EnCase

EVIDENCE FILE FORMAT

Technical Specification - Version 2

GUIDANCE SOFTWARE is now opentext™

CONTENTS

OVERVIEW

The existing EnCase evidence file has performed well for over a decade. It is court-accepted, well-known, and adopted in the industry. Despite its effectiveness, some limitations remain that can only be overcome with an updated evidence file format.

This document outlines the technical details of the updated EnCase evidence file format version 2 (Ex01) so that developers can customize their applications to integrate with the new format. It describes the details, data structures, and algorithms behind Ex01.

The intended audience of this document is a technical reader with a forensic background and familiarity with C-style binary structure layout and algorithms.

COMPARISON OF E01 AND EX01 FORMATS

Many of the central design principles of the E01 format have been retained; implementers familiar with the E01 structure will find the Ex01 format similar. The Ex01 format still stores data in blocks that are verified with an individual 32-bit CRC, and all of the source data stored in the file is hashed with the MD5 and/or SHA-1 algorithms if requested by the user. The Ex01 enhancements do not affect features of the file such as these that many courts have relied on to rule that the file is an accepted container of original evidence; the additions merely facilitate the ability to track and handle new characteristics of the stored data.

EX01 CAPABILITIES

The new Ex01 format introduces the following capabilities:

- Support for encryption of the data.
- Ability to use different compression algorithms.
- Improved support for multi-threaded acquisitions, where sectors can be out of order.
- Efficient storage and handling of sector blocks that are filled with the same pattern (such as 00-byte fills).
- Alignment considerations to improve efficiency and performance.
- Improved support for resuming acquisitions.
- Internal improvements of the data structures.

While some of this new functionality is not yet fully leveraged in the current version, all necessary data is stored, the data structures support expansions, and subsequent versions will use this new format to its fullest.

EVIDENCE FILE STRUCTURE

THE Ex01 file format is built via segments, which are a type of chunked, binary file. An evidence file may be comprised of one or more segments.

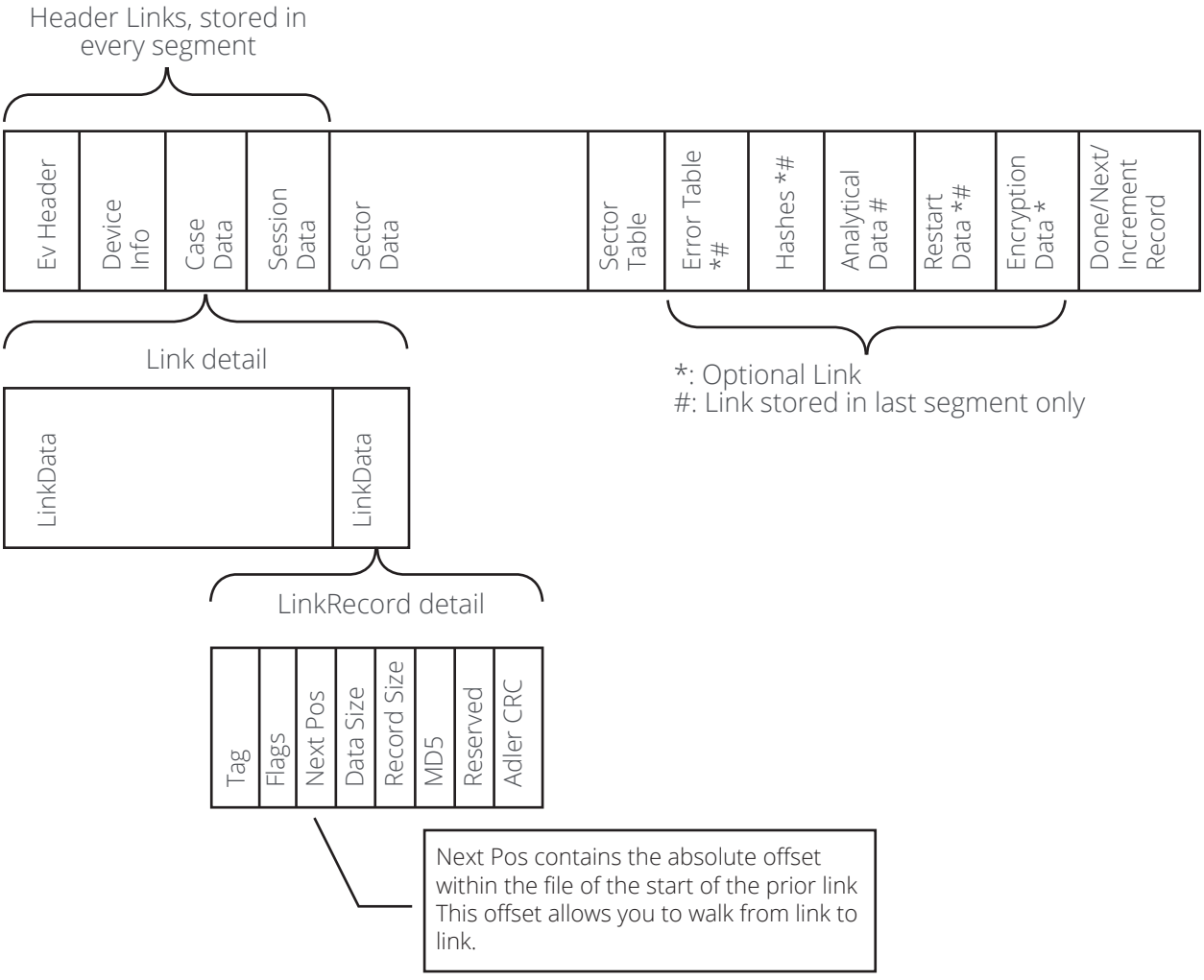
Segments are structured as follows:

