

# Buyer’s Guide to Log Management: Comparing On-Premise and On-Demand Solutions

On initial inspection, log management appears a straight forward and fairly basic feature of infrastructure management. It has long been understood as an operational best practice and security measure required for troubleshooting, performance management and security incident response. However, in the last five years both governmental and industry specific regulations have included log management as a required control within an infrastructure. These regulations have also expanded the scope of what log management is by specifying a higher frequency of manual log review, not just “spot use” of the information on an as-needed basis.

This white paper examines and compares two solutions to log management - traditional on-premise log management managed in-house by the infrastructure team and a Software-as-a-Service (SaaS), or on-demand, log management solution, which moves the infrastructure footprint for log management to a physically and organizationally separate company. By examining each element of log management and comparing on-premise and on-demand, it will be clear that a SaaS solution for log management is more efficient, effective and costs less to own than a traditional on-premise solution.

## Contents

Infrastructure .....	2
Architecture Design.....	2
Installation.....	3
Professional Services .....	3
Monitoring .....	4
Backups and Fault Tolerance .....	5
Retention.....	5
Restoration.....	5
Time to Recovery.....	6
Transport.....	7
Highly Distributed Networks .....	7
Firewalls.....	7
Security .....	7
Log Review's Compensation	
Controls.....	9
Manual Review Issues .....	9
Automated Log Review.....	9
Conclusion .....	10
About Alert Logic.....	11

**Alert Logic, Inc.**

1776 Yorktown, 7th Floor, Houston, TX 77056 | 877.484.8383 (toll free) | 713.484.8383 (main) | 713.660.7988 (fax) | [www.alertlogic.com](http://www.alertlogic.com)

Alert Logic and the Alert Logic logo are trademarks, registered trademarks, or service marks of Alert Logic Inc. All other trademarks listed in this document are the property of their respective owners.

THIS DOCUMENT AND THE SOFTWARE DESCRIBED IN THIS DOCUMENT ARE FURNISHED UNDER AND ARE SUBJECT TO THE TERMS OF A LICENSE AGREEMENT OR A NON-DISCLOSURE AGREEMENT. EXCEPT AS EXPRESSLY SET FORTH IN SUCH LICENSE AGREEMENT OR NON-DISCLOSURE AGREEMENT, ALERT LOGIC, INC. PROVIDES THIS DOCUMENT AND THE SOFTWARE DESCRIBED IN THIS DOCUMENT "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. SOME STATES DO NOT ALLOW DISCLAIMERS OF EXPRESS OR IMPLIED WARRANTIES IN CERTAIN TRANSACTIONS; THEREFORE, THIS STATEMENT MAY NOT APPLY TO YOU.

This document and the software described in this document may not be lent, sold, or given away without the prior written permission of Alert Logic, Inc., except as otherwise permitted by law. Except as expressly set forth in such license agreement or non-disclosure agreement, no part of this document or the software described in this document may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, or otherwise, without the prior written consent of Alert Logic, Inc. Some companies, names, and data in this document are used for illustration purposes and may not represent real companies, individuals, or data.

This document could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein. These changes may be incorporated in new editions of this document. Changes or improvements may be made to the software described in this document at any time.

© 2010 Alert Logic, Inc., all rights reserved.

U.S. Government Restricted Rights: If the software and documentation are being acquired by or on behalf of the U.S. Government or by a U.S. Government prime contractor or subcontractor (at any tier), in accordance with 48 C.F.R. 227.7202-4 (for Department of Defense (DOD) acquisitions) and 48 C.F.R. 2.101 and 12.212 (for non-DOD acquisitions), the government's rights in the software and documentation, including its rights to use, modify, reproduce, release, perform, display or disclose the software or documentation, will be subject in all respects to the commercial license rights and restrictions provided in the license agreement.

Invision Security and Alert Logic are trademarks or registered trademarks of Alert Logic, Inc. or its subsidiaries in the United States and other jurisdictions. All other company and product names mentioned are used only for identification purposes and may be trademarks or registered trademarks of their respective companies.

