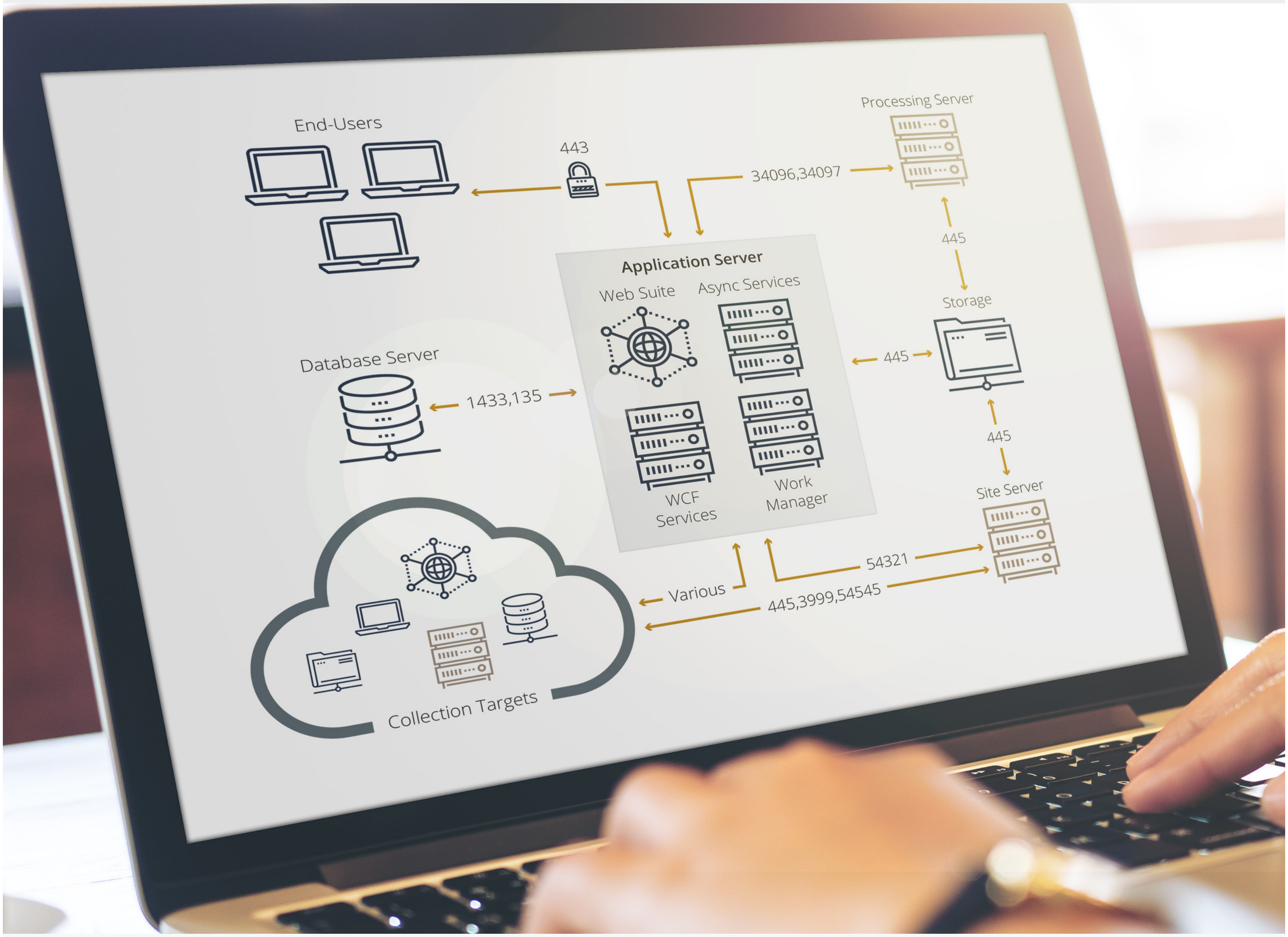
WHITE PAPER



AccessData® AD eDiscovery® Agent

Version 6.3

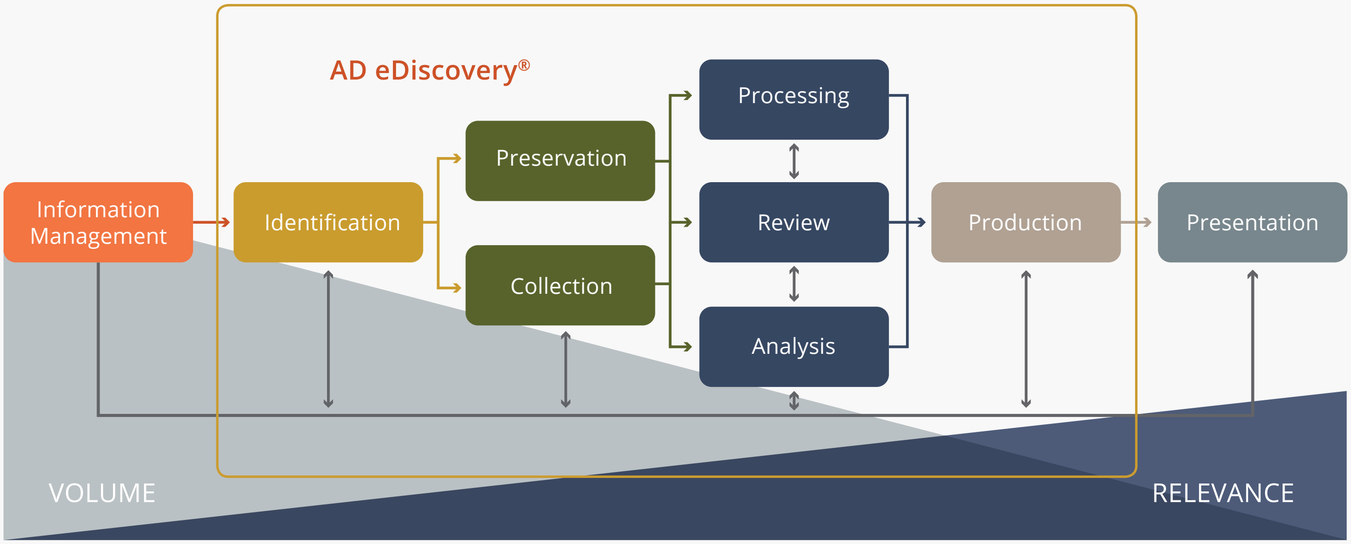
Introduction

This document is intended to describe the expected behavior of the 6.3 version of the AccessData Agent when used in conjunction with the AccessData AD eDiscovery® software solution in situations involving the interruption of collections being performed against computer endpoints. However, many of the behaviors described herein are generally applicable to the AccessData Agent, regardless of version.

AD eDiscovery Architecture

AD eDiscovery is a fully integrated software platform designed to help organizations preserve, collect, process, review, and produce ESI (Electronically Stored Information) in accordance with EDRM model and FRCP rules regarding evidence preservation.

The solution is entirely web-based and scalable to meet even the most demanding e-discovery challenges. AD eDiscovery is comprised of a series of functional



AD eDiscovery® and the EDRM

components that allow the solution to be customized to meet the customers' collection, processing and production requirements. All components can be installed on a single server or distributed in various ways across multiple servers depending on the organization's needs and desired workflow.

The following section contains a brief explanation of each of the components involved in the collection of data from targeted computer endpoints and its role within the solution.

Agent

The AccessData Agent is a modular application that can be deployed to targeted computer endpoints and performs secure forensic-level access, analysis, and preservation of a target endpoint's static data.