- The evidence file header (fixed length).
- The header is followed by a series of variable-size structures (LinkData), and their associated fixed-size LinkRecord structures. LinkRecords are footers that follow the variable-size data. A LinkRecord contains a LinkTag identifying the type of data contained within the Link. Together, the LinkData and LinkRecord are called a Link.
- Every segment has a copy of the Device Information, Case Information, and Session Table (optional, if the evidence is a CD/DVD) link). This data is duplicated for recovery purposes.
- These links are followed by the Sector Data links.
- At the end of every segment, after the sector data, are the following links:
    - Sector Table.
    - If this is not the last segment, then the following links are found:
        - Encryption Data (optional, if evidence is encrypted)
        - Next Record

- If this is the last segment, then the following links are found:
  - Error Table (optional, if errors are encountered)
  - Hash Records (optional, if hashes are calculated)
  - Analytical Data
  - Restart Data (optional, if the evidence file is restarted)
  - Encryption Data (optional, if evidence is encrypted)
  - Done Record

Although there is a header at the beginning of the link, LinkRecords follow the data they represent so that the file can be written out more efficiently. Each LinkRecord contains the offset of the start of the prior LinkRecord; therefore, EnCase reads evidence files starting from the end of the segment.

EVIDENCE FILE VERSION 2 - EVIDENCE FILE SEGMENT LAYOUT



DATA STRUCTURES

Data structures used within the evidence file are all packed with 1-byte packing order. They are padded at the end to multiples of 16 bytes due to encryption considerations. Specific data types used are described in the Data Types section of this document.

EVIDENCE FILE HEADER

```
struct EvHeader {
  uint64    Signature;          // 0x00  ASCII string "EVF2\r\n\x81"
  byte      MajorVersion,       // 0x08  EVMAJOR = 2
            MinorVersion;       // 0x09  EVMINOR = 1
  uint16    CompressionEngine;  // 0x0A  The code for the compression engine (CompressionCodes)
  uint32    Series;             // 0x0C  The evidence number in the series (starting at 1)
  GUID      EvGuid;             // 0x10  Random V4 UUID (Unique)
                                // 0x20  Structure size
};
```

Note: EvGuid is the same GUID as in the Device Information Record.

The compression codes currently in use are:

```
enum CompressionCodes {
  COMPRESSION_NONE    = 0,
  COMPRESSION_LZ      = 1,
  COMPRESSION_BZIP2   = 2
};
```

Note: As some links in the evidence file are always compressed, COMPRESSION_NONE will never be used.

