IBM Content Collector
Version 4.0.1.12

Readme

IBM Content Collector 4.0.1 Fix Pack 12 Readme

Readme documentation for IBM® Content Collector 4.0.1 fix pack 12 including installation-related instructions, prerequisites and co-requisites, and information about fixes, changes, and limitations for IBM Content Collector for Email, IBM Content Collector for File Systems and IBM Content Collector for Microsoft SharePoint.

Readme file for:
>IBM Content Collector

Product / Component release:
>Version 4.0.1.12

Update name:
>Fix Pack 12

Fix ID:
>4.0.1.12-IBM-ICC-FP012

Publication date:
>13 May 2021

Last modified date:
>13 May 2021

Contents

New in this fix pack

1. **Upgraded Microsoft Redistributables and .Net Framework**

   The code is compiled with VC++ 2015-2019 and VC++ 2013 redistributables along with .Net Framework 4.7.2. Thus, removing the dependency on the VC++ 2005, VC++ 2008 and VC++ 2010 redistributables, which are announced for end of support (EoS).

   Notes:

   1. VC++ 2013 redistributables are required for compatibility with:

      a. IBM GSKit 8.0.55.17 (upgraded and provided along with this fix pack), and

      b. Content Collector Notes Client Extension

   2. VC++ 2008 redistributables are required for compatibility with Content Collector Legacy Support component in this fix pack. (Note that the support for the said component is deprecated in this release and will be removed in future releases.)

2. **Upgraded InstallAnywhere to 2020**

   You no longer need to run the fix pack installers on compatibility mode as the new installers are now built with InstallAnywhere 2020.

   Due to the InstallAnywhere 2013 incompatibility with Java 8, the earlier releases (except fix pack 11 Content Collector Server installer) had to be run on compatibility mode on the latest operating systems.

   The new installers remove the old and obsolete VC++ 2005 and VC++ 2010 redistributables and add the new VC++ 2015-2019 and VC++ 2013 redistributables.

3. **Content Collector Email Connector's Transport protocol for LDAP connection is now configurable**

   Content Collector securely connects to Active Directory over LDAP using the available "ex.activeDirectory.ForceSSL" variable. However, you cannot configure the transport protocol. Starting with this release, you can set the desired transport protocol using a new variable "ex.activeDirectory.Protocol".

   Usage:

   a. Open the Configuration Manager.

   b. Create a variable ex.activeDirectory.Protocol with value as desired (TLSv1.2, TLSv1.1, TLSv1) in Configuration Manager > Connectors > Email > Variables.

      Note: The values TLSv1.2, TLSv1.1, TLSv1 are case-sensitive.

   c. Create another variable ex.activeDirectory.ForceSSL and set its value to 'true' in Configuration Manager > Connectors > Email > Variables.

   d. Save the changes.

   e. Restart Task Route service.

4. **Upgraded Third Party dependencies libraries**
   - IBM Global Security Kit (GSKit) to 8.0.55.17
   - Oracle Outside In Technology (OIT) to 8.5.5
   - IBM whitney_core.jar to 5.5.4.0

Deprecated in this fix pack

The support for the Content Collector Legacy Support component is deprecated in this release and will be removed in future releases.

Download location

The following files are available for download from the IBM Fix Central:

Download Content Collector 4.0.1 fix pack 12 fix files
https://www.ibm.com/support/fixcentral/swg/downloadFixes?parent=Enterprise%20Content%20Management&product=ibm/Information+Management/Content+Collector&release=4.0.1.12&platform=All&function=fixId&fixids=4.0.1.12-IBM-ICC-FP012&includeRequisites=1&includeSupersedes=0&downloadMethod=http