# Infrastructure

---

In this section, we will compare the preparation for, installation and continuing maintenance of the infrastructure that supports the log management solution for the two scenarios - on-premise and SaaS. In geographically dispersed networks, an n-tier architecture is required for on-premise solutions. The tiers include at minimum a collection appliance (servers that collect logs from multiple sources would also be considered a collection appliance) and a centralized infrastructure to support the increased processing resources required for log analysis and data retention resources required for long term archive. An additional tier may be a log aggregation point to more efficiently transmit data back to the central infrastructure site for log analysis and archive.

## Architecture Design

Architecture design is comprised of physical hardware placement, understanding log collection and aggregation points, and traffic requirements and patterns. This is a required first step no matter what solution is implemented, on-premise or SaaS. However, there are common issues that arise for an on-premise design that do not in a SaaS design. First is the cost of the physical space consumed by the infrastructure to support the n-tier solution. A SaaS solution is comprised of a single or multiple 1U rack mounted appliances that can be dispersed among many geographical sites. (1U is 1.7" or 43.66mm.) An on-premise solution will typically require a 3-4 server (at minimum) concentration of infrastructure within the same datacenter, server room and typically even the same server rack in order to operate effectively. This requirement is a by-product of the solution application being sensitive to inter-component communication latency. This concentration increases the cost of the on-premise solution for a variety of reasons, such as moving the current infrastructure and rewiring the newly added hardware, purchasing additional rack space, and downtime during the move. The change control process also introduces delays for implementing the moves. Additionally, with the increasing cost of electricity, simply running the equipment, and the additional cooling required to ensure safe equipment operation, should be considered. A SaaS solution houses the entire log management solution infrastructure off-site at the vendor's datacenter, except for the appliance that collects logs from the sources (sources such as operating systems, applications, routers, firewalls, etc). Once the appliance has collected the logs, the information is securely sent to one or more off-site datacenters that house the physical infrastructure required to analyze and redundantly archive the log data.

With either solution, design includes understanding how the solution will affect current infrastructure implementations. For log management, this is specifically of interest since large amounts of log data are sent across production networks for centralized analysis and archival. The first step in this analysis is understanding how much data will be sent from the log sources to the collection appliance. Assessing the amount of data can be done by using either real world variables collected by monitoring log sources, or by using averages published by industry or vendor sources. Real world variables are obviously the best; however few IT departments have the resources to gather this information and thus depend on the averages. The additional bandwidth usage does introduce risk into the infrastructure environment. Within the on-premise solution, the risk introduced is greater than with the SaaS model. The on-premise traffic pattern must transverse multiple routing points, often within a single site and assuredly within a geographically dispersed network. These routing points are frequently the cause of network

latency since they are natural traffic aggregation points and may become overwhelmed with the increase in traffic flow.

This puts production and business traffic at risk of not reaching its destination, causing business application slowness and affecting business efficiency. The SaaS model helps reduce this risk by removing the need to aggregate traffic within a site and especially back to the central analysis and archive infrastructure. The collection appliance traffic is bound solely for the internet gateway of the network and reduces the number of routing resources required to centrally collect the data.

In summary, the architectural design of a log management solution seems fairly simple, but with the introduction of regulations that require many more log sources to be collected, the amount of log data to be gathered has increased. This increased data creates stress on the current infrastructure that becomes apparent as the design process is begun. The SaaS model reduces the infrastructure resource requirements for implementation and transport of data for log collection, analysis and archive making the SaaS solution more efficient and less costly.

## Installation

After determining the physical and logical placement of an installation, and the impact the new solution will have on current infrastructure, the next step is to implement the solution. The scope of the installation work is significantly different for an on-premise solution compared to a SaaS solution. The physical installation of the SaaS solution would be limited to log collection appliances and would require logical changes only to access the Internet – logical changes such as a firewall rule set change, etc. The on-premise solution will include physical installation, the supporting infrastructure - such as servers or switches, training on the configuration and interoperability of the solution, and monitoring of all installed equipment and software to ensure consistent performance. It is also during installation that the vendor's professional services group usually becomes involved in the implementation, incurring additional and sometimes unforeseen cost.

