CHALMERS

An Analysis of Security Information and Event Management Systems

The Use of SIEMs for Log Collection, Management and Analysis

Master of Science Thesis in the Programme Secure and Dependable Computer Systems

HENRIK KARLZÉN

Department of Computer Science and Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Göteborg, Sweden, January 2009

An Analysis of Security Information and Event Management Systems – The Use of SIEMs for Log Collection, Management and Analysis

En Analys av Säkerhetsinformations- och säkerhetshändelsehanteringssystem – Användandet av SIEMer för Logginsamling, Hantering och Analys

HENRIK KARLZÉN

Examiner: ROGER JOHANSSON

Department of Computer Science and Engineering
Chalmers University of Technology
SE-412 96 Göteborg
Sweden
Telephone + 46 (0)31-772 1000

Department of Computer Science and Engineering
Göteborg, Sweden December 2008

Abstract

In today's computer network environments huge amounts of security log data are produced. To handle this data and provide an increased level of information security and centralised log management and analysis Security Information and Event Management Systems (or SIEMs) can be used. SIEMs can help organisations that struggle with the various compliance regulations that exist and reduce the risk of intrusions into the network. SIEMs collect and aggregate log data from various devices and applications through software called agents, filter uninteresting data and normalise to a proprietary format, analyse through correlation using contextual information and alert administrators in case of attack. Log data is stored using special security mechanisms in so called write-once-read-many media for compliance reasons. In this paper special attention is also given to security at the log source. An overview of the market is detailed as are suggestions on how to organise the environment around the SIEM and what log data that is worthy of analysis. It is forecasted that compliance will continue to be the most important motivator for procuring SIEMs. The usability and scalability is anticipated to increase as the market continues to grow rapidly and standardisation will become a key factor. More focus will be on incorporating contextual information into the analysis process, especially for identity and access management. Supported types of log sources will increase in number and policy oriented automated response capabilities will be developed.

Sammanfattning

I dagens datornätverksmiljöer produceras enorma mängder säkerhetsloggdata. För att hantera denna data och tillhandahålla en ökad informationssäkerhetsnivå och central logghantering och analys kan så kallade Security Information and Event Management Systems (SIEMer) användas. SIEMer kan hjälpa organisationer som kämpar med att efterfölja de olika regler och lagar som finns och reducera risken för intrång i nätverket. SIEMer samlar in och aggregerar loggdata från olika enheter och program genom mjukvara som kallas agenter, filtrerar ointressant data och normaliserar till ett proprietärt format, analyserar genom korrelering som använder områdesspecifik information och larmar administratörer vid attack. Loggdata sparas med hjälp av speciella säkerhetsmekanismer i så kallade write-once-read-many-medium av regelefterföljdnadsskäl. I den här uppsatsen ges säkerhet i loggkällan särskild uppmärksamhet. En överblick av marknaden presenteras liksom förslag på hur miljön runt SIEMen kan organiseras och vad för loggdata som är värd att analyseras. Det förutsägs att regelefterföljdnad kommer att fortsätta att vara den viktigaste anledningen till att ackvirera SIEMer. Användarvänligheten och skalbarheten förutses öka när marknaden fortsätter att öka snabbt och standardisering kommer att bli en nyckelfaktor. Mer fokus kommer att läggas på att införliva områdesspecifik information i analysprocessen, särskilt för identitets- och tillgångshantering. Antalet typer av stödda loggkällor och policyorienterad automatisk reaktionsförmåga kommer att utvecklas.

Keywords

Security Information and Event Management, Security Information Management, Security Event Management, SIEM, SIM, SEM, logs, log collection, e-discovery, forensics, user monitoring, identity management, policy monitoring, incident management, real-time response, security.

Preface

In many organisations there is comprehensive work being done to make the large amounts of log data from various IT systems more lucid and easier to handle. For instance, in the case of security intrusions (internal or external), it must be possible to analyse the intrusions (and order such analysis according to priority) to determine what happened and to secure the information and possibly forward it to the authorities. It is important to have a central log server where all relevant logs are stored and analysed from the specific demands and criteria of the organisation.

In this paper one outcome of such work, Security Information and Event Management Systems (or SIEMs), which can provide an increased level of information security (administrative, physical and IT security [13]) and centralised log analysis as well as prioritisation, are described in detail. First a short introduction and a small glossary to clear up some terminology confusion are given after which some motivators for why SIEMs are useful (including some notes on cost analysis) are detailed. A more in-depth look at the different parts of a SIEM follows as does some advice on what an organisation procuring such a product should look for and think about, including the deployment stage. Some specific security issues are also given some thought.

It is not enough to simply deploy a product in the security arena, much policy writing and configuration to adhere to said policies must usually be undertaken and this is not an exception. Therefore, focus on what log data that should be collected to the SIEM and reacted on is detailed. A section on what kinds of roles that comes into question with some additional notes on policy which must be defined clearly stating how the SIEM is to be used is also included.

To simplify for interested organisations (and others), an overview of some specific vendor's products follows. Please note that the author of this paper has no reason to be biased towards any product but that he has only tested a select few. Some special attention is then given to complying with data regulations and laws and the retention issues that arise. Finally, some notes on log formats and standardisation and a section where some forecasting is done and conclusions made. There is not much log source system or log format specific technical information. The reason for this is that different systems are highly diverse and the great amount of different systems prohibits them all to be covered.

This Master's thesis paper was written in conjunction with assisting business unit Saab Microwave Systems, of Swedish company Saab AB, creating a logging security solution and thus procuring a SIEM. While the author was given an opportunity for practical experience the organisation received some help on what has been called "arguably among the most complex and highest profile information security projects undertaken today" [1]. This work included researching SIEMs in general and some specific products as well as some testing of the latter. A conscious choice was made to write this paper as independently of this fact as possible, to provide an objective (somewhat "tainted" by selective vendor contact) and full picture of SIEMs, and it is primarily based on research of academic and community papers.

Acknowledgements

I, the author, would like to thank:

- My supervisor at Saab Microwave Systems, Styrbjörn Johansson for giving me a unique opportunity and all his help along the way.
- My examiner at Chalmers University of Technology, Roger Johansson, for helping out at short notice.
- My friends and colleagues of the project at Saab Microwave Systems.
- The research community, without which this paper could never have been written.

Contents

1 Introduction

Security Information and Event Management Systems are systems that provide centralised log handling, by collecting logs (primarily those related to security) from various devices and applications of a network, as well as analysis and storage of these logs. If the system detects an attack it can react through its incident management channels which include alerting personnel and even initiating counter measures. A SIEM can also help an organisation comply with regulations pertaining to data retention and the latter can be helpful in cases of e-discovery (also known has litigation preparation) and forensics. The system can also, to some extent, help with network diagnostics. Other use cases include user and policy monitoring and identity management. For detailed information on how SIEMs work see Section How SIEMs Work. Note finally that there are many different names for SIEMs, a glossary can be found in Appendix B.

1.1 Motivation for SIEMs

In this section the reasons behind the interest in SIEMs will be covered.

According to a survey by the information security institute SANS [7], 23% of their members planned to purchase log management in the coming year (April 2007-2008). Additionally forensics tools were planned to be procured by 14 %, identity and access management (IAM) by 13% and compliance automation by 12 %. A SIEM normally provides all of this (although only a part of IAM) and thus there should be considerable interest in them [8]. This is indeed indicated by Information Security's survey Priorities 2008, in which 17% of respondents had implemented SIEMs while 28% were planning to evaluate or implement them [9]. So no doubt there is interest in SIEMs and many organisations are getting one but what are the motivators? That is the topic of this section and some different motivators will be given and at the end of the section also some information on market value and a discussion of cost in security (with emphasis on SIEMs).

1.1.1 Compliance

According to a survey by Ernst & Young, complying with data audit regulations and laws is the most important factor for pushing information security forward. It is certainly the most important reason for the increase in the SIEM market of the last few years. In fact, compliance is the main reason for 80% of all SIEM projects [2]. Certainly it is no surprise that organisations do not want to risk being non-compliant with penalties ranging from lost business opportunities (e.g. credit card transaction privileges revoked) to criminal prosecution (for instance when not complying with SOX – see Compliance). While many organisations no doubt use the additional benefits of a SIEM, and not just the compliance automation, Ernst & Young recognise that there is a risk of compliance receiving too much attention while security is neglected to be included in the overall business strategy [10].

1.1.2 Insider threats

Another very important motivator for SIEMs is to stop insider threats. With increased user monitoring and improved IAM everyone from entry level to super administrator to executive can be traced. Especially administrators with their high privileges and possibly consultants who are outsiders inside the organisation constitute a possible danger to the organisation. With organisations growing in size also layoffs will most likely get bigger and thus there will be more people "out for revenge" and a higher risk for insider attacks. SIEMs can lower the risk of these insider threats. With the forensics capabilities of SIEMs the cost of after the fact determining what happened and what was damaged (this is often a major contributor to total incident cost) can be lowered considerably.

1.1.3 Incidents are costly

One, more general, reason for security and specifically SIEMs is that security incidents are normally very costly for the organisation. Not only can information leaks mean billions of dollars in losses and many man-months of investigations and repairs, the organisation's popular conception (reputation) can be substantially damaged. Of course if computer systems have been taken down by the attacker this leads to further cost since normal operations cannot continue. With real-time security alarms SIEMs can stop attacks earlier and the improved incident handling capabilities and forensics will make it easier to determine the cause and to repair systems. Additionally SIEMs store logs in a structured way that ensures the integrity of the logs which means SIEMs contribute to litigation preparedness and e-discovery.

1.1.4 Complex problem needs multi-faceted solution

The amount and complexity of devices and products increase in networks leading to more complex network environments and more information (and logs) [11]. The nature of threats is similarly becoming more complex and blended and attacks use multiple vectors [12]. Traditionally new types of threats have been met with new types of tools (such as antivirus software against viruses and spyware leading to anti-spyware tools emerging) and there are now a vast number of different tools requiring some kind of centralised administration. One example solution that provides such administration is of course SIEMs which also provide means to handle the increased information on the network and the diversity mentioned in the beginning of the paragraph [5]. Another centralised solution is that of special application layer firewalls called UTMs which are used to administer (and in fact replace) existing security like antivirus, spam filter, IDSs and web filtering [12].

While the network and security pictures are getting more complex security and network operations are also becoming more tightly intertwined, as seen for instance by the fact that IT companies have been spending more on security and have been acquiring security companies to a larger extent [9]. One such example is data storage manufacturer EMC which acquired security vendor RSA and SIEM vendor Network Intelligence for $175M (now subsidiary RSA is responsible for the company's SIEM – see more in Products) [16]. Another example is networking giant Novell which acquired one of the first SIEM-vendors [17] e-Security in 2006 [18].

