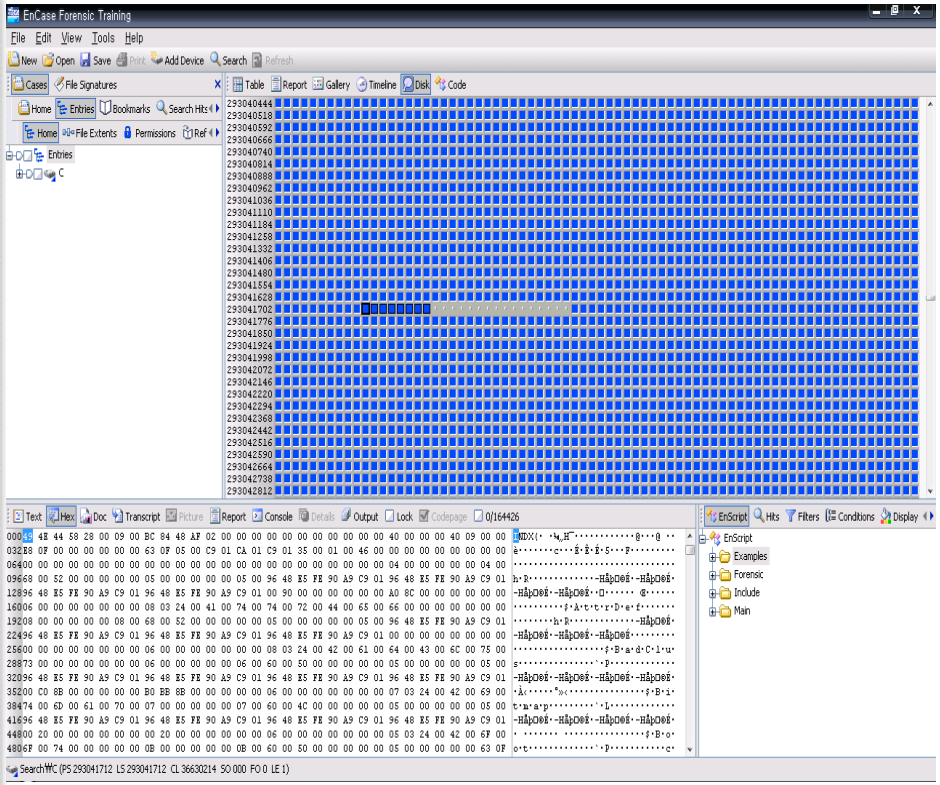
로그 정보	설 명
Host	클라이언트의 호스트 이름이나 IP 주소
Authuser	인증이 필요한 경우 사용자 이름 기록
Date	접속한 시간과 날짜 기록
Request	클라이언트가 요청한 메시지
Status	요청한 것에 대한 서버의 처리사항 (상태 코드)
Bytes	전송된 Bytes의 크기 (헤더 제외)

시각화 기술

- 보이지 않는 것을 일정한 형태로 나타내거나 가려져 있던 어떤 현상이나 실체가 눈에 띄게 드러나게 하는 것으로 추상적인 자료를 사람이 인지할 수 있는 형대로 만드는 과정



Open Knowledge Viewer

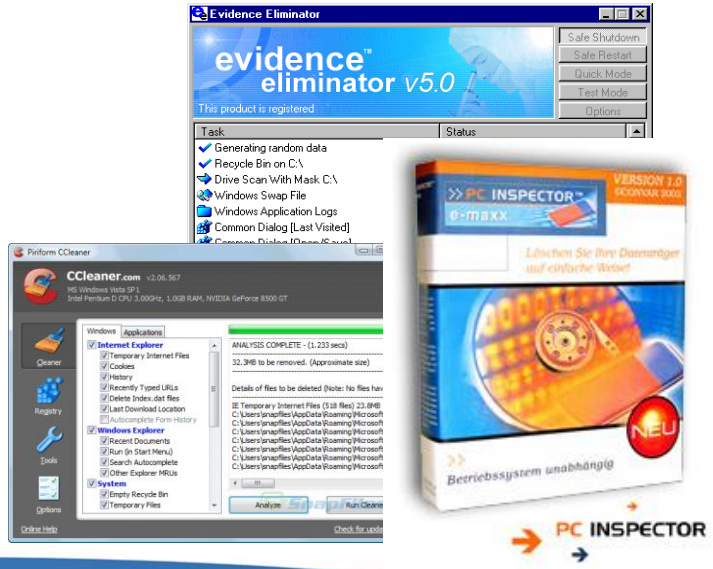


EnCase의 디스크 할당 상태 시각화 기능

안티 포렌식 대응 기술

- 안티 포렌식 **(Anti-Forensic)**이란?
  - 포렌식 기술에 대응하여 자신에게 불리하게 작용할 가능성이 있는 증거물을 차단하려는 일련의 활동
  - 과거에는 증거가 될 수 있는 자료들을 수동으로 처리하였지만, 최근에는 추적 및 증거물 획득을 원천적이고 자동화된 방법으로 막아주는 전문 제품들이 등장하고 있음

