Splunk for Microsoft Active Directory Domain Services
Executive Summary

Active Directory has become one of the most ubiquitous technologies in enterprise environments with approximately one billion users logging on to access trillions of objects daily. Few technologies occupy such a critical role for security, compliance, application management and operational intelligence. However, Active Directory is notoriously difficult to manage given its complexity, flexibility and distributed deployment architecture. Most system administrators have also found that native monitoring and auditing tools are inadequate—and third-party tools are often difficult to deploy and expensive. In contrast, Splunk’s flexible architecture, scalability and powerful search language can transform Active Directory from a management headache to a wealth of business insights and operational intelligence.

This paper touches upon the general architecture of Active Directory, its monitoring and auditing requirements and Splunk’s ability to meet those requirements with unprecedented support.

Overview

Microsoft Active Directory Domain Services (or Active Directory) is a foundational technology that has experienced near ubiquitous adoption across all sizes of customers in all industries and public sector verticals around the world. Approximately eighty-seven percent of all enterprises use Active Directory for managing access to, authorization for, and usage of a variety of IT assets, including servers, applications, networks and messaging systems.

Flexibility and coherent policy management has driven this broad and extensive adoption. However, this combination of benefits also creates havoc for IT professionals focusing on security, compliance and accessibility. Active Directory is hierarchical and constantly changing—a single domain controller can manage up to 2 billion objects in its database. The associated object metadata contributes to even more complexity. Throw in multiple domain controllers and Active Directory’s use of various types of trusts and you end up with a chaotic soup of activity that is difficult to manage, troubleshoot or secure.

Fortunately, Active Directory generates a variety of data that documents its inner workings. For the IT professional who can collect, aggregate and correlate all of this data, Active Directory can become not only manageable but an extremely valuable source of information about operations. Since Active Directory is involved in almost all IT operations, visibility into Active Directory usage and changes can produce phenomenal operational intelligence about your organization.

Splunk is perfectly suited for monitoring and auditing Active Directory logs because it matches the flexibility of Active Directory and can scale linearly as your Active Directory environment grows. Splunk can manage and analyze any data from any source type without requiring connectors. In addition, Splunk not only manages Active Directory’s huge amount of data for trending and compliance requirements, it can also handle complex event processing for real-time monitoring and alerting.

Active Directory as a Foundation Technology

Active Directory is a central component of the Windows Server operating system. It functions as a distributed, replicated, multi-master and fault-tolerant directory service. It is used as the repository for storing user identity data, computer and application configuration information, system configuration and security policies.

Active Directory acts as a distributed authentication mechanism for the Windows infrastructure, providing single sign-on for all systems, users and applications with access to the directory. Coupling distributed authentication with policy dissemination allows Active Directory to increase enterprise-wide security by allowing for tightly managed systems which can ensure that corporate-wide compliance with established security policies and best practices is maintained.

Active Directory can contain a vast amount of information related to a plethora of objects (applications, computers, users, groups, distribution lists, network devices, etc.). As such, it can act as a central point for managing these objects. Business management structures often possess a hierarchical structure where the individuals at the highest levels of that structure delegate responsibilities and tasks to those at the lower levels. Active Directory possesses a similar structure where information about objects within the directory are organized according to a hierarchy.

Active Directory is built on open standards such as LDAP. It can act as an integration point for other enterprise applications and systems requiring directory or authentication services or requiring other data for processing related to the Windows infrastructure and its users. Many enterprise applications are now being written specifically to take advantage of Active Directory, including email systems, public key infrastructures, remote access providers, human resource systems, customer relationship management applications and more. These applications act as the mainstay of many day-to-day operations in enterprise environments. As a result, if Active Directory can be properly monitored and audited, the additional Windows workloads can benefit immensely.

The Importance of Monitoring and Auditing Active Directory

Active Directory and directory services in general can be an extremely far-reaching technology component in an enterprise infrastructure. Disruptions to the directory service, including outages, non-optimal configurations, latencies and inconsistent data, have a widespread impact on business operations. The impact may come in the form of lost productivity, disabled...
manufacturing systems, lost sales and more—all ending in additional costs to the organization.

The advent of Active Directory and its adoption for use as the network operating system (NOS) and enterprise directory service has driven the importance of proper management and monitoring of Windows-based systems. In fact, Active Directory features, which allow for very distributed and complex deployments, are in many cases so fault-tolerant that when a true outage occurs the system may go completely down.