1.1.5 Hard to measure cost and benefit

Apparently there are many diverse reasons for an organisation to buying a SIEM. However it is difficult to put numbers on the different reasons and to estimate the "cost of not buying the SIEM" and in fact the cost of buying it. It is a general problem in security to determine return on investment (ROI) [19] and there has been some debate on whether the term ROI is even applicable and what should be used instead (like loss prevention) [22, 63]. Even with some proper methods (these are hard to find – see below) for determining at least some of the cost most organisations will not even bother with financial calculations as reported by a big logging survey by SANS [7]

When considering a new security measure it is of course important to analyse if the measure will in fact save the organisation money (or more specifically reduce risk as defined by probability of an event multiplied by the impact (severity) of the event [21]). It is interesting to note that while security measures can mean a reduced availability (while availability is a part of security it can be reduced while the confidentiality and integrity are improved) for users (due to restrictions) they can also lead to improvements due to the security now being good enough (to allowing e.g. remote access). However, since SIEMs are not only about security but also compliance the risk as a whole (e.g. protection against lawsuits) may decrease but security may not necessarily be improved. Generally security risk management is a hard task since data is lacking and the long-term effects are especially hard to forecast. An organisation can not know how effective the measure will turn out to be (will they be hacked anyway?) and they also don't know what the risk of not implementing the measure is (how often will they be hacked without it?).The organisation must close up all security holes while an attacker only needs to find one still open. It seems reasonable to assume that the more holes that are closed the less likely the attacker is of finding a way in but whether this risk reduction is worth the investment in closing the holes (such as buying a SIEM) is not easy to determine [22].

1.1.6 Market value and cost

There are some figures of course on cost of SIEM (and on doing without it). A high ranking IT manager/CIO in the banking industry, estimated in 2002 that real-time monitoring of logs (without SIEM assistance) of 30 devices would require nine full time employees making it "cost-prohibitive" at $110 000 (or twice the cost of a netForensics SIEM at the time). Even with a SIEM one or two full time employees devoted to log management and related aspects may be necessary resulting in yearly SIEM spending of $45-175 000 (with an initial cost of $100-500 000) [23].

With all these motivators for SIEM and such high costs it is interesting to see what the how the market is doing and what it is worth [24]. gives such figures, estimating a market value of $15M (in 2002), forecasting the market to be worth four times that in 2005. However this was a big underestimation of the 2005 market value which ended up at $284M according to [25] which in turn forecasts a tripled market value in 2010. Note the decreased growth rate forecast which indicates that the market is reaching saturation (while the high figures makes it quire mature). However [26] from April this year (2008) provides an estimated growth rate of $400M/year which is even better than the rate given in 2002 [2]. estimated the market value to be at $800M

already (2007). For a further discussion of the outlook of SIEMs - see A Future Outlook in Conclusions.

1.2 How SIEMs Work

It is important of course to understand how the SIEMs work. There are naturally differences between different SIEMs of different vendors but there are some general parts/concepts that remain the same. The fundamental parts of a SIEM are described by collection, analysis/aggregation and retention (or CAR for short). Log data is collected from the various sources and, since there are so many different formats of the data, it is first aggregated (from the various devices) and then usually normalised into a proprietary format. This process is known as consolidation. The data is then analysed by aggregating data from the different devices and correlated by putting together different parts of an attack into a complete picture. In this stage contextual information about the network environment and common threats is very useful. Alerts and reports are generated as an output of the analysis. Log data is usually stored online on the SIEM for a few hours at the very least after which it is moved to an archive for instance for complying with regulations (see Compliance) and to save data that might be important for forensics/e-discovery. This functionality can also be expressed as the "five Cs" (Gartner) [27]:

- Collection
- Consolidation
- Correlation
- Communication
- Control

Each of these important parts of a SIEM will be given in more detail below. An abstract view is provided by Figure 1.

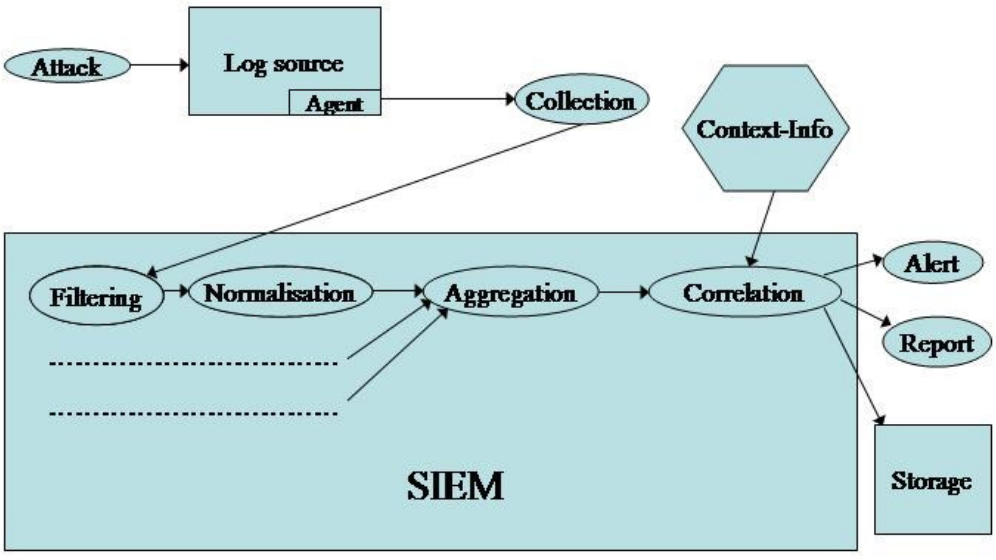

Figure 1 – An overview of a SIEM.

1.2.1 Collection

SIEMs collect log data from a large amount (typically dozens or hundreds) of different kinds of devices (this is device density – see some specifics in Section Products). Transportation from log source to SIEM might need to be confidential, authenticated (to protect against false logs) and reliable. Different protocols for collection include syslog (both reliable and encrypted and non-secure versions exist), SNMP, SFTP, IDXP and OPSEC (a Checkpoint standard [28]). See also Section Log Formats and Normalisation below. For log sources that do no support these, or related standard collection protocols, a so called agent must be used. This is a piece of software installed on the log source that translates (normalises) source log data to a format the SIEM understands. The use of agents usually means longer deployment times for the product although they could also be used for normalisation and other pre processing to distribute the work [28]. Most products need to use agents for Windows Event Logs and one of the most commonly used agents is Snare (formerly known as Backlog) by Intersect Alliance which translates to syslog and is used even in SIEM products by Snare's competitors [28].

There are two different approaches to collection. Either the SIEM decides when collection should happen or the source system. The former, called pull, means that the SIEM logs in to the source device or agent and pulls the logs while the latter, called push, has the source/agent transfer the logs to the SIEM at its convenience. Of course "pull" means the SIEM is more in control (and indeed has some sort of access to each log device!) and thus more important. Of course, the centralised approach means more processing power is necessary in the SIEM. When log data must be collected depends on whether real time analysis and alerting is required. If not, then the best idea may be to transmit only when the network is not overly busy [33].

Depending on what the SIEM is to be used for it may not be essential to collect all the log data. Instead some "uninteresting" data can be filtered upon collection (or indeed not be generated at all) to reduce demands on network bandwidth, storage and SIEM processing power. This means lower cost (in fact the two most important price indicators are number of devices to collect from and number of events per second or EPS) and less risk of denial of service attacks. Of course it is not always obvious at that point what constitutes interesting and uninteresting data since the latter can suddenly seem very interesting after the processes of aggregation and correlation [29]. If the SIEM is used to adhere to certain regulations or laws it may not be an option to not collect all data and even if no present regulation demands it a future one might so it may be a good idea to be "future proof" and collect everything (design for change).

There are typically some restrictions on log sources that the SIEM will communicate with. To prevent attackers from adding false log entries (e.g. an attacker adding a device that generates logs the SIEM will collect), which could be problematic for several reasons (e.g. input faulty fiscal figures etc), some SIEMs prevents logging from devices that are attached without a special procedure. Of course this can be a problem in the case of "benevolent" removable devices such as USB disks and DHCP devices (which receive IP addresses dynamically and thus seem to be different devices for SIEMs that assign identity on an IP-basis) that are often disconnected and reconnected.

1.2.2 Consolidation or Normalisation and Aggregation

With many types of log formats most SIEMs normalise data to a proprietary format. For inter-operability issues ideally this is a standard format (see Log Formats) but normally this is not the case. Usually the diversity of source formats mean that a lot of normalised data will be included under a general heading like one labelled something like “Miscellaneous” or “String” which may not be very helpful for search purposes. It is important to note that for compliance (see Section Compliance) it may be important to store the data in its original form. While it is important to normalise, among other things, the date and time of the log data, it is also vital to make sure that the time is correct in relation to a common clock. Usually temporal normalization like this uses a method like the one provided by the Network Time Protocol or NTP (see more under Log Events).

After normalising, the process of aggregation starts. Its purpose is to put together different events that are of the same type to help “seeing” log data as less separate and more related. This is similar to correlation which however puts together different pieces (different types of events) of one attack. Since the processes are quite alike, there is some confusion on the correct terminology in the literature. The definitions used here are as in [24] and [27] but not as in [28]. The term “collation” is not used here but in [5]. The importance of aggregation is best illustrated by an example:

There are several log messages that can be indicatory of intrusion attempts and/or other security issues. For example it is a good idea to monitor logins that take place very early in the morning or late in the evening or on weekends when there should normally be no one logged on. Also, an aggregate of messages can show things that separate ones cannot. For instance, a malicious attacker trying to get in to a specific user's account might try and guess the password. Since failed logons are normal (misspelling one's password is common since good passwords are likely to contain many special characters and mixed characters case) log messages are not enough separately. Instead many failed logon attempts in rapid succession is more likely to be intrusion attempts. Of course skilled attackers know all this and will try and spread out their attempts enough to fool the intrusion detection mechanisms in place. As always there is a constant battle and one must keep in mind that security is a process and not a product.

1.2.3 Correlation and Contextual Information

In the process of correlation different log events (of different types) are put together to form an attack or incident (sometimes the word entry is used for an uncorrelated piece of log data while event is reserved for correlated data). This process is quite advanced and processing intensive since it must understand what an attack is (and is not) which is usually accomplished by downloading threat information from online databases and using contextual information (a knowledge base) to understand the specific network environment to a smaller or larger extent. This information can include user directories and device priority and location (logical or even physical). This can potentially be used to map devices to mediate the problem of DHCP as described above [28]. Ideally the SIEM can also learn from the events it receives data on and update the contextual info (like “this server seems to get attacked a lot – I will

remember that"). Of course this means more computational overhead. Ideally context-info should be incorporated into the SIEM and updated automatically when for instance a new vulnerability scanning has been performed. Context-info is a hot topic in SIEMs – see more in A Future Outlook.

There are several ways to judge whether something constitutes an attack or not and how severe it is. Two approaches commonly used in IDSs are anomaly based detection (ABD) of an attack and misuse (or signature) based detection (MBD). The former reacts on anything that is not specified as "good" behaviour while the latter reacts when something known to be bad occurs. ABD obviously requires a lot of policy writing and so it will most likely react on some good behaviour (because that kind of behaviour was forgotten when writing the policy). However it can detect abuse of privileges, deviations from normal usage patterns and insider attacks (insiders have more access and knowledge and thus may not require well known "signature" vulnerabilities in their attack) and is really the hardcore way of security (also known as default deny). With the slightly more relaxed approach, MBD (which maintains a signature list updated e.g. online like above), some bad stuff will instead be missed [33].

1.2.4 Communication or Alerting/Reporting

There are three ways that administrators are informed by the SIEM that attacks or strange behaviour are taking/have taken place. Either the SIEM sends out an alert as soon as it realises that something is wrong, or sends a report at a pre determined time. The third option is of course that the administrator is actively monitoring the SIEM in real time (normally through a web based GUI supporting different user account roles). It may seem strange that alerts of attacks (i.e. intrusions) are not left to IDSs but a SIEM is quite obviously more intelligent, giving a lower false positive rate, and may catch things an IDS does not (and since logs are generated by other types of devices too the SIEM will also catch thing an IDS simply cannot). Of course the alerts should contain necessary information for further action or they risk being simply ignored.