## Professional Services Costs

Professional services is an often overlooked cost that may not be realized as necessary until late in the solution evaluation - when the stakeholders are already emotionally committed to the proposed implementation. Professional services can, at best, be defined as the paid knowledge transfer between the vendor's solution experts and the infrastructure team supporting the log management solution. Professional services needs are almost exclusively the domain of on-premise solutions. SaaS is specifically designed to require little to no significant customization or development work on the solution before the customer can use it. A characteristic of SaaS solutions are their narrow focus on a problem to solve for the customer. This ensures their product is Customer Needs Driven, meaning significant thought has gone into solving the customer's problem. An on-premise solution will usually have more depth, also known as more complicated features, than a SaaS solution. More complicated features entail more professional services to build the internal knowledge base needed to ensure that the on-premise solution is operational and efficient.

# Monitoring

## Log management system infrastructure

The infrastructure that supports the log management solution requires monitoring and the ability of the infrastructure team to respond quickly to detected issues. The risk involved with discovering an infrastructure problem that has persisted for a period of more than a day is the loss of log data. This loss is significant when attempting to maintain compliance over time. An auditor will typically allow for some loss of data during an infrastructure problem; however, if the problem is persistent, undiscovered for a long time or not responded to with the appropriate urgency, then the log management solution could be considered non-compliant or ineffective. The responsibility of monitoring and ensuring continued infrastructure availability within the SaaS solution is shifted entirely to the SaaS vendor. Even the appliance that collects the log data should be monitored for availability and health by the SaaS log management vendor. The on-premise solution requires that the infrastructure team monitor and appropriately respond to performance issues within the solution. The loss of log data would be the sole responsibility of the infrastructure team.

## Log source reliability

Log source reliability is defined as the level of confidence that the sources configured to send logs to a collection appliance are currently and have been consistently sending logs to the appliance. Log source reliability is one of the key areas within which any solution, on-premise or SaaS, can fail. The essential feature for any solution is to ensure that if a log source stops sending logs to the collection appliance for a pre-defined amount of time (30 seconds to 5 minutes depending on the sensitivity of the log source) that an alert is generated to which the infrastructure team can respond. In too many audits, log management reports show long lags of time (sometimes days) where log sources went offline and there was no awareness of it within the infrastructure management team. This lack of data is often the source of failed or only marginally acceptable audit results.

# Backups and Fault Tolerance

---

The ability to access log data is the most critical and most difficult aspect of log management. Collection can be tricky, but if aggregate log data is not accessible for analysis and review then the entire solution has failed. Access to recent and relatively small amounts of data isn't difficult; access for old and increasingly larger amounts of data is. This section will explore the multifaceted controls required to ensure older log data can be accessed. However, this difficulty is exclusively the domain of an on-premise solution. In a SaaS solution, the SaaS vendor is solely responsible for ensuring that log data is always accessible regardless of the log format, the age, the backup software, format or media used. The economies of scale are irrefutably in favor of a SaaS model for large data storage. It is difficult to cost justify an on-premise solution when those economies are considered.

## Retention

Retention is defined as keeping log data for as long as it is needed or required. For HIPAA, log data retention can be one year or seven years depending on the type of data. For PCI, log data must be retained for one year. The key variable in considering retention is the size or amount of data that needs to be kept. This size will determine how much backup material (such as tapes) is needed, how fast the backup drives are needed to write data within the backup window timeframe, how much off-site media storage space will be used, and what kind of network or fiber switch will be needed to transfer data from the log data retention server to the backup device or server. A complicating factor in retention is restoration of data. As technology changes, being able to restore data can be an issue.

