Veri Elde Etme (Data Acqusition)

Dosya Sistem Analizi Dersi Hafta 3

Yrd. Doç. Dr. Erhan AKBAL

Genel Veri Toplama Prosedürü

- Bir depolama aygıtından veri toplamak için genel ve sezgisel yöntemler kullanılır.
- 1 baytlık veriyi kaynaktaki orijinal depolama cihazından hedef bir depolama aygıtına kopyalayarak bu işlemi veri bitene kadar tekrarlamaktır.
- Bu normal bir belgenin nokta virgülüne kadar elle kopyalanarak aynısının elde edilmesine benzerdir. Fakat çoğumuz karakter karakter kopyalama yapmayız. Kelimeleri aklımızda tutabileceğimizden birden fazla kelimeyi tek seferde kopyaya aktarabiliriz.
- Bilgisayarlar aynı şeyi yapar ve şüpheli sistemlerden gelen verileri 512 bayttan binlerce bayta değişen bir miktarda kopyalar. Her seferinde aktarılan veri parçaları genellikle 512 baytın katları yada diskteki sektör boyutunda kopyalar. Şüpheli sürücüden veri okurken edinme aracı bir hata ile karşılaşırsa, birçok araç hedefe sıfır yazacaktır

Veri Toplama Katmanları

- Uçucu olmayan veri edinimi ile ilgili genel teori, kanıt içerebileceğini düşündüğümüz her byte veriyi kurtarmaktır.
- Verilerin disk, volüm, dosya ve uygulama katmanları gibi farklı katmanlarda yorumlanabileceğini görmüştük.
- En temel kural, kanıt olacağını düşündüğümüz en düşük katmandaki verileri toplamaya çalışmaktır. Çoğu durumda inceleme uzmanı diskin tüm sektörlerinden veri toplamaktadır.
- Ayrıca her sektörün sadece içeriğini aldığımızda veri kurtarma uzmanlarının gereksinim duyabileceği verileri kaybedebiliriz.

Veri Toplama Katmanları

- Niçin genelde disk seviyesinde veri edindiğimizi göstermek için bazı senaryoları değerlendireceğiz.
- Birim (Volume) seviyesinde bir disk aldık ve her bölümdeki (Partition) her sektörün bir kopyasını aldık. Bu, her bölümdeki silinen dosyaları kurtarmamızı sağlayacaktır, ancak bölümlere ayrılmayan sektörleri analiz etmemizi engelleyecektir.
- DOS bölümlerine sahip bir disk 1 - 62 arasındaki sektörleri kullanmayabilir ve gizli veriler içerebilir.
- Volume düzeyinde veri edinirsek, gizli veriler kaybolacaktır.
- Yedekleme programları sadece bölümlendirilmiş dosyaları kopyalamaktadır. Bu durumda silinen dosyalara, bölüm ve dosya sistemi veri yapılarına, gizlenmiş verilere erişilemeyebilir.

Örnek Senaryo - Yedekleme

- Bazen sadece yedeklenmiş veriden incelemenin yapılması gerekebilir.
- Bu nedenle yedeğin inceleme uzmanı tarafından en iyi şekilde kullanılması gerekir.
- Bir yedeklemenin kritik olacağı bir senaryoda, bir sunucunun diskleri sıfırlarla silinip sonra yeniden başlatıldığı için yanıt vermediği kurumsal bir ortam olduğu düşünülürse;
- Sistemin son yedeklemeleri, kimin sisteme erişimi olduğunu ve bir saldırganın onu tehlikeye atıp atmadığı konusunda yapılacak inceleme ipucu sağlayabilir.

Örnek Senaryo - IDS

- Bazı sistemler için, kanıt olacağını düşündüğümüz seviyede veri toplama konusundaki temel kuralımız, yalnızca dosyaları kopyalamanız gerektiğini ifade eder.
- Saldırıya karşılık tutulan log kayıtlarını içeren bir Saldırı Tespit Sistemi (IDS) bulunan bir saldırı araştırması düşünelim.
- IDS'in ele geçirildiğini düşünmüyorsak, sistemdeki tek kanıt dosya düzeyindedir ve gerekli logları kopyalayıp uygun koruma adımlarını atabiliriz.
- IDS'in ele geçirildiğini düşünürsek, tüm verileri analiz edebilmemiz için disk düzeyinde inceleme yapmamız gerekir.