Some SIEM products can even take protective (reactive) actions (like deleting malware, closing ports) through the connected devices (compare IPS) [5]. Of course this requires the SIEM to be in control of the devices and not just a passive listening device which means it is an even more important node in the architecture (see Security Issues). This also means it is not just a security product but also network management. According to a paper from the year 2007 ([1]) this kind of remediation is not yet very well and comprehensively implemented in SIEMs.

Reports are usually scheduled to be generated regularly but can also be made on the fly. There are normally hundreds of template reports for standard needs which speed up deployment of the SIEM greatly. A typical report details the login activity of the past night (when there should likely not be a lot of activity) and most SIEM products provide visualisations of common statistics. These kinds of reports are ideal for quick overview for an administrator and perfect for management (CEO etc).

According to some researchers (most notably [29]) SIEMs are not very useful if not constantly monitored, although with evolving products (they have more use cases than simply real time event monitoring) these opinions may change. While Schneier is

right that we still have problems with e.g. viruses despite anti-virus software we would certainly have many more without it and the same is true for SIEMs. Since real-time monitoring "requires five fulltime employees" or even more it would most likely not be cost-beneficial to use SIEMs according to Schneier. He also means that outsourcing the SIEM monitoring would be a wise move but this is doubtful since the context info (like company-specific knowledge) would be sorely missed [28]. He further calls for simpler tools, a call which seems to have been heard by the vendors [2].

1.2.5 Control or Storage

While being analysed, data is normally stored online and when no longer readily availably needed it is usually archived. The data can be stored normalized (and aggregated) to speed things up when (if) used again but it must be stored in more or less original form (raw data) if to be used as legal evidence or compliance. Usually the data is stored in compressed and possibly encrypted form but since this does mean slightly less security (encryption keys can be lost, compression corrupting) it may not be. Usually storage of SIEM appliances is in the TB range enabling storage of up to thousands of millions of events according to netForensics [23], a SIEM vendor. One could opt for keeping more interesting data (what that is depends of course) online longer than less interesting data [28]. Since huge amounts of data are usually involved it is very important to use really good indexing and management for the online storage. See more under Section Retention Time, Site.

2 Analysis and Results

There are many aspects relating to SIEMs that are of interest. Some analysis of the more important ones are given in 2.1-2.10 below.

2.1 Procurement of a SIEM

As with any costly product there must be a decision making process, with plans, policies, business case and exit strategy being constructed, before deciding on what specific product to buy (if any). Since SIEMs are quite complex and have dependencies on a lot of pre existing products in the network environment they require extensive analysis prior to purchase. Some factors and processes that come into question are described in this chapter. The bigger the network environment and the more agents (software) needed the more complex the deployment and the impact on things like bandwidth and productivity must be considered. Other factors impacting the deployment time are detailed in Section Procurement. Also to make the most of the product personnel may have to receive training. Of course, if the analysis proves (indicates) that buying a SIEM is not cost-beneficial then it should not be purchased and perhaps a log management tool is enough. Of course, making the product in-house is potentially cheaper but also very difficult. In the selection phase several different people are involved, including CIO, CTO, security officers,

management, system owners etc, further complicating the process (see more in Section Roles and Policy). Since SIEMs are quite new it may be a good idea to read up on procurement for related security products like IDSs.

[37] gives three types of factors that come into consideration when procuring a SIEM: organisational ones (policies, roles and operation), those relating to the product (functional like correlation as well as non-functional requirements: interoperability, usability, security) and to the vendor (reputation, contracts on other products – will Microsoft stop working with you if you buy Apple?).

It is also important to consider the aspects of having a vendor-homogenous security environment versus a vendor-heterogeneous one. Buying a SIEM from a previously used vendor likely means higher compatibility with existing devices. However vendors usually use similar security measures in all of their products meaning vulnerability in one can be the same as vulnerability in many. Basically hackers need to know only one "vendor language". Also, most vendors are probably really good at only one specific part of the overall security sphere [1]

Before even considering what product is best it is usually a good idea to organise the current environment. Implementing a proper system of classifying data according to sensitivity is one thing to help the structuring. An asset inventory is important, to make sure that it is clear what exists and thus what needs to exist. Simplifying the network environment in various ways reduces of course the complexity also of the SIEM deployment and may well lead to a lower EPS rate or device amount meaning lower cost. This process also helps understanding what is currently logged by.

2.1.1 Network Environment Fit

Of course, the product must fit the buying organisation (the latter should be include in the target market – see Products) and the buying organisation's network environment. The focus of the SIEM should be the same as the focus of the organisation (one or more of log collection, e-discovery, forensics, user monitoring/identity management, policy monitoring, incident management and/or real-time response) while also having functionality for (possibly through upgrade) other of said aspects. Device density should be appropriate (more importantly the device types in question should be supported) as should the supported collection/search EPS rate (usually these parameters are only a matter of cost). Future developments (like increase of size, more types of devices) of the organisation should be supported by the product. The SIEM must of course reduce risk to the acceptable level cost-effectively (that is, the cost of the product should be less than the benefit of loss prevention and the return on investment). The lifetime of the product (end-of-support) should be at least ten years for most organisations.

Depending on the specific demands of the environment of the buying organisation there are other requirements that come into question: The SIEM should be able to handle the information security (sensitivity) classification levels of the organisation. Virtual environment logs must be supported if the organisation uses such environments (these kinds of systems may pose special problems [14]). The organisation must decide if all data must be collected by the SIEM without filtering or normalisation or if the two are actually beneficial for improved performance. In

general, a vendor should not be chosen if such a selection interferes with future procurements. Additionally, a vendor the organisation has previously (successfully) used is a plus since this not only makes communication simpler but provides a vendor-homogenous security environment meaning better compatibility although there is an increased risk of a single-point-of-failure. A vendor specialised in SIEM is a plus.

Another factor that needs consideration is the collection architecture. The organisation should consider if the SIEM should be as agent less as possible or if installing agents is not a problem. Also, an appliance leads to further shorter deployment times and may be considered. Some other factors pertain to performance and shortened deployment times. Appropriate log storage (external from appliance if used should be supported) should be offered by the SIEM and bandwidth requirements should not be too high. Pre-packaged content in the form of template reports (especially for compliance) and alerts is usually quite important for quick deploy times. It is usually important that the reliability of the product should be high and no "log gaps" should persist. If a device gets offline and the back online the SIEM should instead get up-to-date. The SIEM should not require, for full benefit, too many devoted fulltime employees when deployed and personnel training should then be offered. Outsourcing of administration should be possible in a secure manner for organisations interested in that.

Of course the usual non-functional requirements like those relating to interoperability, manageability, usability, scalability and security (like role-based access to the SIEM) must be met. Some kind of testing of the product should naturally be possible prior to purchase. Certifications (e.g. Common Criteria, [39]) are good since they provide "free testing". For information on some certifications of specific products, refer to Section Products. Customer cases/references are also usually very helpful. Of course a product for which many flaws have been found should not be considered. Instead, the vendor reputation should be good, having manufactured other good security products a plus. In fact, the product roadmap of the vendor should be compatible with the roadmap of the organisation. Since the organisation may in the future want to interoperate the SIEM with another SIEM, or even replace it, it should be possible to transfer log data to a third party system.

Some final issues are described here. The level of support (including for installation) offered should be sufficient and cost-effective and deploying the product may introduce new security issues. These must be detailed by the organisation and remediated before purchase [37]. This includes effects on physical security and personnel privacy.

2.1.2 Common Mistakes

There are some issues that span across several SIEM topics, e.g. not only procurement but development and roles and policy, some of which are described here. One well known SIEM expert (currently working at LogLogic) by the name of Andrei Chuvakin, has compiled a list [11] of the most common mistakes when procuring/deploying a SIEM. Chuvakin is like Schneier (see Collection in Section How SIEMs Work) against filtering/prioritisation before/at collection. In fact he believes that no filtering at all should be performed in some cases (even in the SIEM).

Another opinion the two researchers to some degree share is the proactive monitoring of the SIEM. Deployment should therefore include role assignment of such responsibility. It is also important to make sure everything that should be logged (according to policy) *is* and that focus is not too much on device logs so that application logs are not ignored according to Chuvakin. The final mistake to avoid [11] is not to dispose of logs too early but make sure that there is proper internal or external available storage space (even if an appliance is chosen) especially since insider attacks are normally discovered later than other attacks leading to longer retention time requirements.

2.2 Deployment

There are many things that must be considered before choosing what system to buy as described in the previous Section on Procurement. Depending on how rigorous and thorough the former process is the deployment time can be a matter of hours or days up to months or in some rather extreme cases even years [1]. Some of the most important factors that improve deployment time will be described here. An agent less SIEM speeds up things as less installations are necessary as does an appliance (few true appliances exist though acc to [1]). For helping with the initial configurations installation wizard are useful. Of course the more vendor support (commonly adding a substantial cost) that is procured the smoother things normally go. Learning how the system works before hand is a good way to hasten things and vendor demonstrations, pilots as well as customer case studies can be almost invaluable. Of course deployment time in general depends on things like how well structured the network is, how many personnel that are involved and how much they invest in the project as well as the general usability (including documentation) and configurability of the product. After the system has been deployed there is also a need for the SIEM to be tuned and, for those that implement some kind of automatic learning, the time until completely tuned may be in the range of years according to one vendor.

There are many different people involved in the deployment of a SIEM. For full details, see Section Roles and Policy but some aspects will be briefly discussed here. For instance, management must be convinced that the SIEM will provide or they will not allow a very broad and deep installation. Perhaps more importantly, the personnel responsible for the systems with the log sources that will be connected to the SIEM will have to be “onboard”. Even if management demands that they cooperate, things will run much smoother if they willingly help. One way to give something back to these administrators/mangers is to provide them with some of the generated reports of the SIEM so they too can see and take advantage of the new system.

Finally, while the deployment is usually over within months, the SIEM must be managed throughout its existence. While most products no longer need extensive database tuning and maintenance they are not all scalable enough to allow for future procured products to be integrated into the current solution. To shorten the time for future deployments it is therefore important to plan far ahead already when procuring and deploying the SIEM.

2.3 Security Issues

SIEMs introduce many different security issues, some of which have been described in other Sections of this paper. Since SIEMs are such an important (and in a way weak) point SIEM access must be restricted in a proper way by using authentication and encrypted control data traffic. Also, the SIEM application/appliance/OS must be hardened [5]. There are also issues about the data as detailed below.

Log data has the same basic data security issues as most sensitive data. The confidentiality must be ensured as must the integrity and availability (also denoted by the result of a successful attack on each – (inadvertent) disclosure, deception, disruption or the combination of all three when an attacker covertly gets control of the system: usurpation [42]). Normally security problems such as transit manipulation of messages are mitigated by the use of encryption and authentication. Of course internal attacks with authorised users acting maliciously is harder to defend against and instead other types of secure logging and consequent auditing must be used. Also it is important to trust the author of the solution used (like the manufacturer of the SIEM) and that one's specification of security is correct. Additionally, there must be security mechanisms in place to deal with the situation that the above security measures fail.