LINK RECORD

```
struct LinkRecord {
  uint32   Tag,              // 0x00 Identifies the type of the data stored (enum; detailed in
                             //      the Link Types section of this document)
           Flags;            // 0x04 Flags for the stored data (FlagBits)
  uint64   Next,             // 0x08 The absolute file offset of the next LinkRecord
           DataSize;         // 0x10 Size of the data stored in this record (excluding RecordSi
  uint32   RecordSize,       // 0x18 Size of the LinkRecord
           PadSize;          // 0x1C Number of pad bytes in data area to reach the 16-byte
                                   boundary
  MD5Class MD5Hash;          // 0x20 MD5 hash of the data as stored (if MD5HASHED is set).
                             //      Used to verify the contents of an encrypted evidence
                             //      file without requiring the decryption key
  uint32   ReservedNull[3], // 0x30 Set to 0
           Adler32;          // 0x3C Adler32 value of the previous fields
                             // 0x40 Structure size
};
```

The flags currently in use are:

```
enum FlagBits {
       MD5HASHED   = 1,
       ENCRYPTED   = 2
};
```

TABLE STRUCTURES

All tables stored in links in the evidence file start with this fixed-size structure prior to the table records:

```
struct TableRecord {
    uint32 Count,           // 0x00 The number of records in this table
           ReservedNull[3]; // 0x04 Unused, set to 0
                            // 0x10 Structure size
};
```

LINK TYPES

Currently available tags are:

- **Device Information: LinkTag (0x01)**
  The Device Information is stored in file object format. A copy of the Device Information is stored in each piece of the evidence file. The end result is stored compressed.
- **Case Data: LinkTag (0x02)**
  Case data is stored in file object format. A copy of the case data is stored in each piece of the evidence file. The link is stored compressed.
- **Sector Data: LinkTag (0x03)**
  Sector data is the core "payload" of the evidence file. Sector data is stored in a series of blocks in this section. Each block must be zero-padded out to 16-byte multiples. Each block is compressed with the compression level stored in the Case Information. If no compression is used on a block, it must have an Adler32 value of the data following the stored data. Compressed data have their own inherent validation.

  NOTE: These Blocks might be out of order.

- **Sector Table: LinkTag (0x04)**
  The Sector Table is a table of BlockOffsets for each sector block in the evidence file. The location in the table is the reference for the block number. This link is not compressed.

```
struct BlockOffset {
   union {
     uint64 FileOffset,  // 0x00 File offset of the start of the block's data.
            FillPattern; // 0x00 uint64 fill pattern if PATTERNFILL is set in Flags
   };
   uint32 Size,          // 0x08 Number of bytes in the file for this block
          Flags;         // 0x0C Flags for this sector block (FlagBits)
                         // 0x10 Structure size
 };
```

The flags currently in use are:

```
enum FlagBits {
  COMPRESSED  = 1, // The data is compressed with the algorithm specified in the file
                   // header
  CHECKSUMED  = 2, //  An Adler32 value of the data is stored at the end
  PATTERNFILL = 4, // FillPattern is the pattern; the data is not stored
 };
```

- **Error Table: LinkTag (0x05)**
  The Error Table is a table of sectors where an error reading occurred from the original device. This link is not compressed.

```
struct ErrorRecord {
    uint64 Sector;         // 0x00 The start sector of the error reading
    uint32 Count;          // 0x08 The number of sectors where there was an error reading
    uint32 ReservedNull;   // 0x0C May be used for error code in the future
                           // 0x10 Structure size
};
```

- **Session Table: LinkTag (0x06)**
  The Session Table is a table of records for storing the track information of a CD-ROM. This link is not compressed.

```
struct SessionRecord {
    uint64 Start;           // 0x00 Start sector of track
    uint32 Flags,           // 0x08 Set to 1 if this record describes a music track,
                            //      otherwise 0
           ReservedNull[5]; // 0x0C Set to 0 for expansion
                            // 0x20 Structure size
 };
```