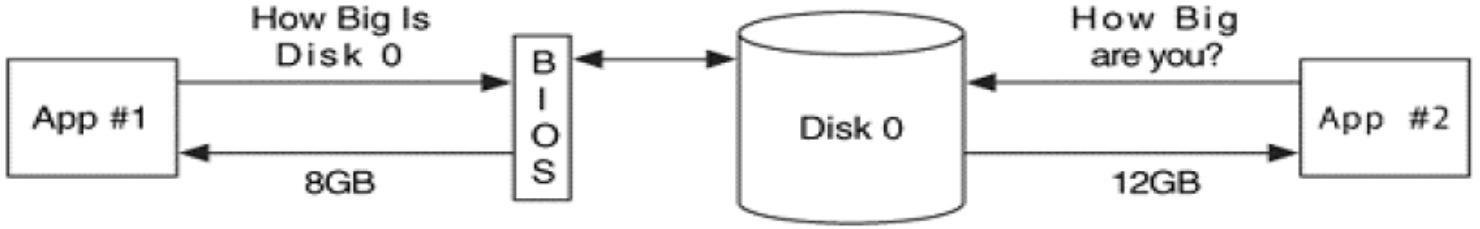
BIOS ile Verilere Erişim

- Bir diskteki verilere erişebilmek için iki yöntem kullanılır.
- Birinci yöntem, işletim sistemi veya veri toplama yazılımı, donanım ayrıntılarını bilmesini gerektiren yazılıma ihtiyaç duymadan sabit diske doğrudan erişir.
- İkinci yöntemde, işletim sistemi veya veri elde etme yazılımı, sabit diske, tüm donanım ayrıntılarını bilen Temel Giriş / Çıkış Sistemi (BIOS) aracılığıyla erişir.
- Bakıldığında bu yöntemler arasında farklılık var gibi gözükmemekte ve BIOS'u kullanmak daha kolay görünüyor çünkü donanım ayrıntılarıyla ilgileniyor. Ne yazık ki, bir inceleme yaparken bu kadar açık değildir.

BIOS ile Verilere Erişim

- BIOS kullanıldığında, disk hakkında yanlış bilgilerin elde edilme riski bulunmaktadır.
- BIOS, bir diskin 8GB olduğunu düşünse de disk gerçekte 12GB'tır, INT13h işlevleri size sadece ilk 8GB'a erişmenizi sağlar. Bu nedenle, disk edinimi yapıyorsanız, son 4GB'lık bir kopyasını kopyalamazsınız. Şekil'de iki uygulamanın farklı yöntemleri kullanarak bir diskin boyutunu belirlemeye çalıştıklarını görebiliriz.

Figure 3.1. Two applications are trying to determine the size of a disk. The BIOS is not properly configured and says that the 12GB disk is only 8GB.



BIOS ile Verilere Erişim

- Bir uygulamanın BIOS tan disk boyutunu isteyebileceği iki yol vardır.
- İlki 8 GB katlarıyla **CHS biçiminde** disk geometrisini kullanarak orijinal INT13h işleviyle boyut döndürmektedir.
- İkincisi ise **LBA formatı** kullanarak genişletilmiş **INT13h** fonksiyonu kullanarak döndürmektedir. NIST tarafından iki yolla testler gerçekleştirilmiş ve genişletilmiş INT13h ile elde edilen değerlerin doğru olduğu ortaya konmuştur.

Canlı Sistemden Veri Edinme

- Bir inceleme uzmanının kapalı veya canlı olmak üzere iki veri toplama seçeneği vardır.
- Şüpheli bir sistemdeki veriler şüpheli işletim sisteminin yardımı olmadan kopyalanması durumunda kapalı ya da ölü bir inceleme yapılmış olur.
- Canlı edinme, şüpheli işletim sisteminin halen çalışmakta ve verileri kopyalamak için kullanıldığı bir veri elde etme sürecidir.