2.3.1 Security of logs at source

It is important of course to secure logs not only when archived (see Retention Time, Site) but also when collected (before archiving) and before collection, in each device. One allegedly fast approach (the authors concentrate on smaller devices than normally used as log source devices in a network environment so speed may be more of an issue) to ensuring the confidentiality and detection of integrity-loss of logs on source devices (i.e. *before collection*) was described by Schneier and Kelsey in [43]. Using this approach logs that have been created and still reside on the machine cannot be read even by an administrator or hacker with administrator privileges but only those with special central log data privileges (or those with the appropriate encryption keys anyway). The hacker can still selectively delete logs but not without detection when the logs are later inspected. For example if the hacker deletes log entry 5 but lets 1-4 and 6-7 remain in place this will be detected on inspection.

In fact even if entries 6-7 are also deleted it may still be possible to know that logs have been deleted (this can perhaps also be suspected if logging timestamps indicate no logs have been generated for an inordinate amount of time which incidentally estimates the time of the break-in). This is because it is always possible to tell for someone with log privileges what number each log entry has since this is implicitly contained in the encryption key used for the particular entry. So, if the hacker deletes entry 6-7 and then lets logging continue (either as normal or maliciously generating false logs) then it will be obvious that the log entry after number 5 is 8 and not 6 and auditors will know something is wrong upon inspection.

Of course, when a hacker has broken in to the log source device he/she can create new (possibly maliciously false) arbitrary logs. If there was no indication in the log of the intrusion and the hacker does not delete any log entries but only add, the break-in may

go unnoticed. Of course this “primary limitation” (Schneier et. al) is not due to a lacking logging system but rather due to lacking intrusion detection mechanisms.

In more detail (although far from all is presented here, see the original paper for the full details): the log authority and log source device start with a common encryption key used to encrypt the first log entry. Then the source device sets its next encryption key to the hash (using an appropriate function) of the current key after which the latter is disposed of in a secure manner. The next log entry is encrypted using the new key which is then hashed and disposed of, in a process that can continue almost indefinitely. This provides the implicit knowledge of log entry number as described above since the log authority can simply use a method of hash-and-error i.e. trying to decrypt with the first key and if unsuccessful than hash the key and try with the new, upon inspection. Incidentally this also gives the rather restrained option of assigning different users different log data privileges. For instance one user can be given the key gained after hashing 10 times meaning he/she can read any log entry with a number higher than 10, so this could be described as temporally-differing privileges. Of course other, more flexible, log access control mechanisms can also be used. This also shows that each key is a major weak point in the system and the secure disposal mechanism must be given much thought and the key in use kept confidential as must the initial key at the log authority.

Of course logs are normally continually sent to the SIEM from the log source but if this communication is cut off or is too slow the method described is good to have. Additionally, logs created after last transmission will be “secured” with the proposed method. Basically the method provides a separation between the log system at the logging part of the source device and the rest of it resulting in increased in-depth security. It would also be possible to use WORM as log storage even at the log source (and not only for central archiving) but this would probably be too costly for most.

2.4 Log Sources

In this section the different log sources that SIEMs will collect log data from will be given. Note that log source means the device or application generating logs that are later delivered to the SIEM. It does not mean the source (reason) of the log data itself (i.e. a hacker for instance or even a primary log source if there are secondary log sources before the SIEM).

There are several different types of logs that are of interest. These include security event logs and administrator logs as well as access logs. There is also a large amount of different kinds of devices and applications producing logs on most network environments in most organisations. For instance Windows generates several different logs one of which is the Security Event Log which is usually the most interesting Windows log from a security perspective. On Linux machines syslog (which however lacks security but is too common to get around according to [44]) logs are usually used. Of course, security measures like firewalls, antivirus and anti-spyware software and intrusion detection/prevention systems all produce logs and these too are highly important for security purposes. Even logs from routers and switches might be interesting and mobile devices such as mobile phones and laptops are not to be

forgotten. Removable media (includes in a way also printers – see Log Events) is also of essence to track for a complete picture of the security situation. While it is usually noted in normal system logs when such media is un-/mounted it is less clear how to monitor if (when) they are taken off the organisation premises. Again, logs should be used to monitor activity and when and by whom the media is taken away. To ensure that information is still secure (confidential) when moved to removable devices special security products can be used. Logging for removable/dynamically named devices is however problematic as described in Section How SIEMs Work.

As mentioned in the Section How SIEMs Work, Windows Event logs do not support standard transport methods. This is due to the fact that they are stored in binary instead of text (as Linux logs which use syslog transfer). To further complicate things, the logs of a Windows domain are distributed with each separate log located in the respective host computer registry. One solution is to use the native viewer for the Event log, the Event Viewer, but this is very basic and contains no real search capabilities [45]. Instead agents are usually used for Windows log collection as described in the mentioned Section of this paper.

Logs from IDSs and IPSs contain, of course, information on attacks while firewalls and routers can give information on for instance outgoing connections to the Internet (which can reveal if computers on the internal network have been compromised and made into so called zombies - doing further work for the attacker). Using antivirus tool logs things like outdated virus signatures and software can be revealed. It is also important to remember logs from authentication servers like Windows Kerberos.

In short, all systems handling confidential information and/or use access control or come into contact with any of the previous must use logging, according to [46]. Of course the level of logging should depend on the level of security required as described below in Section Log Events. For some organisations even physical security logs will be sent to the SIEM.

2.5 Log Events

There are some details that should almost always be logged (here meaning logged in the source and then collected by the SIEM) and analysed from each device/application. These include security alerts of IDSs/IPSs, logon/logoffs (both failed and successful ones and especially remote ones and ones for inactivated user names) and start/shut down of the machine. Also, due to integrity reasons, any changes to policies and use of/changes to important products should be logged. This includes generation and collection of log messages if logging is shut down or logs (successfully/unsuccessfully) modified, privileges are escalated or modified (or unsuccessfully attempted to be) or sensitive information declassified. User accounts that are inactive for long periods of time should also be logged (providing for instance the employer the possibility of determining if an employee does not work although this has privacy implications) and attempts to logon after an inactive period should be logged under a special heading. Of course log data fields containing strange (e.g. out-of-bounds) values or e.g. factory user names should be focused on as well [52].

Depending on the required security level there are other things that may be necessary to log. These include starting/stopping applications/processes, changing or adding/removing applications, new open ports and network connections in general and the use of critical operating system commands as well as access to sensitive files, the use of removable media and traffic to unusual destinations. Additionally, at least for network management reasons, resource usage and hardware/software failures should be logged and collected by the SIEM. Also events that occur at irregular times should be logged as should system errors. What irregular times include depends of course. Normally this will be after hours, weekends and holidays (and for a specific employee – whenever the employee is known not to have any business to attend to). Also operations related things like disk space and CPU usage should be considered to be monitored [7].

Of course all logs should include information on who (what user account - that the account holder is actually the one using the account is another matter (Personal Identity Verification) although this can be ascertained by using e.g. biometric authentication) or what caused the log entry to be created and how. What happened (with the type of action detailed in a general sense) and whether it "succeeded" as well as on what system/application it happened must be detailed. Where the acting entity (e.g. attacker) resided (logically/physically) is essential information so identification details like IP and MAC address, computer name, file name, user name must be included in the logs when applicable. For further investigation domain specific details like error codes are also necessary (see Appendix A for an example of such codes) [7]. Finally, it is of course highly useful to know *when* the incident occurred. For this reason it is important that all log clocks are synchronised using a central time server. The preferred kind uses the Network Time Protocol which can be used on both Unix and Windows 2003 or later or Windows 2000 with some loss of precision [47] (see more about temporal normalisation in Section How SIEMs Work). Of course the time server must then also be included in the list of required systems to log.

When sensitive data is printed to paper a log entry is also required with indication of what data that was printed (e.g. accomplished through reacting on certain keywords). Of course log administrators that are not normally authorised to read the data should not have access to such logs and the log entries have to be carefully created if not too much is to be given away.

2.5.1 Further Considerations

Some events may be infeasible to keep logged automatically and manual logs must then be used [48]. Such logs create new security issues and have new dis-/advantages. For instance, manual logs are much easier to edit and who edited the manual log must be logged (although that would require further logs in a circular reasoning but in practice the hierarchical nature of organisations usually eliminate this problem). Also, there must be extensive documentation, that is available to the relevant individuals, that details how the logs should be handled, how retention/archiving is to be done/is done, how often to analyse the logs and how to act on the result and who is responsible for what log activities [48]. For more details on this consult the next Section, Roles and Policy.

It may be hard to decide exactly what to log at each source device and what can be disregarded either by turning logging off or filtering it on its way to the SIEM (see Consolidation in Section How SIEMs Work for a discussion on where and if filtering should occur). Logging "everything" would introduce performance issues but would allow for a proactive discovery approach with analysts watching the SIEM in real-time for new types of attacks.

Finally, when handling a specific incident the logging requirements may be more extensive. For instance, in the case of events under active internal investigation suspected users should be especially monitored. What exactly this entails depends on the specifics of the incident under investigation and on the network environment as a whole.

2.6 Roles and Policy

For an important part of security with many stakeholders it is a good idea to assign roles to various personnel. What roles and how many depend largely on the size of the organisation and the importance of the part of the business involved. For the log part of security it is important to clarify who the SIEM system owner is. The System Owner (or Information System Owner) is responsible for ensuring that logs are kept and that the logging system/method is working as required (e.g. through log review) by the security logging policy which should be an available document detailing required security logging practices. It is important to document what individuals that have access to each log and the SIEM if one exits. Not only must it be detailed who can read the logs but also who can modify or delete them (and who the owner of the logs is – the Information Owner). The System Owner needs to also make sure that these access rights are honoured [49]. Of course also the security logging policy document needs an owner. The Chief Information Officer is ultimately responsible for policy and regulations compliance and is the direct manager of the Information Security Officer who makes sure personnel receives proper training and systems proper testing. The ISO is also responsible for making risk analysis and implementing cost-effective security [49]. The Incident Handler is one of the most important actors and will be detailed in paragraphs below. Depending on the level of security required some other roles that may be important for log policies/systems are: Intrusion/Forensics Analyst, Company Internal/External Auditor and more general security roles like Network/Security Administrator and Penetration Tester. Of course the more different roles that are required the more people must be hired which introduces higher cost. In fact, even if only one individual is hired as log system responsible, the cost of operation (which includes the new employee's salary) may exceed the cost of acquiring and deployment and with time it no doubt will.

An Incident Handler (IH) must be knowledgeable in the area of forensics to be able to determine the nature of an incident. The IH must have general knowledge on topics like OSs, common network protocols and hacking and must have more detailed knowledge on anti-forensic techniques (ways to conceal/destroy data) and also have a good idea of the organisation network infrastructure. Since the IH is sometimes collecting evidence that will be used in court he/she needs to be prepared to testify if necessary. Of course the IH must be able to delegate some tasks that are better

handled by experts like security administrators or IT support/helpdesk personnel. In fact, some tasks may require such highly specialised experts that outsourcing to another organisation is necessary or enlisting the help of the SIEM vendor. This does introduce a security risk but on the other hand it may be good to have an external actor helping if the source of the incident originates from the inside of the organisation requiring internal investigation.

