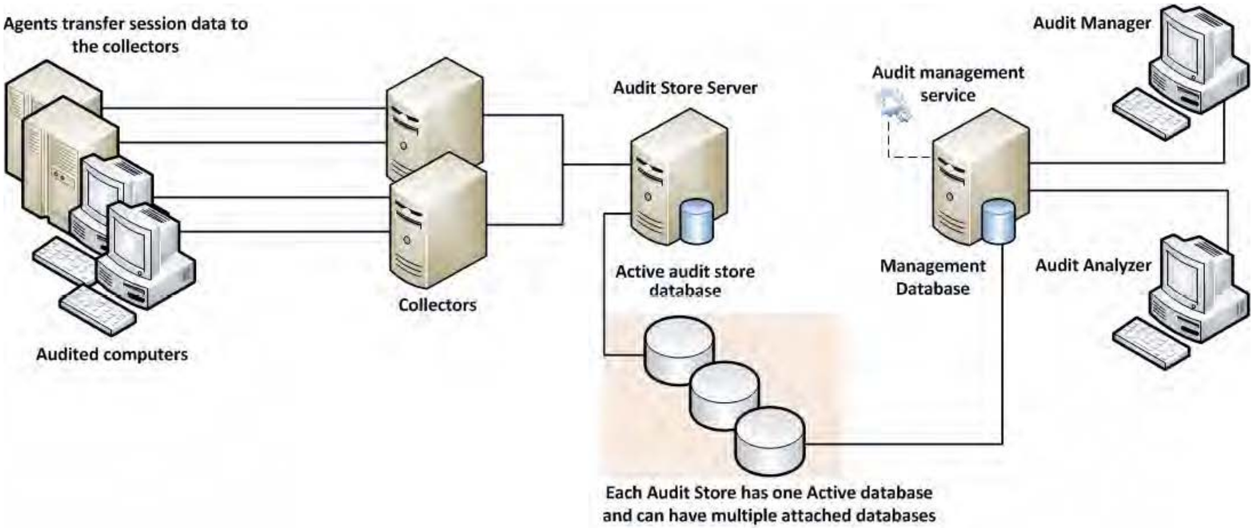
WHITE PAPER

Centrify DirectManage Audit Sizing Considerations and Best Practices

WWW.CENTRIFY.COM

Contents

Introduction


Agents transfer session data to the collectors
Audited computers
Collectors
Audit Store Server
Active audit store database
Each Audit Store has one Active database and can have multiple attached databases
Audit management service
Management Database
Audit Manager
Audit Analyzer

(Fig. 1 – Centrify DirectManage Audit Architectural Overview)

A typical deployment of Centrify DirectManage Audit consists of a number of components such as one or more Audited Systems (Unix/Linux or Windows), one or more Collectors, Audit Management Server, Management database, one or more Audit Store databases and Consoles (viz. the Audit Manager and the Audit Analyzer consoles) which all communicate with each other. Given the complexity of this communication and number of components involved, good planning is important for a successful deployment of the product. When planning a deployment, some of the most common questions that we asked are below:

1. Will just one installation of DirectManage Audit suffice? Or are multiple installations needed or recommended?
2. How many Audit Stores need to be provisioned in each of the installations and how should their scope be configured?
3. How many Collectors will be needed and what kind of hardware is recommended for each of them?
4. What is the recommended version/edition of SQL Server and what kind of hardware is recommended to host this SQL Server?

You must take into consideration a number of factors when deciding how DirectManage Audit deployment will be planned and configured, and what kind of hardware will be needed to deploy the key components. This whitepaper will help you understand these factors in detail and come up with answers to such questions.

Planning a Centrify DirectManage Audit deployment

System Integrators often rely on the number of Audited Systems to estimate the hardware requirements and to come up with the overall strategy of DirectManage Audit deployment. For example, an environment with 100 Audited Systems may look like a small setup and one may incorrectly conclude that it's a small scale deployment that won't require a powerful hardware to support it. Once setup however, such assumptions may turn it into a deployment that seldom scales and often produces poor performance, both when capturing the audit activity and when querying the already captured audit data.

Given below are a few factors that one must consider before making any deployment decisions,

1. SQL Server - Out of all the components in the DirectManage Audit ecosystem, SQL is the most heavyweight and will share most of the burden when it comes to workload. Using a properly equipped and optimally configured SQL Server is very important. The version and edition of SQL Server being used (such as Express or Standard or Enterprise) or the type of machine being used to host the SQL Server (such as a virtual or physical machine) can noticeably improve the overall performance. On the contrary, a poorly configured SQL Server may produce a very poor performance no matter how powerful the underlying hardware is.