Work Manager

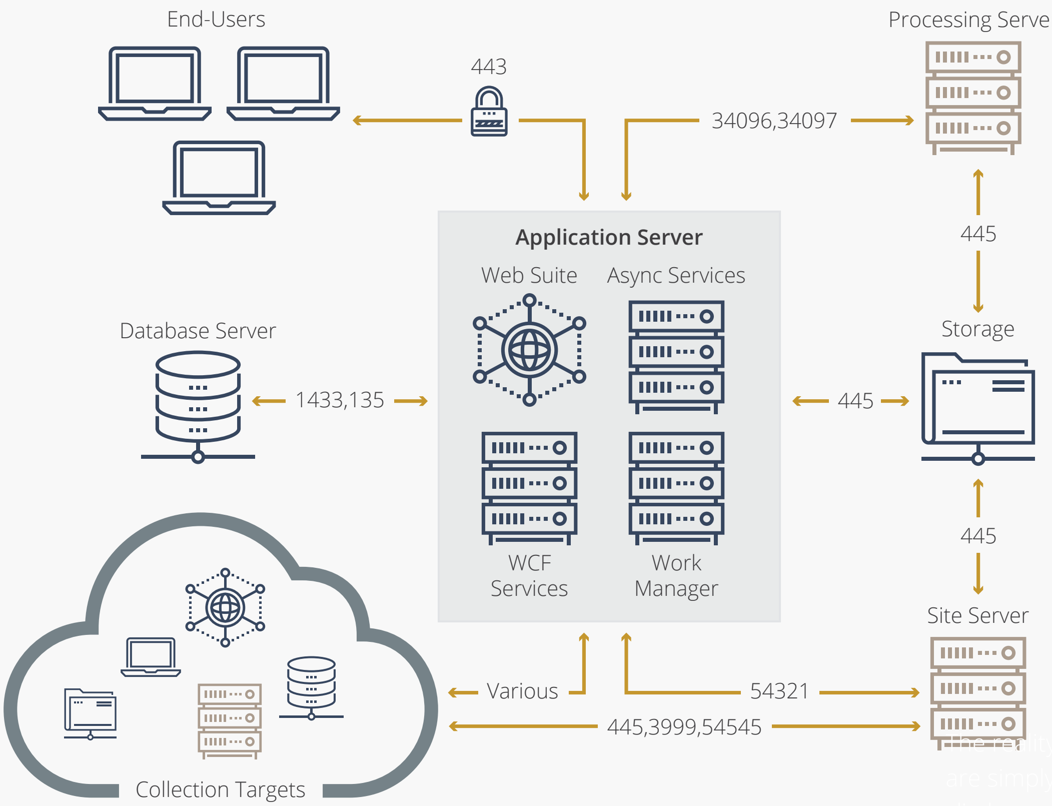
The AccessData Work Manager governs the flow of work to the Processing Engines and Site Servers, as well as performing collections of structured data sources such as Microsoft® Exchange and Microsoft SharePoint®.

Site Server

The AccessData Site Server is generally responsible for managing communication from the Work Manger through to the Agents and for performing collections from Network Share sources. The Site Server may be configured to perform one of three primary roles:

Root Site Server—The Root Site Server is generally responsible for managing the flow of information from the Work Manager down to the appropriate collection endpoints and of collected data up from the appropriate collection endpoints to the Work Manager. The Root Site Server may be configured to control one or more Site Servers that have been configured in a child relationship. AD eDiscovery requires the use of a Root Site Server if Agent or Network Share collections are to be performed.

Private Standard Site Server—The Private Standard Site Server manages the collection activity of specified endpoints within a local network. A Private Standard Site Server must be configured to operate under a parent Site Server and may be configured to operate with one or more Site Servers that have been configured in a child relationship. AD eDiscovery does not require the use of Private Standard Site Servers, but they are highly recommended in geographically-disparate environments and environments expecting to perform a significant number of concurrent Agent collections.



Simplified AD eDiscovery Network Communication Diagram

The reality is that data volumes are simply going to continue to climb, so you need to be prepared with respect to storage capacity.

Private Protected Site Server—The Private Protected Site Server manages the collection activity of specified endpoints within a local network. This configuration is identical to the Private Standard Site Server, but a Private Protected Site Server will not communicate up to its parent without first receiving a communication down from its parent.

Public Site Server—The Public Site Server manages collection activity initiated by Agents located outside the local network. A Public Site Server must be configured to operate under a parent Site Server. AD eDiscovery only requires the use of a Public Site Server in environments that desire the ability to collect from computer endpoints while they are connected to the Internet, but are not connected to the local network.

Collection Communication

SSL Certificates

Secure Sockets Layer (SSL) Certificates are a global standard technology that are commonly leveraged to encrypt network traffic. SSL Certificates employ public-key cryptology; an asymmetric encryption technology involving two mathematically related strings, a Public Key and a Private Key. These two related Keys are commonly referred to as a Key Pair.

Unlike symmetric encryption technologies that rely on one key to encrypt and decrypt data, each key in a Key Pair performs a unique function. The Public Key is used to encrypt data and the Private Key is used to decrypt data that has been encrypted using the Public Key.

