Bilgi Güvenliği ve Log Yönetimi Sistemlerinin Analizi

Ertuğrul AKBAŞ

Anet Yazılım, İstanbul
e-posta: ertugrul.akbas@anetyazilim.com.tr

1. Log Analizi

Bilgisayar ağlarında kullanılan ağ cihazları olaylar hakkında kayıt yapma özelliğine sahiptirler. Bu kayıtlar sayesinde ağ üzerinde güvenlik olaylarının belirlenmesi ve önlem alınması sağlanmaktadır. Buna Log Analizi denilmektedir.
Log analizi sayesinde sisteme girmeye çalışan kişilerin adres bilgilerine ulaşılmaktadır. Ayrıca sistem içinde bulunan kullanıcıların yaptıkları (dosya kaydetme, yazıcıdan çıktı alama gibi) işler kontrol edilmesi mümkün olmaktadır.
Büyük firmalarda ise internet ortamında kullanıcıların hangi siteye girdikleri, hangi aşamada terk ettikleri, hangi sayfada daha çok / az zaman harcadıkları gibi bilgilere kolaylıkla ulaşmaları sağlanır[1].

2. Log Yönetimi

Log Yönetimi hiç olmadığı kadar önem kazanmıştı. FISMA, HIBAA, SOX, COBIT, ISO 27001 gibi uluslararası standartlar log yönetimini zorunlu kılmaktadır. PCI veri güvenliği standartı da log yönetimini zorunlu kılan standartlara örnek verilebilir. PCI DSS Standardı 6 başlık altında 12 gereksinim ister. Bunlardan biride log yönetimine aittir[2] .Kanunlar ve standartlar tüm yaptırımlardan her zaman daha etkin bir role sahipdir. Ayrıca, 04.05.2007 tarihli 5651 sayılı kanunda internet suçlarını önlemeye yönelik olarak kurumların log yönetimi ile ilgili yükümlülükleri belirlemiştir.
Log sistemleri karakterleri itibari ile dağıtık yapıya sahip sistemlerdir (Şekil 1).
Log Yönetimi: Log Toplama, Log Normalleştirme, Log Indexleme, Korelasyon, Filtreleme/Raporlama ve Alarm yönetimi katmanlarından oluşur.
Bu özelliklere sahip sistemlerin özellikleri:

- Logların merkeze toplanması
- Log Saklama
- Verilere hızlı erişimi ve gösterimi
- Desteklediği Log formatının çokluğu
- Veri analizi
- Kayıtların saklanması
- Arşivleme ve geri getirme
- Verilerin yetkiler ve ilişkiler seviyesinde erişimi
- Veri bütünlüğünün sağlanması
- Loglara erişim auditlerinin tutulması

Sistemlerden pek çok kaynaktan log toplanabilir.

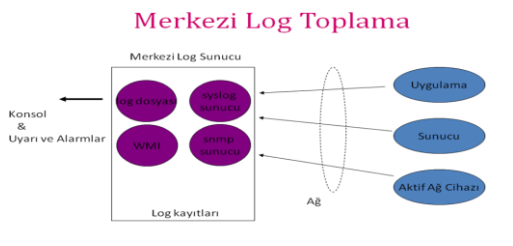
İşletim Sistemleri:
Windows XP/Vista/7, Windows Server 2000/2003/2008/R2, Unix/Linux Türevleri, Nas Cihazları (NetApp) vs..
Uygulamalar:
DHCP, IIS 6/7/7.5 (W3C), Apache (Syslog), Text-Based Log (Csv/Tsv/W3C/Txt/Custom), Dansguardain, Postfix vs..
***Firewall/Proxy*:**
ISA/TMG Server, BlueCoat, 3Com, Astaro, CheckPoint, Cisco Systems, Clavister, CyberGuard, D-Link, Fortinet, FreeBSD, IPCop, Juniper, Drytek, Kerio, Lucent, McAfee-Secure Computing, NetApp, NetFilter, Snort, SonicWALL, Netopia,Network-1, St. Bernard Software, Sun Microsystems, WatchGuard, Zywall, Anchiva, Applied Identity, ARKOON, Aventail, AWStats, Cimcor, DP Firewalls, Electronic Consultants, Global Technologies, Ingate, Inktomi, Lenovo Security Technologies, NetASQ, Websense vs..
***Network Cihazları*:**
Syslog Gönderen Cihazlar, SNMP Trap Gönderen Cihazlar, Cisco router,Cisco switch vs..
***Email*:**
Exchange 2003, Exchange 2007, Exchange 2010, IIS SMTP, SendMail/Qmail ve Bezeri *nix Tabanlı Sistemler vs..
***Veritabanları*:**
Oracle, MSSQL, MySQL, Sybase vs..