2. Number of concurrently audited users - Relying on the number of Audited Systems is not always a good assumption. For example, an environment may have just a handful of systems but may have a large number of users logging into these systems on a daily basis. A jumpbox scenario such as Citrix XenApp Server is a perfect example. When planning, you should plan for the number of concurrently audited users, not just the total number of Audited Systems. User activity patterns and behaviors also play an important role in overall performance and storage requirements. For example, the audited data will be much smaller in an environment where no logins are expected most of the time as compared to a network control systems wherein audited users are logging on and logging out throughout the day. The sizing guidelines specified in the later section of this whitepaper have all been based on workload simulations for the exact same reason.

3. What needs to be captured - What's being captured controls the overall workload on various components. Capturing video is more expensive than not doing so in terms of disk usage and load on Collectors and SQL Server. Similarly, capturing interactive sessions is always going to produce more audited data when compared to capturing a handful of commands thus putting system under more pressure. Capturing large quantities of data has another side effect; it slows down database backups and other maintenance processes which is not always liked by the database administrators.

4. Who needs to be audited - Who is being audited is equally important. Under default settings, Centrify DirectManage Audit Agent audits everything and everybody and this may not be a practical solution in many large environments. In production environments, it's very common to see processes or scheduled tasks that periodically monitor Unix/Linux or Windows systems for their health by remotely executing certain commands (System Monitoring and Management software, such as BMC Patrol that periodically runs vmstat or iostat command on each of the Unix/Linux systems is a good example). Activities like these needlessly generate thousands of Audited sessions on a daily basis and in many cases create tremendous load on an entire DirectManage Audit system.

5. Unix/Linux vs. Windows - Type of system being audited influences the amount of data that'll be captured from that system and the overall CPU load on Collectors. e.g. A Windows Audited System almost always generates more data per day compared to a Unix Audited System with comparable number of concurrent users. This also means that an environment with Windows Audited Systems will most likely be more demanding (in terms of hardware resources) compared to an environment with same number of Unix/Linux Audited Systems.

6. Query performance - Query performance is one factor that often gets ignored. Capturing user activity and storing it in the database in a reasonable time is important. What's also important is to be able to search these records in a predictable time frame irrespective of the combined size and number of all the databases in the Centrify DirectManage Audit system.

7. Audit data retention policy – Audit data retention policy dictates how many days of data should be online and readily available for querying purpose and this number varies from one enterprise to another. Pay special attention to data retention policy requirements in the target environment. A longer retention policy typically results in large databases which also suffer from poor query performance if databases are not well maintained. On the contrary, too frequent rotation will also result in poor query performance if you keep too many inactive databases attached to the Audit Store.

8. System overheads - Keep in mind the overhead that is caused by the DirectManage Audit system itself; there are a number of background jobs carried out by various components of the DirectManage Audit system, including the Audited Systems themselves, Collectors, and the Audit Management Server. This includes activities such as sending the Audited System's heartbeat to the database (via Collector), sending the Collector's heartbeat to the database, processing active sessions list, processing and synchronizing information of Audit Roles with Active Directory Group criteria, calculating effective size of audited sessions, storing license usage information in Active Directory, and many more.

9. Latency – Geography/Network topology play an important role as it introduces latency. For example, an environment may well have just a handful of Audited Systems but if they're not geographically co-located, you may see delays in getting the audited user activity to its final destination (the database server); the same may happen if Audited Systems are not connected to Collectors by a network link with reasonable bandwidth. A general rule of thumb is to group together Audited Systems, Collectors and databases that are connected by a high speed network using the concept of Audit Store.

Best practices

The previous section listed out a number of factors that may affect how a DirectManage Audit system will be deployed. Below is a set of best practices that are derived from these factors. Follow these practices for planning any DirectManage Audit deployment (large or small). You can also refer to the last section of this whitepaper that discusses how to tweak settings in an existing environment to improve performance.

1. Plan based on concurrently audited users - When planning, always focus on the number of concurrently audited users, not just the total number of Audited Systems. Take into consideration user sessions that might be generated as a result of automated monitoring activity from System Monitoring and Management software, such as BMC Patrol etc.

2. Avoid single box deployment – Always avoid installing key components such as SQL Server, Collectors and Audit Management Server on the same system, especially in environments with heavy workload. Keep in mind that a Collector's workload is CPU intensive and SQL Server's workload is CPU, IO, and memory intensive. If both a Collector and SQL Server are installed on the same system, they'll slow each other down.