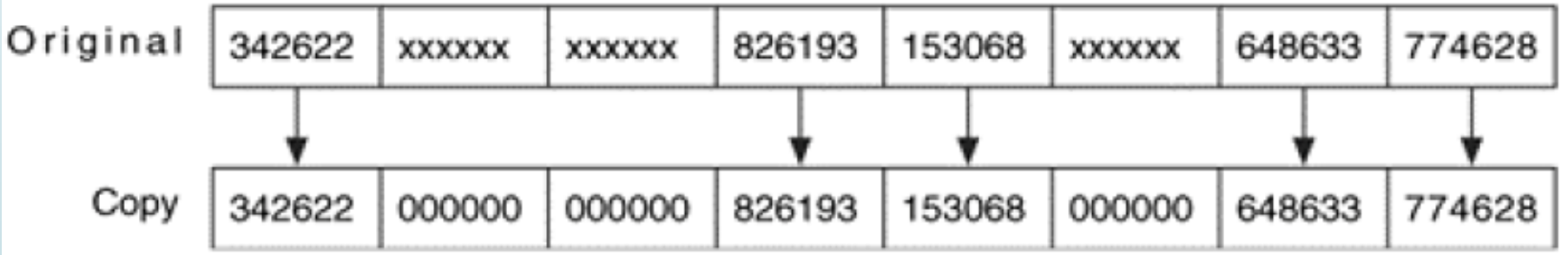
Canlı Sistemden Veri Edinme

- Saldırganlar, genellikleri uzlaştıkları sistemlere kök setleri adı verilen araçları yükler ve bir kullanıcıya yanlış bilgi döndürürler.
- Kök setleri, dizindeki bazı dosyaları veya çalışan işlemleri gizler. Saldırganlar sistemin güvenliğini aştıktan sonra yükledikleri dosyaları gizler.
- Bir saldırgan işletim sistemini de değiştirebilir, böylece veri elde edildiği sırada diskin belirli sektörlerindeki verilerin yerine geçer. Ortaya çıkan görüntü, değiştirildiğinden olayın kanıtına sahip olunmayabilir.
- Mümkün olduğunda, tüm kanıtların güvenilir bir şekilde toplanabilmesi için canlı edinimden kaçınılmalıdır.
- Bir inceleme uzmanının, yapılandırılmış güvenilir bir Linux CD'si kullanarak önyükleme ile şüpheli bir sistemin sürücülerine bağlanmadan yada verileri değiştirmeden incelemek yaygın yöntemdir.
- Teknik olarak, şüphelinin, güvenilir bir işletim sisteminde bile sahte veri döndürmesi için donanımlarını değiştirmesi mümkündür, ancak bu, işletim sisteminin değiştirilmesinden çok daha az olasıdır.

Hata İşleme

- Kötü bir sektörle karşılaşıldığında genel kabul gören davranış, o adresi loglara kaydetmek ve okunamayan veriler için 0 yazmaktır.
- 0 yazma, diğer verileri doğru konumda tutar. Eğer sektör 0'lar yazmak yerine yok sayılırsa sonuçtaki kopya çok küçük olur ve analiz araçları bu kopyayı açamaz.

Figure 3.2. The original has three errors in it that have been replaced by 0s.



HPA

- Bir ATA diskten veri alırken, gizli verileri içerebileceğinden diskin Kullanıcı Korumalı Alanı'na (HPA) dikkat etmek gerekir. Bir imaj alma aracı HPA alanlara bakamıyorsa imajı doğru alamayacaktır.
- İmaj programı iki ATA komutunun çıktısını karşılaştırarak bir HPA algılayabilir.
- Gerekli ATA komutlarını yürütecek bir araca erişiminiz yoksa, veri elde etme işlemi sırasında kopyalanan sektör sayısını, disk etiketinde belgelenen sektörlerin sayısıyla karşılaştırmanız gerekebilir.
- HPA'lı bir diskle karşılaşırsanız ve gizli verilere erişmek istiyorsanız, disk yapılandırmasını değiştirmeniz gerekir. Bir HPA, maksimum kullanıcı adresli sektörü diskte maksimum sektör olacak şekilde ayarlayarak kaldırılır.
- Bu sabit disk kapatıldığında yapılandırma değişikliği kaybolacak şekilde uçuculuk biti kullanılarak yapılabilir.