As it is computationally infeasible to compute the value of a Private Key, the Public Key may be freely disseminated without fear of compromise. This provides a safe and reliable method for the secure encryption of data, which can only be decrypted by the owners of the Private Key.

The AD eDiscovery software solution uses a Key Pair to authenticate and encrypt communications between a Site Server and an Agent. This Key Pair must contain the complete chain of trust, but does not require a signature from a trusted authority. For this reason, the use of an internally generated SSL Certificate is recommended.

AD eDiscovery will accept the following certificate formats:

- Public: .CER, .CRT, .P7B
- Private: .P12, .PEM, .ADP12[1]

The Agent only supports the use of a single communication Key Pair. In environments that contain multiple Site Servers (or in which, for example, FTK® Enterprise is also being used to collect from Agents), each Site Server and Agent must be configured to use the same Key Pair. If each Site Server is configured to use a different Key Pair, each Site Server will be able to communicate only with Agents that possess its Public Key.

Communication between the Work Manager and Site Server

Communication between the controlling Work Manager and the Root Site Server will run in one of two modes depending on the architecture of the environment:

- **Local Unsecure**—When the Work Manager and the Root Site Server are both installed on the same server, the two components will communicate on an unsecured socket that is bound to local host only.

- **Remote Secure**—When the Work Manager and the Root Site Server are installed on separate servers the two components will communicate on a SSL Certificate encrypted socket that is bound to all interfaces.

Communication between Site Servers

In an environment that includes multiple Site Servers, the Site Servers will be configured in a series of parent-child relationships under the Root Site Server. Communication between Site Servers is conducted on two separate SSL Certificate encrypted sockets. One socket is used to pass data down from the parent to the child and one socket is used to pass data back up to the parent.

If a connection cannot be established between a parent Site Server and a child Site Server, each Site Server will continue to execute collection operations and store any data awaiting transfer on the local file system for later communication when a connection can be reestablished.

Communication between Site Servers and Agents

When a Site Server receives a collection Job that falls within its configured authority, the Site Server will attempt to securely contact the relevant Agent. If this communication is successful, the Agent will attempt to retrieve any necessary collection software modules from the Site Server and execute the collection Job. As the collection Job is executed, the collected data is passed back to the Site Server for staging until the collection Job completes.

Communication between Agent and Site Server

By default, the Agent will not initiate any network communication with any other component. Instead, the Agent waits and listens for incoming authenticated communication from authorized components. However, when the Agent “heartbeat” feature is enabled, the Agent does initiate communication to the Site Server.

Agent Collection Jobs

The AD eDiscovery software solution features Job functionality that is designed to facilitate the remote collection of data from computer endpoints. This functionality supports the collection of complete hard drive images (“Full Disk Acquisition”) or of a subset of files that are responsive to one or more user-configured filters (“Filtered Collection”).

Full Disk Acquisition

Full Disk Acquisition allows for the collection of a targeted computer endpoint’s physical or logical drive space.

Physical Disk acquisition is capable of scanning and collecting data from the entire physical hard drive of a targeted computer endpoint, including both allocated and unallocated space. This method will collect all sectors of a targeted computer endpoint’s hard drive. Consequently, the resulting collection will be larger, relatively slow to complete, and limited by the available resources of the workstation and the network bandwidth back to the server collecting the data.

Logical Disk acquisition is capable of scanning and collecting from allocated space on a targeted computer endpoint’s hard drives.

[1] The .ADP12 file format is an AccessData-generated, protected, and encrypted .P12 certificate format. The Site Server component will automatically generate an ADP12 file if supplied with a .PEM or .P12 certificate.

Filtered Collection

Filtered Collection allows for the granular collection of specific files from a targeted computer endpoint through the use of inclusionary and exclusionary Filters. A Filter may include one or more Filter Options and a collection Job may include one or more Filters.

A Filter that includes multiple Filter Options will only return a positive match on a file if the file matches all of the Filter Options (i.e., logical AND behavior).

- If a collection Job includes more than one inclusion or exclusion Filter, the Job will return a positive match on a file if the file matches any of the Job's Filters (i.e., logical OR behavior).