3. Control the amount of data – It's always a good practice to establish rules to avoid capturing unnecessary data. This typically includes blacklisting commands such as `top` or `tail` (which generate large outputs and seldom contain any meaningful user activity) or enable per-command auditing instead of session auditing. Also, compile a list of users that do not really need to be audited and add them to the non-audited user's list. This often includes user accounts that are used to run automated jobs from System Monitoring and Management software, such as BMC Patrol and so forth.

4. Scope the Audit Stores efficiently – Always visualize the flow of traffic, not just when audited activity is being captured but also when it's being searched and replayed. It's better to avoid traffic over slow links by splitting the Audited Systems into multiple Audit Stores based on their geographic location, even if it may mean that you'll be deploying more Collectors and SQL Servers. In certain cases, splitting Audited Systems into multiple Audit Stores may not be sufficient enough and you may even need to consider provisioning multiple DirectManage Audit installations. When audited data is being queried, all calls are routed to the Audit Store databases via the Management database. If the Management database is not connected to the console or to the Audit Store databases via a fast network link, the queries will always return the results slowly no matter how good the performance of SQL Server is.

5. Estimate storage requirement based on pilot data – No two customers are the same and you can never accurately predict how much data will be collected over a period of time in each environment. Hence, it's important to analyze existing data in a customer's environment (from pilot project) to predict the future data growth. A pilot testing is an effective way to help you understand a number of things such as the following factors:

a) Understand workload patterns and come up with an overall configuration strategy that determines how the Audit Stores will be scoped, which users should or should not be audited, which commands should be blacklisted etc.
b) Database storage requirement – Roughly, how much data will be collected over the retention policy period? This will also help you establish the active Audit Store database rotation policy.
c) What kind of hardware will be needed for the SQL Server to serve the production workload?
d) How many collectors will be needed in each Audit Store (this number is especially important when auditing Windows systems)?

The DA Data Analysis tool (see KB-4496) can be very helpful to understand data trends. If the DA Data Analysis tool reveals that more than anticipated amount of data is being captured, you can always use the database rotation to keep the active Audit Store database's size in control thus controlling the storage requirements for all attached databases.

6. Maintain databases periodically – Apart from taking regular backups, it's also important to keep the databases healthy by maintaining them periodically. This includes activities such as reorganizing or rebuilding indexes; these tasks must be done by a customer's DBA periodically. Centrify recommends reorganizing indexes if they are 5% to 30% fragmented and rebuilding indexes if they are more than 30% fragmented.

7. Control the size of active databases – A large active Audit Store database often results in poor performance as a result of fragmented indexes, lengthy backups, and out of date database statistics, especially when the databases are not maintained periodically. Centrify recommends keeping the active Audit Store database size between 250GB-500GB (as of Suite 2016). Consider rotating databases whenever the size exceeds the recommended thresholds. You can rotate databases programmatically by using either the Centrify DirectManage SDK or the Centrify Audit PowerShell Module, or manually using the Audit Manager console). It's also a good practice not to keep too many Audit Store databases attached to an Audit Store, because doing so affects query performance.

8. Plan database rotation based on retention policy – Always try to align the audit data retention policy with the active Audit Store database rotation. For example, if the audit data retention policy requires last 90 days of data to be online, try to rotate the active Audit Store database every 90 days. This strategy makes it easy to find achieved data if it's ever needed for reviewing purpose in the future. One exception to this strategy is an environment where the audit data retention policy is so long that the active Audit Store database is guaranteed to exceed the recommended maximum size of the active Audit Store database (as mentioned in the previous section). In such cases, you can divide the entire retention policy period into small periods (e.g. one database for each month) and continue to rotate the active Audit Store database at the recommended intervals. Irrespective of which strategy you choose and implement, it's always recommended to detach all Audit Store databases that contain data outside of the retention policy period. This not only improves the query performance but also reduces the disk usage on the database server.

9. Configure SQL Server optimally – Centrify recommends setting the SQL Server machine's power plan settings (Control Panel > Power Options) to High Performance.

SQL Server has a setting called Max Server Memory that controls the maximum amount of physical memory that can be consumed by the SQL Server's buffer pool. An incorrectly configured Max Server Memory may either result in the SQL engine causing high IO or OS/other programs starving for more memory. It's critical to configure the Max Server memory correctly based on the amount of total physical memory available. Refer to the *Centrify DirectManage Audit Administrator's Guide* and always configure this value as recommended before deployment begins.