Sometimes some work will have to be done by different kinds of investigators like those belonging to the legal department, physical security staff, (business) management, financial auditors or those of the human resources or the Personnel Security Officer (e.g. in the case of an internal investigation) or even the appropriate authorities [51, 49]. Since the IH may have to quickly get in contact with different personnel it is important that emergency contact information is readily available for the persons "on call" [51]. Since SIEMs involve logging personnel activities in various ways it may be important to introduce/involve also a Chief Privacy Officer. Due to the extensive nature of the SIEM solution it may also be a good idea for a kind of Chief Enterprise Architect, acting as a coordinator, to be involved [49]. Since so many different people can be involved a crystal clear hierarchy of personnel must exist to avoid confusion and contradictory decisions. In the end the idea of a SIEM is to decrease the number of so called man hours required and not increasing it but since proper security is often not in place before the introduction of the SIEM the total benefit may exceed the cost.

2.7 Products

According to [2] there are 20 competitive SIEM vendors (with one SIEM product each in most cases). In this chapter some of these products are described (the reason why some were left out is that they were not well represented in papers in the research and there seemed to be no need to be "complete" anyway), some in more detail than others. The level of detail depends primarily on the fact that some of the systems where specially investigated for possible deployment on behalf of the company at which this thesis was written. Since this paper is however not on specific products but on SIEMs in general even these products will not be described very meticulously.

Some key points are given in Table 1 (note that some sources are dated and the information may not be accurate for present versions of the products). As with any product the usability is very important. If the graphical user interface (GUI) of a product is flawed the user of the product might make mistakes and in a security system this is obviously of a lot of importance. Related areas are manageability and support. Products requiring more of these factors are usually targeted for larger enterprises which typically already have suitable infrastructures in place. The required effort of deployment is also given in some cases and this is usually a quite important differentiating factor, as some SIEMs are deployed in hours or days while some are reported to take months (in an installation environment of standard complexity). Usually the deployment time is tightly related to the agent (described in How SIEMs Work) system used and not only if it is agent based or agent less. It is also important to choose a system which can grow (i.e. which has a high scalability) and that fits the buying organisation (the latter meets the description under the heading "Target"). Of

course the price of a product varies with number of devices and required sustained EPS but the information is usually not given up-front by the vendors and only after some time of procurement talks with an interested organisation. Some of the more important factors included in miscellaneous are the number of templates and the device density (i.e. the number of supported devices). The former affects effort of deployment to some extent while a built-in ticket-system for resolving incidents can improve incident management. While only one product is known to be Common Criteria certified others are undergoing such evaluation according to vendor representatives.

Some of the more major SIEM vendors left out are Cisco, LogLogic (partly due to being too focused on log management although this is true for others that are included) and Quest with its product inTrust (too low device density). These all do quite well according to [2], which puts emphasis not only on functionality but also on business strategy and marketing, but have not been investigated here. Not

Table 1 – Comparison of different vendors' SIEMs.

	Usability	**Incident management**	**Deployment /manageability /support**	**Misc**
ArcSight	Web-GUI [58].		Requires much database configuration and is hardware demanding [2] (although with the new appliance this may have changed).	Good correlation and policy/compliance, good context-info [26]. Signature-based detection [58]. Automated response based on policy [90]. Has user identity focused normalisation fields [55]. Supports 275+ devices through "connectors" (see agent-based column). Logger collects 100k EPS, searches at 3M EPS. Complies with logger tool standard NIST 800-92 (reference [52]) [54]. The only publicly traded SIEM company [2]. Excellent event

				management [GartnerCricitalCaps].
CA		Good [26].	Flexible policy definitions [26]. Needs better deployment for targeting smaller organisations [2].	Does not support log management functions very well. Focused on and great at user monitoring [2].
CheckPoint[*]	Good GUI [59].		Not easy adding unsupported devices [59].	Strong forensics capabilities [59].
IBM	Good web-GUI [71].	Better than Novell and netForensics [87]. Built-in ticket- system [71].	Long, complex deployment. [71]	Product previously neuSecure [71]. Product is split in several [2]. Great user and resource access analysis [GartnerCriticalCaps].
Intellitactics			Requires a lot of management although it has been significantly improved from a previous release [2].	
LogRhythm			High configurability [31]. Good documentation [31]. Limited support needed [2].	Very good forensics. Good compliance [31]. Uses anomaly-detection [2].
netForensics			Short [24]. Better support than netIQ and IBM [87]. Not much extra support and management required [2].	Not as good real-time [71]. Early product [5]. Certified EAL2 of CC (see Compliance) [53].
netIQ		Better than Novell and netForensics [87].		Windows-focused [90]. Primarily *not* for compliance and event handling. More networks than

				security focused [2].
Novell[**]	Nice GUI, incorporates workflow, wizard for creating rules [26].	Automated response capabilities [2].	Long [24]. "Bit time-consuming", robust [26]. Better support than netIQ and IBM [87]. High scalability [2]. Requires a lot of database tuning [2].	Early product [5]. Big hardware requirements, broad device support, tracks not only devices but users [26]. Good rules creation engine [59]. Lacking log management, excellent event management [Gartner08, GartnerCriticalCaps].
OP5[*, ***]	Web-GUI.		Very fast deployment.	Lacking templates and event management capabilities, reducing it to log management. Open-source.
OpenService			Difficult to deploy according to [32] but easy to deploy according to [2]. Scalable [32].	Broad device support [32]. Less good user tracking and compliance functionalities [2].
Q1 Labs	Good GUI using Java [50].		Complex (i.e. long learning curve). Appliance optional. Robust [50]. Auto-tuning rules [58].	Built-in reports. Anomaly-detection. Excellent device support [50]. Offers network behaviour analysis [2].
RSA[***]	Web-GUI [36]. GUI needs and will get a big overhaul according to company representatives.		Easy to deploy (4h) and configure [41]. Short deployment time. Typically a few days with special start support package [2].	Device density is approaching 400. 1100 template reports, knowledge base of 10 000s known log messages [40]. XML "Universal Device Support" engine for easy integration of

	Also the GUI does not support normal hot keys [41].			legacy devices and home grown applications [38]. Not as complex and capable as e.g. ArcSight but very widely used [2]. Does not support mobile/DHCP devices. Collects all data without filtering. Searching all logs for specific user name takes a few minutes in a typical environment. DISA Gold disk certified (a type of hardware hardening certification [56]). Uses threat info from CVE according to company representatives.
SenSage			Easy through wizards [58]. Scalability slightly low [2].	Anomaly-based detection [58]. Primarily auditing focused and has support for widely used business application SAP [2]. Lacking event management [GartnerCriticialCaps].
Snare[*, ***]	Confusing and messy Web-GUI.		Simple and quick deployment (note: of test product).	Sister product of a popular Windows agent.
Symantec		Good. Ticket-system [26].	Easy deployment [26, Gartner08].	Windows-based. Good compliance. Anomaly-detection. Incorporates vulnerability-data from vendor's own network [26].
TriGeo	Good	Automated	Quick	Not very broad

	GUI [35].	response (like blocking IPs and shutting down systems/applications through input devices) through agent and IDS capabilities through open-source IDS Snort [35, 2].	deployment appliance and support [12, 2]. Graphical rules creation [34].	device (100 different) support [90]. Not very good context-info support [35]. Correlates in memory instead of using a database. Only 64-bit appliance SIEM [34]. 100s of built-in filters [35].
	Target	**Archiving**	**Agent-based?**	**Reporting**
ArcSight	Very enterprise-focused [90].	Compressed in proprietary format. Can store information in raw or normalised [26].	Yes (called connectors) [26].	
CA	Large enterprise [2].			
CheckPoint				
IBM	Large enterprise [71].			Strong reporting capabilities [2].
Intellitactics	Well suited for large enterprise [59, Gartner08].	Compressed proprietary [2].		Great visualisations [59].
LogRhythm	>=$20 000 [31]. Primarily midsize organisations [2].		Offers both agent based and agent less [2]	Good visualisation [31].
netForensics				Great [71].
netIQ	Those needing user monitorin			

	g [57].			
Novell	Large enterprise [87, Gartner08]		Yes [87]	
OP5	Those needing only log managem ent.	Configurable archiving.	Yes.	
OpenService	Expensive [32].			
Q1 Labs	Large enterprise [2].			
RSA	Price: $ 36 – 500 000 according to company representa tives.	Proprietary which stores data encrypted and in native format (raw) and instead of database uses metadata for searches [36, 55]. 20 times less storage required and 10 times faster than db [38]. Typical compression rate 0.2-0.3 according to company representative s.	No.	Nice visual reports.
SenSage	Large enterprise [2].	Proprietary [58].		
Snare			Yes.	Very good visual reports.
Symantec		Good. Stores data normalised and/or raw [26].	Yes [26]	Lacking good reporting capabilities [2].
TriGeo	Cheap	Log	Yes [35]	

	[12, 2]. For small to medium businesses [90, 2].	management functions are based on Splunk [2].		

* Product not mentioned in [2].
** Product formerly of e-Security.
*** Product tested (to a smaller extent) as part of work at company thesis was written at.

2.8 Compliance

There is a wide range of regulation standards and laws pertaining to logging. Some are interesting for organisations in a certain field, some are nation specific. Sometimes organisations are required by law or agreements to comply while sometimes SIEM products with certain compliances are chosen because of a perceived competitive edge. Compliance puts requirements (although exactly what these entail is not always obvious [63]) on how log data is treated in the entire process from collection to storage (both site and longevity) as well as review and response. As more regulations and improved functionality have arrived more interest in SIEMs has been generated. Some of the more common standards, some of which are mentioned by [64], are detailed below, as well as some other relevant aspects.

The Sarbanes-Oxley Act of 2002 (abbreviated SOX or SOA) constitutes the response to lacking financial control of US public companies like Enron [65]). According to Section 404 of said act publicly traded US companies are required to keep detailed documentation of financial control/administration mechanisms like specific logging of related computer systems for at least seven years. Since companies may normally spend a substantial part of their revenue (2.55% for companies with revenue <100$M, 0.16% for a revenue of 1-5$B [66]) on implementing SOX a SIEM product with such compliance may be highly cost-beneficial. The EU is currently working on a similar regulation and there is also work on telecommunications data retention laws [67]. In Sweden financial logs must be retained and kept readily available for ten years for private companies based in Sweden (§7:2 of [68]). Such logs must normally be stored in Sweden and in any of the Scandinavian languages or English (§7:1, §7:3, §1:4 in [68]).

The Payment Card Industry Data Security Standard (PCI DSS or simply PCI) is a non-nation-specific standard for ensuring the security of payment card transactions and must be followed by all companies using the latter. Specially Requirement 10 of the standard is related to logging and specifies among other things that there must be daily review of authentication and IDS logs. The audit trails must be preserved for a year with three months online storage. (10.5, 7) [69]

The Health Insurance Portability and Accountability Act (HIPAA) is a US Congress act of 1996 aimed at secure patient confidentiality and thus regulates what and how health data is to be logged and how long (minimum 6 years) [70].

The Gramm-Leach-Bliley Act (GLBA) of the US Congress 1999 opened up the possibility for a bank to offer, among other things, insurance services. It also introduced some new demands on security for any organisation in possession of financial institution customer personal financial information including demands on logging [65] and it is interesting to note that it has been accused of being one of the reasons for the current (2008) financial crisis [72].

The Federal Information Security Management Act (FISMA) of the US Congress in 2002 consists of a number of different standards relating to computer security for federal agencies. One such standard is the Federal Information Processing Standard 200: Minimum Security Requirements for Federal Information and Information Systems [73] which specifies security requirements, on for instance System and Information Integrity, and how to select controls for meeting said requirements. Security controls are detailed in National Institute of Standards and Technology publication 800-53 [74].