- 안티 포렌식 기법
  - 주로 데이터 암호화 등을 통한 복구 기법 회피, 중요 증거 데이터의 증거 자동 삭제, 데이터 은닉 제품 등이 있음
  - 데이터 영구 삭제
    - Disk Wiping, Degausser
    - 증거 자동 삭제
  - 데이터 암호화
    - 압축파일, 문서파일 등의 암호화 등
  - 데이터 은닉
    - 스테가노그래피



안티 포렌식 대응 기술 - 데이터 영구 삭제

- **Disk Wipe**
  - 하드디스크의 기존 데이터를 완벽히 제거하고 모든 Sector의 내용을 0으로 만드는 과정
- 디가우저 (소자, **Degausser**)
  - 하드디스크나 테이프에 강력한 자기장을 노출시켜 기록된 데이터를 파괴하고 복구가 불가능하도록 하는 장비

[Disk Wipe S/W]

안티 포렌식 대응 기술 - 데이터 영구 삭제

- 데이터 복구 기법 회피 기술
  - 디스크 덮어쓰기
    - 삭제된 파일의 데이터 중 물리적으로 디스크에 남아있는 부분을 덮어쓰고 삭제하는 과정을 반복하면 데이터 복구 기법을 회피할 수 있음
    - 美 국방성(DoD)에서는 기밀 자료를 삭제하기 위한 표준 (DoD5220, 22-M)을 다음과 같이 제시하고 있음

      1. 임의의 문자로 데이터를 덮어 씀
      2. 첫 번째 문자의 보수로 덮어씀
      3. 다시 임의의 문자로 데이터를 덮어씀
      4. 이 과정을 7회 반복

안티 포렌식 대응 기술 – 데이터 복구

삭제 파일 복구 기술의 필요성

- 비할당 영역에서 삭제 파일을 복구함으로써 심도 있는 컴퓨터 사용 흔적 조사가 가능
- 용의자 및 범죄자는 증거 인멸을 위해 저장 매체를 물리적으로 파괴, 훼손하거나 디지털 증거를 삭제할 가능성이 높으므로, 이를 원래의 상태로 복구하는 기술이 필요함

파일시스템의 영역

- **메타데이터 영역** : 파일의 이름, 만든이, 날짜, 크기 등을 저장
- **데이터 영역** : 파일의 실제 데이터 스트림 저장

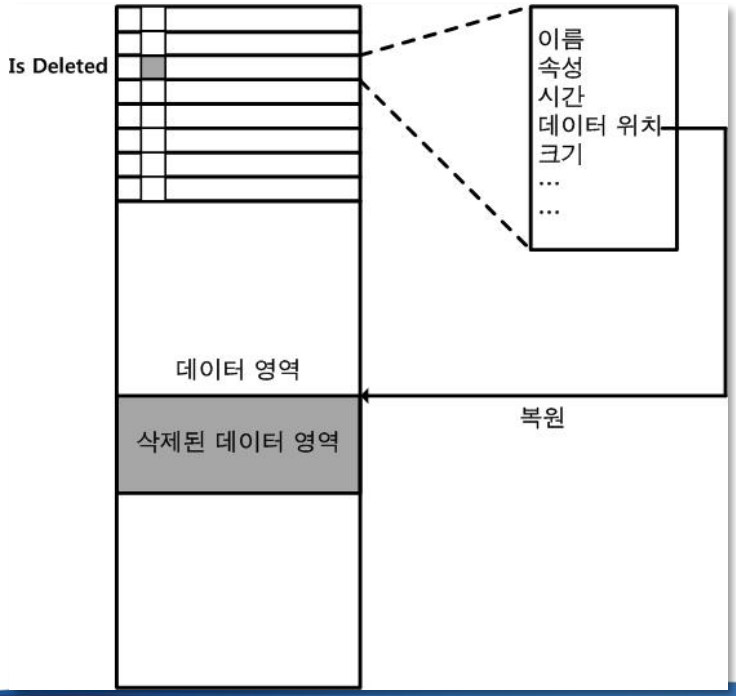
저장매체 복구 기술

- **물리적 복구** : 저장매체의 물리/전자적 복구. 물리적 또는 전자적 단/합선으로 훼손된 저장매체를 정상 상태로 복구하는 기술
- **논리적 복구** : 삭제/훼손된 파일 및 파일 시스템을 복구하는 기술

