ControlPoint

Software Version 5.6.1

Microsoft Windows

Remote Analysis Agent Technical Note

Document Release Date: December 2018
Software Release Date: December 2018

Legal notices

Copyright notice

The only warranties for products and services of Micro Focus and its affiliates and licensors (“Micro Focus”) are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. Micro Focus shall not be liable for technical or editorial errors or omissions contained herein. The information contained herein is subject to change without notice.

Documentation updates

The title page of this document contains the following identifying information:

- Software Version number, which indicates the software version.
- Document Release Date, which changes each time the document is updated.
- Software Release Date, which indicates the release date of this version of the software.

You can check for more recent versions of a document through the MySupport portal. Many areas of the portal, including the one for documentation, require you to sign in with a Software Passport. If you need a Passport, you can create one when prompted to sign in.

Additionally, if you subscribe to the appropriate product support service, you will receive new or updated editions of documentation. Contact your Micro Focus sales representative for details.

Support

Visit the MySupport portal to access contact information and details about the products, services, and support that Micro Focus offers.

This portal also provides customer self-solve capabilities. It gives you a fast and efficient way to access interactive technical support tools needed to manage your business. As a valued support customer, you can benefit by using the MySupport portal to:

- Search for knowledge documents of interest
- Access product documentation
- View software vulnerability alerts
- Enter into discussions with other software customers
- Download software patches
- Manage software licenses, downloads, and support contracts
- Submit and track service requests
- Contact customer support
- View information about all services that Support offers

Many areas of the portal require you to sign in with a Software Passport. If you need a Passport, you can create one when prompted to sign in. To learn about the different access levels the portal uses, see the Access Levels descriptions.

Contents

Introduction 4
Prerequisites 4
Limitations 4

Use the Remote Analysis Agent 5
Command line options 7

Send data for analysis 9

Import ControlPoint RAA information for analysis 10

Clarify obfuscated data 12

Send documentation feedback 13

Introduction

The ControlPoint Remote Analysis Agent (RAA) collects data from file shares and generates Micro Focus IDOL index (.IDX) files.

You can then send the index files to ControlPoint for subsequent data analysis by ControlPoint Legacy Data Cleanup.

You can *obfuscate* (disguise) confidential information during collection. Later, when you receive the analysis report, you can use a mapping file to map the obfuscated information back to the actual data.

Prerequisites

The Remote Analysis Agent and Statistics De-Obfustication utilities are supported on the following Windows operating system versions:

- Windows Server 2008 R2
- Windows Server 2012

The Remote Analysis Agent and Statistics De-Obfustication utilities are supported on a server that has the following software installed:

- Microsoft .Net Framework 4.5.

Limitations

The Remote Analysis Agent is not supported for use with an ControlPoint environment deployed with FIPS security. For more information, see the *ControlPoint Installation Guide*.

Use the Remote Analysis Agent

This section describes how to use the ControlPoint Remote Analysis Agent.

> **IMPORTANT:** Ensure that the RAA utility executable is on a machine on the same network as the file shares whose data you want to collect.

To collect data from file shares

1. Navigate to the following directory:

   install\ControlPoint Utilities\Remote Analysis Agent

   where

   install is the ControlPoint installation directory.

2. Run the RAA executable file by entering:

   ControlPoint.Remote.Analysis.Agent.exe

   The Remote Analysis Agent for File Shares dialog box opens.

3. Specify the following information.

Field	Description
Name	Enter the name of the repository that should contain the collected content.
Description	(Optional) Enter a description for the repository.
Output Directory	Enter the output directory to contain the collected information on your remote server. Enter or browse to the destination directory for the .IDX file. For example: C:\MyFiles\analysis **NOTE:** The Remote Analysis Agent overwrites any previously generated files in the specified output directory. If you run the utility more than once, ensure that you specify different output directories each time.
Network Paths	Enter or browse to the file share directory paths from which to collect data. To add another file share directory path, click **Add**.