## Restoration

Restoration of data is defined as moving log data from one media to another, most commonly from backup media (such as tape) to a production system or server. Restoration seems straight forward and mundane but when considered over time, it can become more complex. The common issues in restoration are software and media changes. When data storage is required over years, the ability to restore data becomes more difficult. Backup files are not standardized, meaning files backed up from a vendor's application must be restored by that same application. Backup applications from the same vendor are not required to be backwards compatible, meaning past versions of the application are not required to restore data from previous or older software versions. It is then required that old application versions be kept to ensure the ability to restore data. However, as operating systems change, they are also not required to be backwards compatible with older applications, so operating system and even system patches must also be kept. All of these issues are specific to an on-premise solution. The SaaS solution removes this issue from the customer's realm and ensures that data is restored and available as part of the log management solution.

## Time to Recovery

Time to recovery is defined as how long it takes to ensure the log management solution is working as well as before the failure. This would include the ability to view and audit archived data as well as collecting current logs for archiving. The time to recovery in a SaaS model is faster than an on-premise solution and does not require additional resources or costs from the customer when a failure does occur. The number one reason for staff burn out is “fire fighting” or the sudden and unexpected demand for increased resources, which reduces the ability to complete business focused projects. In the SaaS model, multiple layers of redundancy have been built into the architecture of the solution resulting in a faster time to recovery. The cost of this redundancy is greatly reduced by the economies of scale inherent in the SaaS model. SaaS delivers predictable cost and resource demands; something an on-premise log management solution cannot provide.

# Transport

---

Transport is defined as moving logs from one location to another. The transport can be from a server to a log collection appliance, or from the appliance to a central repository located locally or at another geographical site. For highly distributed networks, designing and implementing the ability to transport logs can be complex and expensive. The SaaS model removes many of the complexities and reduces the expense of implementing a log solution in a highly distributed network.

## Highly Distributed Networks

Moving large amounts of log data across a highly distributed network can prove to be a formidable task. The most common issues are bandwidth utilization and firewall rule sets. Bandwidth utilization within a geographically centralized network is usually not an issue; wire speed in that model is usually not below 100Mbps. However, when considering traffic flow over WAN links, which are often at or below 1.5Mbps, the need to reduce the size of log data and manage how much bandwidth is utilized by log transport is crucial. Log transport traffic can flood a WAN link and affect business production traffic negatively. The SaaS model architecture can reduce this potential issue by shortening the path from the log source or appliance to the central repository and by securely utilizing less expensive WAN paths such as internet connections.

## Firewalls

Firewall rule sets are complex and sensitive in nature; a single incorrect statement in a rule set can expose a business to great risk. For that reason, firewall changes are infrequent, often difficult to have approved by a change management organization, and handled by specialized staff. A log management solution for a highly distributed network will require passing log data through several firewall appliances. Ensuring the traffic is flowing correctly requires testing and troubleshooting. In an on-premise log management solution, this transport path through the firewalls can take considerable time. In the SaaS solution, the log management appliances are able to use common secure protocols that are typically already allowed through a firewall. This removes the need to engage additional firewall resources and the change management organization

## Security

Security is defined as controls designed to allow authorized users to view data and ensure unauthorized users cannot view data. For log management, the content of log messages should be considered highly sensitive. The sensitivity is based on log information revealing critical information about the network resources, including IP addresses, domain names, user names, applications, services and patches installed. For either the on-premise or the SaaS model, care should be taken to request and test the controls designed to maintain security.

## Server to collector

Collection of logs from the source to the log management appliance is commonly accomplished using a service called “syslog,” short for system logging. Syslog traffic is sent unencrypted or in clear text. That means, if captured, it is easily read. For that reason, it is recommended that either source to appliance traffic be encrypted using a different transport service than syslog (such as “syslog-ng”) or that the source and appliance are geographically and logically on the same network (within the same VLAN for instance). This is a consideration for either the on-premise or SaaS log management model.

## Collector to central repository

Transport from the log collection appliance back to the central repository should also be considered and secured. In the SaaS model, the central repository is a data center outside of the local network via the Internet on a common secured protocol, SSL on port 443 for example. For the on-premise model, the secured traffic would need to be routed using intra-company WAN links or site-to-site VPNs. Either case is more complex than the SaaS model.

