Hitachi Content Intelligence

Hitachi Content Intelligence 1.4.1 Release Notes

Final Build Number: 1.4.1.14
© 2016, 2019 Hitachi Vantara Corporation. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including copying and recording, or stored in a database or retrieval system for commercial purposes without the express written permission of Hitachi, Ltd., or Hitachi Vantara Corporation (collectively, “Hitachi”). Licensee may make copies of the Materials, provided that any such copy is: (i) created as an essential step in utilization of the Software as licensed and is used in no other manner, or (ii) used for archival purposes. Licensee may not make any other copies of the Materials. "Materials" mean text, data, photographs, graphics, audio, video, and documents.

Hitachi reserves the right to make changes to this Material at any time without notice and assumes no responsibility for its use. The Materials contain the most current information available at the time of publication.

Some of the features described in the Materials might not be currently available. Refer to the most recent product announcement for information about feature and product availability, or contact Hitachi Vantara Corporation at https://support.hitachivantara.com/en_us/contact-us.html.

**Notice:** Hitachi products and services can be ordered only under the terms and conditions of the applicable Hitachi agreements. The use of Hitachi products is governed by the terms of your agreements with Hitachi Vantara Corporation.

By using this software, you agree that you are responsible for:

1. Acquiring the relevant consents as may be required under local privacy laws or otherwise from authorized employees and other individuals; and
2. Verifying that data continues to be held, retrieved, deleted, or otherwise processed in accordance with relevant laws.

**Notice on Export Controls.** The technical data and technology inherent in this Document may be subject to U.S. export control laws, including the U.S. Export Administration Act and its associated regulations, and may be subject to export or import regulations in other countries. Reader agrees to comply strictly with all such regulations and acknowledges that Reader has the responsibility to obtain licenses to export, re-export, or import the Document and any Compliant Products.

**EXPORT CONTROLS -** Licensee will comply fully with all applicable export laws and regulations of the United States and other countries, and Licensee shall not export, or allow the export or re-export of, the Software, API, or Materials in violation of any such laws or regulations. By downloading or using the Software, API, or Materials, Licensee agrees to the foregoing and represents and warrants that Licensee is not located in, under the control of, or a national or resident of any embargoed or restricted country.

Hitachi is a registered trademark of Hitachi, Ltd., in the United States and other countries.

AIX, AS/400e, DB2, Domino, DS6000, DS8000, Enterprise Storage Server, eServer, FICON, FlashCopy, IBM, Lotus, MVS, OS/390, PowerPC, RS6000, S/390, System z9, System z10, Tivoli, z/OS, z9, z10, z13, z/VM, and z/VSE are registered trademarks or trademarks of International Business Machines Corporation.

Active Directory, ActiveX, Bing, Excel, Hyper-V, Internet Explorer, the Internet Explorer logo, Microsoft, the Microsoft Corporate Logo, MS-DOS, Outlook, PowerPoint, SharePoint, Silverlight, SmartScreen, SQL Server, Visual Basic, Visual C++, Visual Studio, Windows, the Windows logo, Windows Azure, Windows PowerShell, Windows Server, the Windows start button, and Windows Vista are registered trademarks or trademarks of Microsoft Corporation. Microsoft product screen shots are reprinted with permission from Microsoft Corporation.

All other trademarks, service marks, and company names in this document or web site are properties of their respective owners.
Contents

New in release 1.4.1 .............................................................................................................. 5
New in release 1.4.0 .............................................................................................................. 5
  Hitachi Content Search improvements ................................................................. 5
  Hitachi Content Monitor improvements .............................................................. 6
New in release 1.3.1 .............................................................................................................. 7
New in release 1.2.2 .............................................................................................................. 7
New in release 1.3.0 .............................................................................................................. 8
New in release 1.2.1 .............................................................................................................. 10
New in release 1.2.0 .............................................................................................................. 11
  Workflow updates ........................................................................................................ 11
  Data connection updates ......................................................................................... 12
Index updates .................................................................................................................. 12
Pipeline updates ............................................................................................................ 13
Search App updates ....................................................................................................... 14
Notification rule updates ......................................................................................... 14
REST API changes ....................................................................................................... 15
Service updates ............................................................................................................ 15
New in release 1.1.5 .............................................................................................................. 17
New in release 1.1.4 .............................................................................................................. 17
System requirements and sizing ........................................................................... 18
  Sizing guidance for Hitachi Content Search ......................................................... 18
    Simple sizing .......................................................................................................... 18
    Detailed sizing ...................................................................................................... 19
Sizing guidance for HCM ......................................................................................... 21
  Minimum hardware requirements ....................................................................... 21
  Determining number of instances ......................................................................... 21
Operating system and Docker requirements ........................................... 25
  Docker considerations ........................................................................ 26
  SELinux considerations ...................................................................... 27
Virtual machine host requirements ....................................................... 27
Networking ......................................................................................... 27
Supported browsers ............................................................................. 34
Hitachi Content Intelligence documentation ....................................... 34
Update notes ....................................................................................... 35
  Update version qualification ............................................................. 35
  Version-specific update prerequisites ............................................... 35
Service REST API and CLI differences from release 1.0 ......................... 36
Note about using the log download tool ............................................... 37
Known issues ...................................................................................... 38
Resolved Issues .................................................................................. 44
  Resolved in release 1.4.1 ................................................................. 44
  Resolved in release 1.4.0 ................................................................. 47
  Resolved in release 1.3.1 ................................................................. 51
  Resolved in release 1.2.2 ................................................................. 56
  Resolved in release 1.3.0 ................................................................. 57
  Resolved in release 1.2.1 ................................................................. 60
  Resolved in release 1.2.0 ................................................................. 64
  Resolved in release 1.1.5 ................................................................. 68
  Resolved in release 1.1.4 ................................................................. 69
New in release 1.4.1

Release 1.4.1 addresses a number of issues found in previous versions of the product. For more information, see "Resolved in release 1.4.1" on page 44.

New in release 1.4.0

Admin Console and workflow component redesign and reorganization

In release 1.4, the Administrator Console has been redesigned with a new look and feel.

Substantially, all pages related to Hitachi Content Intelligence workflows have been moved to a new and separate Workflow Designer application that includes its own CLI, REST API, and browser UI.

Hitachi Content Search improvements

Release 1.4 includes various improvements for working Hitachi Content Search through the new Workflow Designer application:

- On the new index Maintenance tab, you can now run query commands to delete documents from an index.

- For HCI Indexes, the schema has been updated with all field types supported by Solr 6.6.

- Pipeline conditional stages now support date/time variables. For example, you can now configure a conditional stage to match only documents that were created within the last month: `HCI_createTime greater_than $(NOW-1MONTH/DAY)`
New in release 1.4.1

- You can now configure a workflow task to allow failed documents to continue being processed through the end of the workflow pipeline. Previously, documents would not be processed by any pipeline stages following the stage on which the document failed.

- In content classes, XML-type content properties now support the inclusion of XPath functions.

New data connections

- **System-wide HCP Anywhere connector** — This data connection allows you to read documents from and write documents to an entire HCP Anywhere system. Previously, Hitachi Content Search provided a data connection for accessing documents belonging to only a single HCP Anywhere user.

- **CIFS** — This data connection allows you to read documents from and write documents to CIFS shares.

New pipeline stages

- **Attach Stream stage** — Allows the addition of new streams to a document. These streams can come from an existing data source or be formed from multiple streams already existing in the document.

- **JavaScript stage** — Executes the custom JavaScript code that you specify on each document processed by the stage.

- **Decompression stage** — Decompresses GZIP, Bzip2, XZ compressed documents.

**Hitachi Content Monitor improvements**

**Storage usage forecasting**

In release 1.4, Hitachi Content Monitor can now make predictions about your HCP system’s future storage needs. After examining at least one week of storage data from your HCP, Hitachi Content Monitor can estimate the number of days until HCP reaches its maximum storage capacity.

**Anomaly detection**

In release 1.4, Hitachi Content Monitor can now alert you of anomalous or abnormal trends in your HCP system’s storage usage and front-end network traffic.
Historical log import

If you've downloaded and stored batches of historical logs from your HCP system, you can now upload those logs to Hitachi Content Monitor and create visualizations and dashboards to view them. Previously, the only way to import log data into Hitachi Content Monitor was to have it actively monitor an HCP system over a period of time.

HCP 8.1 monitoring support

Hitachi Content Monitor can now monitor HCP systems running release 8.1.

New in release 1.3.1

Release 1.3.1 includes a number of bug fixes and improvements. For information, see "Resolved in release 1.3.1" on page 51.

Additionally, the Email Notification stage can now be configured to send HTML-formatted emails, instead of only plain text.

New in release 1.2.2

Note: Version 1.2.2 was released after 1.3.0 and was not generally available. Versions 1.3.1 and later include all fixes from 1.2.2.

Release 1.2.2 contains improvements to the HCP MQE data connection and Email Expansion pipeline stage. For information, see "Resolved in release 1.2.2" on page 56.
New in release 1.3.0

Hitachi Content Monitor

Release 1.3.0 introduces the Hitachi Content Monitor application (HCM). This app lets you monitor the performance, activity, and health of your Hitachi Content Platform (HCP) systems. You can view historical HCP data collected over a period of time, or even monitor your HCP in real time. The app includes over 20 types of charts and graphs that you can customize and filter to show only the information you’re interested in.

With the introduction of this application, newly installed Hitachi Content Intelligence systems can run either HCM or this existing search application, Hitachi Content Search, but not both simultaneously. Additionally, systems running HCM must be sized differently than those running Hitachi Content Search. For information, see "Sizing guidance for HCM" on page 21.

Jobs

Release 1.3.0 adds jobs, which are a new means for Hitachi Content Intelligence to perform work. Jobs allow services to temporarily devote resources to completing some transient unit of work. Like services, jobs are run in Docker containers on system instances. Unlike services, when a job completes, its container exits.

Jobs are run automatically by services. You cannot manually start and stop jobs, but you can configure where and when they are allowed to be active.

Hitachi Content Intelligence workflows have been updated to run as jobs. This introduces new benefits:

- You can specify specific workflows to run on specific instances. Previously, you could only specify which system instances were allowed to run workflows at all.

- You can use jobs as means of running multiple workflows simultaneously.

- Workflows can make use of the new volumes feature.

Note: With the introduction of jobs, the Workflow-Agent service has been removed.
Volumes

With release 1.3.0, storage space for services and jobs can now be managed by using volumes. Volumes are properties of services and job types that specify where and how a services and individual jobs store their data.