안티 포렌식 대응 기술 – 데이터 복구

파일시스템 메타데이터기반 복구

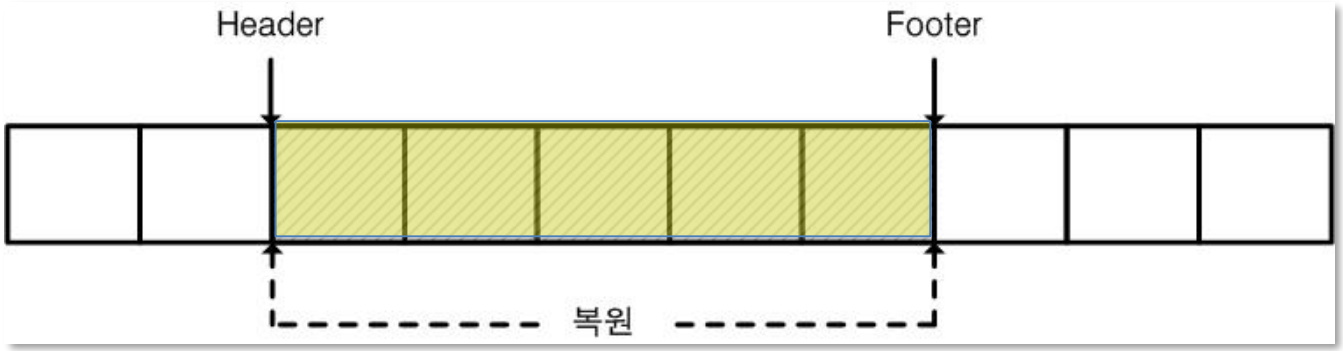
- 파일의 메타 데이터의 **'삭제 플래그'** 를 참조하여 복구
  - 파일 삭제 시 데이터 영역이나 메타데이터 영역을 덮어쓰지 않기 때문에 가능
  - 다른 파일로 덮이지 않았다면 복구 가능



안티 포렌식 대응 기술 – 데이터 복구

파일 시스템 정보를 얻을 수 없는 경우의 복구

- 파일 시스템에서 얻을 수 있는 정보 없이 **'파일 자체 정보'** 기반 복구
  - 즉, 파일의 고유한 특성이 있는 파일만 복구 가능
- 연속적으로 존재하는 파일에 대한 복구는 대부분 가능, 조각난 경우는 어려움
- 추출된 파일이 올바른 파일이라는 보장이 없음
- 많은 시간이 소요됨

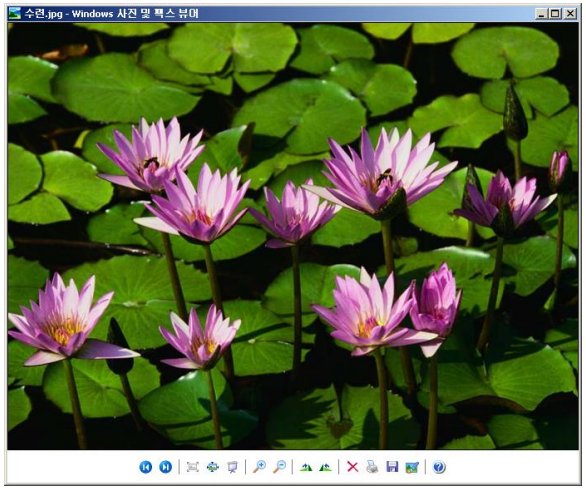


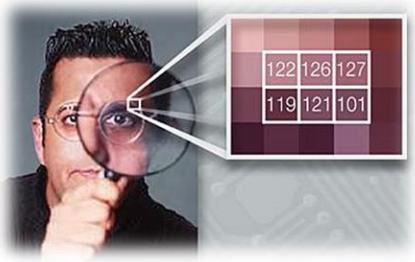
안티 포렌식 대응 기술 - 데이터 암호화

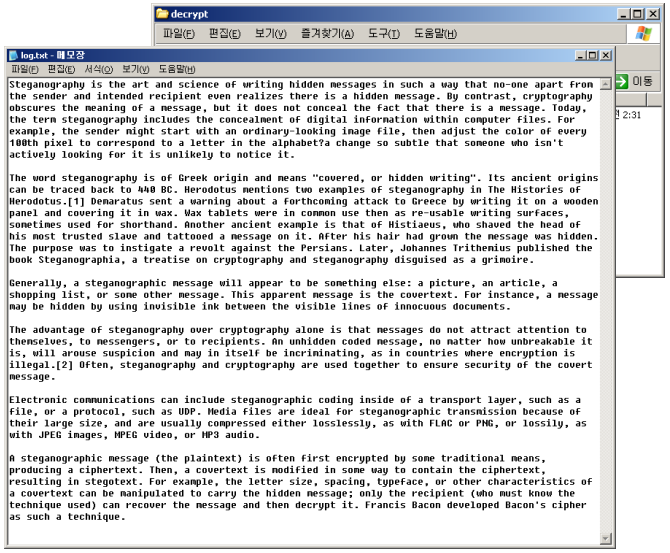
- 데이터 암호화
  - Zip, Rar 등과 같은 압축파일에 암호화 기법을 적용하여, 증거확보를 어렵게 함
  - MS 오피스 및 한글 파일 등과 같은 문서를 암호화하여, 정보를 은폐하는데 활용되고 있음