Product / Component name	Platform	File name
Content Collector 4.0.1 fix pack 12 Release readme file https://www.ibm.com/support/pages/node/6452129	Not applicable	4.0.1.12-IBM-ICC-FP012.pdf
Content Collector 4.0.1 fix pack 12 Server	Windows	4.0.1.12-IBM-ICC-Server-FP012.zip
Content Collector 4.0.1 fix pack 12 Notes Client Extension	Windows	4.0.1.12-IBM-ICC-NotesClient-FP012.zip
Content Collector 4.0.1 fix pack 12 Outlook Extension (32-bit)	Windows	4.0.1.12-IBM-ICC-OE-32Bit-FP012.zip
Content Collector 4.0.1 fix pack 12 Outlook Extension (64-bit)	Windows	4.0.1.12-IBM-ICC-OE-64Bit-FP012.zip
Content Collector 4.0.1 fix pack 12 SharePoint Extension	Windows	4.0.1.12-IBM-ICC-SharePoint-FP012.zip
Content Collector 4.0.1 fix pack 12 ICC App for SharePoint Online	Windows	4.0.1.12-IBM-ICC-SharePointOnline-FP012.zip
Content Collector 4.0.1 fix pack 12 OWA Support	Windows	4.0.1.12-IBM-ICC-OWA-FP012.zip
Content Collector 4.0.1 fix pack 12 CM8 Text Search Support	Windows	4.0.1.12-IBM-ICC-CM8Text-WIN-FP012.zip
	Linux	4.0.1.12-IBM-ICC-CM8Text-Linux32-FP012.tar.gz
	zLinux	4.0.1.12-IBM-ICC-CM8Text-Linuxz-FP012.tar.gz
	AIX	4.0.1.12-IBM-ICC-CM8Text-AIX-FP012.tar.gz

Product / Component name	Platform	File name
Content Collector 4.0.1 fix pack 12 P8 Content Search Services Support	Windows	4.0.1.12-IBM-ICC-P8CSS-WIN-FP012.zip
	Linux	4.0.1.12-IBM-ICC-P8CSS-Linux-FP012.tar.gz
	zLinux	4.0.1.12-IBM-ICC-P8CSS-Linuxz-FP012.tar.gz
	AIX	4.0.1.12-IBM-ICC-P8CSS-AIX-FP012.tar.gz

Prerequisites and corequisites

The server package of this fix pack requires the server package of Content Collector 4.0.1 as a prerequisite. The fix pack includes all the fixes till date.

For information about hardware and software compatibility, see the detailed system requirements document at:

Prerequisites for IBM Content Collector 4.0.1 - https://www.ibm.com/support/pages/node/614445

1. .Net Framework 4.7.2

Starting with this fix pack, all the dependant Content Collector components are upgraded to run on .Net Framework 4.7.2 or later. If this minimum version is not installed, you will be notified during the installation process.

2. IBM Content Collector Component matrix

The following is a list of the latest versions of other components which can be installed with this fix pack:

Components	Content Collector version	What's new
Outlook Web Application (OWA)	4.0.1.12-IBM-ICC	.Net Framework 4.7.2 - Addressed compatibility mode limitation by upgrading InstallAnywhere to 2020
SharePoint Extension	4.0.1.12-IBM-ICC	.Net Framework 4.7.2 - Addressed compatibility mode limitation by upgrading InstallAnywhere to 2020
SharePoint Online App	4.0.1.12-IBM-ICC	.Net Framework 4.7.2 - Addressed compatibility mode limitation by upgrading InstallAnywhere to 2020

Components	Content Collector version	What's new
Outlook Extension (32-bit or 64-bit)	4.0.1.12-IBM-ICC	• VC++ 2015-2019 • .Net Framework 4.7.2 • Addressed compatibility mode limitation by upgrading InstallAnywhere to 2020
CM8 Text Search support	4.0.1.12-IBM-ICC(Windows, Linux, AIX, zLinux)	• VC++ 2015-2019 • OIT 8.5.5 • Addressed compatibility mode limitation by upgrading InstallAnywhere to 2020
Notes Client	4.0.1.12-IBM-ICC	• VC++ 2013 • .Net Framework 4.7.2 • Addressed compatibility mode limitation by upgrading InstallAnywhere to 2020
P8 Content Search Services support	4.0.1.12-IBM-ICC (Windows, Linux, AIX, zLinux)	• P8/CSS Indexer - IBM whitney_core.jar upgraded to 5.5.4.0 • Addressed compatibility mode issue by upgrading InstallAnywhere 2020
Legacy support	4.0.1.1-IBM-ICC-FP001	

Notes:
1. The following components do not have functionality changes and may be installed only in case of fresh installation:
    o SharePoint Extension
    o SharePoint Online App
    o CM8 Text Search Support
    o Notes Client
    o P8 Content Search Services support
    o Legacy support

2. The P8 Content Search Services Support component for Solaris platform is not available because IBM Java 8 32-bit JRE is not available for Solaris.