Volumes can be configured only during system deployment. The services and job types built into Hitachi Content Intelligence do not support volume configuration changes once the system is up and running.

Volumes can be used to allow jobs and services to store their data in external storage systems, outside of the system instances. This allows data to be more easily backed-up or migrated.

Volumes can also allow services or jobs to store different types of data in different locations. For example, a service may use two separate volumes, one for storing its logs and the other for storing all other data.

**Note:** For release 1.3.0 volume configuration is supported for product services. Volumes should not be used for System services on a production system; leave the default bind-mount setting for those services in volume.config when installing a new 1.3.0 system.

New data connections

Release 1.3.0 adds new data connections that return information about the performance and storage usage of your Hitachi Content Platform (HCP) system:

- HCP Monitoring data connection
- HCP Syslog Kafka Queue data connection

Index action

In earlier releases, documents could be indexed only after passing through an entire workflow pipeline and being sent to the workflow’s outputs. With release 1.3.0, you can now index documents as they pass through a pipeline. You do this by adding an Execute Action stage to a pipeline and configuring it to run the new Index action. As of release 1.3.0, all Hitachi Content Intelligence-supported index types (except for HDDS) support the Index action.
Workflow aggregations

Release 1.3.0 adds these new aggregations to workflows:

- Standard Deviation Aggregate
- Variance Aggregate

Data connection improvements

- Kafka Queue data connection — The new Initial timestamp setting lets you collect only the messages that were added to a Kafka queue after a specified time.

Service changes

With the introduction of the Jobs feature, the Workflow-Agent service has been removed.

Docker and OS requirement changes

With release 1.3.0, the minimum required Docker version is now 1.13.1. Additionally, Hitachi Content Intelligence is no longer being qualified on Fedora 24.

For more information, see "Operating system and Docker requirements" on page 25.

New in release 1.2.1

Release 1.2.1 resolves a number of issues, mainly involving system updates. For more information, see "Resolved in release 1.3.0" on page 57.

Required properties in REST API and CLI

The Hitachi Content Intelligence REST APIs and CLIs now enforce that you provide values for required properties in model objects. This affects any scripts or applications you've written to use the REST APIs or CLIs. For more information, see "Service REST API and CLI differences from release 1.0" on page 36.

Updated Docker considerations

The Hitachi Content Intelligence documentation has been updated with sizing recommendations for running Hitachi Content Intelligence with the Docker devicemapper storage driver. For information, see "Docker considerations" on page 26.
New in release 1.2.0

This topic describes newly added features and improvements to existing features made for Hitachi Content Intelligence release 1.2.0.

Workflow updates

Aggregations and Triggers

Release 1.2.0 adds two new features to workflows that can be used together to monitor workflow performance:

- **Aggregations** — Aggregations keep track of information across all documents that a workflow processes. For example, you can create aggregations that track the total number of PDF documents processed, or that track the average size of all documents processed.

- **Triggers** — Triggers monitor aggregation values. When an aggregation value meets a condition that you specify, the trigger is activated and sends a special workflow status document through a trigger pipeline. You can configure your trigger pipeline to include stages that send you notifications.

Workflow pipeline execution modes

With release 1.2.0, pipelines added to workflows can now run in either of these execution modes:

- **Workflow-Agent** — Document processing for the pipeline is distributed among the system instances that run the Workflow-Agent service. This allows a pipeline to process multiple batches of documents at the same time. This is the default execution mode.

- **Preprocessing** — Document processing for the pipeline occurs on a single instance. This may result in pipeline performance increases when using the Read Lines stage to process log or csv files.

Historic metrics

With release 1.2.0, the Metrics page for each workflow task now includes graphs that show changes in various workflow performance metrics since the workflow was first run.
New in release 1.4.1

Data connection updates

New data connections

- PostgreSQL JDBC data connection — Connects to PostgreSQL databases using the Java Database Connectivity (JDBC) API and uses SQL queries to retrieve documents from the database tables you specify.

- MySQL and MariaDB JDBC data connection — Connects to MySQL and MariaDB databases using the JDBC API and uses SQL queries to retrieve documents from the database tables you specify.

- Solr JDBC data connection — Connects to specified Solr indexes using the JDBC API and uses SQL queries to retrieve documents from those indexes.

- Internal Index JDBC data connection — Retrieves documents from an internally-managed Solr index. This data connection is equivalent to the Solr JDBC data connection, but has been automatically configured for connecting to internal indexes.

- Hadoop File System data connection — Allows access to files in Hadoop Distributed File Systems (HDFS).

- Kafka Queue data connection — Allows messages to be read from and written to Apache Kafka message queues.

Improvements to existing data connections

The Local File System can now perform actions on documents and collects POSIX metadata for the files it reads.

Index updates

Elasticsearch index support

Hitachi Content Intelligence can now create, manage, and index data to Elasticsearch indexes. You can use these indexes as the outputs for your workflows and your users can search them using the Hitachi Content Intelligence Search App.

The Elasticsearch indexes that Hitachi Content Intelligence manages are stored on your Elasticsearch servers, not within Hitachi Content Intelligence.
Solr 6 support

Hitachi Content Intelligence now supports Solr version 6 indexes:

- All internally-managed Hitachi Content Intelligence indexes are now Solr 6 indexes.
  
  In previous versions of Hitachi Content Intelligence, internally-managed indexes used Solr 5. Upon update to version 1.2.0, existing internally-managed indexes are automatically updated to use Solr 6.

- The external Solr indexes that Hitachi Content Intelligence manages can exist on either Solr 5 or Solr 6 servers.

SQL query support

With Hitachi Content Intelligence 1.2.0, you can now use SQL queries to search Solr 6 indexes.

Solr debugging information

Solr debugging information is now returned when you query a Solr index from the Hitachi Content Intelligence Admin App.

Pipeline updates

New built-in stages

- **DICOM Metadata Extraction** — Extracts fields from images formatted in DICOM, a standard file format for medical images.

- **Replace** — Uses regular expressions to find and replace text within field values.

- **Syslog Notification** — Sends notifications to syslog servers about documents processed by the pipeline.

- **Email Notification** — Uses an SMTP server to send email notifications about documents processed by the pipeline.
Changes to conditional statements

Conditional statements have a number of new comparators that perform lexical comparisons for field values. For example, you can use the `lexically_less_than` comparator to check whether one field value precedes another alphabetically.

A number of existing comparators, such as `equals`, have also been updated to require that the field values being compared be of compatible types (for example, integers can be compared to longs, but not to strings).

Search App updates

Search Visualization

Facets can now be displayed visually. You can enable visualization under Query Settings in the Admin App. When enabled, facet information shows in the Search App as either a pie graph or a bar graph.

Requesting fields to be returned

With release 1.2.0, users of the Search App can now specify which metadata fields are returned with each search result.

Additional query parameters

With release 1.2.0, users of the Search App can now add additional parameters to their queries. For example, when searching a Solr index, a user can specify the Solr `rows` query parameter to control how many rows of documents are displayed on a page.

Download search result lists

With release 1.2.0, users of the Search App can now download lists of search results. The lists are downloaded as comma-separated value (.csv) files. These lists contain only the metadata, not the full content, for each search result.

Notification rule updates

Syslog Notification rules

Hitachi Content Intelligence can now send log messages to syslog servers when system events occur.
REST API changes

General REST API changes

The Hitachi Content Intelligence administrative and search REST APIs have been updated to version 1.2.0.

Changes to service APIs

A number of non-backwards compatible changes have been made to the administrative REST API endpoints used for managing and gathering information about services. For more information, see "Service REST API and CLI differences from release 1.0" on page 36.

Service updates

New service: Clustered-File-System

Release 1.2.0 adds the HDFS service to Hitachi Content Intelligence. This service provides storage space to other services in an Hitachi Content Intelligence system. Hitachi Content Intelligence uses this space for storing temporary files extracted from archive files by Preprocessing pipelines.

Important: The Clustered-File-System service is offered as a technology preview. Do not use it on a production system.

Service network types

With release 1.2.0, newly installed Hitachi Content Intelligence systems can optionally support two network types, internal and external. You can use these two networks to isolate all system-internal network traffic to one network and reserve the other for use by your end users.

When installing a new Hitachi Content Intelligence 1.2.0 system, you can configure which network type each service listens on for incoming traffic. For example, you can have the Logging service use the internal network and the Admin-App use the external network.

You can configure service networking only during initial system deployment.
Configurable service ports

When installing a new Hitachi Content Intelligence 1.2.0 system, you can configure the ports that each service listens on for incoming network traffic.

You can configure service ports only during initial system deployment.

Floating services

With release 1.2.0, some Hitachi Content Intelligence services are now floating services.

With a non-floating (or persistent) service, each instance of the service can run on only once instance in the system. If that service instance fails, it is not restarted on a different system instance.

With floating services, you define a pool of eligible instances on which the service can run. If the service fails on one of those instances, the system can automatically restart it on another instance in the pool.

Support for services with multiple types

With release 1.2.0, Hitachi Content Intelligence services can now have multiple service instance types. Each type handles a unique subset of the service’s functionality.

For example, a service with two types could run on three separate instances. In this case, one service instance could be of a different type, and thereby perform different functions, than the other two service instances.

Scheduling and Logging are now configurable

With release 1.2.0, the Logging and Scheduling services are now user-configurable; you can now change the settings for these services and specify the instances that they run on.

Service operation changes

With release 1.2.0, you no longer need to construct a queue of operations when you want to scale or reconfigure services. Now, each service operation you define begins running as soon as you click on the Update Service button.
Service log directory changes

With release 1.2.0, the system log directories have been renamed to match the underlying plugin name for each service. For example, the directory that contains logs for the Admin-App service is now named com.hds.ensemble.plugins.service.adminApp.

New in release 1.1.5

Note: Release 1.1.5 was not generally available. Releases 1.2.0 and later include all fixes from 1.1.5.

Release 1.1.5 contains improvements to a number of built in pipeline stages:

- **Text and Metadata Extraction stage:**
  - Now supports extracting headers from all sub-components of journaled emails, including the list of BCC'd email addresses.
  - Now has an option for including raw, unprocessed content fields in extracted metadata.

- **Email Expansion stage:**
  - Now adds attachment names on to the original email as metadata.
  - Now has an option to select whether attachment names are multivalued or comma separated.

- **PST Expansion stage** — Now has an option to select whether attachment names are multivalued or comma separated.

- **Filter stage** — Now supports filtering out fields by pattern matching using regular expressions.

For information on issues resolved in release 1.1.5, see "Resolved in release 1.1.5" on page 68

New in release 1.1.4

This release includes these workflow task performance improvements:

- Workflow task metrics are now updated more frequently as a task runs.
• Workflow tasks now run faster when processing large data sets (that is, those that include millions of documents).