안티 포렌식 대응 - 스테가노그라피

- **Steganography**
  - 메시지가 전송되고 있다는 사실 즉, 통신의 존재를 숨기는 기술로서 이미지 및 오디오 파일과 같은 다양한 디지털 매체를 통해 메시지를 은닉하여 전송하는 기술
    - 모르는 사람이 보면 평범한 사진에 불과하지만 약속된 수신자는 그 안에 메시지를 확인할 수 있는 기술

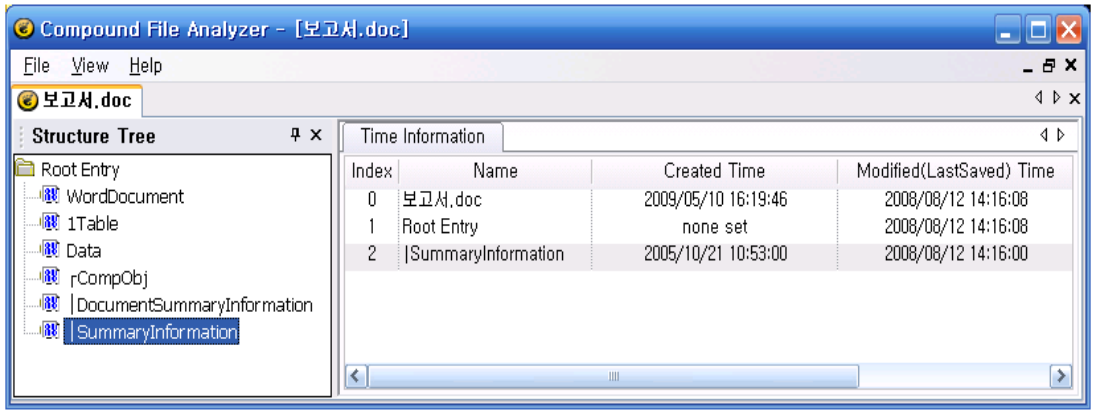






안티 포렌식 대응 기술 – 파일 내부의 시간 정보 분석

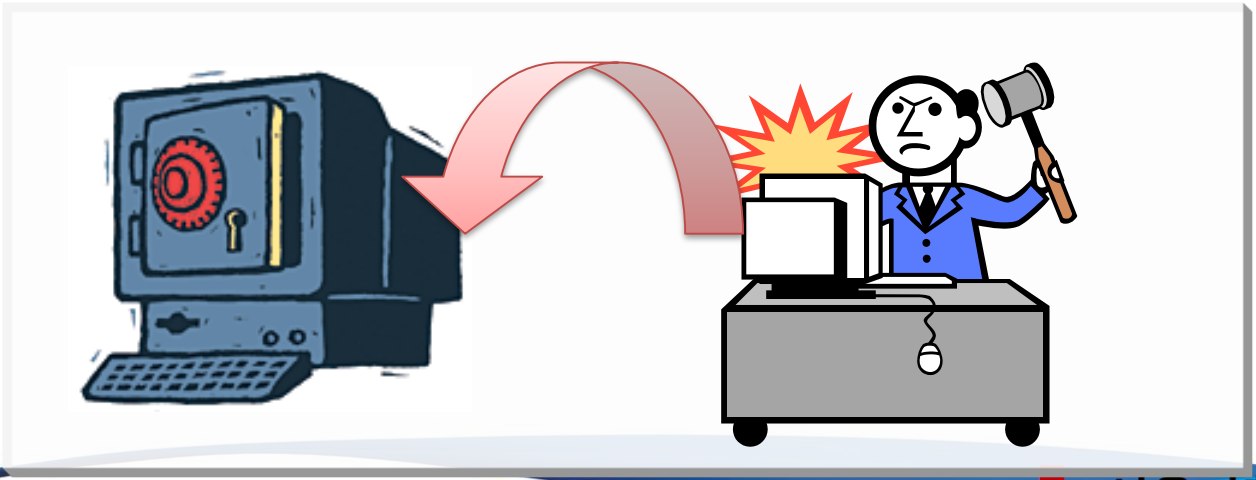
- 파일 시스템 상에 저장되는 시간 정보는 변조가 용이
- 응용 프로그램에 의해 파일 내부에 저장되는 시간 정보는 해당 파일의 저장 형식을 알지 못하면 변경하기 매우 어려움