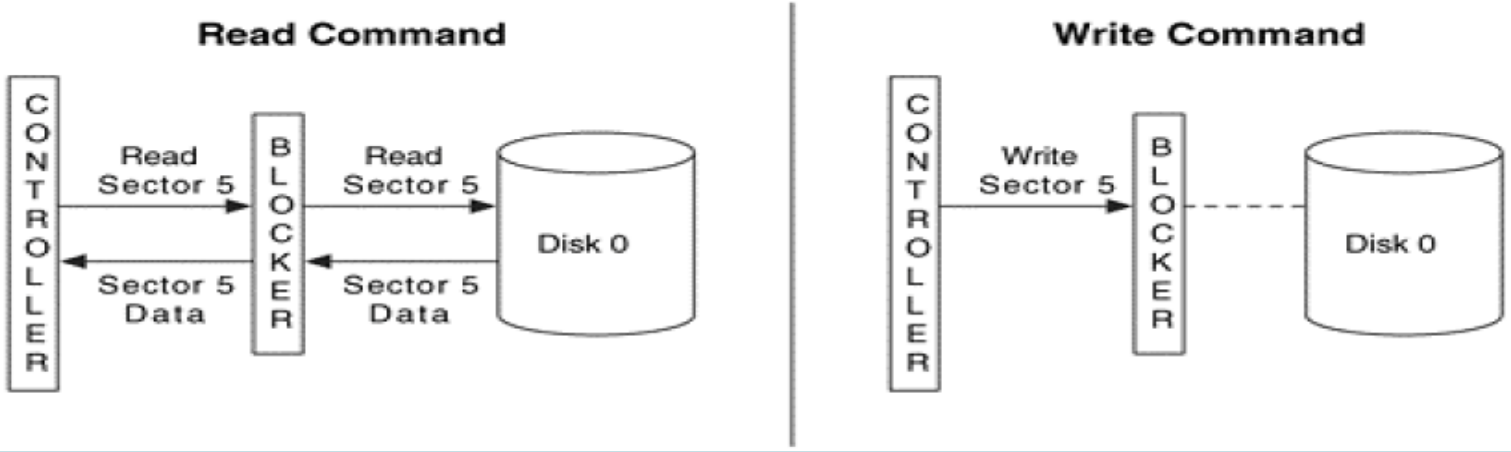
DCO

- Yeni teknoloji bir ATA diskten veri elde ederken, diskin gerçekten daha küçük görünmesine neden olabilecek bir Aygıt Yapılandırma Kaplaması (DCO) varmı yokmuya bakmak gerekir.
- Bir DCO, iki ATA komutunun çıktısı karşılaştırarak tespit edilir. Bir DCO var ve tüm veriler alınacaksa DCO nun kaldırılması gerekmektedir.
- Kullandığımız birçok yazma koruma aracı DCO'yu algılar ve kaldırır.

Donanımsal Yazma Koruyucular

- Donanımsal yazma koruyucu, bir bilgisayar ve bir depolama aygıtı arasındaki bağlantıda bulunan bir aygıttır.
- Verilen komutları izler ve bilgisayarın depolama aygıtına veri yazmasını engeller.
- Yazma engelleyici, ATA, SCSI, Firewire (IEEE 1394), USB veya Seri ATA gibi birçok depolama arabirimini destekler.

Figure 3.3. The read request for sector 5 is passed through the write blocker, but the write command for the same sector is blocked before it reaches the disk.