This release also includes fixes for a number of other issues. For more information, see "Resolved in release 1.1.4" on page 69.

**System requirements and sizing**

This section lists the hardware, networking, and operating system requirements for running an Hitachi Content Intelligence system with one or more instances.

**Sizing guidance for Hitachi Content Search**

**Simple sizing**

This table shows the minimum and recommended hardware requirements for each instance in an Hitachi Content Intelligence running Hitachi Content Search.

<table>
<thead>
<tr>
<th>Resource</th>
<th>Minimum</th>
<th>Recommended</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAM</td>
<td>16 GB</td>
<td>32 GB</td>
</tr>
<tr>
<td>CPU</td>
<td>4-core</td>
<td>8-core</td>
</tr>
<tr>
<td>Available disk space</td>
<td>50 GB</td>
<td>500 GB</td>
</tr>
</tbody>
</table>

**Important:**

• A large number of factors determine how many documents your system can index and how fast it can process them, including: the number of documents to be indexed; the contents of those documents; what search features (such as sorting) the index supports; the number of fields in the index; the number of users querying the system; and so on.

Depending on how you use your system, you may require additional hardware resources to index all the documents you want and at the rate you require. See "Detailed sizing" on the facing page.

• Each instance uses all available RAM and CPU resources on the server or virtual machine on which it's installed.
Detailed sizing

If you are installing Hitachi Content Intelligence to run Hitachi Content Search, you should size your system based on the number of documents you need to index and the rate at which you need documents to be processed and indexed.

To determine the system size that you need:

1. Determine how many documents you need to index.

2. Based on the number of documents you want to index, use the following tables to determine:
   - How many instances you need
   - How much RAM each instance needs
   - The Index service configuration required to support indexing the number of documents you want

For example, if you need to index up to 150 million documents, you need at minimum a 4-instance system with 64 GB RAM per instance.

3. Determine how fast you need to index documents, in documents per second.

For example:
- If you want to index 100 million documents in 2 days, you need an indexing rate of 578 documents per second.
- If you want to continuously index 1 million documents every day, you need an indexing rate of 12 documents per second.
4. Determine the base indexing rate for your particular data set and processing pipelines:
   a. Install a single-instance Hitachi Content Intelligence system with that has the minimum required hardware resources.
   b. Run a workflow with the pipelines you want and on a representative subset of your data.
   c. Use the workflow task details to determine the rate of documents processed per second.

5. To determine the number of cores you need per instance, replace Base rate in this table with the rate you determined in step 4.

<table>
<thead>
<tr>
<th>Number of instances you require</th>
<th>Cores per instance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4 (minimum required)</td>
</tr>
<tr>
<td>1</td>
<td>Base rate</td>
</tr>
<tr>
<td>4</td>
<td>300% Base rate</td>
</tr>
<tr>
<td>8</td>
<td>600% Base rate</td>
</tr>
<tr>
<td>More than 8</td>
<td>Contact Hitachi Vantara for guidance</td>
</tr>
</tbody>
</table>

For example, if you had previously determined that:

- You need a 4-instance system,
- You need to process 500 documents per second,
- The base processing rate for your data and pipelines is 100 documents per second,

You require 8 cores per instance.

6. Multiply the number of instances you need times the number of cores per instances to determine the total number of cores that you need for your system.

7. Once your system is installed, configure it with the index settings you determined in step 2.
For information on index shards, Index Protection Level, and moving the Index service, see the Administrator Help, which is available from the Admin App.

**Sizing guidance for HCM**

**Minimum hardware requirements**

If you are installing Hitachi Content Intelligence to run HCM, each instance in the system must meet these minimum hardware requirements.

<table>
<thead>
<tr>
<th>Resource</th>
<th>Per instance requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU cores</td>
<td>24</td>
</tr>
<tr>
<td>RAM</td>
<td>64 GB</td>
</tr>
<tr>
<td>Disk space</td>
<td>At least 2 TB</td>
</tr>
</tbody>
</table>

**Determining number of instances**

The number of instances that your HCM system requires is based on:

- Whether you need the system to remain highly available.
- The number of documents being produced by the HCP system you want to monitor. In this case, each document represents a single piece of data about the HCP system. A more active HCP system will produce more documents than a less active one.
- The total number of documents you want HCM to store.
Number of instances: Simple procedure
If you're monitoring a typically-active HCP system (roughly 75 operations per second per node), you can use this table to determine the number of HCM instances you need. This table lists the number of HCM instances you need based on the number of nodes in your HCP system and the number of days that you want your HCM system to retain the data it receives from HCP.

If your system is more active, see "Number of instances: Detailed procedure" below.

<table>
<thead>
<tr>
<th>HCP nodes</th>
<th>Data retention time on HCM</th>
<th>HCM instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up to 8</td>
<td>Up to 30 days</td>
<td>1*</td>
</tr>
<tr>
<td>Up to 8</td>
<td>Up to 60 days</td>
<td>2*</td>
</tr>
<tr>
<td>Up to 16</td>
<td>Up to 30 days</td>
<td>4</td>
</tr>
<tr>
<td>Up to 24</td>
<td>Up to 60 days</td>
<td>8</td>
</tr>
</tbody>
</table>

* A HCM system must have a minimum of 4 instances to maintain high system availability.

Number of instances: Detailed procedure
1. Determine whether you require your HCM system to maintain high availability. If so, you need a minimum of 4 instances.

2. Determine the number of documents per second being produced by the HCP system you want to monitor. You can easily do this if you already have an HCM system up and running:
   a. Go to the Monitor App:
      
      https://<system-hostname>:6162
   
   b. Add the HCP system as a source. For information, see the help that's available from the Monitor App.
   
   c. Go to the Hitachi Content Intelligence Admin App:
      
      https://<system-hostname>:8000
   
   d. Go to Workflows > Monitor App Workflow > Task > Metrics.
   
   e. View the value for the Average DPS field.
Tip: Let the workflow run for a while to get a more accurate measure for the Average DPS field.

Otherwise, you can get an average documents per second value by doing this:

a. Select a time period.

b. Download the HCP Internal Logs for this time period. For more information, see the help that's accessible from the HCP System Management Console.

c. In the downloaded logs for each node, count the number lines logged during the selected time period.

d. Add the line value for each node and then divide the sum by the number of seconds in the time window you selected.

3. Use this table to determine the number of instances required based on the number of documents per second produced by your HCP system.

<table>
<thead>
<tr>
<th>Documents per second</th>
<th>Instances required</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up to 3200</td>
<td>1</td>
</tr>
<tr>
<td>3201 to 3800</td>
<td>2</td>
</tr>
<tr>
<td>3801 to 4400*</td>
<td>4</td>
</tr>
</tbody>
</table>

* This is the maximum documents per second that HCM currently supports.

4. Based on your data availability requirements, determine the number of instances you need.

<table>
<thead>
<tr>
<th>Data availability requirement</th>
<th>Index replicas required</th>
<th>Instances required</th>
<th>Impact on total documents stored</th>
</tr>
</thead>
<tbody>
<tr>
<td>No failure tolerance</td>
<td>1</td>
<td>1</td>
<td>None</td>
</tr>
<tr>
<td>Survive 1 failed replica</td>
<td>2</td>
<td>2</td>
<td>2x</td>
</tr>
</tbody>
</table>
### Data availability requirement vs. Index replicas required vs. Instances required vs. Impact on total documents stored

<table>
<thead>
<tr>
<th>Data availability requirement</th>
<th>Index replicas required</th>
<th>Instances required</th>
<th>Impact on total documents stored</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survive 2 failed replicas</td>
<td>3</td>
<td>3</td>
<td>3x</td>
</tr>
<tr>
<td>Survive 3 failed replicas</td>
<td>4</td>
<td>4</td>
<td>4x</td>
</tr>
</tbody>
</table>

An index with multiple copies remains available in case of an instance outage. For example, if an index has 2 copies stored on 2 instances and one of those instances fails, one copy of the index remains available for servicing requests.

5. Use this formula to determine the total number of documents your HCM system must be able to store:

\[
\text{documents per second from step 2} \\
\times 3600 \text{ seconds in an hour} \\
\times 24 \text{ hours in a day} \\
\times \text{number of days you want to store data (default is 30)} \\
\times \text{Impact from the data availability table in step 4.} \\
= \text{Total document count}
\]

For example, if your HCP system produces 1500 documents per second, you want to store data for 30 days, and you want to maintain 2 copies of each index containing the stored data, your system must have enough instances to be able to store roughly 8 billion documents:

\[
1500 \\
\times 3600 \\
\times 24 \\
\times 30 \\
\times 2 \\
= 7,776,000,000
\]
6. Use this table to determine the number of instances required based on the total number of documents your HCM must store.

<table>
<thead>
<tr>
<th>Total document count</th>
<th>Instances required</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 billion or less</td>
<td>1</td>
</tr>
<tr>
<td>4 billion or less</td>
<td>2</td>
</tr>
<tr>
<td>6 billion or less</td>
<td>3</td>
</tr>
<tr>
<td>8 billion or less</td>
<td>4</td>
</tr>
</tbody>
</table>

7. Take the highest number of instances from steps 2, 3, and 6. That's the number instances you need.

**Monitoring smaller HCP systems**
If the HCP system you are monitoring is not generating a lot of traffic, a single, less-powerful instance may be sufficient. Use the following guidelines:

<table>
<thead>
<tr>
<th>Documents per second</th>
<th>Cores</th>
<th>RAM (GB)</th>
<th>Disk (GB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up to 1200</td>
<td>8</td>
<td>28</td>
<td>600</td>
</tr>
<tr>
<td>1200-1600</td>
<td>12</td>
<td>32</td>
<td>800</td>
</tr>
<tr>
<td>1600-2000</td>
<td>16</td>
<td>40</td>
<td>1000</td>
</tr>
<tr>
<td>2000-2400</td>
<td>18</td>
<td>48</td>
<td>1400</td>
</tr>
<tr>
<td>2400-2800</td>
<td>20</td>
<td>56</td>
<td>1700</td>
</tr>
<tr>
<td>2800-3200</td>
<td>24</td>
<td>64</td>
<td>2000</td>
</tr>
</tbody>
</table>

**Operating system and Docker requirements**
To be a system instance, each server or virtual machine you provide:

- Must run a 64-bit Linux distribution
- Must have Docker version 1.13.1 or later installed

**Important:** Install the current Docker version suggested by your operating system, unless that version is earlier than 1.13.1. The system cannot run with Docker versions prior to 1.13.1.
This table shows the operating systems and Docker and SELinux configurations that Hitachi Content Intelligence has been qualified with. For more information, see "Docker considerations" below and "SELinux considerations" on the facing page.