문서 파일 내부 시간 정보 분석 도구 : CFA

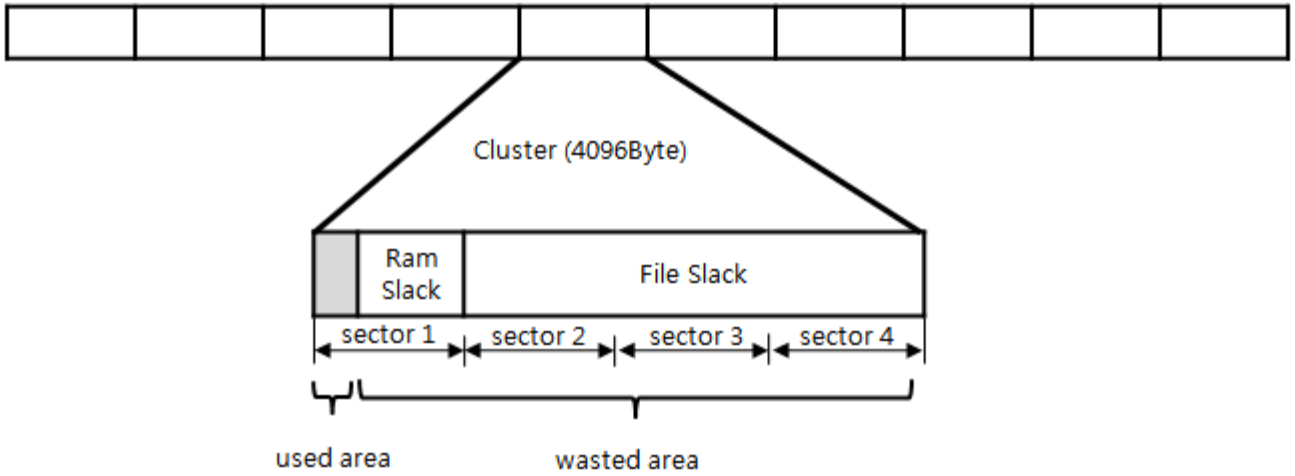
안티 포렌식 기술 대응방안

- **Anti-Forensic**에 대한 대응
  - 프라이버시 및 개인 정보 보호라는 긍정적인 측면도 있지만, 범죄자가 범행직후 증거를 없애는 용도로 사용하는 경우에는 컴퓨터 범죄 수사에 많은 어려움을 초래할 수 있음
  - Anti-Forensic 기술은 앞으로도 계속 발전되고 대중화 될 것이 예상되며, 이에 대응할 수 있는 기술과 정책적 기반이 필요함



파일 시스템의 이해 – 클러스터

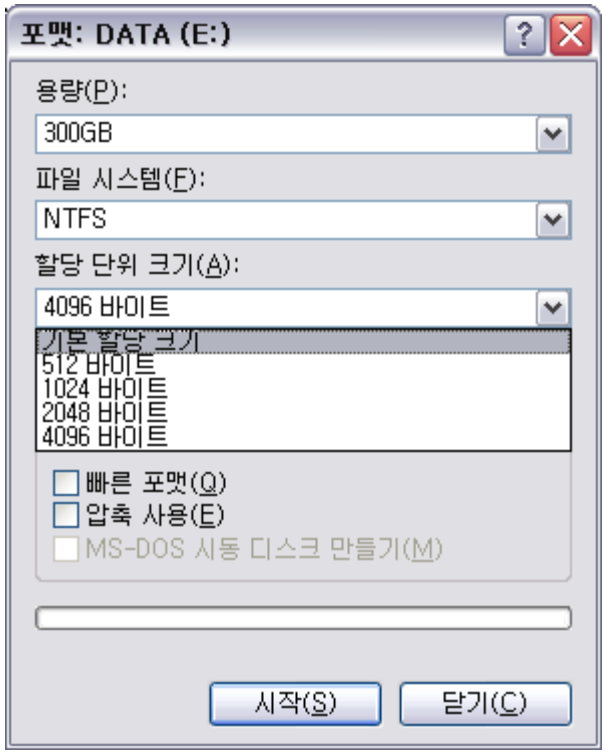
- **Cluster (클러스터)**
  - 클러스터 = 여러 개의 섹터(하드디스크의 물리적 최소 단위)를 묶은 단위
  - 섹터단위로 입출력 처리하면 시간이 오래 걸리므로 여러 개의 섹터를 묶어 한번에 처리