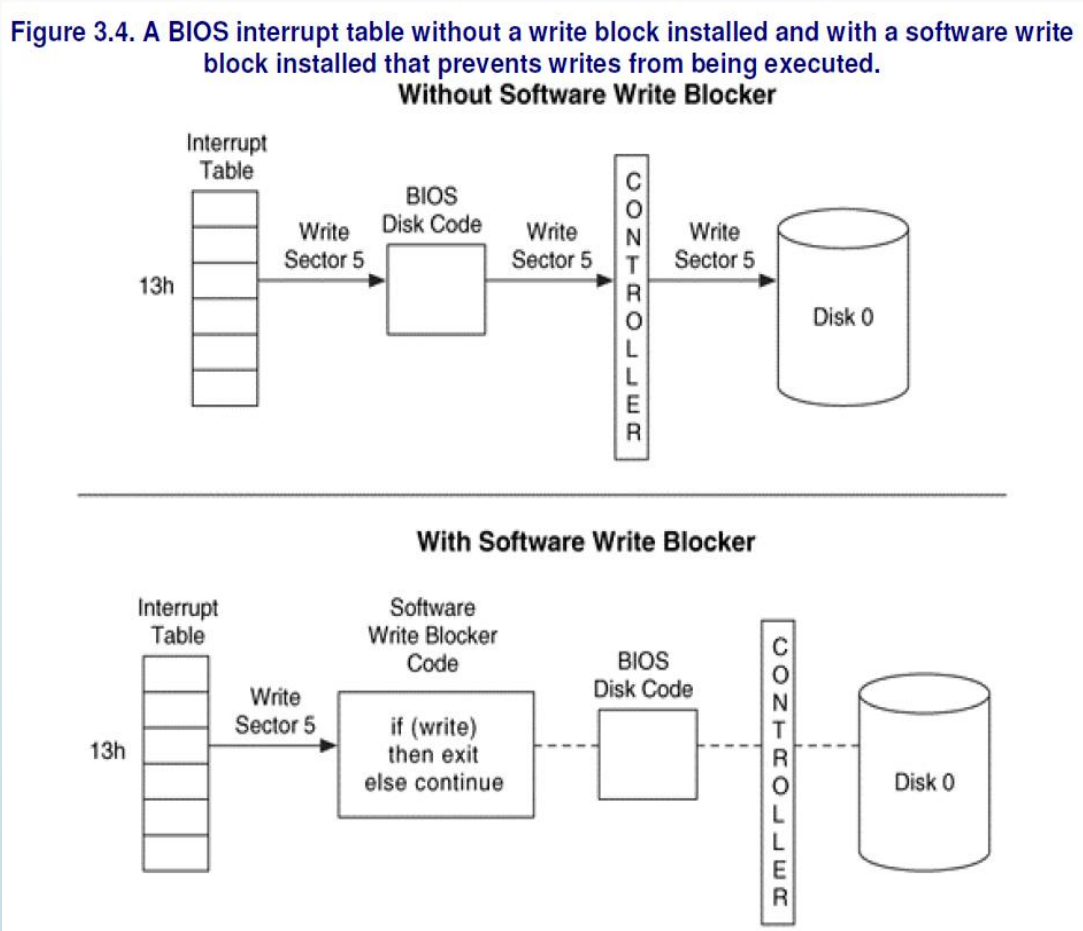
Yazılımsal Yazma Koruyucular

- Donanım yazma engelleyicilerine ek olarak, yazılım yazma engelleyicileri de kullanılmaktadır.
- Yazılım yazma engelleyicileri, belirli bir BIOS hizmetinin kodunu bulmak için kullanılan kesme tablosunu değiştirerek çalışırlar.
- Kesme tablosunda, BIOS'un sağladığı her hizmet için bir girdi bulunur ve her girdi hizmet kodunun bulunabileceği adresi içerir.

Yazılımsal Yazma Koruyucular

- Örneğin, INT13h girişi diskteki veya diskten veri okuyacak veya yazacak kodu gösterecektir.
- Bir yazılım yazma engelleyicisi, kesme tablosunu değiştirir, böylece kesme 0x13 için tablo girişi, BIOS kodu yerine yazma engelleyicisi kodunun adresini içerir.
- İşletim sistemi INT13h'yi çağırdığında, yazma engelleyici kodu çalıştırılır ve hangi işlevin talep edildiğini inceler.
- Bir yazma engelleyicisi, isteği doğrudan orijinal INT13h BIOS koduna geçirerek bir yazmaz işlevin yürütülmesine izin verir.

Figure 3.4. A BIOS interrupt table without a write block installed and with a software write block installed that prevents writes from being executed.



Yazılımsal Yazma koruyucu uygulamaları inceleyiniz.