<table>
<thead>
<tr>
<th>Operating System</th>
<th>Docker Version</th>
<th>Docker Storage Configuration</th>
<th>SELinux setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fedora 27</td>
<td>Docker 1.13.1-58.git87f2fab.el7.x86_64</td>
<td>direct-lvm</td>
<td>Enforcing</td>
</tr>
<tr>
<td>Red Hat Enterprise Linux 7.4</td>
<td>Docker 1.13.1-58.git87f2fab.el7.x86_64</td>
<td>direct-lvm</td>
<td>Enforcing</td>
</tr>
<tr>
<td>Ubuntu 16.04-LTS</td>
<td>Docker 17.03.0-ce</td>
<td>aufs</td>
<td>N/A</td>
</tr>
<tr>
<td>CentOS 7.4</td>
<td>Docker 18.03.1-ce</td>
<td>overlay</td>
<td>Enforcing</td>
</tr>
</tbody>
</table>

**Docker considerations**

- The Docker installation directory on each instance must have at least 20 GB available for storing the Hitachi Content Intelligence Docker images.

- Make sure that the Docker storage driver is configured correctly on each instance before installing Hitachi Content Intelligence.

After installing Hitachi Content Intelligence, changing the Docker storage driver requires a reinstallation of Hitachi Content Intelligence.

To view the current Docker storage driver on an instance, run:

```
docker info
```

- If you want to enable SELinux on the system instances, you need to use a Docker storage driver that supports it. The storage drivers that SELinux supports differ depending on the Linux distribution you’re using. For more information, see the Docker documentation.

- If you are using the Docker devicemapper storage driver:
  
  - Make sure that there's at least 40 MB of Docker metadata storage space available on each instance. Hitachi Content Intelligence requires 20 MB to install successfully and an additional 20 MB to successfully update to a later version.

  To view Docker metadata storage usage on an instance, run:
On a production system, do not run `devicemapper in loop-lvm mode`. This can cause slow performance or, on certain Linux distributions, Hitachi Content Intelligence may not have enough space to run.

**SELinux considerations**

- You should decide whether you want to run SELinux on system instances and enable or disable it before installing Hitachi Content Intelligence.

  Enabling or disabling SELinux on an instance requires you to reboot the instance.

  To view whether SELinux is enabled on an instance, run:

  ```
  sestatus
  ```

- If you want to enable SELinux on the system instances, you need to use a Docker storage driver that supports it.

  The storage drivers that SELinux supports differ depending on the Linux distribution you're using. For more information, see the Docker documentation.

**Virtual machine host requirements**

To deploy the Hitachi Content Intelligence example OVF, your virtual machine host must run VMware ESXi version 6.0 or later.

The Hitachi Content Intelligence example OVF has been qualified on these virtual machine host platforms:

- VMware ESXi 6.0
- Hitachi Unified Compute Platform (UCP) 4.1.0

**Networking**

This topic describes the network usage and requirements for both system instances and services.

**Note:** You can configure the network settings for each service when you install the system. You cannot change these settings after the system is up and running.
**Note:** If your networking environment changes such that the system can no longer function with its current networking configuration, you need to reinstall the system. See "Handling IP address changes" in the administrator help, which is accessible from the Admin App.

**Instance IP address requirements**

All instance IP addresses must be static. This includes both internal and external network IP addresses, if applicable to your system.

**Important:** If the IP address of any instance changes, see "Handling IP address changes" in the administrator help, which is accessible from the Admin App.

**Network types**

Each of the Hitachi Content Intelligence services can bind to one type of network, either **internal** or **external**, for receiving incoming traffic. If your network infrastructure supports having two networks, you may want to isolate the traffic for most system services to a secured internal network that has limited access. You can then leave only the Search-App and Admin-App on your external network for user access.

You can use either a single network type for all services or a mix of both types. If you want to use both types, every instance in your system must be addressable by two IP addresses; one on your internal network, and one on your external network. If you use only one network type, each instance needs only one IP address.

**Allowing access to external resources**

Regardless of whether you're using a single network type or a mix of types, you need to configure your network environment to ensure that all instances have outgoing access to the external resources you want to use, such as:

- The data sources where your data is stored
- Identity providers for user authentication
- Email servers that you want to use for sending email notifications
- Any external search indexes (for example, HDDS indexes) that you want make accessible through Hitachi Content Intelligence
Ports

Each service binds to a number of ports for receiving incoming traffic.

Before installing Hitachi Content Intelligence, you can configure the services to use different ports, or use the default values shown in the tables below.

Note: If you reconfigure service ports, make sure that each port value you assign is unique across all services, both System services and Hitachi Content Intelligence services.
### System-external ports

The following table contains information about the service ports that users use to interact with the system.

On every instance in the system, each of these ports:

- Must be accessible from any network that requires administrative or search access to the system
- Must be accessible from every other instance in the system

![Warning Icon] **Note:** Debug ports are accessible only when debug is set to true in /<installation-directory>/config/cluster.config

<table>
<thead>
<tr>
<th>Default Port Value</th>
<th>Service</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>8000</td>
<td>Admin-App</td>
<td>Access to administrative interfaces:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Administration App</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Administrative REST API</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Administrative CLI</td>
</tr>
<tr>
<td>6162</td>
<td>Monitor-App</td>
<td>Access to the HCM application, which is used to monitor the health of HCP systems.</td>
</tr>
<tr>
<td>8888</td>
<td>Search-App</td>
<td>Access to search interfaces:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Search App</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Workflow Designer</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Search REST API</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Workflow Designer REST API</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Search CLI</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Workflow Designer CLI</td>
</tr>
</tbody>
</table>
**System-internal ports**

This table lists the ports used for intra-system communication by the services. On every instance in the system, each of these ports:

- Must be accessible from every other instance in the system
- Should not be accessible from outside the system

You can find more information on how these ports are used in the documentation for the third-party software underlying each service.

<table>
<thead>
<tr>
<th>Default Port Value</th>
<th>Used By</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>5601</td>
<td>Dashboard service</td>
<td>Primary port for communicating with the Dashboard service.</td>
</tr>
<tr>
<td>7000</td>
<td>Database service</td>
<td>TCP port for commands and data.</td>
</tr>
<tr>
<td>7199</td>
<td>Database service</td>
<td>Port for JMX connections to Database service.</td>
</tr>
<tr>
<td>9042</td>
<td>Database service</td>
<td>Primary port for communicating with the Database service.</td>
</tr>
<tr>
<td>8020</td>
<td>Clustered-File-System service</td>
<td>Port used for file system metadata operations.</td>
</tr>
<tr>
<td>8480</td>
<td>Clustered-File-System service</td>
<td>HTTP port JournalNodes.</td>
</tr>
<tr>
<td>8481</td>
<td>Clustered-File-System service</td>
<td>HTTPs port for JournalNodes.</td>
</tr>
<tr>
<td>8485</td>
<td>Clustered-File-System service</td>
<td>Port for the JournalNode RPC server.</td>
</tr>
<tr>
<td>9000</td>
<td>Clustered-File-System service</td>
<td>Port used for file system metadata operations.</td>
</tr>
<tr>
<td>50010</td>
<td>Clustered-File-System service</td>
<td>Port for DataNode data transfers.</td>
</tr>
<tr>
<td>50020</td>
<td>Clustered-File-System service</td>
<td>Port for DataNode IPC server.</td>
</tr>
<tr>
<td>Default Port Value</td>
<td>Used By</td>
<td>Purpose</td>
</tr>
<tr>
<td>--------------------</td>
<td>----------------------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>50070</td>
<td>Clustered-File-System service</td>
<td>Port for the web UI used to view the current status of the service and explore the clustered file system.</td>
</tr>
<tr>
<td>50090</td>
<td>Clustered-File-System service</td>
<td>Secondary HTTP port for NameNodes.</td>
</tr>
<tr>
<td>50091</td>
<td>Clustered-File-System service</td>
<td>Secondary HTTPS port for NameNodes.</td>
</tr>
<tr>
<td>50075</td>
<td>Clustered-File-System service</td>
<td>Port for the web UI used to access the status and logs for DataNodes.</td>
</tr>
<tr>
<td>50470</td>
<td>Clustered-File-System service</td>
<td>HTTPS port for NameNodes.</td>
</tr>
<tr>
<td>50475</td>
<td>Clustered-File-System service</td>
<td>HTTPS port for DataNodes.</td>
</tr>
<tr>
<td>5003</td>
<td>Index service</td>
<td>Debug port used by the Index service.</td>
</tr>
<tr>
<td>8983</td>
<td>Index service</td>
<td>Primary port used to communicate with the Index service.</td>
</tr>
<tr>
<td>10000</td>
<td>Index service</td>
<td>Port used by the Index service for graceful shutdowns.</td>
</tr>
<tr>
<td>5123</td>
<td>Monitor-App service</td>
<td>The debug port used by the Monitor App.</td>
</tr>
<tr>
<td>6175</td>
<td>Monitor-App service</td>
<td>The port used by the Monitor App for graceful shutdowns.</td>
</tr>
<tr>
<td>7203</td>
<td>Message Queue service</td>
<td>Port for JMX connections to Message Queue service.</td>
</tr>
<tr>
<td>9092</td>
<td>Message Queue service</td>
<td>Primary port for communicating with Message Queue service.</td>
</tr>
<tr>
<td>9600</td>
<td>Logging service</td>
<td>Primary port for communicating with Logging service.</td>
</tr>
<tr>
<td>9601</td>
<td>Logging service</td>
<td>The port used to receive syslog messages.</td>
</tr>
<tr>
<td>9200</td>
<td>Metrics service</td>
<td>Port used to communicate with the Metrics service cluster.</td>
</tr>
<tr>
<td>Default Port Value</td>
<td>Used By</td>
<td>Purpose</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>9201</td>
<td>Metrics service</td>
<td>Port used to communicate with an individual Metrics service node.</td>
</tr>
<tr>
<td>9301</td>
<td>Metrics service</td>
<td>Port that nodes in the Metrics service cluster should use when communicating with each other.</td>
</tr>
<tr>
<td>8081</td>
<td>Scheduling service</td>
<td>Primary port for communicating with the Scheduling service.</td>
</tr>
<tr>
<td>5002</td>
<td>Search-App service</td>
<td>Debug port used by the Search-App service.</td>
</tr>
<tr>
<td>8006</td>
<td>Search App service</td>
<td>Port used by the Search App service for graceful shutdowns.</td>
</tr>
<tr>
<td>5007</td>
<td>Sentinel service</td>
<td>Debug port used by Sentinel service.</td>
</tr>
<tr>
<td>8007</td>
<td>Sentinel service</td>
<td>Port used by the Sentinel service for graceful shutdowns.</td>
</tr>
<tr>
<td>8889</td>
<td>Sentinel service</td>
<td>Primary port for communicating with Sentinel.</td>
</tr>
<tr>
<td>18889</td>
<td>Sentinel service</td>
<td>Sentinel service internal communication.</td>
</tr>
<tr>
<td>5001</td>
<td>Admin-App service</td>
<td>Debug port for Admin-App service.</td>
</tr>
<tr>
<td>8005</td>
<td>Admin-App service</td>
<td>Port used by Admin-App for graceful shutdowns.</td>
</tr>
<tr>
<td>18000</td>
<td>Admin-App service</td>
<td>Admin-App internal communication.</td>
</tr>
<tr>
<td>5555</td>
<td>Watchdog service</td>
<td>Port for JMX connections to Watchdog service.</td>
</tr>
<tr>
<td>9091</td>
<td>Network-Proxy service</td>
<td>Primary port for communicating with Network-Proxy.</td>
</tr>
<tr>
<td>8080</td>
<td>Service-Deployment service</td>
<td>Primary port for communicating with Service-Deployment.</td>
</tr>
<tr>
<td>18080</td>
<td>Service-Deployment service</td>
<td>Service-Deployment internal communication</td>
</tr>
<tr>
<td>5050</td>
<td>Cluster-Coordination service</td>
<td>Primary port for communicating with Cluster-Coordination.</td>
</tr>
<tr>
<td>Default Port Value</td>
<td>Used By</td>
<td>Purpose</td>
</tr>
<tr>
<td>--------------------</td>
<td>----------------------------------</td>
<td>-----------------------------------------------------</td>
</tr>
<tr>
<td>15050</td>
<td>Cluster-Coordination service</td>
<td>Cluster-Coordination internal communication</td>
</tr>
<tr>
<td>5051</td>
<td>Cluster-Worker service</td>
<td>Primary port for communicating with Cluster-Worker.</td>
</tr>
<tr>
<td>2181</td>
<td>Synchronization service</td>
<td>Synchronization service client port.</td>
</tr>
<tr>
<td>2888</td>
<td>Synchronization service</td>
<td>Synchronization service internal communication.</td>
</tr>
<tr>
<td>3888</td>
<td>Synchronization service</td>
<td>Synchronization service leader election.</td>
</tr>
<tr>
<td>4040</td>
<td>Workflow jobs</td>
<td>Spark UI port.</td>
</tr>
<tr>
<td>5005</td>
<td>Workflow jobs</td>
<td>The port to use for debugging the job driver.</td>
</tr>
<tr>
<td>5008</td>
<td>Workflow jobs</td>
<td>The port to use for debugging the job executor.</td>
</tr>
<tr>
<td>31000-34000</td>
<td>Cluster-Coordination and Cluster-Worker services</td>
<td>High ports used by both Mesos and Docker.</td>
</tr>
</tbody>
</table>