Known issues / Limitations

1. **Notes Client Extension component**

    Notes Client Extension requires VC++ 2013 Redistributables as the ones built for VC++ 2015/2017/2019 were found incompatible with Notes Client Application.

2. **IBM Content Collector CM8 Text Search Support component**

The afuIndexer command throws a version mismatch exception on all supported non-windows platforms Linux / zLinux /AIX. This issue can be overcome by performing the following workaround steps:

a. Navigate to the bin directory within ICCIndexer. Default path: /opt/IBM/ICCIndexer/bin

b. Edit the <itemtype>.ini file and add the following line under [SETTINGS] section
*JniCheckLibraryVersion=false;

c. Save and close the <itemtype>.ini file.

d. Re-run the afuIndexer command.

Installation information

You can either apply this fix directly on Content Collector 4.0.1 or on any of the 4.0.1 fix packs from 1 to 11.

Preparing to install

Before you install the fix pack in a production environment, make sure that you have a complete backup or snapshot of your system so that you can roll back the fix pack installation if a problem should arise. If you have modified any configuration files, take a backup copy of the modified files including any configuration changes in the AFUWeb directory such as **afumessage.properties** and **custom_label.properties**.

On Windows: The firewall might prompt you to block the program alerts (pop-up windows). Mark them 'Trusted' and add them into the network access trust list.

Installing the IBM Content Collector server

1. Back up your configuration store.

2. If you are migrating from fix pack 10 interim fix 7 or earlier, back up the Liberty keystores (p12 files) present in the following location:

   ICC_Install_Dir\AFUWeb\wlp\usr\servers\AFUWeb\resources\security

3. Make sure all the Content Collector services are stopped.

4. Install Content Collector server 4.0.1 fix pack 12.

   Important:

   a. Do not change the host name of the Content Collector Web Application during installation. Changing the host name will make all previously created stub links unusable.
   b. Do not run the Initial Configuration wizard.

5. If you run the Content Collector in a cluster and use the Cluster Tool afuClusterTool.exe file to configure the cluster, run afuClusterTool.exe file with the appropriate options to make sure that the cluster is configured to fit your environment. For example, ensure the first and secondary configuration hosts are set correctly.

For verifying whether SolidDB is set up successfully in Content Collector cluster, refer to the following:

SolidDB setup in an IBM Content Collector cluster

https://www.ibm.com/support/pages/node/261685

6. Restart the Task Routing Engine and the Content Collector Web Application service.

Note: In case you face any issues during the configuration process; for example, during search or preview or cluster setup, go to the **Configuration Web Service** section in the **General Settings** tab of the Configuration Manager and click **Reconfigure** again.

With the recent Apache Struts 2.5.22, the Content Collector Web Application Liberty service (20.0.0.6) does not auto start because the installation script extracts the required web applications and is not dependent on the Content Collector Web Application Liberty service (20.0.0.6). Hence, it is not started automatically post the installation.

For all the other components, install the respective fix pack package.

Post-installation

If you are migrating from fix pack 10 interim fix 7 or earlier, restore the Liberty keystores (p12 files) you backed up in the 'Preparing to install' section of this document, in the following location:

ICC_Install_Dir\AFUWeb\wlp\usr\servers\AFUWeb\resources\security

Installing Microsoft Azure AD OAuth App for Content Collector for Email

Note: If the Microsoft Azure AD OAuth App has been registered before upgrading to fix pack 12, deregister the existing app (refer to section B below) and re-register the app (refer to section A below) for encryption changes to take effect.

A. Prerequisites and Registering

Refer to IBM Content Collector OAuth Support for Microsoft Exchange Online:
https://www.ibm.com/support/pages/node/6259361

B. Deregistering

1. Navigate to the following directory:
   <ICC Installation Directory>\tools\ExchangeOnlineAzureAdApp
2. Open **PowerShell** with Administrator privileges.
3. Run the **cleanup.ps1** script to deregister the Azure AD application.
4. Sign in when prompted.

The Azure AD application will be deleted from the tenant (if it exists) and the corresponding properties will also be deleted from the **ews.properties** file.

Fixes and changes

This section provides details of the fixes and vulnerabilities covered in this fix pack.