USB Write Blocker for ALL Windows, SAFE Block, Linux Software write blocker, SWBT

http://www.cftt.nist.gov/software_write_block.htm

Çıkış Verilerinin Yazılması

Hedef Konumu Belirleme

- Verileri kaydettiğimizde bunları doğrudan bir disk veya bir dosya olarak yazabiliriz.
- Analiz yazılımına geçmeden önce inceleme uzmanı şüpheli sistem yada diskleri kendi analiz sistemine takar. Verilerin doğrudan başka bir diske kopyalayarak incelemeye başlar.
- Başka bir deyişle kaynak diskin 0. Sektörü ile hedef diskin 0. Sektörü aynıdır. Ortaya çıkan sonuç disk kopya disk yada klon disk olarak adlandırılır.
- Hedef disk kaynak diskten büyük olduğunda sorunlar çıkabilir. Çünkü kopyanın nerede bittiğini tam olarak söylemek zor olabilir.
- Doğrudan disk üzerinden veri elde edilirken veri edinmeden önce diskin sıfırlarla silinmesi önerilir.

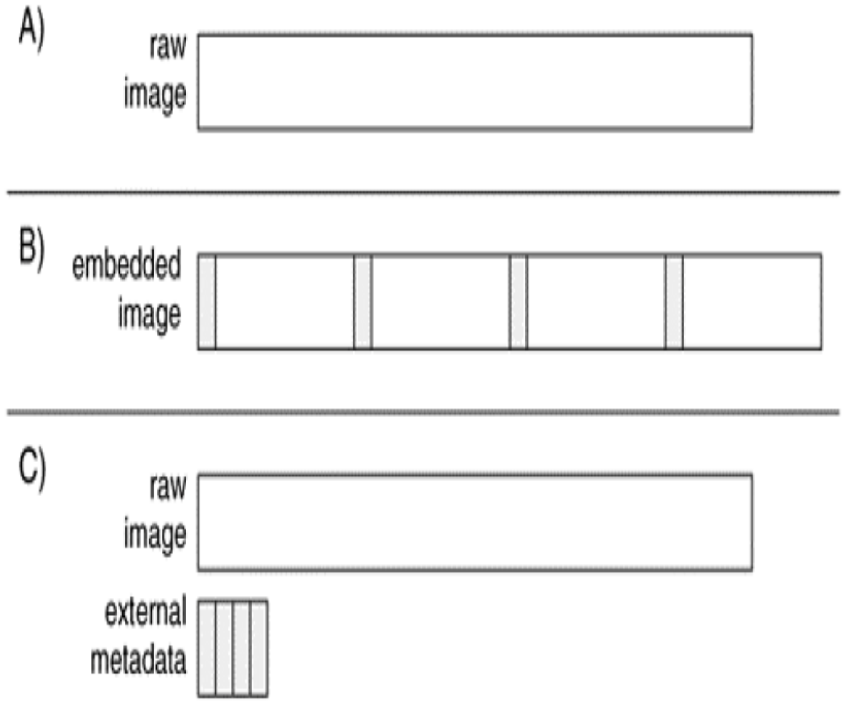
Hedef Konumu Belirleme

- Veri edinme ile ilgili ikinci sorun Windows gibi bazı işletim sistemlerinde diski takmaya çalışırken ya da veri edinme kopyalaması yapılırken veride değişiklikler olabileceğidir.
- Ayrıca orijinal disk ile hedef disklerin disk geometrilerinin farklı olması durumunda zorlukla karşılaşılabilir. Çünkü bazı veri yapıları diskte konumları belirlemek için geometriyi kullanır.

İmaj Dosya Formatları

- Verileri bir dosya olarak kaydedersek, imajın hangi formatta olacağını seçme hakkına sahip oluruz.
- Raw imaj, yalnızca kaynak aygıttan gelen verileri içerir ve imajı kaynak verileri ile karşılaştırmak kolaydır.
- Gömülü bir imaj, kaynak aygıttan alınan verileri ve hash değerleri, tarihleri ve saatler gibi veri edinme ile ilgili ek tanımlayıcı verileri içerir.
- Bazı araçlar ham imaj oluşturur ve ek açıklayıcı verileri ayrı bir dosyaya kaydeder.
- MD5 ve SHA-1 gibi hash değerleri verilerin bütünlüğünü göstermek için kullanılmaktadır.

Figure 3.5. Examples of (A) a raw image, (B) an embedded image with meta data interleaved in the raw data, and (C) an image with the data stored in a raw format and the meta data stored in a second file.