Log Yönetimi ile ilgili pek çok açık kaynaklı sistem bulunabilir. Bunların en bilinenleri OSSIM, LASSO, SNARE-AGENT, SPLUNK sayılabilir [1].

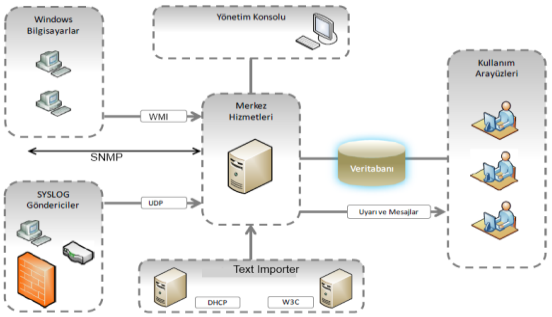
Sistemlerden loglar ya etmen(agent) kurarak ya da log sunucu tarafından uzaktan çekilerek toplanır. Her ikisinin de avantaj ve dezavantajları vardır.

***Ajanlı Yöntem*:**
Avantajları: Log sunucu kapalı da olsa veri Kaybı olmaz,Log sunucu ne zaman toplayacağına karar verebilir,Log sunucu istemcinin durumunu tespit edebilir.
Dezavantajları:Bütün makinelerin önceden konfigurasyona ihtiyacı olması,Sistem ele geçirilirse ajanın log göndermesi engellenebilir.

***Ajansız Yöntem*:**
Avantajları: Kurulum ve konfigürasyonu çok kolay, çok esnek ve çok büyük sistemler için ölçeklenebilir
Dezavantajları: Syslog UDP temelli bir protokol ve veri kaybı olabilir, Log sunucu istemcileri takip edemez.



Şekil 2: Merkezi Log Toplama



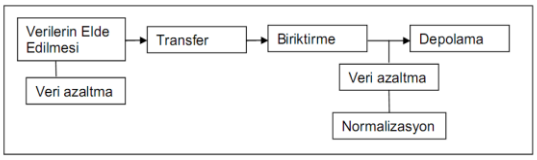
Şekil 1: Dağıtık Log Toplama Mimarisi

3. Log Analizi ve Bilgi Güvenliği

Bilgi sistemleri içerisinde çalışmakta olan tüm ağ ve güvenlik bileşenleri her gün çok sayıda log üretir. Ancak bunlara bir de sunucular ve istemcilerin logları da eklendiğinde hareketlere ve trafiğe ilişkin değerli olabilecek birçok bilginin süzülmesi, başka hareket ve trafik bilgileri ile ilişkilendirilmesi, analiz edilmesi, takibi ve anlamlı sonuçlar verebilecek şekilde raporlanması neredeyse olanaksız bir hale gelmektedir. Bu tür bilgilerin yeterince etkili şekilde değerlendirilemediği durumlarda, kontrol gerçek anlamda sağlanamamakta ve kontrol edilemeyen bilgi sistemleri alt yapılarına yönelik yatırımlar da, verimli kullanılamamış olacaktır[3].