Centrify recommends storing the transaction logs and data files that are associated with any SQL Server database on two separate volumes. For more information, see the Microsoft Knowledge base article https://support.microsoft.com/en-us/kb/2033523.

10. Other recommendations – Centrify recommends deploying at least two Collectors per Audit Store for redundancy purpose.

11. Understand that any hardware has its limits – It's entirely possible that even after following all the best practices, the Centrify DirectManage Audit system continues to perform poorly. In such cases, you must consider splitting the workload by deploying additional SQL Servers or Collectors, depending on where the bottleneck is. Deploying an additional SQL Server will almost always result in reconfiguring scope of the Audit Stores (in order to redirect some traffic to the new SQL Server) and it must be done with careful planning.

Guidelines for determining hardware configuration

The overall performance of the Centrify DirectManage Audit ecosystem ultimately depends on the performance of SQL Server and the Collectors. To come up with guidelines for hardware, we have created a test environment wherein the SQL Server hardware configuration has been categorized into three variants: a low end SQL Server, a high end server SQL Server, and a mid-level SQL Server. Below are the test environment configuration details:

	Low end hardware specification	**Mid-level hardware specification**	**High end hardware specification**
Physical machine	DIY PC	S5000 Intel Xeon	Dell R730
Physical memory	8 GB (2x4GB)	16 GB (2x8GB)	32 GB (2x16GB)
CPU	Intel i5-650, 3.2 GHz	E5420 (2.5 GHz)	2xIntel Xeon E5-1620 v3 (2.4 GHz, 8C/16T)
HDD	1x1TB (7200 rpm SATA)	1x1TB (7200 rpm SATA)	1x1TB (7200 rpm SAS 6Gbps)

* The Hardware configuration depicted in the above table reflects the sizing test environment. Centrify cannot make specific recommendations (such as physical memory, CPU frequency, or CPU type) for purchasing hardware; use these numbers only as a guideline.

The table below lists the test conditions along with the outcome of tests, and this roughly indicates the recommended number of Audited Systems that can be supported in this test environment.

	Unix Agent (session auditing)	**Unix Agent (command auditing)**	**Windows Agent (video enabled)**	**Windows Agent (video disabled)**
Test conditions	60% agents are idle 35% agents are running simple commands 5% agents are running tail command	5% agents are idle 2% agents are running “su” sessions 93% agents are running “dzdo” command sessions	60% agents are idle 40% agents are active	100% agents are active
Low end SQL Server	1100	1800	400	1300
Mid-range SQL Server	1500	3600	400	2400
High end SQL Server	2000	4500	640	3000

- The numbers depicted in the above table reflects the outcome of a sizing test in a very specific test; use these numbers only as a guideline.
- Refer to the table in the next section for actual recommendations.

Based on these test results, Centrify recommends using the table below when planning a deployment of Centrify DirectManage Audit. Please note that the recommended SQL Server configuration is only applicable to the SQL Server hosting the Audit Store database. For recommendations on the SQL Server that hosts the Management database, see the *Centrify DirectManage Audit Administrator's Guide*. It's generally a good practice to host the Management database on the same SQL Server where the other Audit Store databases are hosted.

Audited System type	Audit Type	Number of Audited Systems	Expected activity	Recommended SQL Server configuration	Recommended number of Collectors	Average response time (ms)
Unix	Command auditing	1800	5% agents are idle 2% agents are running "su" sessions 93% agents are running "dzdo" command sessions	Low end	2	83
Unix	Command auditing	3600	5% agents are idle 2% agents are running "su" sessions 93% agents are running "dzdo" command sessions	Mid-range	2	60
Unix	Command auditing	4500	5% agents are idle 2% agents are running "su" sessions 93% agents are running "dzdo" command sessions	High end	4	102
Unix	Session auditing	1100	60% agents are idle 35% agents are running simple commands 5% agents are running tail command	Low end	2	87
Unix	Session auditing	1500	60% agents are idle 35% agents are running simple commands 5% agents are running tail command	Mid-range	2	76
Unix	Session auditing	2000	60% agents are idle 35% agents are running simple commands 5% agents are	High end	4	104