Fixed APARs

1. **HE12955**: Content Collector 4.0.1 fix pack 10 Outlook Extension cannot be installed remotely due to compatibility mode.

   Component: Outlook Extension

   For installers shipped with earlier fix packs like fix packs 9 and 10, you had to set the compatibility mode to 'Compatibility mode for Windows 8' explicitly. This prevented you from using any software distribution tool to install Content Collector Outlook Extension on multiple systems parallelly. With the new installer of the fix pack 12, setting the Compatibility mode is not required.

2. **HE12960**: Content Collector for File System: Value for missing tag in XML metadata gets set with the next same tag value.

   Component: Server (File System Connector)

   When archiving documents with an XML Metadata file in a multiple documents scenario, if there was a missing XML tag for any single or multi-valued property, the value for that missing tag used to get filled with the next identical tag value, set in the next document's entry in the Metadata file. With fix pack 12, this issue is fixed.

3. **HE12967**: Content Collector for File System: All the tag values of a multi-value property in XML metadata get combined and applied to all the documents in the XML.

   Component: Server (File System Connector)

   When archiving documents with an XML Metadata file in a multiple documents scenario, if there is a multi-valued metadata property in the XML, all its tag values across all documents listed in the XML, were getting combined and set in the associated P8 property for each document. With fix pack 12, this issue is fixed.

4. **HE12975**: Task Routing service memory utilization keeps on increasing until it reaches a certain memory threshold.

   Component: Server (File System Connector)

   When multiple instances of the warning 'The implemented MetadataSource does not support the property' were seen logged during processing of documents from File System as the source, the Content Collector Task Routing Engine service's memory utilization also increased, reached a certain memory threshold and then it went into a hung state. With fix pack 12, this issue is fixed.

5. **HE13015**: Expiration manager tool throws parse data error when the stipulated date parameter is used.

   Component: Server (Email Connector)

   Content Collector's Expiration Manager tool gave a parse data error when the user passed "stipulatedDate" as a parameter on command for the count. With fix pack 12, this issue is fixed.

6. **HE13018**: Content Collector for Email: Content Collector email archival fails when Notes API NIFOpenCollection fails during the crawling phase.

    Component: Server (Email Connector)

    The fix pack 12 handles the issue by checking and logging the expected failure code returned from NIFOpenCollection and returns normally when NIFOpenCollection API fails on a Private view, instead of allowing an exception to be thrown that halts the archival process.

7. **HE13026**: No proper error message recorded in Content Collector fix pack 10 Audit logs.

    Component: Server (P8 Connector)

    Whenever a P8 runtime exception is thrown during archival, the error details are now also recorded in Audit logs along with P8 Connector logs.

8. **HE13035**: Document ingestion fails when an empty string passed to a metadata property that is associated with a choice list in P8.

    Component: Server (File System Connector)

    File archiving used to fail, if a metadata property tag was found empty in an XML file and if that tag value was mapped to a P8 property which is associated with a choice list. With fix pack 12, this issue is fixed.

9. **HE13041**: Content Collector Web Application: In Content Collector 4.0.1 fix pack 5 interim fix 2 onwards, the date range warning message displayed on the search page flickers on any change in other fields of the search form.

    Component: Server (Web Application)

    With fix pack 12, the flicker of date range warning message displayed on the search page is fixed by triggering background validations only for change in the date fields.

10. **HE13044**: Content Collector for Email: If the encrypted notes email with mimetype cannot be decrypted and has a non-empty body item value an exception is thrown.

    Component: Server (Email Connector)

    With fix pack 12, the exception can be prevented by setting the environment variable AFU_IGNORE_NONEMPTY_BODY_ON_FAILED_DECRYPTION to a value of 1.

Internal defect

69993: **Microsoft Exchange Online OAuth Client ID / Secret encryption**
When the app creation script is run, it copies the app properties (clientID, client_secret, tenantID) in the ews.properties file as plain text. As the app credentials are in plain text format, it is insecure and can easily be misused. We have now ensured that the app properties in the ews.properties file are encrypted.

Security vulnerability

Cross-Site Scripting.
Cross-Site Scripting (XSS) attacks occur when data enters an API / Web application through an untrusted source. This vulnerability is now fixed.