Log yönetiminde kullanılacak araçların çeşitli yönleri ile incelenmesi de bu araştırmanın içerisinde yer alırken, ortamın ihtiyaçlarına göre bunun hangi sistemler için gerçekleştirileceği ve hangi amaçlara hizmet edebileceği, ayrıca bunun sonuçları üzerinde nasıl bir değerlendirme yapılabileceği de araştırma kapsamına girmiş, bu konuda işleyen bir organizmanın çeşitli birimleri ile görüşülerek bilgi toplanmaya çalışılmıştır. Yasal düzenlemelere göre, bilgi sistemlerinde denetim izlerinin bir çoğu, sistemler için tutulan log kayıtlarını işaret eder. Log yönetim çalışmalarında ele alınması gereken ilk konunun, bilgi sistemlerinin hangi süreçlerinde, hangi log verilerinin log yönetimi amacı ile değerlendirilmesi gerektiğine karar verilmesidir.

Log yönetimi, bilgi güvenliği yönetiminin önemli bir bölümü veya bileşeni, bilgi güvenliği yönetimi ise ağın yönetimi ile oldukça yakın ilişkideki bir yönetim sistemidir.(Network güvenliği ve yönetimi, bilgi güvenliğinin sağlanması için gereken sistemlerden yalnızca biridir.) Ağ güvenliği yönetimi için, dikkat edilmesi gereken önemli bir mesele, network üzerindeki atakların saptanabilmesi ve bunları doğru olarak tanımlayabilmektir. Aslında bu durum çok kullanıcılı ve çok çeşitli sistem-ağ yapısına sahip ortamlarda, samanlıkta iğne aramaya benzetilebilir. Bilgi güvenliğini sağlamak üzere ihtiyaç duyulabilecek en önemli veriler güvenlik raporlarıdır. Güvenlik Raporları birer birer sistemlerin kendisinden alınabilecek raporlardan çok, farklı veritabanlarında bulunan veriler kullanılarak; otomatik olarak oluşturuldukları takdirde, güvenlik durumuna ait daha kapsamlı genel bir görüntü sunabilecek ve farklı bakış açılarına ait verileri bir araya getirecektir. Yine bu sonucu elde etmek üzere toplanan olay bilgilerinin, log kayıtları şeklinde depolanarak kullanılması, log yönetimindeki amaçlardan biridir.



4. Log Toplama

Log toplama süreci aşağıda gösterilen bir formül ile ifade edilmiştir. Denklemdesistemler (domain) R ile simgelenmiş, bu sistem içindeki tüm cihazlar D ilegösterilmiştir. Denklem 1 log kayıtları kümesi, 2 bilgi sistemleri cihazlarının farklılog tiplerini, 3 ise her sistem cihazının farklı olay logları kümesine sahip olduğunugöstermektedir.

$$R = \{D1\ D2, ..., Dn\} \quad (1)$$

Her sistem cihazının kendi loglarının olması, B log türleri olmak üzere,

$$Di = \{Bi1, Bi2, ..., Bim\},\ i \in [1,n] \quad (2)$$

Her indeksin bir cihazı temsil etmesi durumunda, her cihazın farklı tip olay kayıtları oluşturması durumu (Örneğin windows sistemlerinde, uygulama (application), sistem (system), güvenlik (security) loglarının ayrı ayrı olması gibi...), e olay tipi olmak üzere; log modellemesini formüle edebiliriz.

$$Bij = \{eij1, eij2, ..., eip\},\ i \in [1,n],\ j \in [1,m] \quad (3)$$

Log Toplama Sitrmlerinde Karşılaşılan Başlıca Problemler

1. Çok yüksek sayılarda ve büyüklüklerde log kayıtları,
2. Log kayıt desenlerinin farklılığı,
3. İçeriklerin oldukça farklı olması

5. Log Normalleştirme

Log kaynakları birbirinden farklı formatlarda log üretirler:

Cisco Router

Jul 20 14:59:32 router 20: *Mar 1 01:39:25: %SYS-5-CONFIG_I : Configured from console by vty0 (10.1.6.160)

Jul 20 15:01:07 router 21: *Mar 1 01:41:00: %SYS-5-CONFIG_I : Configured from console by vty0 (10.1.6.160)

Jul 20 15:10:43 router 22: *Mar 1 01:50:36: %RCMD-4-RSHPORTATTEMPT: Attempted to connect to RSHELL from 10.1.6.165