- **Increment Data: LinkTag (0x07)**
  An Increment Data record is placed near the end of the last evidence file, if the acquisition is not completed. The Increment Data record stores information to allow a restart of the acquisition and show that this is not the last file in the set. The data is stored in file object format. Note that this record is written only if the acquisition is explicitly terminated via cancellation or device dropout and not in the event of a crash during acquisition. The link is stored compressed.
- **MD5 Data: LinkTag (0x08)**
  This is the MD5 hash of the sector data (16 bytes), followed by an Adler32 value of the MD5 data. This link is not compressed.
- **SHA1 Data: LinkTag (0x09)**
  This is the SHA1 hash of the sector data (20 bytes), followed by an Adler32 value of the SHA1 data. This link is not compressed.
- **Restart Data: LinkTag (0x0A)**
  Restart Data is a list of times the acquisition has been restarted. This information is stored in file object format. The link is stored compressed.
- **Encryption Keys Data: LinkTag (0x0B)**
  Encryption Keys Data contains a copy of the keys in each file of the set. Refer to the document outlining the encryption support for Ex01 for further detail.
- **Memory Extents Table: LinkTag (0x0C)**
  The Memory Extents Table is a table of records containing the extent information for an acquisition of a process. This record is stored in the first file of the set. This link is not compressed.

```
struct MemoryExtentRecord {
  uint64 Start, // 0x00 Start page of in use memory.
         Count; // 0x08  Count: Number of pages for this extent.
                // 0x10 Structure size
};
```

- **Next Record: LinkTag (0xD)**
  A LinkRecord with the Next tag is placed at the end of all but the last evidence file of the set. No data follow this record.

- **Final Info: LinkTag (0xE)**
  This is currently unused.
- **Done: LinkTag (0xF)**
  A LinkRecord with the Done tag is placed at the end of the last evidence file of the set. No data follow this record.
- **Analytical Data: LinkTag (0x10)**
  Analytical Data is stored in file object format. A copy of the Analytical Data is stored in each piece of the evidence file. The end result is stored compressed.

FILE OBJECT BASED STORAGE

Some objects, especially ones that are stored in a tree, are serialized as a File Object. File objects serialize data in an objectoriented manner. The data is written out as UTF-16 with a BOM. At the beginning, every field has a short Unicode "tag" to identify its type. This is not repeated for multiple objects of the same type. U+0009 (tab) and U+000D (linefeed) are used to delimit fields and objects. This is the same system that was used in E01.

ESCAPING CONTROL CHARACTERS

Because U+0009 (tab) and U+000D (linefeed) are used to delimit data, these characters must be escaped when they occur in text data. U+000A (carriage return) is also escaped. The following substitutions are used:

- Tab = 0x03
- CR = 0x02
- LF = 0x01

FILE OBJECT DATA

The beginning of the complete File Object contains a UTF16 BOM (0xFF 0xFE). The first item is a number indicating how many file objects are to be written. Each file object has a unique tag to identify it, followed by a linefeed, then the data.

RECORD FILE OBJECT

This is a file object that contains a tab-delimited list of records beneath it. The first line of the text is the ordering of the fields by tag (a unique string for each object), followed by a tab-delimited list of data, ending with a linefeed. The end of the records is specified by an empty linefeed.

TEXT FILE OBJECT

Contains the UTF16 text of the data.

INTEGER32

The integer of the data, in text, between 0 and 4,294,967,295 without commas.

INTEGER64

The integer of the data, in text, between 0 and 18,446,744,073,709,551,615 without commas.

DATE

An Integer32 with the number of seconds since January 1, 1970.

INTEGER ARRAY

A space-separated list of integers, without commas.

ENUM FILE OBJECT

A single character representing an enumeration.

BOOL PROPERTY FILE OBJECT

A single character: "1" if the property is true, blank otherwise.
An example of a file object in binary form is provided in the File Object Detail section for Case Data: LinkTag (0x02)

FILE OBJECT DETAIL

DEVICE INFORMATION: LINKTAG (0X01)

The Device Information Link contains information specific to the device being acquired.

Device Information: Record File Object – Tag("main")

- Parent object
- Members

Name	File Object Type	Tag	Description
SerialNumber	Text File Object	sn	Serial number of the drive acquired.
Model	Text File Object	md	Model number of the drive acquired.
TotalSectors	Integer64	ts	Total number of sectors of the acquired device.
HPASectors	Integer64	hs	Total number of sectors in the HPA protected section of the drive.
DCOSectors	Integer64	dc	Total number of sectors in the DCO protected section of the drive.
RamSectors	Integer32	rs	Total number of sectors for a Palm RAM device.
LogSector	Integer32	ls	Total number of sectors in the SMART or general logs.
ProcessId	Integer32	pid	The process ID of acquired memory.
DriveType	Enum File Object	dt	The type of the device stored in the evidence file. **r** = Removable **f** = Fixed **c** = CD-ROM **a** = RAM Disk **p** =Palm **l** = Logical evidence **m** = Memory
Label	Text File Object	lb	Description of device.
Physical	Bool Property File Object	ph	True, if the acquired device is physical.
BytesPerSector	Integer32	bp	The number of bytes in each sector of the device.