Active Directory has become a leverage point for an increasing number of enterprise-wide applications and services, many of which have a global scope. As they rely on Active Directory, it must function with a distributed global environment in mind. Considering that enterprise applications may need to both read and write to the directory, Active Directory acts in a multi-master configuration, allowing changes and queries to be made to any domain controller. As such, objects may be created on different domain controllers spread around the world. Active Directory, via replication, will consolidate the state of these objects—new, modified and removed—but due to the network latency and the potentially distributed nature of the directory data in the Active Directory is seen as loosely consistent or not immediately convergent. Applications and services that rely on Active Directory (such as Exchange 2010, SharePoint 2010 and Office 365) typically expect to get the same data regarding objects regardless of the request’s origin within the infrastructure. As consistency waivers, applications may behave incorrectly. Therefore, ensuring that you know about all changes within the Active Directory infrastructure is essential, particularly given that replication latency is key when deploying business applications on top of the directory.

Given that all users interact with Active Directory to access all objects across the enterprise, user activity monitoring and auditing is extremely important for security and compliance use cases. System administrators and security specialists need audit reports and alerts for:

- User management activities
- Last logon to workstations
- Security and distribution groups
- Membership changes
- User logon activities
- Logon failures
- Domain controller, member server or workstation logon activities
- Passwords changed / set
- Enabled / disabled users
- Account lockouts
- Creation, modification or deletion of AD accounts
- Domain policy changes

These reports should be easily consolidated and formatted into reports used for regulatory compliance, as well as making logs easily accessible for reference during audits or forensics investigations.

Log Monitoring for Active Directory in Conjunction with Windows Server

The Event Log service is automatically started when a Windows machine starts. All users can view application and system logs. However, only administrators can gain access to security logs. Security logging is turned off by default. To ensure that a security log is available it should be turned on by the administrator. Therefore, system administrators will need to turn on security event logs for monitoring Active Directory.

Windows and Active Directory have several different log types that should be monitored:

- Application log these are events logged by applications.
- Security log this log contains records of valid and invalid logon attempts and events related to resources use, such as creating, opening, or deleting files or other objects. This log is also customizable.
- System log contains system component event. Driver failures and hardware issues.
- Active Directory Domain controllers have two extra logs for the directory service.
- File Replication service log containing windows File Replication service events. Sysvol changes are recorded in the file replication log.
- DNS machines also store DNS events in the logs.

Each log can contain a wide variety of data: errors, warnings, information, success audits and failure audits. On any busy machine or on any busy network many hours are logged and megabytes of data are generated making it difficult or impossible to monitor all of the data of networked computers by traditional means.

Logging Requirements for Active Directory

There are a standard set of requirements that Windows system administrators have agreed upon over the years, particularly with Windows Server 2003, Windows Server 2008 and Windows Server 2008 R2. These requirements are derived from security and compliance use cases primarily, but can also be associated with application management and change management use cases. The following requirements come from WindowsSecurity.com. The table in the Appendix is a comparison of native tools and third-party tools and the relative weaknesses of both.

- Real-time monitoring and notification. It’s important that a security event is immediately flagged and notification is sent to an organization’s security team. (Windows is unable to notify security professionals of triggered events.)
- Audit trail unconsolidated in Windows. Since individual machines hold isolated event logs, viewing event logs is extremely difficult and time consuming. It is much easier to look at one event log to get a current network status than to look at multiple event logs and miss information because of the vast amount of entries that have not been filtered. So it is ideal to have a central log monitoring system that security professionals can use at a glance.
• **Security logs monitored remotely.** This means that when intruders attempt to use local accounts to log into the machine the audit trail is limited to the local security logs.

• **Critical events described in plain english.** In the normal Microsoft tradition of critical event notification “event 12345%$#” means your server was rebooted or the like. This data is typically cryptic and misleading. Consolidation and remote log reading applications have alerts that can be preprogrammed for specific events to make the administrators life much easier than deciphering the misleading logs.

• **Data archived for better systems management.** Institutions such as banks are required in most countries to keep audit logs for over 7 years and even longer in some circumstances. The typical Windows default configuration is set to overwrite over the logs when a certain file size is reached. The other issue is that users have to physically archive and clear the logs. Automating and centralizing this process increases productivity on a large network environment by reducing support calls and letting administrators see what is happening locally on a user’s machine.

• **Log file integrity maintained.** Files stored on a user machine have less integrity as the user can clear the logs quickly or an intruder after gaining access can cover his tracks by clearing the event logs. Intruders sometimes produce an excessive amount of event-triggering actions to fill up security logs to camouflage their activity. By using a product like Splunk, the security professional can be alerted to this phenomenon and can react to it immediately; furthermore the logs are stored remotely so the user or intruder cannot erase them.