Control Objectives for Information and related Technology (COBIT) is a framework of best practices for IT management created in 1996 by ISACA, an international association for IT governance. COBIT governs IT-processes in general including managing data and ensuring security as well as regulatory compliance (how to comply with other standards) [75].

Basnivå för informationssäkerhet (BITS or Basic levels of Security), developed by the Swedish Emergency Management Agency (SEMA or KBM in Swedish), is a set of recommendations for a minimum level of security for organisations critical to the infrastructure of the society. BITS was constructed to be compatible with ISO 27001 (below) [76]. According to BITS security logs normally need to be saved for two years [48].

The International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) have published a more security specific best practices standard by name of ISO/IEC 27002 which covers information security management. ISO/IEC 27001 is highly related and puts more detailed demands on the implementation and operation of information security management systems [77].

The Information Technology Infrastructure Library (ITIL), developed by the United Kingdom Office of Government Commerce, is another best practices framework which partially implements ISO 27001 while putting much focus on physical security issues [78].

The Basel II Accord of the Basel Committee on Banking Supervision in 2004 is a recommendation on banking laws related to operational risks of banks. It, among other things, puts constraints on how banks log financial transactions [79].

The Common Criteria for Information Technology Security Evaluation (or CC or ISO/IEC 15408) is a certification standard on computer security. The Evaluation

Assurance Level (EAL) of a product is a measure of how well it meets the CC demands on the specification, implementation and specially evaluation processes. For instance, to receive an EAL2 the product must be "structurally tested" meaning the manufacturer must provide the evaluation team with already produced documents pertaining to the product. An EAL7, which is the highest, means the product has been very thoroughly evaluated through a normally very costly process.

If the organisation handles information relevant to national security especially firm rules may apply for logging. For instance in Sweden such logs must be stored for at least 25 years since this is the period of prescription for the most serious crimes (e.g. murder). Additionally, there has been some talk on removing the statute of limitations for the most serious crimes requiring some organisations to store their logs indefinitely [80].

The truth and integrity of logs are normally not contested ([81]) but in the case of screenshots used as evidence in trials some controversies exist, e.g. [82].

2.8.1 Open Records and Ethics

There are also other aspects concerning logging and the law that may need to be considered. For instance sometimes it is not entirely clear what logs of public sector actors that are protected against disclosure by secrecy and when legislation concerning freedom of information takes the upper hand. See for instance a case in Swedish law: [83]. Additionally companies using logs to comply with regulations may have to supply that information to the suitable part of the government for auditing possibly making the information public property. Such information would likely contain sensitive information (possibly detailing business as well as security practices). Another complex matter is that of copyright. Since it is illegal (depending of course on the the applicable legislation) to keep copies of copyrighted information it may be problematic if logs contain such information.

It is important to note that many of these regulations overlap. Organisations that implement these manually risk doing a lot of redundant work and it is important to use an automated tool like a SIEM that helps fulfil all required standards. In September 2006 ¾ of finance/IT companies used manual processes according to ControlPath, one company apparently wasting $0.5 mil/year [84].

It is also important to consider the privacy issues related to logging. It must be ensured that the SIEM is not used unethically (like spying on employees) and normally that users of the network log source devices are informed of the centralised logging in place (unless this means a significant security set back).

2.9 Retention Time and Site

Another important issue is about log storage, more specifically about how long logs should be stored and how. The time for log retention obviously depends on how far in the future the log information might be needed (by the organisation or by regulation).

Since it is not always easy to foresee the need of log entries it is possible to instead focus on how important (sensitive) the logs are which of course depends on how important the system/user creating the logs is. Some logs will then of course be discarded quickly since usually many log entries that are of no interest to the organisation are created. Furthermore, logs of some but limited interest may be stored for a couple of months while more important logs relating to more sensitive data or created by users with higher privileges should perhaps be saved for at least a year. Some logs may have to be stored for decades or perhaps for the entire existence of the organisation (and beyond). Usually logs relating to specific products in development need to be saved for the entire duration of the product development/lifetime. The specific times are often given by the relevant regulation (see Compliance). Due to increased time to uncover insider attacks logs should be retained longer than for external attacks (see the discussion on a paper by Chuvakin at the end of Section Procurement).

Of course important data should always be kept in secure locations and so log backups/archives should be kept securely (of course distributing the logs is slightly ironic since they were collected centrally from distributed sources in the first place!). Redundancy is often a good idea for very important data and protection against physical attacks (for instance an attacker simply walking up to the log server and ripping out the disk or locally formatting the drives). If anyone can edit the logs they are obviously more or less useless and the integrity must be ensured by denying ordinary users access to the logs and carefully developing policies about who may edit the logs. In fact, in some cases logs should be made completely impossible to modify (within reason). To make sure that it is possible to tell if the logs have been changed, cryptographic (digital) signing may be used as recommended by many government agencies like the Swedish Emergency Management Agency [48].

2.9.1 WORM

A log which is a couple of months old is usually not needed anymore in a day-to-day basis and can be archived to a medium which provides less speed and is less costly. To ensure the integrity of archived logs (for confidentiality issues and change-detection before archiving and collecting see Section Security Issues), to be certain that they have not been changed by an external or internal attacker, special data retention storage systems called fixed content storage or Write Once, Read Many (WORM) media can be used. With WORM it is not possible (this is not strictly true in practice, see more below) to manipulate or destroy already written data. Examples of WORM are DVD-R and hard disk drives using special software. It is interesting to note that the property of write-once which is so restricting for most uses is highly useful for achieving integrity. The technique can be implemented in hardware (physical WORM) or software (logical WORM). At least one SIEM product, RSA's enVision uses both a logical part (being manufactured by RSA itself) and a physical (Centera made by parent company EMC).

In the paper [15] one of the first logical WORM systems, Venti, is described. It is designed for high integrity archiving (not necessarily of log data). Although the system design is quite complex the basic idea is not. When data is written to Venti the address of the data will be the cryptographic hash or fingerprint of the data which means that it is only possible to overwrite the data with the very same data for a

collision-free hash function. This is known as Content Addressed Storage or Content-Addressable Storage abbreviated CAS. This means that probably not all the disk space will be used but disk space is usually considered quite cheap nowadays and this is not a major issue according to [15]. Furthermore, this approach means that duplicate data will only be stored in one location actually saving some disk space. When the system is writing to Venti it saves the fingerprint to be able to retrieve it if necessary. Metadata like time and date are saved with the fingerprint. Since many fingerprints will need to be stored they are also written to Venti in a bundle with the source system storing only the fingerprint of the fingerprint file or root fingerprint locally. This may save further disk space since several source systems can use the same archive system without duplicate entries or knowing what the other systems have stored.

Of course it is important that systems authenticate themselves before writing to Venti to make sure that only trusted systems can write data. Also it may be important to keep track of who wrote what for non-repudiation. A simple solution would be to save the root fingerprint together with the authenticated user name of the transaction. It is important to note that the root fingerprints constitute a weak point of the system.

EMC Centera from 2001 is the first commercial WORM CAS system [60] and the main driver behind its success is likely the increased demand for solutions for compliance. Other products include IBM DR550 (which uses a slightly different method from the others), NetApp SnapLock and one from a company called iTernity. iTernity is also offered through IBM and so is Centera together with another product [61].

Although it is important for compliance reasons that the integrity of the data can be ensured the time period of retention required is usually limited. Since in CAS each address can only store specific data there is no storage-space specific reason for wanting to purge data after the compliance time has elapsed but often organisations want to make sure sensitive data that is no longer needed is deleted. This is especially true for organisations operating in the defence vector where there may even be special compliances demanding such deletion. For this reason WORMs usually store a timestamp in the metadata of the stored data and use software to allow deletion only after a certain period of time has elapsed after the timestamp. Alternatively some WORM systems can trigger deletion to be allowed when a special event occurs (this is similar to the triggering of time/logic bombs used by e.g. hackers) [62]. In Centera it is in fact possible to delete data even before the required time has passed but only through a "tightly audited channel" allegedly to comply with EU personal privacy regulations [60]. Of course these features mean new potential attack vectors for intruders attacking the archived log data. Since the system relying on an elapsed time must have access to an accurate clock it could be vulnerable to attacks on the system clock or in the case of power loss. Of course, if duplicates are stored they will have different time stamps and deletion should only be allowed after an appropriate time has passed since the latest timestamp.

2.10 Log Formats

As with any type of software product, SIEM manufacturers use different designs. For instance the way events are displayed to the user and stored, i.e. the log formats, differ between products. This means there might be problems if two SIEMs communicate with one another. Additionally, users that come in contact with different SIEMs (e.g. external auditors) may need to learn a new format for each product. The applicable standardization organisation, the Internet Engineering Task Force (IETF) has been developing a standard format, the Intrusion Detection Message Exchange Format (IDMEF). Although primarily for IDS events it can likely be used for general events without problems. The IDMEF bundles events part of the same attack together (usually done by the aggregation part of the SIEM) and messages of this type has fields for time of attack, target IP, attack severity level among other things. In contrast information on newly discovered threats published with various formats on different websites (such as Bugtraq, CVE, and OSVDB) is more general in nature.
While there are many ways to present logs there are also various methods for transfer between log source and SIEM and between SIEMs. Again the IETF are developing a standard, this one called the Intrusion Detection Exchange Protocol (IDXP) which uses the so-called BEEP framework over TCP making it an application protocol with reliable transfer. To establish end-to-end security the IDXP can use Transport Layer Security (TLS) through a BEEP security profile (compare HTTP over TLS over TCP). Normally the log messages being transported with IDXP uses the IDMEF described above. The IDMEF itself provides very limited security in the form of optional checksums [85].

Of course log formats differ between source devices too (UNIX's syslog being among the more common) but these formats are more similar and fewer than those in IDSs due to the lower context dependencies in the former and so there is less need for standardisation (like IDMEF). However, there was an organisation called the Open Security Exchange with the purpose of among other things standardising security management in general (see [86]) [5]. Of course more standardisation in log source means less CPU time "wasted" on normalisation in the SIEM.

According to [88] IP devices such as routers, switches of large Internet Service Providers must provide logging in a standard format or a format providing the same information (fields) and reliable log transmission to remote servers. The document also puts demands on using NTP for time stamping the logs.

3 Conclusions

SIEMs are complex tools and they have a lot of interdependencies with different systems and security measures. They come into contact with the lowest levels of security when processing the data it collects from a large number of different network devices and the highest when applying policies. In the first generation of SIEMs little more than centralised collection was offered [89] and highly trained experts were the only ones to be able to handle the systems. As compliance became a more important motivator the second generation of SIEMs was born. The large amount of different compliance regulations turned SIEMs into more or less necessary tools to avoid much redundant work.

3.1 A Future Outlook

The market (for figures on market value and growth – see Motivation) is now an "adolescent" [89], i.e. not only efficient but also practical and now they offer more than just centralised log collection and storage. Previous challenges like speed and manageability have been overcome [90] while some things remain (this is probably why the market was not termed an "adult"). For example, incorporating more contextual information (like geographical information and network directory listings as well as network topology details) is something that will most likely given more focus in the future versions of the different vendors' products. Further demands on increased usability, deployment and standardisations will probably drive development. Additionally, major vendors will no doubt start offering more diverse product lines targeting not only large enterprise customers but also medium organisations with slimmed down (both with respect to functionality and price) products. It is furthermore likely that increased focus will be given to policy oriented automated response capabilities approaching the final dream of every security administrator. More pre defined template rules and visualisation reports is forecasted to arrive as management will get more involved while compliance will be continuing to be the driving factor. Device density in most products will continue to rise.