**Supported browsers**

The Hitachi Content Intelligence web applications support these web browsers:

- Google Chrome latest
- Mozilla Firefox latest
- Microsoft Edge latest

**Hitachi Content Intelligence documentation**

The following documents contain additional information about the system:

- **Administrator Help** (MK-HCI000) — Contains information on administering an Hitachi Content Intelligence system.

  A subset of the information contained in this document is accessible from the Admin App.
• **Search Help** (MK-HCI001) — Contains information on using a system to search for files.

This document is accessible from the Search App.

• **Installing Hitachi Content Intelligence** (MK-HCI002) — This document contains information on installing an Hitachi Content Intelligence system, either on physical servers or virtual machines that you provide.

This document is contained within the Hitachi Content Intelligence-<version-number>.tgz installation package and is also available from within the Administrator Help.

• **Deploying the HCI Example OVF** (MK-HCI003) — This document contains information on deploying a Hitachi Content Intelligence system for testing and evaluation using the example OVF installation package provided by Hitachi Vantara.

This document is contained within the Hitachi Content Intelligence-<version-number>.tgz installation package.

• **Workflow Designer Help** (MK-HCI007) — Contains information on using Hitachi Content Intelligence workflows to process your data and make it searchable.

This document is accessible from the Workflow Designer App.

---

**Update notes**

**Update version qualification**

Hitachi Content Intelligence updates have been qualified from version 1.2.1 to all later versions.

An HCI 1.0 system cannot be updated to any later version.

**Version-specific update prerequisites**

• Before updating to version 1.2.0 or later, you need to configure the mmap count setting on all instances. To do this, add the following line to the `/etc/sysctl.conf` file:

```bash
vm.max_map_count = 262144
```
Before updating to version 1.4.0 or later, check the version of Docker running on each instance in your system:

```
docker version
```

If the Docker API version is earlier than 1.26, you need to update Docker before you can update HCI. To do this:

1. Stop the run script from running. You do this using whatever method you're currently using to run the script.

2. Update Docker to the most up to date version supported by HCI. For information on how to do this, see the Docker website.

3. Restart the run script.

**Service REST API and CLI differences from release 1.0**

With release 1.2, a number of changes were made to the administrative REST API and CLI for managing services. These changes are not backwards compatible with version 1.0 of the Hitachi Content Intelligence administrative REST API.
<table>
<thead>
<tr>
<th>Removed</th>
<th>Replacement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>REST API Endpoint</strong></td>
<td><strong>CLI Command</strong></td>
</tr>
<tr>
<td><strong>GET /services</strong></td>
<td>listServices</td>
</tr>
<tr>
<td><strong>GET /services/configurable</strong></td>
<td>listUserConfigurableServices</td>
</tr>
<tr>
<td><strong>GET /services/configure/recommend</strong></td>
<td>getServiceConfigRecommendation</td>
</tr>
<tr>
<td><strong>GET /services/detailed</strong></td>
<td>listServicesDetailed</td>
</tr>
<tr>
<td><strong>GET /services/{id}</strong></td>
<td>getService</td>
</tr>
<tr>
<td><strong>GET /services/{id}/config</strong></td>
<td>getServiceConfig</td>
</tr>
<tr>
<td><strong>GET /services/{id}/detailed</strong></td>
<td>getServiceDetails</td>
</tr>
<tr>
<td><strong>GET /services/{id}/status</strong></td>
<td>getServiceStatus</td>
</tr>
<tr>
<td><strong>GET /services/deploy</strong></td>
<td>getDeployTaskStatus</td>
</tr>
<tr>
<td><strong>POST /services/deploy</strong></td>
<td>deploySystem</td>
</tr>
<tr>
<td><strong>POST /services/configure</strong></td>
<td>updateServiceConfig</td>
</tr>
</tbody>
</table>

**Note about using the log download tool**

When using the log_download tool, if you specify the --output option, do not specify an output path that contains colons, spaces, or symbolic links. If you omit the --output option, you cannot run the script from within a directory path that contains colons, spaces, or symbolic links.
## Known issues

<table>
<thead>
<tr>
<th>Issue</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
</table>
| ENS-4051| Services, system deployment, update  | **Service fails to start during system deployment and update**  
Rarely, a system deployment, service management operation, or system update may fail because a service fails to start. When this happens, the Admin App is inaccessible from the instance where the failure occurred.  

The logs in the watchdog-service log directory contain this error:

```
Error response from daemon: Conflict. The name "<service-name>" is already in use by container <Docker-container-id>. You have to remove (or rename) that container to be able to reuse that name.
```

For information on collecting logs, see **Retrieving logs and diagnostic information**.

**Workaround**

Restart the Docker service on the instance where the service failed. For example, if you are using systemd to run Docker, run:

```
systemctl restart docker
```

After restarting Docker, retry the system deployment, service management operation, or system update.
<table>
<thead>
<tr>
<th>Issue</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
</table>
| ENS-5045  | Metrics service | **Moving multiple instances of the Metrics service requires Index Protection Levels to be increased**  
The Metrics services maintains a number of indexes used for storing system events and workflow task data. To successfully move multiple instances of the Metrics service at one time, you need to increase the Index Protection Level setting for every index that the service manages.  

**Workaround**  
To move all instances of the Metrics service without having to increase Index Protection Levels, move only one Metrics service instance at a time. First add new instances of the service, then remove the service from the system instances where it is currently running. |
| ENS-6176  |               | **Node does not start due to port already being in use**                                                                                                                                                   |
| ENS-6489  | Workflows     | **Workflow task halts with InvalidQueryException: Batch too large**  
Rarely, a workflow task may halt with this task error:  

    com.hds.ensemble.cassandra.CassandraConnector  
    [CassandraConnector.java:184] Retries exceeded (1)  

    com.datastax.driver.core.exceptions.InvalidQueryException:  
    Batch too large  

This occurs when a document, as retrieved from a data connection, contains too much metadata for the Database service to store for a single document. |
| ENS-6629  | Index service | **Errors when changing index IPL while a service instance is down**  
If you change the Index Protection Level for an index while any instances of the Index service are unavailable, the change appears to succeed but the index's replicas are not actually affected. |
<table>
<thead>
<tr>
<th>Issue</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
</table>
| ENS-6662| HCP data connections        | **HCP and HCP MQE data connections differ in handling case for some settings**  
When a user configures the regular HCP data connection in the Admin App, the values specified for the **HCP Tenant Name** and **HCP Namespace Name** settings are internally converted to lower case letters. When a user configures the HCP MQE data connection, the values entered for these fields are taken as-is; they are not converted to all lower case.  
This issue affects the values for these document fields:  
- HCI_id  
- HCI_namespace  
- HCI_URI  
**Workaround**  
When configuring the HCP MQE data connection, specify only lower case letters for the **HCP Tenant Name** and **HCP Namespace Name** fields. |
| ENS-7397| System update               | **On update to version 1.3.0, jobDriver logs are not rotated or downloadable**  
With release 1.3.0, the system no longer stores logs in the com.hds.ensemble.plugins.service.jobDriver directory. Upon updating to this version, no changes are made to any existing logs in this directory. These logs are not rotated and cannot be downloaded using the log_download script. |
| ENS-7957| System update               | **Network types cannot be configured for new services before system update**  
Before starting an update, you are prompted to specify network configuration for any new services included in the version that you’re updating to. However, you can specify only the port numbers for the new service. You cannot specify the network type (that is, internal or external) for the service to use. Each new service gets the default network type, which is determined by the service itself. |
<table>
<thead>
<tr>
<th>Issue</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
</table>
| ENS-7962 | System update | **Uploading an update package fails after the failure and recovery of a system instance**  
If a system instance enters the Down state, when you attempt to upload an update package, the upload fails. However, after the system instance recovers, when you try again to upload an update package, the upload again fails, even though the system is in a healthy state.  