  - 클러스터 크기를 4,096바이트(4KB)로 지정 했을 때, 100바이트의 데이터를 저장하는 경우로 클러스터의 크기만큼 할당됨
    - 3996바이트가 낭비되지만 그럼에도 불구하고 디스크 입출력 횟수를 줄이기 위해 클러스터 단위를 사용
    - 4MB(4,096KB)파일 저장할 때 4KB크기의 클러스터 사용 = 1,024번 입출력 수행
    - 4MB(4,096KB = 4,194,304B)파일 저장할 때 512B크기의 클러스터 사용 = 8,192번 입출력 수행

파일 시스템의 이해 – 클러스터

- **Cluster (클러스터)**
  - 윈도우 시스템에서 디스크 포맷할 때 클러스터 크기 지정



FAT32에서의 클러스터 크기

볼륨 크기	클러스터 크기
32MB – 8GB	4KB
8GB – 16GB	8KB
16GB – 32GB	16KB
32GB	32KB

NTFS에서의 클러스터 크기

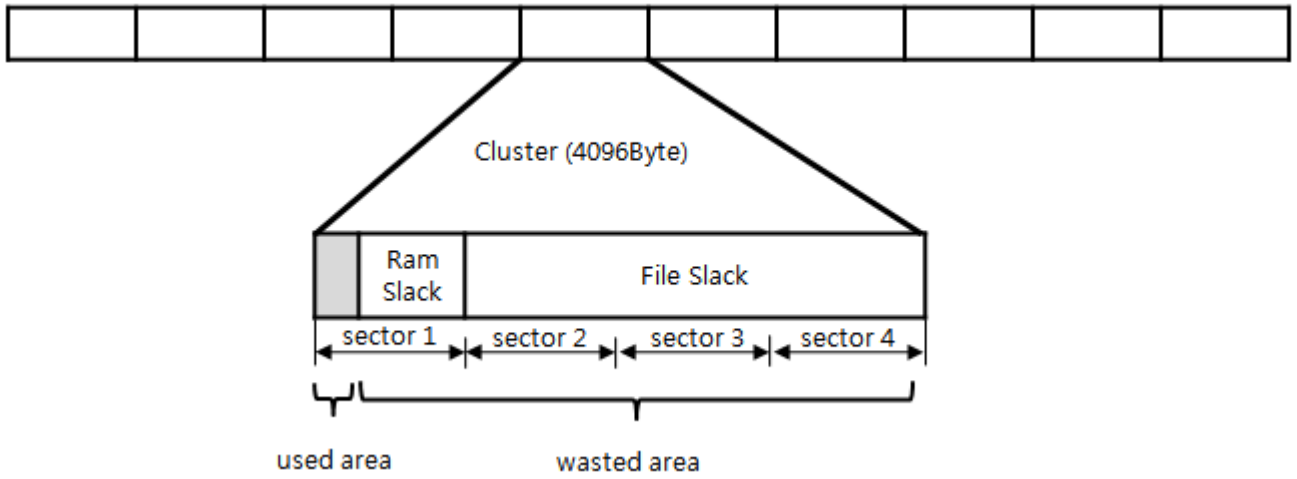
볼륨 크기	클러스터 크기
512MB 이하	512Byte
513MB – 1GB	1KB
1GB – 2GB	2KB
2GB 이상	4KB

파일 시스템의 이해 – 슬랙 공간

- **Slack Space**
  - 물리적인 구조와 논리적인 구조의 차이로 발생하는 낭비 공간
  - 물리적으로 파일에 할당된 공간이지만 논리적으로 사용 할 수 없는 낭비 공간
  - 디지털 포렌식 관점에서 - 정보 은닉 가능성, 파일 복구와 삭제된 파일의 파편 조사
    - RAM Slack (Sector Slack)
    - File Slack (Drive Slack)
      - 이전에 사용한 데이터가 존재, 흔적 조사에 활용
    - File System Slack
    - Volume Slack