4. Click **Next**.

The data collection settings dialog box opens.

5. In the Data Collection section, specify how to handle the various file properties.

Field	Description
Name	Specify the handling of the Name field. • **Collect** the file property in the generated .IDX file. • **Ignore** the file property and omit it from the generated .IDX file. • **Obfuscate** the file property by disguising the information in the generated .IDX file. **NOTE:** If you obfuscate file properties, you must use a generated mapping file to interpret the analysis report you receive from Micro Focus. This data can be de-obfuscated at a later time.
Location	Specify whether you want to **Collect**, **Ignore**, or **Obfuscate** the Location field.
File Owner	Specify whether you want to **Collect**, **Ignore**, or **Obfuscate** the File Owner field.
File Type	Specify whether to **Collect** or **Ignore** the file type.
File Size	Specify whether to **Collect** or **Ignore** the file size.
Creation Date	Specify whether to **Collect** or **Ignore** the file creation date.
Last Modified Date	Specify whether to **Collect** or **Ignore** the file's last modified date.
Last Accessed Date	Specify whether to **Collect** or **Ignore** the file's last accessed date.
Duplicate Identification Hash	Specify whether to **Collect** or **Ignore** the file's duplicate identification coverage hash values. If you select **Collect**, a confirmation dialog box opens. Click **Continue.** The RAA requires all files to be opened during generation of duplicate identification coverage hash values. The Agent will attempt to preserve the Last Access timestamps, subject to having write access to each file.
Duplicate Identification Coverage	Move the slider to the desired percentage of file sampling used to detect duplication. During duplicate identification, RAA copies data from file shares to the server on which it runs.

Field	Description
	The accuracy of duplicate identification increases as the sampling percentage increases, however, so does the time required for data collection and analysis. You can reduce the sampling percentage for faster data collection; just know that accuracy also decreases.

6. Click **Start**.

   RAA collects the data. A summary dialog box appears, listing the indexed directories, location of the output directory, errors and messages.

7. Click **Finish**.

   The RAA closes.

The Remote Analysis Agent output directory

A Remote Analysis Agent directory is created in the output directory.

For example, for an output directory of `C:\MyFiles\analysis`, the following files and directories are created:

- `Log.txt`. The data collection log file.
- `Obfuscation.csv`. A mapping file for any obfuscated data.
- `\INDEXED` directory contains the following information.
  - `IndexDetails.xml` contains the name and description of the indexed data.
  - Index file subdirectory contains the generated .IDX files. The directory name is a random, unique identifier by default.

Command line options

You can run the ControlPoint RAA from the command line instead of using the interface.

To run the utility from the command line

Run `ControlPoint.Remote.Analysis.Agent.exe` with the following parameters.

Parameter	Description
`/name`	The name of the index.
`/description`	(Optional) A description of the index.
`/output`	The destination directory for the .IDX file. **NOTE:** The Remote Analysis Agent overwrites any previously generated files

Parameter	Description
	in the specified output directory. If you run the Agent more than once, ensure that you specify different output directories each time.
`/input`	A comma-separated list of the file share directory paths from which to collect data.
`/duplicate`	A number indicating the percentage of file sampling to use to detect duplication.
`/ignore`	(Optional) A comma-separated list of document details to omit from the output. You can specify the following values. • `Name` • `Location` • `DocumentOwner` • `DocumentType` • `FileSize` • `CreateDate` • `LastModifiedDate` • `LastAccessedDate`
`/obfuscate`	(Optional) A comma-separated list of document details to obfuscate in the output. You can specify the following values. • `Name` • `Location` • `DocumentOwner` • `DocumentType` • `FileSize` • `CreateDate` • `LastModifiedDate` • `LastAccessedDate` `Name`, `Location`, and `DocumentOwner` are obfuscated by default.
`/meta`	(Optional) A comma-separated list of Micro Focus IDOL fields to include in the output, in `name=value` format.