İmaj Dosya Formatları

- Adli kopya alınırken sonuç çıktı dosyasının türünün seçilmesi gerekmektedir.
- Yazılım ve donanımlar farklı çıktı dosya türlerini desteklemektedir. En sık kullanılan dosya türleri DD(RAW), E01, AFF ve SMART formatlarıdır.
- **DD**: Raw imaj türü de denir. Kopyası alınacak donanımdan verilerin bit-bit olarak herhangi bir ekleme çıkarma işlemi yapılmadan kopyanın alınmasıdır. Kopya boyutu ile kopyası alınacak diskin boyutu eşittir. Ayrıca alınan kopyada herhangi bir metadata verisi saklamaz. Dosya uzantısı DD dir.
- **E01**: Expert Witness Format (EWF) olarak Encase tarafından standart kullanılan kopya türüdür. Bu format disk üzerinde parçalı kopyalar alınmasına izin verir. Kopya alınırken veri parçalara bölünür. Bölünen parçalar için bir kontrol değeri hesaplanarak veriye eklenir. Kopya dosya ile birlikte kontrol ve doğrulama bilgileri de elde edilir.
- **AFF**: Advancede File Format olarak bilinmektedir. Kopyası alınacak veri ile veriyi tanımlayan başlık bilgileri birleştirilerek saklanır.
- **SMART**: Linux tabanlı açık kaynak kodlu geliştirilmiş SMART programının elde ettiği kopya türüdür. Kopyası alınacak veri ile birlikte başlık ve doğrulama verilerini de beraber saklamaktadır.

İmaj Dosyalarının Sıkıştırılması

- Verileri bir dosyaya yazdığımızda, dosyayı daha az depolama alanı kaplayacak şekilde sıkıştırma seçeneğine sahip olabiliriz.
- Örneğin, veriler 10.000 ardışık 1 lere sahipse, sıkıştırılmış bir format 10.000 bit yerine birkaç yüz bit olarak tanımlayabilir. Veriler rastgele seçiliyse, çok az tekrarlama olacak ve sıkıştırma etkili olmayacaktır.
- Bir imaj dosyası sıkıştırıldığında, kullandığınız herhangi bir analiz aracı sıkıştırma tipini desteklemelidir.
- İmajı incelemeden önce imajı tekrar açarak asıl dosyaya erişmeniz gerekir.

Sıkıştırmanın Dezavantajları

- Biçimi destekleyen analiz araçları sayısı sınırlı olabilir.
- Yazılımın sıkıştırma işlemini gerçekleştirmesi gerektiği için veri edinme daha uzun sürebilir.
- Analiz aracı, veriyi okuduğu zaman imajın açılması gerekir, bu nedenle analiz daha yavaş olabilir.

Ağ Tabanlı Veri Edinme

- Temel veri edinme kuramında, bir ağ kullanarak bir uzaktaki bilgisayarda bir imaj dosyası oluşturmanıza olanak tanır.
- Bu durumda, veriler kaynak diskten okunur, hedef sunucuya bir ağ üzerinden gönderilir ve bir dosya olarak yazılır.
- Şüpheli diske erişemiyorsanız veya şüpheli disk için doğru bağdaştırıcılara veya arabirime sahip değilseniz, bu veri edinme yöntemi uygundur.
- Birçok mevcut araç, ölü ve canlı sistemlerin ağ tabanlı incelenmesini desteklemektedir. Bazıları ağ üzerinde gizlilik sağlamak için şifreleme sunar.
- Sıkıştırma, iletim hızı yavaş bir ağ üzerinden gönderilen veri miktarını azaltması için yararlı olabilir.

Doğrulama Fonksiyonları

- Dijital delillerin değiştirilmezliğini ve bütünlüğünü garanti etmek için doğrulama fonksiyonları kullanılmaktadır.
- Adli kopya alma yazılım ve donanımları doğrulama fonksiyonları kullanarak dosya, dizin, bölüm ve disk bazlı hash değerleri üretebilmektedir.
- Sık kullanılan hash fonksiyonları MD2 (Message-Diegest), MD4, MD5 ve SHA-0, SHA-1, SHA-256/224, SHA-384 VE SHA-512 serisi algoritmalardır