# Log Review's Compensating Control

---

Automated review of log data is a relatively new and very quickly expanding technology. Policy or rule based log analysis will soon be more sophisticated to include anomaly detection based on behavioral heuristics. The key to automated review is knowing before hand what to look for, no matter how the review is being done - be it manual or automated. The more policies and log analysis that is done, the more processing resources will be needed. Processing power is a declining cost resource; however, the operating system hardware, software and often applications (think database applications) have much higher licensing costs based on the number of processors on the server. The risk of underestimating the required processing needs is exclusively an on-premise solution issue. A SaaS solution shifts the responsibility to ensure effective resource allocation and availability to the SaaS solution vendor, allowing greater flexibility and zero risk for the solution buyer.

## Manual Review Issues

Log management solutions have been developed to solve two major issues - storage and review, or analysis. Most government regulations require some log review and industry regulations are as specific as daily review of all log data. Taking a general average of 30 minutes to review the logs for a single server's operating system logs, any environment larger than sixteen servers would require hiring a dedicated person for just log review. Not only is this a terrible waste of human energy, but also an ineffective one. Computers excel at analyzing large amounts of data - humans do not. This contributes to the difficulty of retaining employees who are tasked solely with log data review. Considering this issue, regulations have now changed to allow a compensating control for manual review - automated review.

## Automated Log Review

Automated log review is defined as analyzing logs using predefined rules or policies. A policy is an individual or combined set of rules that have inter-dependencies. Again, the issue with automated review is knowing what symptoms or issues to look for in log data so a rule can be written to detect them. Typical rules include threshold monitoring. An example would be how many times a log message is received within a specific period of time; for example, five "access denied" log messages for the same user name of five different servers within a five minute period. Another type of rule is time based; if an action is taken during a timeframe that is unusual or unexpected. An example would be a user logging in at 3:00AM. Not necessarily an issue, but unexpected behavior that warrants further investigation. Knowing what an auditor will expect in a log review determines how to construct the automated review rules. Additional rules are helpful to detect performance and hardware issues. Once a rule is triggered and an event is created, there still exists the need to investigate and resolve the event if it is proven to be an incident. Ensuring that you track the management of incidents is a universal issue to log management solutions and would not be managed differently depending on the log management solution model; however, you must ensure that your log management solution does allow work flow management. This will provide essential auditor facing reports to prove effective log incident management.

## Conclusion

---

When comparing a SaaS log management solution to an on-premise solution, SaaS is clearly more efficient, effective and costs less to own. Because the SaaS solution requires less bandwidth for log transport, less labor and fewer hardware resources for installation, and all monitoring is part of the SaaS solution, it is a more efficient solution than an on-premise implementation. With a shorter time to recovery, fewer firewall rule set changes and automated log review, SaaS proves to be a more effective log management solution. And finally, because SaaS eliminates software and hardware related costs, no professional services are required and log data retention issues are no longer the responsibility of the customer, SaaS also costs less than an on-premise solution. The distributed nature of log management, the required long term data storage and uptime requirements for both the log collection appliance and central repository are features and requirements easily met by the SaaS model. Choosing a solution for something as complex and critical as log management is difficult and requires careful consideration. As illustrated in this paper, approach log management business needs on a case-by-case basis and this will help determine the solution best suited for the business and regulation(s) under consideration.

## About Alert Logic

---

Alert Logic's patented solutions are the smartest choice for over-regulated businesses with underfunded IT departments to secure networks and ensure compliance. Its cloud-powered managed solutions combine intrusion protection, vulnerability assessment, log management and 24x7 threat surveillance, and are designed to maximize revenue and profit opportunities for service providers and hosting partners. Enterprises experience a solution that addresses network security and compliance requirements at a low price point, with little dependency on IT resources. Alert Logic is based in Houston, Texas and was founded in 2002. More information about Alert Logic can be found at <http://www.alertlogic.com>.

[CONTACT US](#)