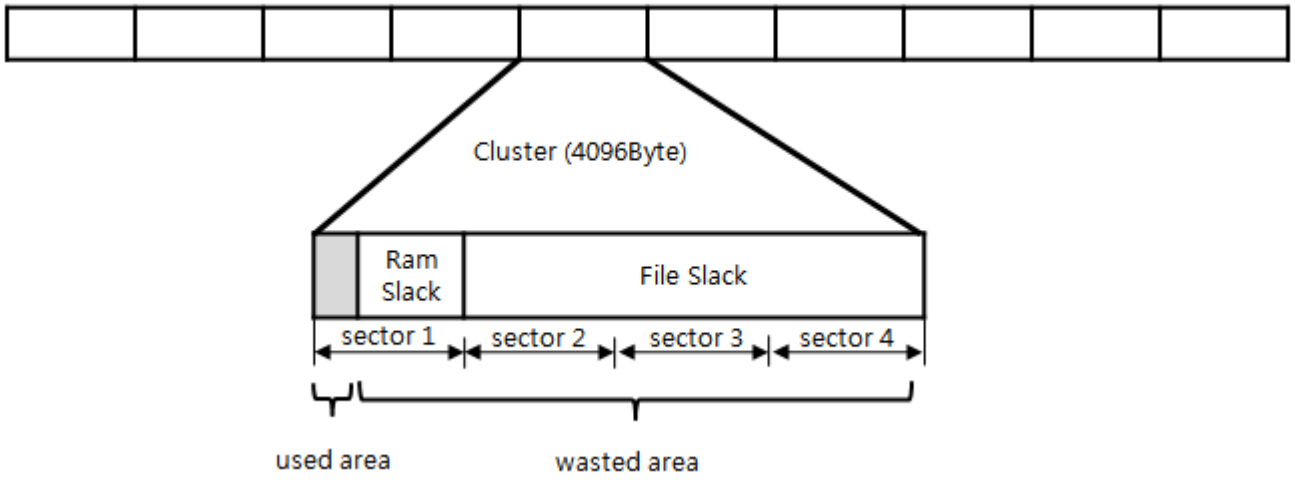
파일 시스템의 이해 – 슬랙 공간

- **Slack Space (RAM Slack & File Slack)**
  - RAM Slack (Sector Slack)
    - 램에 저장 되어있는 데이터가 디스크에 저장될 때 512 바이트씩 기록되는 특성 때문에 발생하는 공간으로 섹터 슬랙(Sector Slack)이라고도 함
    - 지정되는 파일 크기가 512 바이트의 배수가 아닐 경우 발생
    - 여분 바이트 0x00 값으로 기록
    - 램 슬랙을 이용하면 파일의 끝을 알 수 있기 때문에 삭제된 파일 복구 시 유용하게 사용



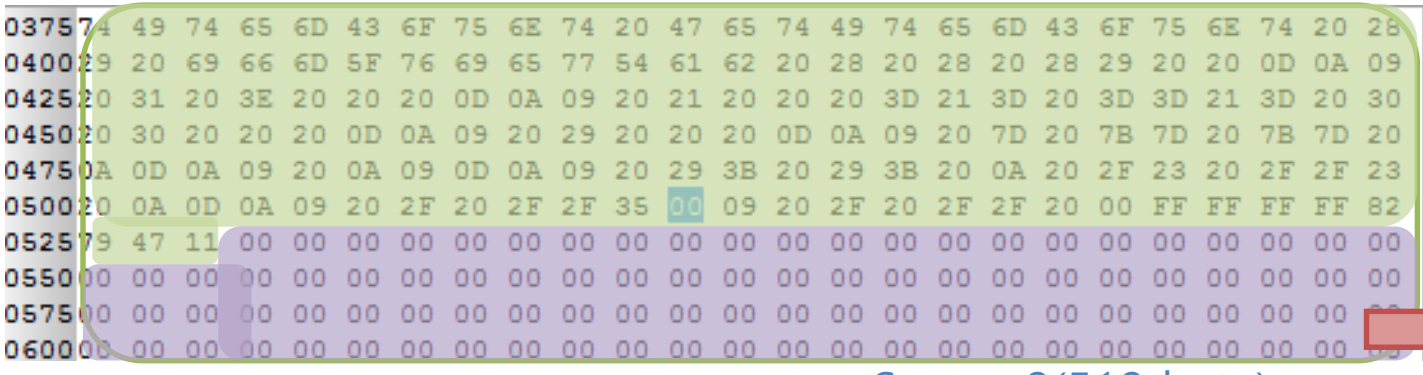
파일 시스템의 이해 – 슬랙 공간

- **Slack Space (RAM Slack & File Slack)**
  - File Slack (Drive Slack)
    - 클러스터의 사용으로 인해 낭비되는 공간 중 램 슬랙을 제외한 부분으로 드라이브 슬랙(Drive Slack)이라고도 함
    - 파일 슬랙을 이용하면 특정 파일이 해당 저장 매체에 존재하였는지 규명 가능
      - 존재 여부를 알아야 할 파일을 클러스터 단위로 나눈 후, 각 클러스터의 마지막 부분과 파일 슬랙 중 일치하는 부분이 있는지 확인
    - 최하단의 디스크 입출력은 섹터 단위로 진행되므로 0x00으로 기록되는 램 슬랙과 다르게 이전의 데이터가 그대로 남아있음(I/O는 섹터단위로 진행)