**Workaround**  
1. In the Admin App, go to the Monitoring > Processes page and click on the Retry Task button for the Upload Plugin Bundle task.  
2. Try uploading the update package again. |
| ENS-7964 | Volumes       | **Volume configuration is not displayed correctly in Admin App**  
During installation, you can configure volumes for System services by specifying different values in the volume.config file on each system instance. Each volume is correctly configured with the settings you specify, but the Monitoring > Services > Service Details page in the Admin App incorrectly shows each volume as having identical configuration. |
| ENS-8666 | Volumes       | **With some Docker versions, services and jobs that use NFS-backed volumes may fail to deploy while SELinux is enabled**  
Services and jobs can be configured to use the Docker local volume driver for storing data on NFS servers. Such services and jobs may fail to deploy when:  
- System instances are running a version of Docker later than 1.12.6, and  
- SELinux is enabled  

**Workarounds**  
- Update Docker to Community Edition (ce) version 18.03 or later  
- Disable SELinux |
<table>
<thead>
<tr>
<th>Issue</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
</table>
| ENS-10148 | System update          | Updating from version 1.2.x to 1.4 fails due to port error  
Performing an update from a 1.2 version to 1.4 does not work when keeping the default value for STOP_PORT in the Search App service.  
**Workaround**  
Check to make sure the port is not in use when updating. If update fails, cancel the update entirely, upload the package again, and modify the STOP_PORT value in the configuration settings. |
| ENS-10310 | Identity providers     | Identical names allowed for groups across identity providers  
Groups in separate identity providers can have the same name. The Document Security stage identifies groups by name instead of UUID and cannot differentiate groups with the same name.  
**Workaround**  
Do not duplicate names when creating groups to be used by the Document Security stage. |
| ENS-10639 | Import historical logs | Log downloads start 30 days prior to download date  
Resource metric logs downloaded from HCP start 30 days prior to the date the logs are downloaded. When importing multiple logs from the same system, data ends up duplicated. |
| ENS-10750 | Volumes                | Volume validation does not work properly for update  
Before initiating a system update, you can specify configuration settings for any new volumes that the update will create on your system. However, the system does not verify the configuration settings you specify are correct.  
**Workaround**  
Use caution when editing volume configuration. Make sure your the instances in your system allow volumes to be created with the configuration you specify. |
<table>
<thead>
<tr>
<th>Issue</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
</table>
| ENS-10793 | System update, workflow tasks | **Workflow task goes into idle state during update**<br>During update, any running workflow tasks go into an idle state instead of pausing.  
**Workaround**<br>Make sure any workflow tasks are either completed or canceled before starting an update. |
| ENS-10804 | Workflow tasks                | **Workflow task halts if data source root directory is changed**<br>When importing logs, an expansion workflow may halt if the HCP root directory is modified after creation.  
**Workaround**<br>Once a source is configured, do not modify it. Delete the source and add a new one with the correct configuration. |
| ENS-10889 | Admin App                     | **Admin App unavailable after removing a master instance on a system with dual networks**<br>The Admin App becomes inaccessible after removing a master instance from a system that has two networks configured. |
| ENS-10919 | Jobs                          | **Anomaly Detection jobs not running after update**<br>After updating from version 1.3.1 to 1.4, Anomaly Detection jobs stop running on system instances. This prevents Analytics from being added to an existing signal source.  
**Workaround**<br>After update, assign instances to the job type before enabling the jobs. |
## Resolved Issues

### Resolved in release 1.4.1

<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCI-486</td>
<td>Workflow Designer</td>
<td><strong>Admin App fails to import package containing a workflow pipeline with conditional groups</strong>&lt;br&gt;Importing a package with conditional groups from a previous version causes issues when the contents are converted to business objects.&lt;br&gt;&lt;br&gt;<strong>Fix:</strong> Packages containing conditional groups now import successfully.</td>
</tr>
<tr>
<td>HCI-396</td>
<td>Hitachi Content Monitor</td>
<td><strong>System log messages fail to process in load balancing environments</strong>&lt;br&gt;In load balancing environments, the DNS resolves an HCP name to the IP address of the load balancer instead of the HCP node IP address, causing system log messages from the node to be ignored.&lt;br&gt;&lt;br&gt;<strong>Fix:</strong> HCP node IP addresses can now be specified in load balancing environments by using a new configuration property available in this release.</td>
</tr>
<tr>
<td>HCI-2</td>
<td>Workflow Designer</td>
<td><strong>Amazon S3 connectors fail when an endpoint contains an IP address</strong>&lt;br&gt;Amazon S3 connectors fail when an IP address is used for an endpoint, instead of a fully qualified domain name.&lt;br&gt;&lt;br&gt;<strong>Fix:</strong> Amazon S3 compatible data connections can now be established with an endpoint that uses an FQDN, or an IP address.</td>
</tr>
<tr>
<td>Reference number</td>
<td>Area affected</td>
<td>Description</td>
</tr>
<tr>
<td>------------------</td>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>HCI-7</td>
<td>Workflow Designer</td>
<td><strong>Amazon S3 connectors fail when an endpoint contains a custom port number</strong>&lt;br&gt;Amazon S3 compatible data connections fail when custom port numbers are used in place of the default HCI port numbers (80/443).&lt;br&gt;&lt;br&gt;<strong>Fix:</strong> Amazon S3 compatible data connections now resolve the IP address of an endpoint, so that both custom and default ports can be used when establishing a connection.</td>
</tr>
<tr>
<td>HCI-113</td>
<td>Search</td>
<td><strong>Query refinements added to original query string produces inaccurate results</strong>&lt;br&gt;When refinements are added to a query, the original query becomes &quot;optional&quot; and results can be inaccurate.&lt;br&gt;&lt;br&gt;<strong>Fix:</strong> When refining a query, the entire query is now handled as &quot;required&quot;, producing more accurate results.</td>
</tr>
<tr>
<td>HCI-388</td>
<td>Workflow Designer</td>
<td><strong>POSIX utime and atime collection not supported</strong>&lt;br&gt;HCP connectors do not support the collection of POSIX utime and atime for an object.&lt;br&gt;&lt;br&gt;<strong>Fix:</strong> The POSIX utime and atime object attributes are now supported.</td>
</tr>
<tr>
<td>HCI-8</td>
<td>Workflow Designer</td>
<td><strong>HCP Anywhere connectors do not include filesystem Type field</strong>&lt;br&gt;HCP Anywhere connectors do not create the filesystem Type field, which help to refine query results.&lt;br&gt;&lt;br&gt;<strong>Fix:</strong> HCP Anywhere connectors now include this field type.</td>
</tr>
<tr>
<td>Reference number</td>
<td>Area affected</td>
<td>Description</td>
</tr>
<tr>
<td>------------------</td>
<td>------------------------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>HCI-28</td>
<td>Hitachi Content Monitor</td>
<td><strong>Monitor History pages display incorrect page count</strong>&lt;br&gt;When data is collected from a signal source and viewed in Monitor History, the total page count for the data is incorrect.&lt;br&gt;&lt;br&gt;<strong>Fix:</strong> Monitor History now shows the correct total page count.</td>
</tr>
<tr>
<td>HCI-206, HCI-489</td>
<td>Workflow Designer</td>
<td><strong>Custom metadata fields beginning with &quot;0&quot; values truncated in Content Class stage</strong>&lt;br&gt;Content classes extract numerical data types that contain leading &quot;0&quot; values (0000123, for example) by interpreting the string as &quot;123&quot;, excluding the leading &quot;0&quot; values.&lt;br&gt;&lt;br&gt;<strong>Fix:</strong> Content classes can now be configured to extract numerical data without guessing the value type.</td>
</tr>
<tr>
<td>HCI-390</td>
<td>Hitachi Content Monitor</td>
<td><strong>Monitor missing option for creating service metrics</strong>&lt;br&gt;Monitors cannot be configured with service metrics that show objects classified as &quot;examined&quot;, &quot;serviced&quot;, or &quot;failed&quot;.&lt;br&gt;&lt;br&gt;<strong>Fix:</strong> These metrics can now be configured when creating a new monitor.</td>
</tr>
<tr>
<td>HCI-293</td>
<td>Hitachi Content Monitor</td>
<td><strong>HCM Dashboard time range resets on auto-refresh</strong>&lt;br&gt;Dashboards containing imported logs with specified time ranges do not display the correct time range when the dashboard is refreshed. Instead, the time range resets to the default time.&lt;br&gt;&lt;br&gt;<strong>Fix:</strong> The correct time range is now displayed for a log when the dashboard is refreshed.</td>
</tr>
</tbody>
</table>
### Resolved Issues

<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
</table>
| HCI-16           | Workflow Designer | **Amazon S3 Connectors do not contain latest list of regions**  
Amazon S3 connectors do not contain the latest list of currently known regions.  
**Fix:** The list has been updated in this release. |

### Resolved in release 1.4.0

<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
</table>
| ENS-4409         | Workflow tasks, Text and Metadata Extraction stage | **Workflow task halts due to error with Text and Metadata Extraction stage**  
Rarely, if you use the Text and Metadata Extraction stage in your pipeline, your workflow tasks may halt.  
**Fix:** These workflow tasks no longer halt. |
| ENS-8026         | Jobs          | **Scheduled jobs stop and restart every day even when there's no break in the schedule**  
If you schedule a job to run from 00:00 to 24:00 on two consecutive days, the job stops and starts again at the beginning of the second day when it should instead continue running.  
**Fix:** Jobs no longer stop and restart when scheduled to run over two consecutive days. |
| ENS-8077         | Email Expansion stage | **Attachments are not extracted from undeliverable emails**  
For undeliverable journaled emails, the Email Expansion stage does not extract the attached files that describe why the email couldn’t be delivered.  
**Fix:** Attachments are now extracted from undeliverable emails. |
## Resolved Issues