The tuning time of SIEMs will have to be addressed and by extension the false positive rates [30]. Data reduction will slowly be reduced as storage gets cheaper and processing performance better, although this is not entirely clear with more and more products being installed in the network environments leading to higher demands. Finally, with the increased focus on user tracking, SIEMs will likely get improved identity and access management capabilities to grow closer to a complete security solution.

References

[1] Hutton N., *Preparing for Security Event Management*, 360 Information Security Ltd, 2007. Retrieved from http://www.windowsecurity.com/uplarticle/NetworkSecurity/360is-prep-sem.pdf on December 9 2008.

[2] Nicolett M. and Kavanagh K. M., *Magic Quadrant for Security Information and Event Management,* Gartner RAS Core Research, 2008.

[3] *Log Management,* Information Security, 2008. Retrieved from http://www.searchsecurity.com on December 9 2008.

[4] *Security Event Management,* Bitpipe.com, 2008. Retrieved from http://www.bitpipe.com/tlist/Security-Event-Management.html on December 9 2008.

[5] Kelley D., *Report: Security Management Convergence via SIM (Security Information Management) — A Requirements Perspective,* Journal of Network and Systems Management, Vol. 12, No. 1, March 2004, Plenum Publishing Corporation.

[6] Desai N., IDS *Correlation of VA Data and IDS Alerts*, SecurityFocus, 2003. Retrieved from http://www.securityfocus.com/infocus/1708 on December 10 2008.

[7] Shenk J., *Fourth Annual SANS 2008 Log Management Market Report Demanding More from Log Management Systems*, SANS Analyst Program, SANS 2008.

[8] Paller A., *Prospective Buyers Want Answers*, Information Security, April 2007. Retrieved from http://searchsecurity.techtarget.com/magazineFeature/0,296894,sid14_gci1257129_idx2,00.html on December 10 2008.

[9] Savage M., *Survey: Security Pros Identify Priorities for 2008*, Information Security, February 2008. Retrieved from http://searchsecurity.techtarget.com/magazineFeature/0,296894,sid14_gci1297408_idx9,00.html on December 10 2008.

[10] *Regulatory compliance takes the lead as the most important driver of information security, surpassing worms and viruses*, Ernst & Young, 2005. Retrieved from http://www.ey.com/Global/Assets.nsf/Hungary/PressRelease_Global_Information_Security_Survey_2005_EN/$file/Global_Information_Sec_Survey2005_EN.pdf on December 10 2008.

[11] Chuvakin A., *Six Mistakes of Log Management*, InfoSec Writers, 2007. Retrieved from http://www.infosecwriters.com/text_resources/pdf/Six_Mistakes_of_Log_Management_AChuvakin.pdf on December 10 2008.

[12] Connolly J., *Security Management Special Report: Under Fire*, CIO Decisions Magazine Archives, 2007. Retrieved from http://searchcio-midmarket.techtarget.com/magItem/0,291266,sid183_gci1256411_idx4,00.html on December 10 2008.

[13] *IT-säkerhetsstandarden Common Criteria (CC), En introduktion*, KBM Rekommenderar 2007:2, Krisberedskapsmyndigheten.

[14] *Log Management and Virtualization: New Meets Old,* LogLogic, 2008.

[15] Quinlan S. and Dorward S., *Venti: a new approach to archival storage*, Bell Labs, Lucent Technologies, 2001(?). Retrieved from http://plan9.bell-labs.com/sys/doc/venti/venti.html on December 10 2008.
[16] Fisher D., *EMC acquires Network Intelligence, closes RSA deal*, SearchSecurity.com, 2006. Retrieved from http://searchsecurity.techtarget.com/news/article/0,289142,sid14_gci1216394,00.html on December 10 2008.
[17] Shipley G., *SIEM tools come up short,* Network World, 2008. Retrieved from http://www.networkworld.com/reviews/2008/063008-test-siem.html on December 10 2008.
[18] Rooney P., *Novell Acquires e-Security*, ChannelWeb, 2006. Retrieved from http://www.crn.com/software/186100156 on December 10 2008.
[19] Payne S., *A Guide to Security Metrics, SANS Security Essentials GSEC Practical Assignment Version 1.2e*, SANS, 2006. Retrieved from http://www.sans.org/reading_room/whitepapers/auditing/55.php on December 10 2008.
[20] *Guide To Snare For Windows,* Snare, Intersect Alliance, 2008. Retrieved from http://www.intersectalliance.com/resources/Documentation/Guide_to_Snare_for_Windows-2.8.pdf on December 10 2008.
[21] Storey N., *Safety-Critical Computer Systems*, Prentice Hall, 1996 (ISBN 0-201-42787-7).
[22] Schneier B. and Ranum M., *Bruce Schneier, Marcus Ranum debate risk management*, Information Security, October 2008. Retrieved from http://searchsecurity.techtarget.com/magazineFeature/0,296894,sid14_gci1332745_idx1,00.html on December 10 2008.
[23] Savage M., *Snapshots of SIMs*, Information Security, 2006. Retrieved from http://searchsecurity.techtarget.com/magazineFeature/0,296894,sid14_gci1257083,00.html on December 10 2008.
[24] Dubie D., *Users Shoring Up Net Security With SIM*, Network World, 2002. Retrieved from http://www.networkworld.com/news/2002/0930apps.html on December 10 2008.
[25] Worldwide Security and Vulnerability Management Software 2008-2012 Forecast and 2007 Vendor Shares: Making Security Smart, Doc #214144, IDC, 2008.
[26] *Readers' Choice Awards*, Information Security, April 2008. Retrieved from http://searchsecurity.techtarget.com/magazineFeature/0,296894,sid14_gci1308627_idx23,00.html on December 10 2008.
[27] Hollows, P., *Security Threat Correlation: The Next Battlefield*, eSecurity Planet, 2002. Retrieved from http://www.esecurityplanet.com/views/article.php/10752_1501001 on December 10 2008.
[28] Murray T., *Getting a Handle on Security Events, GIAC Practical Assignment v.1.4b*, 2003. Retrieved from www.giac.org/practical/gsec/Sean_Murray_GSEC.pdf on December 10 2008.
[29] Schneier B., *Security Information Management Systems (SIMS)*, Schneier on Security, 2004. Retrieved from http://www.schneier.com/blog/archives/2004/10/security_inform.html on December 10 2008.
[30] Rothman M., *Security information management finally arrives, thanks to enhanced features*, Network Security Tactics, 2007. Retrieved from

http://searchsecurity.techtarget.com/tip/0,289483,sid14_gci1247084,00.html on December 10 2008.
[31] Moser G., *Product Reviews, Log Management, LogRhythm*, Information Security, 2008.
[32] Henderson T., *Security Information Management*, Information Security, April 2005. Retrieved from http://searchsecurity.techtarget.com/magazineFeature/0,296894,sid14_gci1257159,00.html on December 10 2008.
[33] Debar H. and Viinikka J., *Intrusion Detection: Introduction to Intrusion Detection and Security Information Management,* France Telecom Division R&D, FOSAD 2004/2005, LCNS 3655, pp.207-236, Springer-Verlag, 2005.
[34] *The Case for Security Information and Event Management (SIEM) in Proactive Network Defense*, TriGeo Network Security, 2008.
[35] Sidel S., *Security Information Management*, Information Security, June 2005. Retrieved from http://searchsecurity.techtarget.com/magazineFeature/0,296894,sid14_gci1257171,00.html on December 10 2008.
[36] Huston B., *Security Information Management*, Information Security, November 2006. Retrieved from http://searchsecurity.techtarget.com/magazineFeature/0,296894,sid14_gci1257287,00.html on December 10 2008.
[37] Grance T. et al, *Guide to Selecting Information Technology, Recommendations of the National Institute of Standards and Technology*, NIST Special Publication 800-36, NIST, 2003. Retrieved from http://csrc.nist.gov/publications/nistpubs/800-36/NIST-SP800-36.pdf on December 10 2008.
[38] *An Introduction to RSA enVision, The Information Management Platform for Security and Compliance Operations Success*, RSA, 2007.
[39] *Validated Products List*, NIAP, 2008. Retrieved from http://www.niap-ccevs.org/cc-scheme/vpl/ on December 10 2008.
[40] *RSA Solution Brief, The RSA enVision Platform, A Single, Integrated 3-in-1 Log Management Solution*, RSA, 2008.
[41] *Product Review: RSA Security's RSA enVision*, Information Security, July 2008. Retrieved from http://searchsecurity.techtarget.com/magazineFeature/0,296894,sid14_gci1321710_idx2,00.html on December 10 2008.
[42] Kaeo M., *Current Operational Security Practices in Internet Service Provider Environments*, Network Working Group, Request For Comments 4478, Informational, IETF, 2007. Retrieved from http://www.rfc-editor.org/rfc/rfc4778.txt on December 10 2008.
[43] Schneier B. and Kelsey J., *Cryptographic Support for Secure Logs on Untrusted Machines*, The Seventh USENIX Security Symposium Proceedings, USENIX Press, pp. 53-62, 1998. Retrieved from http://www.schneier.com/paper-secure-logs.pdf on December 10 2008.
[44] Strom D., *Log management reins in security and network device data*, Information Security, October 2007. Retrieved from http://searchsecurity.techtarget.com/magazineFeature/0,296894,sid14_gci1274439_idx7,00.html on December 10 2008.
[45] Garbrecht F., *Practical Implementation of Syslog in Mixed Windows Environments for Secure Centralized Audit Logging, GSEC Practical*

Assignment Version 1.4 Option 2, SANS, 2002. Retrieved from http://www.sans.org/reading_room/whitepapers/casestudies/713.php on December 10 2008.

[46] *SANS Consensus Project Information System Audit Logging Requirements*, SANS, 2007. Retrieved from http://www.sans.org/resources/policies/info_sys_audit.pdf 46 on December 10 2008.

[47] Mills D. L., *Network Time Protocol (Version 3) Specification, Impl*, Network Working Group, Request For Comments 1305, IETF, 1992. Retrieved from http://www.rfc-editor.org/rfc/rfc1305.txt on December 10 2008.

[48] *Basnivå för informationssäkerhet (BITS), KBM Rekommenderar 2006:1, Utgåva 3,* Krisberedskapsmyndigheten, 2006.

[49] Bowen P. et al., *Information Security Handbook: A Guide For Managers, Recommendations of the National Institute of Standards and Technology*, NIST Special Publication 800-100, NIST, 2006. Retrieved from http://csrc.nist.gov/publications/nistpubs/800-100/SP800-100-Mar07-2007.pdf on December 10 2008.

[50] Huston B., *Hot Pick*, Information Security, May 2006. Retrieved from http://searchsecurity.techtarget.com/magazineFeature/0,296894,sid14_gci1257236,00.html on December 10 2008.

[51] Kent K. et al., *Guide to Integrating Forensic Techniques into Incident Response, Recommendations of the National Institute of Standards and Technology*, NIST Special Publication 800-86, NIST, 2006. Retrieved from http://csrc.nist.gov/publications/nistpubs/800-86/SP800-86.pdf on December 10 2008.

[52] Kent K. and Souppaya M., *Guide to Computer Security Log Management, Recommendations of the National Institute of Standards and Technology*, NIST Special Publication 800-92, NIST, 2006. Retrieved from http://csrc.nist.gov/publications/nistpubs/800-92/SP800-92.pdf on December 10 2008.