Audited System type	Audit Type	Number of Audited Systems	Expected activity	Recommended SQL Server configuration	Recommended number of Collectors	Average response time (ms)
			running tail command			
Windows	Video disabled	1300	100% agents are active	Low end	2	91
Windows	Video disabled	2400	100% agents are active	Mid-range	3	67
Windows	Video disabled	3000	100% agents are active	High end	4	100
Windows	Video enabled	400	60% agents are idle 40% agents are active	Low end	5	85
Windows	Video enabled	400	60% agents are idle 40% agents are active	Mid-range	5	88
Windows	Video enabled	640	60% agents are idle 40% agents are active	High end	8	113

- Expected activity is based on 8 hours of work every day. Results may vary if the target environment has a different pattern for user activity/behavior, different workload/ratio of idle to active systems compared to the test environment.
- Average response time is the total time taken in milliseconds to send a unit of data from Audited System to the SQL Server via Collector.
- All recommended numbers are based on the assumption that the target environment is stable in terms of performance of individual components and network throughput. Intermittent transient errors are expected and typically do not impact the sizing assessments.
- Windows Audited System generates large amount of audit data when video capture is enabled and such environments require high performance SQL Server storage. This is the primary reason why the number of agents supported between the low and medium SQL Server configuration are similar. The artificial load generated by the test simulators is also higher than the expected daily activity in a typical production environment. With high performance storage, the total number of Windows Audited Systems supported will likely be higher compared to the numbers recommended in this whitepaper.
- When monitoring both Windows and Unix/Linux Audited Systems in the same environment, use the Windows numbers as a guideline.

Identifying typical deployment issues

It's fairly easy to identify scaling/performance issues with a Centrify DirectManage Audit system that are typically a result of poor planning or deployment. Below are some of the most common deployment issues.

1. Large spool files on Audited Systems - A healthy DirectManage Audit system should be able to keep up the pace with users' audited activity. When the system cannot keep up the pace, it means either the user's audited activity is generating too much data (such as when a user runs the `cat` command on a very large file) or the Centrify DirectManage Audit system components (such as Collectors and databases) are not able to process and store the generated data fast enough. In such cases, you'll typically see large spool files on the Audited Systems that often need more time to get despooled completely.

2. Constant high CPU on Collector/SQL Server - It's perfectly normal to see high CPU activity on Collector and SQL Server machines during peak hours as this is the time when data is continuously getting pumped from the Audited System to the Collector and finally to the database. However, when you see similar activity during off-peak hours (especially when it doesn't correspond to the number of active users in that environment at that time), it indicates that the DirectManage Audit system is getting backlogged.

3. Low despool rate - The despool rate largely depends on the type of data being captured, the speed of network/latency between Audited System and Collector, the speed of the network/latency between the Collector and the database, and ultimately the performance of the SQL Server itself. Because of these factors, there's no ideal value or range for the despool rate. However, you should not see a despool rate that's significantly lower than the rate of data capture, especially when there are no known issues related to network speed or SQL Server performance.

4. False "Agent disconnected" alerts - Each Agent periodically sends its heartbeat to the database (via Collector) and the Audit Manager console relies on this ping to determine if the Agent is connected or not. If there are deployment issues with Centrify DirectManage Audit, the Agent heartbeat may not get registered even if the Agent is online, and this may raise false alarms as the system will be shown as disconnected in Audit Manager Console. Whenever you see such contradicting information regarding the status of the Audited System, it typically is indicative of underlying deployment issues.

5. Too many SQL Server tasks in queue – SQL Server has a fixed set of worker threads that it can use to perform its job and this number depends on the CPU architecture, such as 32-bit or 64-bit, and the total number of CPUs on the SQL Server. If SQL Server is given more tasks than it can finish, they'll end up waiting at the bottom of this queue, thus consuming memory and degrading overall system performance. Always consult the DBA to confirm if the environment is consistently showing a lot of tasks in the worker queue; this can indicate that the workload is too much for this SQL Server to handle. For more information, see the Microsoft article https://msdn.microsoft.com/en-us/library/ms177526(v=sql.105).aspx.

Settings to adjust for performance improvement

In an environment where DirectManage Audit is already deployed and experiencing scalability/performance issue, it's not always possible to re-architect the deployment or make significant configuration changes (such as re-scoping the Audit Stores or adding a new SQL Server may not be practical); this is true especially in large environments. The table below that lists some key settings that you may try to change in order to improve the overall performance of various DirectManage Audit components.