- If a collection Job includes both inclusion and exclusion Filters, the Job will return a positive match on a file only if the file matches at least one inclusion Filter and does not match any exclusion Filter (i.e., logical AND behavior).

Agent Collection Workflow

Collection Jobs are executed by the AccessData Agent ("Agent"), a modular application that is installed on computer endpoints and performs the secure forensic-level access, analysis, and collection of computer endpoint data.

The activities of an Agent are managed by two specific AD eDiscovery components, the AccessData Work Manager ("Work Manager") and the AccessData Site Server ("Site Server").

- The Work Manager component is responsible for managing the overall flow of collection Jobs, including both Full Disk Acquisition and Filtered Collection jobs against computer endpoints.

- The Site Server component is responsible for communication with the Agent component as well as for the management of Job telemetry information and any collected data received from an Agent in response to the execution of a collection Job.

A collection Job that has been approved and executed is initially sent to the Work Manager component for execution. The Work Manager will initiate a secure connection with the environment's Site Servers and broadcast the Job to them. The first Site Server capable of executing the Job will accept it and attempt to initiate a secure connection with the targeted computer endpoint's Agent. Once an Agent begins executing a collection Job, relevant data is securely transmitted back to the controlling Site Server where it is packaged into an AD1 image. Following the completion of the collection Job's execution by the Agent, the controlling Site Server will signal the controlling Work Manager, which will generate collection reports to complete the collection Job.

Interrupted Agent Collection Job Behavior

The AccessData Agent is designed to facilitate the resiliency of collection activities. By design, the Agent will either restart (start over from the beginning) or resume (pick up where it left off from the point of interruption). The Agent was designed to "resume" interrupted collections in as many situations as possible.

Interruptions or disruptions to a workstation Agent's ability to communicate back to its controlling platform (whether the result of the target computer being shut down or a disruption of network communication), will display an "Interrupted" message in the Job Target Results panel and the Tasks section of the Site Server Console's Tasks pane. Depending on what is being targeted, these collections can be expected to restart from the beginning or resume from the point of disruption once the network communication has been restored.

In practice, any of a number of factors may result in the interruption of an Agent collection, including:

- The shutdown or restart of the targeted computer endpoint during a collection.

- The disruption of network communications between the controlling Site Server and the targeted computer endpoint during a collection.

- The IP address of the targeted computer endpoint changing during a collection.

- An internal software exception involving the targeted computer endpoint or any server hosting the controlling AccessData software.

- A hardware error or failure involving the targeted computer endpoint, any server hosting the controlling AccessData software, or the storage associated with these systems.

As can be reasonably expected, the more extreme examples of these factors (e.g., storage hardware errors, etc.) are more likely to result in a collection entering an unrecoverable interruption state.

It should be noted that an Agent collection Job, whether involving interruptions or otherwise, that has not completed within the configured Job Expiration period will be terminated and marked as “Failed.” The Job Expiration period defaults to a three-day (72 hour) duration, but may be reconfigured as desired for each Job.

Once the collection status has been set to “Interrupted,” the Site Server application component will attempt to reach out to the target every 5 minutes. The system will attempt to validate the IP address of a target with the DNS server(s) every time an active collection job is interrupted.

AccessData Collection Terminology Definitions

Full-disk Agent Collection—is a collection Job of an endpoint workstation hosting the AccessData agent. The resulting output of the collection Job is a physical or logical image of all physical disks attached to the workstation.

Filtered Agent Collection—is defined as a collection against an endpoint workstation’s logical file system, but with a filter applied to target a certain set of file extensions, date ranges, hash values, etc.

Restart—is defined as automatically restarting the collection from the beginning (discarding any data that had been collected prior to the interruption).

Resume—is defined as automatically picking up where it left off before it was interrupted.

Heartbeat—is an Agent-level configuration feature that causes the Agent to autonomously update the Site Server when it has been configured with a new IP address and has network connectivity to communicate with the Site Server. This functionality also allows an agent to automatically populate a “computer” record in the “Data Sources” section of the application.

AD eDiscovery Collections

The table below describes the expected "resume" or "restart" behavior depending on the type of collection that has been interrupted.