<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENS-8078</td>
<td>Email Expansion stage</td>
<td><strong>Attached emails are not extracted</strong>&lt;br&gt; If a journaled email contains another email as an attachment, the Email Expansion stage does not extract the attached email.&lt;br&gt;&lt;br&gt;<strong>Fix:</strong> Emails attached to other emails are now extracted as separate documents.</td>
</tr>
<tr>
<td>ENS-8422</td>
<td>Product installation</td>
<td><strong>Product installation fails due to kernel configuration</strong>&lt;br&gt;Product installations fail on Linux distributions where the kernel is configured with the CONFIG_MODIFY_LDT_SYSCALL setting disabled:&lt;br&gt;&lt;br&gt;[ \text{CONFIG_MODIFY_LDT_SYSCALL} = \text{n} ]&lt;br&gt;&lt;br&gt;<strong>Fix:</strong> Product installations now succeed regardless of how this setting is configured.</td>
</tr>
<tr>
<td>Reference number</td>
<td>Area affected</td>
<td>Description</td>
</tr>
<tr>
<td>------------------</td>
<td>--------------------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>ENS-8701</td>
<td>Syslog components</td>
<td><strong>If multiple syslog components exist, the settings from only one are used</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>HCI includes several components that you can use for sending syslog events:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• In the Admin App:&lt;br&gt;  ○ Syslog notification rules, for system events&lt;br&gt;  ○ Syslog Notification pipeline stages, for document events</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• In the Monitor App, syslog senders, for notifications about triggered monitors</td>
</tr>
<tr>
<td></td>
<td></td>
<td>If you create more than one of these components in a single location, the system only uses the configuration settings from one of them. For example, if you add multiple Syslog Notification stages to a workflow pipeline, each one logs the same events to the same syslog server, even if you've configured the stages differently. <strong>Fix:</strong> Multiple syslog components can now exist simultaneously, with each uses its own configuration settings.</td>
</tr>
<tr>
<td>ENS-8706</td>
<td>Scheduling jobs</td>
<td><strong>Scheduled jobs end days later than their scheduled stop times</strong>&lt;br&gt; If you schedule a job to run until 24:00 on any given day, the job does not actually stop until 24:00 two or three days later. For example, if a job is scheduled to end at 24:00 on Sunday, it does not actually stop until 24:00 on Tuesday. <strong>Fix:</strong> Scheduled jobs now end at the correct time.</td>
</tr>
</tbody>
</table>
### Resolved Issues

<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
</table>
| **ENS-8684**     | Workflow tasks    | **Workflow task operations fail with 504 errors**  
After HCI is updated to version 1.3.1 or later, all workflow task operations fail with 504 errors. This includes attempts to run, pause, resume, delete, or clear workflow tasks.  

**Fix:** This issue no longer occurs.  |
| **ENS-7921**     | Deleting workflows| **Workflow deletion fails because job deletion fails**  
Occasionally, a workflow may fail to be deleted because its underlying job fails to be deleted.  

**Fix:** This issue no longer occurs.  |
## Resolved in release 1.3.1

<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
</table>
| ENS-7998         | Documentation   | **List of installed software is incorrect in Deploying the HCI Example OVF**  
The list of software in the About the HCI Example OVF topic in the document Deploying the HCI Example OVF is incorrect. The correct list is:  

```
cracklib*  
docker  
java-1.8.0-openjdk-devel  
nfs-utils  
ntp  
perl  
samba  
samba-client  
unzip  
wget  
```

**Fix:** The document has been updated with the correct list. |
| ENS-8217         | Index service   | **Index shard replicas are incorrectly placed on the same instance**  
When you increase the Index Protection Level (IPL) setting for an HCI Index, the system creates replicas of each of the index’s shards. However, multiple replicas of the same shard may be placed on the same system instance, meaning that increasing the IPL may not actually make the index more resistant to an instance failure.  

**Fix:** When you increase an index's IPL, shard replicas are now always placed on different system instances. |
| ENS-8218         | Search-App      | **Search-App timeout not long enough for complex queries**  
Requests to the Search-App may time out if they contain very complex queries.  

**Fix:** The Search-App request timeout limit has been increased. |
<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
</table>
| ENS-8219         | Hitachi Content Monitor | **Metrics collection halts when using syslog signal**  
If you've enabled the syslog signal for an HCP source, metrics collection for the source can halt due to improperly formatted syslog messages reported by the HCP system. While metrics collection is halted, visualizations for the source are not updated.  
**Fix:** Metrics collection no longer halts in this situation. |
| ENS-8223         | Admin App              | **Tables on the job monitoring pages are not sortable**  
In the Admin App, the tables on these pages cannot be sorted:  
- Monitoring > Jobs > Job Type  
- Monitoring > Jobs > Job Type > Job Details  
**Fix:** These tables are now sortable. |
| ENS-8236         | log_download script    | **log_download script does not work when Hitachi Content Intelligence is down**  
The log_download script does not collect logs if the Hitachi Content Intelligence system is down, even when run with the `--offline` option.  
**Fix:** If the Hitachi Content Intelligence system is down, running the log_download script with the `--offline` option collects the logs from the instance on which you are running the script. To collect all logs, run the script with that option on each system instance. |
<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
</table>
| ENS-8237         | Database service | **Database service has incorrect replication factor after being scaled**  
The Database service has an internal configuration setting that determines the number of copies to maintain for each database row. When you run the Database service on 3 or more system instances, the replication factor is set to 3. When you scale the service down to 1 or 2 instances, the replication factor is set to 1. However if you scale the service back up to 3 or more instances, the replication factor is not reset to 3. This can cause performance issues for running workflows.  
**Fix:** When you scale the Database, its replication factor is correctly readjusted. |
| HCISUS-55        |              |             |
| ENS-8254         | Hitachi Content Monitor | **HCM fails to monitor HCP systems that use signed SSL certificates**  
HCM cannot monitor HCP systems that use CA-signed SSL certificates, but can monitor those that use self-signed certificates.  
**Fix:** HCM can now monitor these HCP systems. |
<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
</table>
| ENS-8256         | Admin App, user permissions | **Monitoring pages do not load for users without the admin:services:create permission**  
If a user does not have the admin:services:create permission, the Monitoring > Overview and Monitoring > Services pages in the Admin App display Error: Could not get services but continuously try to load.  

**Fix:** For users with incorrect permissions, the Monitoring pages no longer load endlessly and now display this error:  

You do not have permission to view this page.  

Additionally, the admin:services:create permission has been split into these new permissions:  

• admin:services:query:create — Grants ability to query for service information. This permission is now required to view service information on the Monitoring pages.  

• admin:services:configure:create — Grants ability to configure services.  

**Note:** The admin:services:read by itself is not sufficient for viewing the Monitoring pages. |
| ENS-8264         | Text and Metadata Extraction | **Subject fields are multivalued for some email documents**  
For some email documents processed by the Text and Metadata Extraction stage, the subject field incorrectly contains multiple subject values.  

**Fix:** The subject field now contains only one value for these email documents. |
<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENS-8715</td>
<td>Floating services, system updates</td>
<td><strong>Update fails with &quot;service under-protected&quot; error, even though services appear healthy</strong>&lt;br&gt;  In some situations, scaling an instance of a floating service can cause that service instance to incorrectly attempt to use the volumes belonging to another instance of that service. This can cause a system update to fail with an error indicating that the service is under-protected, even though the service appears to be healthy.&lt;br&gt;&lt;br&gt;<strong>Fix:</strong> When floating services are scaled, they now access the correct volumes. Systems on which floating services have been scaled can now be successfully updated.</td>
</tr>
</tbody>
</table>
## Resolved in release 1.2.2

**Note:** Version 1.2.2 was released after 1.3.0 and was not generally available. Versions 1.3.1 and later include all fixes from 1.2.2.

<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENS-8032</td>
<td>HCP MQE data connection</td>
<td><strong>Documents missed when reading from a tenant with multiple namespaces</strong></td>
</tr>
<tr>
<td>HCISUS-49</td>
<td></td>
<td>A workflow can fail to read some documents when:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• The workflow has the Check for Updates setting enabled, and</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• The workflow has an input HCP MQE data connection that reads from an HCP tenant with multiple</td>
</tr>
<tr>
<td></td>
<td></td>
<td>namespaces.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>In this case, the workflow fails to detect a small percentage of documents that were added to the</td>
</tr>
<tr>
<td></td>
<td></td>
<td>tenant at the same time the workflow was checking for newly added documents.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Fix:</strong> Workflows no longer miss newly added documents in this situation.</td>
</tr>
<tr>
<td>ENS-8134</td>
<td>HCP MQE data connection</td>
<td><strong>Documents fail to be processed with &quot;Connection pool shut down&quot; error</strong></td>
</tr>
<tr>
<td>HCISUS-53</td>
<td></td>
<td>If an HCP MQE data connection is configured to read documents from more than five HCP namespaces,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>documents can fail to be processed by a workflow with this error:</td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>java.lang.IllegalArgumentException: Connection pool shut down</code></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Fix:</strong> These failures no longer occur.</td>
</tr>
</tbody>
</table>
## Resolved Issues

<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENS-8185</td>
<td>Email Expansion stage</td>
<td><strong>Incorrect field values for journaled emails that have other emails as attachments</strong>&lt;br&gt; If a journaled email contains another email as an attachment, when that parent email is processed by the Email Expansion stage, the resulting document incorrectly contains field values for the attached email. This issue affects fields for email metadata: <strong>From</strong>, <strong>Date</strong>, and <strong>Subject</strong>.&lt;br&gt;&lt;br&gt;<strong>Fix:</strong> Documents for these emails now include the correct values for the affected fields.</td>
</tr>
</tbody>
</table>

### Resolved in release 1.3.0

<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENS-4196</td>
<td>Monitoring service</td>
<td><strong>Admin App does not report when System services are offline</strong>&lt;br&gt; If a System service such as Service-Deployment fails, the Admin App Monitoring does not show this information.&lt;br&gt;&lt;br&gt;<strong>Fix:</strong> The Monitoring pages in the Admin App now show when System services are offline.</td>
</tr>
<tr>
<td>ENS-4568</td>
<td>Updating Hitachi Content Intelligence</td>
<td><strong>Search App is unavailable during system update</strong>&lt;br&gt; When updating from release 1.1 to release 1.1.1, the Search App UI becomes unavailable for a period of time.&lt;br&gt;&lt;br&gt;<strong>Fix:</strong> The Search App remains available during an update.</td>
</tr>
<tr>
<td>Reference number</td>
<td>Area affected</td>
<td>Description</td>
</tr>
<tr>
<td>------------------</td>
<td>---------------</td>
<td>-------------</td>
</tr>
</tbody>
</table>
| ENS-5407         | Workflow tasks| **Workflow task incorrectly retries instead of halting**  
If a workflow task has insufficient resources, it may halt with this error message:  

```
java.lang.Exception: Task halted due to a failure.  
Please confirm the Driver Heap limit setting is tuned appropriately in the task settings under Memory. If running on Fedora OS, also be sure to enable swap memory on all instances. Restart the task. If the problem persists, contact your authorized service provider.
```

A task in this situation will not run successfully until the user follows the steps in the error message. However, the system unnecessarily retries the task multiple times before finally halting it.