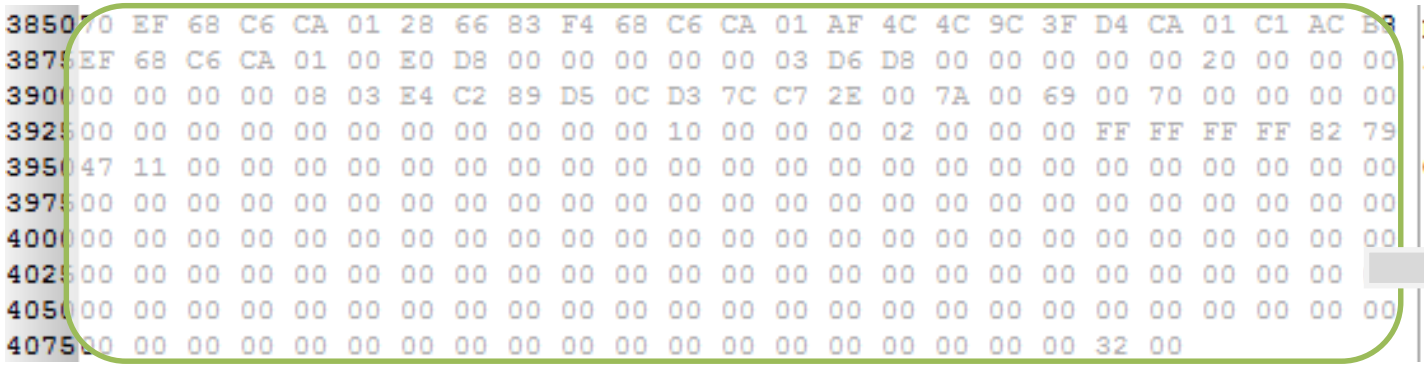


파일 시스템의 이해 – 슬랙 공간

Sector 1(512 byte)



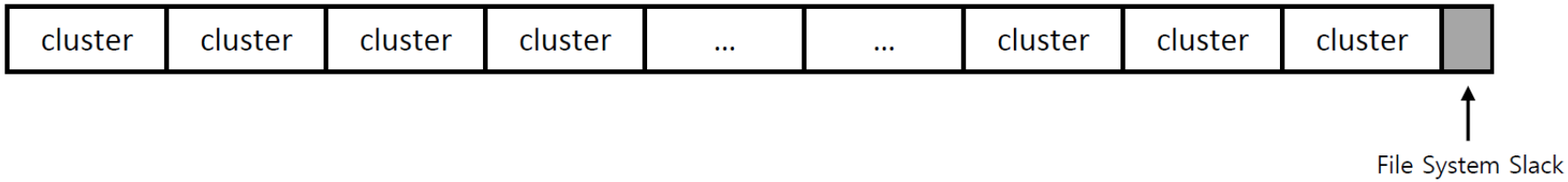
Sector 2(512 byte)



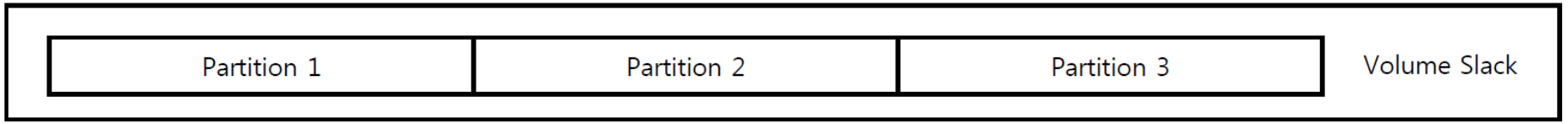
Sector 8(512 byte)

파일 시스템의 이해 – 슬랙 공간

- **Slack Space (RAM Slack & File Slack)**
  - File System Slack
    - 파일시스템 할당 크기와 볼륨 크기간의 차이로 인해 발생되는 공간
      - 1,026KB 볼륨에 4KB 클러스터 사용하는 파일 시스템 구성하면 마지막 2KB이 파일 시스템 슬랙이 됨



  - Volume Slack
    - 전체 볼륨 크기와 할당된 파티션 크기의 차이로 인해 발생되는 공간



참고문헌

- 대검찰정 과학수사본부 : 디지털포렌식센터
  - https://www.spo.go.kr/spo/major/forensics/forensics01.jsp
- 경찰청 사이버 안전국 : 디지털포렌식센터
  - http://cyber.go.kr/bureau/sub4.jsp?mid=040401
- 사이버포렌식협회
  - http://www.cfpa.or.kr/intro2.htm
- 한국포렌식학회
  - https://kdfs.jams.or.kr/co/main/jmMain.kci
- 한국디지털포렌식전문가협회
  - http://fka.kr/

Q & A