• **Log filtering.** Data overload is a huge issue log monitoring applications have the ability to filter out irrelevant noise events that take up time and space and only display the pertinent logs.

• **File access and monitoring for important files.** This can be achieved by auditing failed access to these files, enabling you to find out if someone is attempting to access the files.

• **Remote alerts.** An application that can alert the security professional by SMS, email and pager prove valuable as the administrator may not be in the proximity of a computer at all times. The administrator can then react or have systems in place the can be remotely activated to stop a potential attack.

• **Web server log monitoring.** This important step is often overlooked by hasty administrators. By using software that monitors your local or remote web server you can add an extra layer of security to your web server. This is where the alerting functionality of log monitoring software is useful because it sometimes is challenging to monitor servers that are on the DMZ.

• **Data logging and storing.** Saving data in a flat file makes it possible to index and search against this data. This methodology is preferred in an enterprise environment, as scalability, speed and ease of use is much more attainable. However, most third-party vendor products still use structured, searchable databases for this unstructured data.

• **Log events categorically sorted into prioritized sections.** Software should allow the security administrator to view high profile security events at a glance while hiding medium-profile or low-profile security events.

• **Data properly monitored and cleared.** Clearing of logs should also be monitored as only the administrator should be able to clear security logs.

• **Log particular events on critical machines.** This is an important function as machines that need to remain secure should be monitored at a more granular level.

### Splunk for Active Directory

Splunk Enterprise takes a different approach to monitoring and auditing Active Directory environments. It combines Windows Security Event logs with data collected by Splunk’s native monitoring utility, Active Directory Monitor. The AD Monitor enables Splunk to monitor changes to AD and maintain a cached view of users, computers and group objects. This provides two distinct benefits. First, Splunk users will be able to apply Splunk search technology on AD’s object lifecycle for monitoring, auditing and administrative purposes. Second, it enables Splunk to have an efficient mechanism to map SIDs and GUIDs from event logs throughout the enterprise into logical names.

Splunk’s Active Directory Monitor captures its state in the local computer registry. Its state consists of the last domain controller name, the invocation GUID of the domain controller and the last USN change. Upon startup, the monitor tries to bind to the last domain controller it monitored and if the domain controller is not available, it will rebind to any available domain controller and start a snapshot process.

If the same domain controller is available then the monitor will start the query process at the last known USN by registering for change notifications at the USN. If the last USN is not available or the invocation GUID is different then a snapshot process is executed as if it’s a different domain controller. This handles the cases of restoring and retiring domain controllers while remaining highly available in the presence of multiple domain controllers.

<table>
<thead>
<tr>
<th>Security Event Log</th>
<th>AD Monitor Event</th>
<th>Report</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-in</td>
<td>Change phone number.</td>
<td>Appropriate person could change the phone number.</td>
</tr>
<tr>
<td>Administration rights change</td>
<td>User and profile of the individual who made the change.</td>
<td>Verify that user has sufficient permissions to make changes.</td>
</tr>
<tr>
<td>User Lock-out</td>
<td>Names of users who were attempting to login or access a particular object</td>
<td>Determine if lock out was caused by unauthorized users or if object was unavailable.</td>
</tr>
<tr>
<td>Account Expired</td>
<td>User status changed to &quot;Terminated.&quot;</td>
<td>A former employee tried to access object units.</td>
</tr>
</tbody>
</table>
Combining change monitoring data with security event log data provides greater visibility than the typical reports generated by other products. The table on the next page illustrates the kind of reports that are possible with Splunk. Two types of data are combined to give you greater visibility for your IT infrastructure and AD deployment.

Splunk Features

Index Any Machine Data, From Any Source
Splunk indexes machine-generated data from any source in real time including your live custom and packaged application logs, stack traces, message queues, database audit trails and even logs, status, configurations and metrics from your hypervisor, OS and network layers. No matter the source or format, Splunk indexes it the same way - without specific parsers or adapters to purchase, write or maintain.

Search and Investigate Anything
With Splunk you can search live streaming data and indexed historical data using the same interface. Familiar Boolean reporting commands let you update transaction counts, calculate metrics and even look for specific conditions within a rolling time window. Search Assistant offers type-ahead and contextual help so that you can leverage the full power of the Splunk search language. You can interact with search results in real time. Zoom in and out on a timeline to quickly reveal trends, spikes and anomalies. Click to drill down into results. Real-time search means you see incidents on streaming data and track live transactions and online activity.