Jul 20 15:18:06 router 24: *Mar 1 01:57:58: %RCMD-4-RSHPORTATTEMPT: Attempted to connect to RSHELL from 10.1.6.165

Jul 20 15:18:06 router 25: *Mar 1 01:57:59: %RCMD-4-RSHPORTATTEMPT: Attempted to connect to RSHELL from 10.1.6.165

Remote Apache logging

Jul 18 16:49:27 mapmin.tonjol.org httpd[779]: [error] [client 202.56.224.218] File does not exist: /htdocs/default.ida

Jul 18 23:53:28 mapmin.tonjol.org httpd[30023]: [error] [client 202.197.181.203] File does not exist: /htdocs/default.ida

Jul 20 00:39:32 mapmin.tonjol.org httpd[21509]: [error] [client 202.101.159.163] request failed: URI too long
Jul 21 05:06:39 mapmin.tonjol.org httpd[11348]: [error] [client 202.138.123.1] File does not exist: /htdocs/scripts/nsiislog.dll
Jul 21 05:06:39 mapmin.tonjol.org httpd[11348]: [error] [client 202.138.123.1] File does not exist: /htdocs/scripts/nsiislog.dll

Normalizasyon, kaynak verilerdeki bütünlüğü bozmayacak şekilde, Log dosyası verilerinden uygun bir bilgi ortaya çıkarmak için, korelasyon aracının bu yeteneğe sahip olması gereğidir. Bu logları alıp ortak bir platformda ifade edip bütün bu veriler üzerinde indexleme, korelasyon ve filtreleme/raporlama yapılabilmesini sağlama işlemine normalleştirme denir.
Örneğin, bir Windows etki alanı denetleyicisi ile bir Windows veritabanına giriş hatalarını gösteren
"giriş-başarısız" olayını aynı adla kaydedemeyiz. Çünkü Windows Account_Expired ve MSSQLSVR gibi daha spesifik detaylı bilgelere sahiptirler ve normalize edilmelidirler. Normalize işlemi logların parse edilmesiyle başlar. Her bir tür log için özel parserler REGEX[5] yardımı ile oluşturulur.
Burada 2 farklı yöntem vardır.
1-Neyin logunun toplanacağının tanımlandığı sistemler
2-Özellikle Log Proxy kullanılan sistemlerdeki auto log discovery özelliği-Logların geldiği anda tipinin otomatik tanımlanması.
Normalleştirme işlemi sırasında zaman bilgisi ve timezone hesaplaması da gerçekleşir.

Toplanan loglar normalleştirme öncesinde sadece veri durumunda iken parse/normalleştirme ile bilgi durumuna gelmiş olur.

5.1. Zaman Bilgisi

Pek çok farklı sistemden loglar toplanıp merkezi bir noktada toplanacağı için logların kendi içerisinde tutarlı ve korelasyon kurallarının anlamlı veriler üretebilmesi için loglardaki zaman bilgisi çok önemlidir. Bu tutarlılığı sağlamak için ağda NTP (Network Time Protocol) aktif edilmelidir.

5.2. Timezone

Sistemlerden toplanan logların özellikle dağıtık bir altyapıda timezone bilgisi ve bu bilginin merkezi log toplama merkezi ile uyumlulaştırılması logların analizi için gereklidir.Log kaynaklarının hepsinin aynı timezone ayarını alması ve bu bilgisi her durumda göndermsinin sağlanması gerekir

6. Log Depolama

Logların deoplanması ve arşivlenmesi özellikle çok büyük sistemlerde kritik olmaktadır. Günde 100 lerce GB logu saklamak ve üzerinden sorgu yapmak için özel depolama sistemleri tasarlanmaktadır[7,8].