Documentation

The following **Content Collector 4.0.1 documentation** topics are updated:

1. Browser support for using the TLS v1.2 protocol

   https://www.ibm.com/docs/en/content-collector/4.0.1?topic=communications-browser-support-using-tls-v12-protocol

2. Configuration worksheets for the Content Collector source systems

   https://www.ibm.com/docs/en/content-collector/4.0.1?topic=worksheets-configuration-content-collector-source-systems

3. Content Collector considerations for GDPR readiness

   https://www.ibm.com/docs/en/content-collector/4.0.1?topic=documentation-gdpr

4. Enabling a higher security level for Content Collector communications

   https://www.ibm.com/docs/en/content-collector/4.0.1?topic=communications-enabling-higher-security-level-content-collector

5. Installing or upgrading Content Collector for Microsoft SharePoint in GUI mode

   https://www.ibm.com/docs/en/content-collector/4.0.1?topic=iuccms-installing-upgrading-content-collector-microsoft-sharepoint-in-gui-mode

6. Installing or upgrading Content Collector for Microsoft SharePoint in console mode

   https://www.ibm.com/docs/en/content-collector/4.0.1?topic=iuccms-installing-upgrading-content-collector-microsoft-sharepoint-in-console-mode

7. Installing or upgrading Content Collector for Microsoft SharePoint in silent mode

   https://www.ibm.com/docs/en/content-collector/4.0.1?topic=iuccms-installing-upgrading-content-collector-microsoft-sharepoint-in-silent-mode

8. Installing Content Collector Outlook Extension

   https://www.ibm.com/docs/en/content-collector/4.0.1?topic=components-installing-content-collector-outlook-extension

9. Installing Content Collector Text Search Support for IBM Content Manager

   https://www.ibm.com/docs/en/content-collector/4.0.1?topic=iic-installing-content-collector-text-search-support-content-manager

10. Prerequisites for installing SharePoint Online component

    https://www.ibm.com/docs/en/content-collector/4.0.1?topic=restrictions-prerequisites-installing-sharepoint-online-component

11. Regular expressions

    https://www.ibm.com/docs/en/content-collector/4.0.1?topic=editor-regular-expressions

12. The installation, upgrade, or removal of Content Collector for Microsoft SharePoint fails

    https://www.ibm.com/docs/en/content-collector/4.0.1?topic=tip-installation-upgrade-removal-content-collector-microsoft-sharepoint-fails

13. Upgrading to Content Collector 4.0.1

    https://www.ibm.com/docs/en/content-collector/4.0.1?topic=collector-upgrading-content-401

14. What's new in Content Collector 4.0.1

    https://www.ibm.com/docs/en/content-collector/4.0.1?topic=overview-whats-new-in-content-collector-401

Reference links

- Download IBM Content Collector 4.0.1 fix pack 12 (readme file)
  https://www.ibm.com/support/pages/node/6452129

- IBM Content Collector Downloads
  https://www.ibm.com/support/pages/node/613127

- IBM Software Product Compatibility Reports
  https://www.ibm.com/software/reports/compatibility/clarity/index.html

- IBM Content Collector 4.0.1 documentation
  https://www.ibm.com/docs/en/content-collector/4.0.1

Notices

This information was developed for products and services offered in the U.S.A. IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

For license inquiries regarding double-byte (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:
Legal and Intellectual Property Law
IBM Japan Ltd.
1623-14, Shimotsuruma, Yamato-shi
Kanagawa 242-8502 Japan

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law:
INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:
IBM Deutschland GmbH
Department M358
IBM-Allee 1
71139 Ehningen
Germany

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this information and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement or any equivalent agreement between us.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurement may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

All statements regarding IBM's future direction or intent are subject to change or withdrawal without notice and represent goals and objectives only.

All IBM prices shown are IBM's suggested retail prices, are current and
are subject to change without notice. Dealer prices may vary.

This information is for planning purposes only. The information herein is subject to change before the products described become available.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrates programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

A.1 Trademarks

IBM, the IBM logo, and ibm.com are trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. A current list of IBM trademarks is available on the web at "Copyright and trademark information" at http://www.ibm.com/legal/copytrade.shtml.

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

The Oracle Outside In Technology included herein is subject to a restricted use license and can only be used in conjunction with this application.

Other product and service names might be trademarks of IBM or other companies.

Document change history

Date	Description of change
13 May 2021	Original version published