Add Knowledge
Splunk extracts knowledge from your machine data automatically at search time so you can start using new data sources immediately. You can also add context and meaning to your machine data by identifying, naming and tagging fields and data points. Splunk even lets you add information from external source asset management databases, configuration management systems and user directories, making the system smarter for all users.

Report and Analyze
Use the report builder to quickly build advanced charts, graphs and dashboards that show important trends, highs and lows, summaries of top values and frequency of occurrences. Create robust, information-rich reports from scratch without any advanced knowledge of search commands. Save reports, integrate them into dashboards and create PDFs on a scheduled basis to share with management, business users or other IT stakeholders.

- Splunk provides a full complement of search commands for reporting. Here’s a sample list:
  - Stat – Find statistical values (min, max, variance, avg, mean, count, etc) for fields
  - Top – Find top values for any field
  - Rare – Find rare values (opposite of top) for any field. It’s good for anomalies
  - Chart – Split statics from one field on the X axis by another on the Y axis
  - Timechart – Create charts that use time on the X axis
  - Table – Get a tabular view of any data using fields that the user specifies
  - Eval – Compute mathematics and assignments to create new fields

- All graphical report commands can output charts on the UI in variety of ways when appropriate using pie charts, columns, line graphs, area graphs, etc.

Create Custom Dashboards
Create live dashboards in a few clicks with the dashboard editor. Dashboards integrate multiple charts and views of your real-time data to satisfy the needs of different users. You can personalize dashboards for different user roles. Dashboards can be viewed live or scheduled for delivery via PDF.

Splunk lets you conduct form searches. Type into a form an Order ID and several concurrent searches can run at once on a dashboard. One will show the events from different logs. Another may show a correlation of the order ID with fields from a database using Splunk’s look up command. Still another may show the history of the Order via a graph. Each of these search results is put on a panel in the dashboard for the user to further drill down. For example, here’s a dashboard asking Splunk to find all information about a source IP address that is involved in many failed login attempts. You will see correlations to external DNS and country look ups, graphs on counts of users trying to log in and the raw data that involves this IP address.

Conclusion
Splunk can help fulfill the monitoring and auditing requirements for Active Directory no matter the scope or complexity of the deployment. While many tools can provide monitoring and auditing for Active Directory their limitations often appear once the deployment grows beyond multiple domains or forests, which typically occurs within enterprise environments. Traditional tools can be difficult to deploy and cannot scale to handle the large data volumes of a centralized logging scheme.

Splunk software lets you search, report, monitor and analyze live streaming and historical data across your IT infrastructure from one place in real time. Splunk offers unique visibility into Active Directory data that represents user transactions, customer behavior, machine behavior, security threats and fraudulent
activity. Use Splunk to troubleshoot application problems and investigate security incidents in minutes instead of hours or days, avoid service degradation or outages, deliver compliance at lower cost and gain new business insights.

Splunk indexes machine-generated data from any source in real time including your live custom and packaged application logs, stack traces, message queues, database audit trails and even logs, status, configurations and metrics from your hypervisor, OS and network layers. No matter the source or format, Splunk indexes it the same way—without specific parsers or adapters to purchase, write or maintain.

Appendix

Traditional Approaches for Monitoring and Auditing Active Directory

<table>
<thead>
<tr>
<th>Capability</th>
<th>Native Tools</th>
<th>Third-Party Tools</th>
<th>Missing Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consolidated audit trails</td>
<td>Provides audit trails only for one forest.</td>
<td>Can consolidate logs from several forests and address replication latency.</td>
<td>Does not correlate security event logs or WMI data of various objects and AD.</td>
</tr>
<tr>
<td>Remotely monitor security logs</td>
<td>No ability.</td>
<td>Collect data from multiple geographic locations.</td>
<td>Set-up of tools tends to be very complex and maintenance is high.</td>
</tr>
<tr>
<td>More obvious description of critical event</td>
<td>No ability.</td>
<td>Limited ability — reports describe error codes.</td>
<td>Changes are not correlated with other events that have occurred in the infrastructure, making root-cause analysis very difficult</td>
</tr>
<tr>
<td>Archiving of logs</td>
<td>Limited ability</td>
<td>Limited ability — SQL server. Scale and data accessibility become issues.</td>
<td>Consolidation of logs across an entire forest becomes challenging to set up, particularly agent based solutions.</td>
</tr>
</tbody>
</table>

Free Download

Download Splunk for free. You’ll get a Splunk Enterprise license for 60 days and you can index up to 500 megabytes of data per day. After 60 days, or anytime before then, you can convert to a perpetual Free license or purchase an Enterprise license by contacting sales@splunk.com.