Kategori	Kritik Olmayan Sistemler	Kritik Sistemler	Çok Kritik Sistemler
Log kayıtları hangi süre ile saklanmalıdır	1 veya 2 hafta	En az bir ay	3 ay 1 yıl, *Yasal Düzenlemelere tabi kuruluşlarda en az 1 Yıl
Log rotasyonu hangi sıklıkla yapılmalıdır (Log Sunucusuna)	Opsiyonel (Her hafta bir kez veya 25 MB)	Her 15 ile 60 dakika arasında veya her 2-5 megabyte veride	Her 5 dakikada
Loglar hangi sıklıkta analiz edilmelidir	Haftada 1 kez	Her 24 saatte bir	Her 12 saatte
Logların Şifrelenmesi	Opsiyonel	Evet	Evet

NIST(2007) tarafından yayınlanan Guide to Computer Security Log Management isimli yayından derlenmiştir

Loglanacak sistemlerin ne kadar log üreteceği ve bu loglar için ne kadarlık disk alanı ayrılacağı önemli bir mühendislik hesabı gerektirir.

Örnek Kapasite Hesabı:
Bir yıllık loglama için ortalama disk alanı = 365 gün x 24 saat x 3600 saniye x 100 (yoğun sistemler saniyede çok daha fazla log üretecektir) x 200 byte = 580~ gigabyte / yıl . Tabi bu değer sıkıştırılmamış ham veridir. İyi bir sıkıştırma ile bu değer küçültülebilir. Eğer loglar için veritabanı kullanıyorsanız bu sayıyı iki üç ile çarpmak gerekir. İlişkisel veritabanı kullanımlarında OS seviyesinde sıkıştırma bu alana geri kazanmanızı sağlayabilir.

7. Korelasyan

Log Yönetimi ve SIEM (Security Information & Event Management) birbirini tamamlayıcı katmanlardır. Genellikle Log Yönetiminden bahsedilirken korelasyondan da bahsedilir ve dolayısı ile SIEM katmanına çıkılmış olur.

Sistemlerin korelasyon yeteneği sayesinde farklı cihazlardaki logları otomatik olarak ilişkilendirir, anlamlı hale getirir ve daha önemli uyarılar üreterek, sistem yöneticilerinin ortaya çıkabilecek bir sorunu önceden görmelerini veya ortaya çıkmış bir sorunu daha kolay çözmelerini sağlar.

Korelasyon işlemi belli kuralların işletilmesi şeklinde oluşur. Örnek kurallar

- İstenmeyen mesajları yok etmek
- Tek mesaj eşleştirme ve aksiyon alma
- Mesaj çiftlerini eşleştirme ve aksiyon alma
- Belirlenen bir zaman dilimi içerisinde olay oluşum sayısına bakarak aksiyon alma

Farklı korelasyon teknikleri mevcuttur.Mesela:

- Farklı olayların korelâsyonu (Mantıksal Korelasyon)

- Olay ve güvenlik açıklarının korelâsyonu (Çapraz Korelasyon),

- Olay ve işletim sistemlerinin – hizmetlerin korelâsyonu (Envanter Korelâsyonu)

8. Alarm Yönetimi

Log yönetim sistemleri loglar ile ilgili alarmlar üretebilmekte sistem yöneticilerini mail, sms, snmp gibi yöntemler ile uyarabilmektedir.

9. Kaynakça

[1] Souppaya, M. and K. Kent, 2006. Guide to computer security log management. White Paper, NIST Special Publication 800-92, Computer Security, http://permanent.access.gpo.gov/lps69969/LPS69969.pdf

[2] http://en.wikipedia.org/wiki/Log_management_and_intelligence, 2011.

[3] http://en.wikipedia.org/wiki/Payment_Card_Industry_Data_Security_Standard, 2011.

[4] http://csrc.nist.gov/publications/nistpubs/800-92/SP800-92.pdf, *2006*

[5] Jacob Babbin, Dave Kleiman, et al**, Security Log Management: Identifying Patterns in the Chaos**, Syngress Publishing, 2006.

[6] http://en.wikipedia.org/wiki/Regular_expression

[7] Ariel Rabkin and Randy Katz., Chukwa: A System for Reliable Large-Scale Log Collection. At LISA 2010, the USENIX conference on Large Installation System Administration. San Jose CA, November 2010.

[8] Ranum M., System Logging and Log Analysis, http://www.ranum.com/security/computer_security/archives/logging-notes.pdf