CASE DATA: LINKTAG (0X02)

The Case Data Link contains information specific to this particular acquisition of the device.
Case Information: Record File Object – Tag("main")

- Parent object
- Members

Name	File Object Type	Tag	Description
CaseNumber	Text File Object	cn	Case number, as entered by the user.
EvidenceNumber	Text File Object	en	Evidence number, as entered by the user.
Name	Text File Object	nm	Name of evidence, as entered by the user.
Examiner	Text File Object	ex	Name of examiner.
Notes	Text File Object	nt	User-entered notes.
AppVersion	Text File Object	av	Program name and version of acquisition program.
OSVersion	Text File Object	os	Version of operating system that acquisition is performed on.
TargetTime	Date	tt	GMT time of machine that acquisition is performed on.
ActualTime	Date	tb	GMT time, as entered by the user.
TotalBlocks	Integer64	tb	Total blocks of evidence file.
Compression	Integer32	cp	Compression used for evidence.
SectorsPerBlock	Integer32	sb	Sectors per block of compressed data.
Granularity	Integer32	gr	Error granularity when reading device.
BlockerFlags	Integer32	wb	Which write blocker is being used during acquisition: 1 = FastBloc 2 = Tableau

This example shows the uncompressed binary format of an example Case Data file object. This data is compressed before being written to the evidence file.

```
ff fe 31 00 0a 00 6d 00 61 00 69 00 6e 00 0a 00 6e 00 6d 00 09 00 63 00 6e 00 09 00 65 00 6e  ÿþ1...m.a.i.n...n.m...c.n...e.n
00 09 00 65 00 78 00 09 00 6e 00 74 00 09 00 61 00 76 00 09 00 6f 00 73 00 09 00 74 00 74 00  ...e.x...n.t...a.v...o.s...t.t.
09 00 61 00 74 00 09 00 74 00 62 00 09 00 63 00 70 00 09 00 73 00 62 00 09 00 67 00 72 00 09  ..a.t...t.b...c.p...s.b...g.r..
00 77 00 62 00 0a 00 46 00 32 00 09 00 09 00 09 00 09 00 09 00 09 00 37 00 2e 00 33 00 2e 00 30 00  .w.b...F.2..........7...3...0.
2e 00 36 00 30 00 09 00 57 00 69 00 6e 00 64 00 6f 00 77 00 73 00 20 00 37 00 09 00 31 00 33  ..6.0...W.i.n.d.o.w.s. .7...1.3
00 32 00 30 00 30 00 37 00 37 00 38 00 35 00 30 00 09 00 31 00 33 00 32 00 30 00 30 00 37 00  .2.0.0.7.7.8.5.0...1.3.2.0.0.7.
37 00 38 00 35 00 30 00 09 00 31 00 31 00 38 00 34 00 30 00 09 00 09 00 36 00 34 00 09 00 36  7.8.5.0...1.1.8.4.0.....6.4...6
00 34 00 09 00 0a 00 0a 00                                                                     .4.......
```

The binary data above represents the data in the following table.

main													
nm	cn	en	ex	nt	av	os	tt	at	tb	cp	sb	gr	wb
F2					7.3.0.60	Windows 7	1320077850	1320077850	11840		64	64	

INCREMENT DATA: LINKTAG (0X07)

Increment Data Link contains information required to restart an acquisition.

Increment Data: Record File Object – Tag("main")

- Parent object
- **StateRecord:** Record File Object – Tag("rc")
  - Parent object - Record for the state of the hash
  - Members

Name	File Object Type	Tag	Description
State0	Integer32	s0	State 0 of the hash.
State1	Integer32	s1	State 1 of the hash.
State2	Integer32	s2	State 2 of the hash.
State3	Integer32	s3	State 3 of the hash.
Count0	Integer32	c0	Count 0 of the hash.
Count1	Integer32	c1	Count 1 of the hash.
Buffer	IntegerArray	buf	The 64-byte buffer for the state of the hash.
Bytes	Integer64	sb	Number of bytes currently hashed.
Key	IntegerArray	sk	Key data for SHA1 Hashing – 64 bytes (array of 16 uint32).
IV	IntegerArray	si	IV data for SHA1 Hashing - 20 bytes (array of 5 uint32).
Options:	Integer32	op	Hashing options (1=MD5, 2=SHA1, 3=Both).