[53] *Certified Product List*, Common Criteria Portal, 2008. Retrieved from http://www.commoncriteriaportal.org/products_OD.html on December 10 2008.

[54] *Product Brief: ArcSight Logger, Simplifying Log Collection, Storage and Analysis*, ArcSight, 2008.

[55] *ArcSight IdentityView, Increasing Security and Compliance with a 360-Degree View of User Activity*, Research 003-111208-01, ArcSight, 2008.

[56] *Defense Information Systems Agency*, Department of Defense, 2008. Retrieved from http://www.disa.mil on December 10 2008.

[57] Nicolett M., *Critical Capabilities for Security Information and Event Management Technology*, Gartner RAS Core Research, 2008.

[58] *Recent Releases*, Information Security, February 2006. Retrieved from http://searchsecurity.techtarget.com/magazineFeature/0,296894,sid14_gci1257223,00.html on December 10 2008.

[59] Snyder J., *Security event management, no strings attached*, Information Security, 2006. Retrieved from http://searchsecurity.techtarget.com/news/article/0,289142,sid14_gci1207811,00.html on December 10 2008.

[60] *EMC Content-Addressed Storage System*, EMC, 2008. Retrieved from http://www.emc.com on December 10 2008.

[61] *FileNet Image Services Connector to Centera*, IBM, 2008. Retrieved from http://www-01.ibm.com/software/data/content-management/filenet-image-services/centera.html on December 10 2008.
[62] Mitra S. and Winslett M., *Secure Deletion from Inverted Indexes on Compliance Storage*, Department of Computer Science, University of Illinois at Urbana-Campaign, StorageSS'06, ACM 1595935525/06/0010, 2006.
[63] Chuvakin A., *LogLogic LogBlog*, LogLogic, 2008. Retrieved from http://blog.loglogic.com/ on December 10 2008.
[64] *The SANS WhatWorks 2007 Log Management Summit*, SANS, 2007. Retrieved from http://www.sans.org/logmgtsummit07 on December 10 2008.
[65] *Gramm-Leach-Biley Act,* An Act, Public Law 107-204, 107th Congress, 2002. Retrieved from http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=107_cong_public_laws&docid=f:publ204.107 on December 10 2008.
[66] *Final Report of The Advisory Committee On Smaller Public Companies To the United States Securities and Exchange Commission*, SEC, 2006. Retrieved from http://www.sec.gov/info/smallbus/acspc/acspc-finalreport.pdf on December 10 2008.
[67] Directive 2006/24/EC Of The European Parliament And Of The Council, Official Journal of the European Union, 2006.
[68] *Svensk Författningssamling (SFS) Bokföringslag (1999:1078)*, Justitiedepartementet, 1999. Retrieved from http://www.riksdagen.se/Webbnav/index.aspx?nid=3911&bet=1999:1078 on December 10 2008.
[69] *About the PCI Data Security Standard (PCI DSS)*, PCI Security Standards Council, 2008. Retrieved from https://www.pcisecuritystandards.org/security_standards/pci_dss.shtml on December 10 2008.
[70] *Health Insurance Portability And Accountability Act of 1996*, An Act, Public Law 104-191, 104th Congress, 1996. Retrieved from http://aspe.hhs.gov/admnsimp/pl104191.htm on December 10 2008.
[71] Shipley G., *Security Information Management Tools: NetForensics Leads a Weary Fleet*, Network Computing, 2002. Retrieved from http://www.networkcomputing.com/1307/1307f2.html on December 10 2008.
[72] Ekelund R. B. and Thornton M., *More Awful Truths About Republicans*, Ludwig von Mises Institute, 2008. Retrieved from http://mises.org/story/3098 on December 10 2008.
[73] FIPS PUB 200, *Federal Information Processing Standards Publication, Minimum Security Requirements for Federal Information and Information Systems*, NIST, 2006. Retrieved from http://csrc.nist.gov/publications/fips/fips200/FIPS-200-final-march.pdf on December 10 2008.
[74] Ross R. et al., *Recommended Security Controls for Federal Information Systems, NIST Special Publication 800-53 Revision 2*, NIST, 2007. Retrieved from http://csrc.nist.gov/publications/nistpubs/800-53-Rev2/sp800-53-rev2-final.pdf on December 10 2008.
[75] *ISACA*, 2008. Retrieved from http://www.isaca.org/cobit on December 10 2008.
[76] *Basic Levels of Security, Fakta*, Swedish Emergency Management Agency, 2007. Retrieved from

http://www.krisberedskapsmyndigheten.se/upload/3044/faktablad_grundlaggande_sakerhetsnivaer_2007_engelsk.pdf on December 10 2008.
[77] *ISO*, 2008. Retrieved from http://www.iso.org on December 10 2008.
[78] ITIL, 2008. Retrieved from http://www.itil-officialsite.com/home/home.asp on December 10 2008.
[79] *Basel II: Revised international capital framework*, BIS, 2008. Retrieved from http://www.bis.org/publ/bcbsca.htm on December 10 2008.
[80] *Bodtröm vill avskaffa preskriptionstiden för mord*, SR, 2005. Retrieved from http://www.sr.se/cgi-bin/ekot/artikel.asp?artikel=698292 on December 10 2008.
[81] Kerr O. S., *Computer Records and the Federal Rules of Evidence, USA Bulletin*, U.S. Department of Justice, 2001. Retrieved from http://www.usdoj.gov/criminal/cybercrime/usamarch2001_4.htm on December 10 2008.
[82] Lindström K., *Hovrätten frikänner fildelningsdömd*, Copmuter Sweden, 2006. Retrieved from computersweden1.idg.se/2.139/1.77882 on December 10 2008.
[83] Regeringsrätten, *Målnummer 6251-98*, Referat/Dom RÅ 1998 ref. 44, 1998. Retrieved from http://www.rattsinfosok.dom.se/lagrummet/index.jsp on December 10 2008.
[84] McGillicuddy S., *Regulation redundancy: Money down the drain*, News Writer, 2006. Retrieved from http://searchcio.techtarget.com/news/article/0,289142,sid182_gci1214109,00.html 84 on December 10 2008.
[85] Wood M., *Intrusion Detection Message Exchange Requirements*, Network Working Group, Request For Comments: 4766, Informational, IETF, 2007. Retrieved from http://www.ietf.org/rfc/rfc4766.txt on December 10 2008.
[86] *Open Security Exchange*, 2004. Retrieved from http://www.opensecurityexchange.org/about.html on December 10 2008.
[87] Dubie D., *NetIQ upgrades security management tools*, Network World, 2002. Retrieved from http://www.networkworld.com/news/2002/0930appsnetiq.html on December 10 2008.
[88] Jones E. G., *Operational Security Requirements for Large Internet Service Provider (ISP) IP Network Infrastructure*, Network Working Group, Request For Comments: 3871, Informational, IETF, 2004. Retrieved from http://www.rfc-editor.org/rfc/rfc3871.txt on December 10 2008.
[89] *SIMs, Information Security*, April 2007. Retrieved from http://searchsecurity.techtarget.com/magazineFeature/0,296894,sid14_gci1257124_idx2,00.html on December 10 2008.
[90] *Providing Clarity,* Information Security, 2007. Retrieved from http://www.sensage.com/English/Collaterals/Editorial_Coverage/2007/20070713_SearchSecurity.html on December 10 2008.

Appendix A - EventIDs

There is a vast amount of different identification numbers or error codes identifying the event in question in the Windows OS logging system. These security event IDs are important for determining the type of the event and some of the more important are identified here. On Windows 2003 IDs 528 and 540 are used for user logons while 538 is a logoff. 522 means that the user has logged in as a different user while 576 indicates a special privilege logon and IDs 672-4 are for the authentication system Kerberos's ticket granting. Failed logons are shown with ID 675, exiting processes set off ID 593 and 680 is a non-Kerberos logon where a different logon server is used. When a new account is created 624 is indicated followed by a 624 for account enabling and a 628 for password reset and finally an "account changed" or 642. For logons there are different categories, called logon types. A logon type 3 simply means a network logon while type 8 indicates that the password was sent in clear text (which may be a security problem). A logon type 7 indicates a user unlocked the computer after taking a break and the rather unusual type 0 indicates a logon using a system account. ID 10 means print job status while 13 means print job deleted [20].

Appendix B - Glossary

As in any topic in computer science there is a large number of abbreviations and other domain specific concepts that are often used. There are also many names and abbreviations for the same thing and some of the alternative names for the product know in this paper as SIEM are given in Table A below. Also some that are sometimes mixed up with SIEMs but actually differ in functionality are also included at the end of the Table with some comments to clear up the potential confusion. Different sources do not agree on the exact definitions of some and hence while some let two names be equivalent others see differences. Also different vendors offer slightly different solutions in their products and they do not agree on a common definition.

Table A – Alternative names for SIEMs and differences to similar tools.

Name	*Comment*
SIEM (Security Information and Event Management)	This is the name used in this paper and the most common moniker. Note that SIEM, like several of the other names below, is often taken to mean Security Information and Event Management *System*. A SIEM is based on a SIM and a SEM which were initially separate products but then become merged as different vendors did so.
SEM (Security Event Management)	The real-time incident response part of the SIEM focusing more on network security data than the SIM part [2].
SIM (Security Information Management)	The part of SIEM that deals with compliance and log storage, e-discovery and policy and threat/incident management with less network security device data than in a SEM [2]. Normally they only account for a small part of

	the cost of the SIEM [3]. It does not need as much correlation abilities which makes it quicker and there is less of a need for filtering and normalisation. Not to be confused with the Society for Information Management.
Log management system	This is usually used as a synonym of SIM.
CSEM	Same as SIEM [1].
CIEM	Same as SIEM [1].
ESM	Same as SIEM [1].
Security Event Log Monitoring	Same as SIEM. Mentioned on [4].
SECA (Security Event Correlation and Aggregation)	The same as SIEM. Name used in [5].
ISM (Information Security Management) aka InfoSec	This is a more general term and not a specific product. It is simply a moniker used for how the organisation deals with the security of information.
Security Incident and Event Manager	The same as SIEM. The "I" has become Incident instead of Information but it is the same thing as a Security Information and Event Management System.
Security Operations Center (SOC)	Usually provides more than just SIEM functionality, such as configuration management, making it more comprehensive.
Intrusion Prevention System (IPS)	The response capabilities of a SIEM differs from an Intrusion Prevention System (IPS) in that the latter has direct control of data flow while the SIEM must command some other tool to act meaning it a slower response.
Intrusion Detection System (IDS)	Although a SIEM was initially little more than a "super-IDS" that collected data from all IDSs to make a more intelligent decision it is now even more intelligent with the integration of other types of log systems than just IDSs.
Identity and Access Management system (IAM)	A SIEM is also different from an IAM which controls user rights and access but does not for instance monitor access attempts.
Alert Management System (AMS)	Alert Management Systems (AMS) provide only part of the alert capabilities of a SIEM. While AMSs (to some extent) aggregate IDS-alerts and add context-info to reduce the number of false positives they offer no (computationally expensive) correlation between single events/log entries [6].
Network Management Systems (NMS)	Network Management Systems are quite similar to SIEMs but do not focus on security or compliance. Due to its composite and diverse nature there are other aspects that make a SIEM stand out, compared to the system from SOC through NMS, as well, e.g. its log management capabilities.