Title	Summary	When to tweak	Component	Additional details
		Agent settings		
Agent heartbeat interval for Unix/Linux Audited Systems (dad.timer.update.agent.status)	Controls the interval for sending Unix/Linux Audited System's heartbeat to the Collector	When SQL activity monitor shows high CPU usage at a predictable interval as a result of heartbeat registration or when Collector logs reveal repeated failures in registering Audited System's heartbeat or when Audit Manager console frequently shows a lot of disconnected Audited Systems even if they are all online.	Unix/Linux Agent (centrifyda.conf)	For more information, see the *Centrify DirectManage Audit Unix Configuration and Tuning Reference Guide*.
Agent heartbeat interval for Windows Audited Systems (SessionPingInterval)	Controls the interval for sending Windows Audited System's heartbeat to the Collector	When the SQL activity monitor shows high CPU usage at a predictable interval as a result of heartbeat registration or when Collector logs reveal repeated failures in registering Audited System's heartbeat or when Audit Manager console frequently shows a lot of disconnected Audited Systems even if they are all online.	Windows Agent (registry setting)	For more information, see the *Centrify DirectManage Audit Administrator's Guide*.
User blacklisting (dash.user.skiplist)	Allows specifying blacklist of users that should not be audited on Unix/Linux systems	Useful in preventing capture of audit activity of users such as BMC Patrol agent or ServiceNow service accounts or users that do not really need to be audited.	Unix/Linux Agent (centrifyda.conf) and also available via Group Policy	For more information, see the *Centrify DirectManage Audit Unix Configuration and Tuning Reference Guide* and the *Centrify DirectManage Audit Administrator's Guide*.

Title	Summary	When to tweak	Component	Additional details
Audited/Non-audited users list	Allows specifying whitelist or blacklist of users that should or should not be audited on Windows systems	Useful in preventing capture of audit activity of unwanted users.	Group Policy	For more information, see the *Centrify DirectManage Audit Administrator's Guide*.
BindingCheckInterval	Controls the interval at which Agent checks if it's connected to the correct Collector or not	When binding check causes load on the Domain Controller as a result of periodic Active Directory calls (e.g. when you notice AD call from each Audited System every 10 seconds)	Windows Agent (registry setting)	For more information, see the *Centrify DirectManage Audit Administrator's Guide*.
		Collector settings		
Agent global heartbeat interval (AgentMinimumUpdateInterval)	Controls the interval for sending Audited System's heartbeat to the Collector at the Collector level (in case it's not practical to tweak this setting on each of the Audited Systems)	When SQL activity monitor shows high CPU usage at a predictable interval as a result of heartbeat registration or when Collector logs reveal repeated failures in registering Audited System's heartbeat or when Audit Manager console frequently shows a lot of disconnected Audited Systems even if they are all online.	Collector (registry setting)	For more information, see the *Centrify DirectManage Audit Administrator's Guide*.
Maximum concurrent SQL connections per Collector (MaxPoolSize)	Controls how many SQL connections (maximum) can be opened by the Collector at a time	In order to reduce the workload caused by Collector on the SQL Server. Reducing the MaxPoolSize will reduce the total number of connections open on the SQL Server but may also reduce the despool rate.	Collector (registry setting)	For more information, see the *Centrify DirectManage Audit Administrator's Guide*.
		Installation level settings		
Command blacklisting	Allows specifying one or more commands whose output is not required to be captured	When you see large audited sessions that are a result of running commands with large output (e.g. commands such as tail or top) and you need to control disk space consumed by such audited activity.	Group Policy	For more information, see the *Centrify DirectManage Audit Administrator's Guide*.
Enable/Disable video audit	Allows enabling or disabling	When video capture is resulting into large sessions	Audit Manager	For more information, see the

Title	Summary	When to tweak	Component	Additional details
	video capture (at installation level or on a per machine basis) when storing audited user activity in the database	consuming a lot of disk space and/or it's not desirable to store the video.	console or Group Policy	*Centrify DirectManage Audit Administrator's Guide*.

- Not all configuration parameters/settings are available in releases prior to Suite 2015.1. Please contact Centrify Support for additional information on older releases.
- Agent heartbeat interval can configured per Audited System or globally by configuring it in Collector's registry setting. Centrify recommends configuration the heartbeat interval on the Collector if you want all the Audited Systems to send their heartbeat at an identical interval.
- Tweaking the configuration settings may not always help or eliminate the deployment issues completely. In such cases, making significant deployment/configuration changes may be the only option. Please contact Centrify Support to evaluate possible solutions.

Conclusion

This document has provided some information as to what factors can affect Centrify DirectManage Audit performance. Keep in mind, however, that every installation is unique and we cannot anticipate every use case. If you continue seeing performance degradation after following the best practices outlined in this document, contact Centrify Support for assistance.