- **Record**: Record File Object Data - Tag("rec")
  - Parent object
  - Members

Name	File Object Type	Tag	Description
Start	Integer64	sr	Start sector of this acquisition.
Stop	Integer64	sp	Stop sector of this acquisition.
Date	Date	d	Acquisition start time.

- **Extra Info**: Record File Object Data - Tag("m") - Record to hold extra information
  - Parent object
  - Members

Name	File Object Type	Tag	Description
Block	Integer64	bl	Block number of written hash state.
TableCount	Integer32	tc	Position in the current sector table of hash state.

RESTART DATA: LINKTAG (0X0A)

This is a list of restart nodes that carry information about when an acquisition was restarted.

Restart Data: Record File Object – Tag("main")

- Parent object – This is the root, and an empty RestartNode itself
- **Restart Node**: Tree File Object – Tag("rl")
  - Parent object - Record for the one restarted acquisition
  - Members

Name	File Object Type	Tag	Description
Props	Integer32	p	Properties of this node
Date	Date	d	Acquisition start time
Start	Integer64	sr	Start sector of this acquisition
Stop	Integer64	sp	Stop sector of this acquisition

ENCRYPTION

E01 had the ability of using a "soft" password, meaning that the data itself was not encrypted. Ex01 encrypts the data symmetrically, using AES-256 by default. The encryption key for this can be protected with:

- A password that generates a symmetric key.
- An asymmetric key pair.
- Both of the above.

The encryption algorithms are not fixed. An algorithmID is stored so that EnCase can support additional encryption algorithms in the future. Please refer to the document outlining the encryption support for Ex01 for further detail.

DATA TYPES

The data types used are:

Term	Description
byte	Single-byte unsigned 8-bit integer (0-based).
uint16	2-byte unsigned 16-bit integer (0-based) little-endian.
uint32	4-byte unsigned 32-bit integer (0-based) little-endian.
uint64	8-byte unsigned 64-bit integer (0-based) little-endian.
GUID	16-byte UUID V4 based Global Unique Identifier.
MD5Class	16-byte MD5 of data stored in the standard format.
Date	32-bit (4-byte) unsigned value of number of seconds since January 1, 1970 (epoch / Unix time).
Text	Unicode data in UTF-16 LE (little-endian) usually 0-terminated.
Enum	An integer with specific values.
IntegerArray	An array of integers.

GLOSSARY

The terms used are:

Term	Description
Link	Binary chunk consisting of the data and a descriptor (LinkRecord).
LinkRecord	Fixed-size footer describing the prior data.
File Object	Text-based storage.
Tag(“main”)	A UTF16-LE string, in this case “main” (case sensitive).
Segment	One on-disk file; part of a piece of evidence. (For example, Ex01, Ex02, and Ex03 would be three segments of one evidence file.)
E01	Extension for the first segment of evidence in the version 1 format. Sometimes the only segment. Also, a term for a piece of evidence in the evidence file format version 1.
Ex01	Extension for the first segment of evidence in the version 2 format. Sometimes the only segment. Also a term for a piece of evidence in the evidence file format version 2. This was developed for EnCase Version 7 and beyond.
Adler32	A 32-bit checksum used for quick validation purposes.

GUIDANCE SOFTWARE

ABOUT GUIDANCE

Guidance exists to turn chaos and the unknown into order and the known-so that companies and their customers can go about their daily lives as usual without worry or disruption, knowing their most valuable information is safe and secure. The makers of EnCase®, the gold standard in forensic security, Guidance provides a mission-critical foundation of market-leading applications that offer deep 360-degree visibility across all endpoints, devices and networks, allowing proactive identification and remediation of threats. From retail to financial institutions, our field-tested and court-proven solutions are deployed on an estimated 35 million endpoints at more than 70 of the Fortune 100 and hundreds of agencies worldwide, from beginning to endpoint.