Example

```
ControlPoint.Remote.Analysis.Agent.exe /name test /output “C:\My Files\analysis” /input “F:\User Records” /duplicate 90 /ignore Location /obfuscate CreateDate,LastModifiedDate /meta itemprop=property
```

Send data for analysis

After you run the RAA, a Remote Analysis Agent directory is created in the output directory. For more information on output directory contents, see The Remote Analysis Agent output directory, on page 7

To send data for analysis

1. Create a .ZIP file that contains the `\INDEXED` directory.
2. Send the .ZIP file to your Micro Focus contact for analysis.

   When the analysis is complete, your contact sends you a report containing a variety of metrics and charts describing the data.
3. If you obfuscated any properties during data collection, the obfuscated values appear as `Value1`, `Value2`, and so on.

   To clarify the obfuscated data, run the ControlPoint Remote Analysis Agent Statistics De-Obfustication utility. For more information, see Clarify obfuscated data.

Import ControlPoint RAA information for analysis

To import RAA metrics into ControlPoint

1. In the ControlPoint dashboard, select **+ Add New Repository**.
2. From the left pane, select **Import** and enter the path to the INDEXED directory created by RAA, then click **Get Details**.

   For example:

   `C:\MyFiles\analysis\INDEXED`
3. The **Name** and **Description** fields appear along with the Properties tab.
   - Modify the name and description, if required.

     By default, the name specified for the RAA output is displayed here and it also appears as the repository name in the dashboard.
   - (Optional) add additional properties to the metadata based on those entered using the Administration dashboard.
4. Click the **Properties** tab and then click **Add**, to specify more properties.
5. Select a property from the drop-down list and one or more of the available property values and click **Save**.

   The property information is updated in the **Properties** field on saving.
6. Restart ControlPoint IDOL.
   a. Navigate to: `\Program Files\Micro Focus\ControlPoint\Indexer\IDOL`
   b. Run the following:

      `_stop_service.bat`

      `_start_service.bat`
   c. Restart ControlPoint application pool; `iisreset`
7. Restart the connector (in this example, the filesystem):
   a. Navigate to: `\Program Files\Micro Focus\ControlPoint\Indexer\FileSystem Connector`
   b. Run the following:

      `_stop_service.bat`

      `_start_service.bat`

      In the Repository dashboard, the new repository will now be available with the name provided while importing the repository

8. Select **Refresh** from the drop-down menu.

   The data size and items are displayed.

   a. Select the repository and change the repository status to **Analyze**.

   b. Select the repository to see the details.

Clarify obfuscated data

When you receive an analysis report from ControlPoint that contains obfuscated values, you can run the Statistics De-Obfustication utility to map the obfuscated values to the original values.

To clarify obfuscated data

1. Navigate to the following directory:

   install\ControlPoint Utilities\Statistics De-obfustication Utility

   where

   install is the ControlPoint installation directory.

2. Run the **Statistics De-Obfustication Utility** (`ControlPoint.Statistics.De-Obfustication.Utility.exe`) as the Administrator.

   The utility opens.

3. Provide the following information.
   - In the **File to De-Obfuscate** box, browse to the name of the obfuscated report file.

     For example:

     `C:\MyFiles\analysis\test1.xls`
   - In the **Obfuscation Folder** box, browse to the `Obfuscation.csv` file.
4. Click **Run**.

   The obfuscated values in the report are replaced with the real values.

Send documentation feedback

If you have comments about this document, you can contact the documentation team by email. If an email client is configured on this system, click the link above and an email window opens with the following information in the subject line:

Feedback on Remote Analysis Agent Technical Note (Micro Focus ControlPoint 5.6.1)

Add your feedback to the email and click **Send**.

If no email client is available, copy the information above to a new message in a web mail client, and send your feedback to swpdl.controlpoint.docfeedback@microfocus.com.

We appreciate your feedback!