Collection Type	Will "Restart" Collection	Will "Resume" Collection	Will "Restart" Collection After IP address change¶	Will "Resume" Collection After IP address change¶
Full Disk Agent Collection	Yes*	Yes*	Yes*	Yes*
Filtered Agent Collection	No	Yes§	No	Yes§
Full Disk Agent Collection (w/ Heartbeat)	Yes*	Yes	Yes*	Yes*
Filtered Agent Collection (w/ Heartbeat)	No	Yes§		Yes§
Network Share Collection	No‡	No‡	No‡	No‡

* Certain failures that require the reinitialization of an Agent have been noted to exhibit the following behaviors:

- Rebooting the target endpoint's operating system does not prevent the collection from resuming where it left off.
- Shutting down the endpoint's operating system and waiting an extended period of time before starting it back up forces the full-disk collection to restart from the beginning.

‡ Network share collections will tolerate up to one minute of network disconnection before it marks the collection job as "Failed." § The "Hits" column in the Jobs tab of the application UI will be set back to "0" even when the collection "resumes" where it left off. All of collected file "batches" that were completed prior to the interruption are written to AD1 and then AD1 image is closed. Upon resume, the agent will begin writing collection results to a new AD1 file.

¶ In cases where the collection is interrupted due to the endpoint Agent system switching from one network interface card (NIC) to another, the collection status will be marked "Failed" and the collection will not resume nor restart.

AD Enterprise Collections

“Auto-Resume of a full-disk acquisition” feature relies on an “agent check-in” process. During imaging, the Enterprise Examiner workstation is communicating with the Enterprise agent and tracking the number of E01 segments it has collected and the progress of the acquisition. If the agent goes offline for whatever reason, the investigator can attempt to “retry” and reestablish the acquisition. When the investigator clicks “retry” the examiner workstation machine will call to the agent, the check-in process will determine what has successfully already been collected, and will pick up collecting the drive at the point in which it left off.

1) Drive acquisitions only—Physical and logical. (Does not work with preview, memory dump, or volatile jobs).

2) The agent must be identified by MachineName, FQDN, or Static IP address.

3) The agent must have a static IP address if targeting the agent by IP address; the IP address cannot be allowed to change or the acquisition will fail.

4) Collections will only resume on the network card/connection they began on. For example, an agent acquisition cannot switch between a WiFi and Ethernet connection. If the acquisition started via the WiFi card it will only resume on the WiFi connection.

5) Collections only resume after the first E01 segment has been created.

6) Only works when imaging directly to the examiner workstation; it will not work when conducting “redirected acquisition” options.

Tips for success:

1) Use FQDN when targeting the Agent.

2) Adjust power settings on Agent endpoint so the machine does not go to sleep.

3) Make sure Agent is online using the “Check agent” connection button prior to starting job.

4) Set Agent usage to “High.”

5) Use the default “Fast transfer” setting and check for any QoC network throttling on the Agent TCP/IP port (port 3999) that may be throttled on the network.

6) When using “Auto Resume,” set the number of retry frequency in minutes to 5 – 10 minutes, and no more than 1000 numbers of retries. Setting the frequency too high affects the examiner workstation, and setting the number of retries to over 1000 is not necessary.

Remote drive imaging is a lengthy, time-consuming process and interruptions are common. The auto-retry and auto-resume features will increase your chances at getting the data you need in a timely fashion. Configuring your Site Servers to use a shared UNC storage location can improve auto-resume success in environments where a targeted computer endpoint may be contacted by more than one Site Server. This behavior is most likely to occur in environments that assign IP addresses using Dynamic Host Configuration Protocol (DHCP), as a targeted computer endpoint that is shut down or loses its connection to the network during an active acquisition Job may receive a new DHCP IP address upon rejoining the network.

ACCESSDATA®

Whether it’s for investigation, litigation or compliance, AccessData® offers industry-leading solutions that put the power of forensics in your hands. For over 30 years, AccessData has worked with more than 130,000 clients in law enforcement, government agencies, corporations and law firms around the world to understand and focus on their unique collection-to-analysis needs. The result? Products that empower faster results, better insights, and more connectivity. For more information, visit www.accessdata.com

Visit us online:
www.accessdata.com

Global Headquarters
+1 801 377 5410
588 West 300 South
Lindon, Utah

North American Sales
+1 800 574 5199
Fax: +1 801 765 4370
sales@accessdata.com

International Sales
+44 20 7010 7800
internationalsales@accessdata.com