**Fix:** In this situation, the workflow task now halts sooner, allowing you to address the problem sooner.

| ENS-6552         | Workflow tasks| **Retry Failures Task icons are displayed incorrectly**  
If you change the state of a Retry Failures task (for example, by pausing or resuming it), the task continues to display the icon relating to its previous state.

**Fix:** Icons for Retry Failures Tasks are now displayed correctly.

| ENS-6680         | Packages      | **Hostname is not included when exporting a package**  
The system hostname setting can be selected when exporting a package, but is not actually included in the exported package.

**Fix:** System hostname is now correctly exported.
<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
</table>
| ENS-6696         | Workflow tasks| **Documents are unnecessarily reread when some workflows are paused and resumed**  
For a workflow that uses list-based data connections and has the Check for Updates setting enabled, the workflow may unnecessarily reread some documents after being paused and resumed.  

**Fix:** These workflows no longer unnecessarily reread documents when being paused and resumed. |
| ENS-6716         | Events        | **Querying for events using the CLI returns a 500 error response**  
Depending on the contents of your system’s event log, running the CLI queryEvents command may return this error response:  

```
{
  "statusCode": 500,
  "errorMessage": "Missing required creator property 'subsystem' (index 8) at [Source: java.io.StringReader@5e81e5ac; line: 1, column: 487] (through reference chain: com.hds.ensemble.codegen.model.Builder["events"]->java.util.ArrayList[0])",
  "errorProperties": []
}
```

**Workaround**  
Use the administrative REST API or Admin App UI to query for events. |
| ENS-6754         | Updates       | **Update fails and Admin App becomes unresponsive on some systems**  
Updates may fail on systems where new version of any Hitachi Content Intelligence built-in plugin have been installed. After the failure, the Admin App may become unresponsive and display this error message:  

```
Reconnecting, please wait...
```

**Fix:** Updates succeed and the Admin App remains usable for these systems. |
## Resolved in release 1.2.1

<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
</table>
| ENS-4344         | Database service            | **The Database service cannot be run on additional instances if any instance of that service is down**  
In these situations, you cannot configure the Database service to run on additional instances:  
- The Database service has a status of **Failed** on one of the instances in the system.  
- One of the instances is down and the Database service is configured to run on that instance.  
**Fix:** In the situations listed above, you can now successfully run the Database service on additional instances without first having to remove failed system instances or Database service instances from the system. |
<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
</table>
| ENS-4523         | Updates      | **Update fails with RMIServer stub error**  
Rarely, an instance may fail to be restarted during a system update and the Admin App may display this error message:  

```java
java.io.IOException: Failed to retrieve RMIServer stub: javax.naming.ServiceUnavailableException
```

In this situation, retrying the update does not cause the update to run successfully.

**Workaround**
To successfully complete the update:

1. Use SSH to log into the instance that failed to restart.
2. Navigate to the installation directory.
3. Run the stop script:

   ```bash
   <install-directory>/bin/stop
   ```
4. Restart your system on the instance. For example:

   - If you ran your system in the foreground, run:
     ```bash
     sudo <install-directory>/bin/run
     ```
   - If you ran your system using systemd, run:
     ```bash
     sudo systemctl start Hitachi Content Intelligence.service
     ```
5. On the **Update** page in the Admin App, click on the **Retry** button to restart the update.

**Fix:** This issue may still rarely occur when updating to release 1.2.1 but no longer occurs when updating from 1.2.1 to a later release.
<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
</table>
| ENS-4613         | run and stop scripts | **The stop script fails when run with a different path from the run script**  
The stop script fails to run if you run it with a different path than the run script.  

For example, you can run the run with a path relative to the Hitachi Content Intelligence installation directory:

```
sudo ./bin/run
```

If you then attempt to run the stop script with a full path, the script fails:

```
sudo /opt/<install-directory>/bin/stop
```

realpath: ./bin/run: No such file or directory

**Fix:** The stop no longer fails in this situation.

| ENS-6437         | Workflow tasks | **Workflow task error message does not report the actual reason why the task halted**  
Rarely, a workflow task that processes a large amount of data may halt with this error message:

```
java.lang.InterruptedException: sleep interrupted
```

This error masks the real reason why the workflow task halted.

**Fix:** When a large workflow task halts, it now reports the correct error message. |
<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
</table>
| ENS-6498         | Events, Updates | **Events are no longer visible after update to 1.2.0**  
On some systems that have been running for a long time, events are not visible after updating to version 1.2.0.  

**Fix:** Events now remain visible after updating from 1.x to 1.2.1.  

**Note:** If you encountered this issue after updating to 1.2.0, contact your authorized Hitachi Content Intelligence support provider. |
| ENS-6517         | Updates       | **Admin App pages are unresponsive after update on some systems**  
On a system where a new version of any Hitachi Content Intelligence built-in plugin has been installed, after a system update, the pages in the Admin App are unresponsive and display this error message:  

    Reconnecting, please wait...  

This can occur, for example, on systems that have installed either the 1.1.3 or 1.1.5 maintenance release plugins.  

**Fix:** The Admin App remains usable on such systems after an update. |
| ENS-6525         | Events        | **Cannot query for events with an offset greater than 9,999**  
If a system has 10,000 or more events, you cannot retrieve any events that were logged after the 10,000th event.  

**Fix:** You can now retrieve these events. |
## Resolved Issues

<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENS-6672</td>
<td>Text and Metadata Extraction stage</td>
<td><strong>Distribution list addresses fail to be extracted from journaled emails</strong>&lt;br&gt;If a journaled email was addressed to a distribution list, the email metadata may contain the names of all email addresses in the distribution list. For such email documents, the Text and Metadata Extraction stage extracts the distribution list name but does not extract the individual email addresses.&lt;br&gt;&lt;br&gt;<strong>Fix:</strong> These email addresses are now extracted and added to the applicable document fields.</td>
</tr>
<tr>
<td>ENS-6681</td>
<td>HCP MQE data connection</td>
<td><strong>Initial HCP username and password are incorrectly reused</strong>&lt;br&gt;HCP MQE data connections always use the HCP username and password that they were initially configured with. If you reconfigure one of these data connections with different user credentials, the data connection continues to use the original ones.&lt;br&gt;&lt;br&gt;<strong>Fix:</strong> When you reconfigure an HCP MQE data connection, it now correctly uses the updated credentials.</td>
</tr>
</tbody>
</table>

## Resolved in release 1.2.0

<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENS-1887</td>
<td>Workflows</td>
<td><strong>Workflow status charts fail to display if the system does not have Internet access</strong>&lt;br&gt;The charts that show information about workflow discoveries (such as the distribution of MIME types on processed documents) do not display in the Admin App if the system does not have Internet access.&lt;br&gt;&lt;br&gt;<strong>Fix:</strong> All workflow charts display even when the system does not have Internet access.</td>
</tr>
<tr>
<td>Reference number</td>
<td>Area affected</td>
<td>Description</td>
</tr>
<tr>
<td>------------------</td>
<td>------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>ENS-4304</td>
<td>Log download script</td>
<td><strong>Log download script fails for instances not physically connected to the network</strong>&lt;br&gt;The log_download script fails to collect logs from an instance when:&lt;br&gt;• The instance has been physically disconnected from the network, and&lt;br&gt;• The <strong>--offline</strong> option is specified. For example:&lt;br&gt;. /log_download --offline&lt;br&gt;&lt;br&gt;<strong>Fix:</strong> The log_download script now collects logs from instances that have been disconnected from the network.</td>
</tr>
<tr>
<td>ENS-4398</td>
<td>Workflow and pipeline tests</td>
<td><strong>Workflow or pipeline test fails with KeeperErrorCode = ConnectionLoss</strong>&lt;br&gt;When a workflow or pipeline tasks attempts to save its results, the task fails with a <strong>KeeperErrorCode = ConnectionLoss</strong> error if the size of the test results exceeds the allowed test result size.&lt;br&gt;&lt;br&gt;<strong>Workaround</strong>&lt;br&gt;When you encounter this error, click on the <strong>Clear Test</strong> button to delete the test results.&lt;br&gt;&lt;br&gt;To avoid this error in subsequent tests, you can:&lt;br&gt;• Use a different document to test your workflow or pipeline.&lt;br&gt;• Reduce the number or size of field values that your pipelines add to the test document.&lt;br&gt;&lt;br&gt;<strong>Fix:</strong> The allowed test result size for workflow and pipeline tests has been increased.</td>
</tr>
</tbody>
</table>
## Reference Issues

<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
</table>
| ENS-4525         | Pipelines     | **Attempts to update a pipeline never succeed and cause Admin App to become unresponsive**  
Attempts to edit a pipeline fail if the size of the pipeline object itself exceeds 4MB. This can happen, for example, if you attempt to add several Tagging stages to the pipeline and each of those Tagging stages is configured to add fields with very large values.  

In this case, the Admin App becomes unresponsive.  

**Workaround**  
If the Admin App becomes unresponsive in this situation, close and reopen your browser and then clear the browser cache.  

To avoid this error, limit the size of field values that your pipeline stages add to documents.  

**Fix:** The Admin App no longer becomes unresponsive in this situation. |
| ENS-4561         | Workflow tasks, updates | **Workflow tasks occasionally halt during system update**  
If a workflow task starts right before a system update starts, the task may halt with this error:  

```java  
java.lang.IllegalArgumentException:  
com.hds.ensemble.plugins.solr.SolrPlugin  
$SolrPluginSession@<id> not a SolrPluginSession  
```

This error does not prevent the update from completing successfully.  

After the update completes, the workflow task can be restarted.  

**Fix:** This error no longer occurs. The workflow now correctly enters the Paused state. Or, if the workflow starts shortly after the update, the workflow displays an error message saying that it cannot be started because an update is in progress. |
<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
</table>
| ENS-4962         | setup script | **setup script incorrectly identifies instance IP address**  
When installing an Hitachi Content Intelligence system or adding an instance to a system, if you omit the `-i <ip-address>` option when running the setup script, the script attempts to automatically detect the IP address of the instance it's running on.  
If the instance runs certain recent versions of either the iproute or iproute2 utilities, the script may incorrectly detect the instance IP address as 0, causing Hitachi Content Intelligence deployment to fail.  
**Workarounds:**  
To avoid this issue, use the `-i` option when running setup.  
If you've encountered this issue, on each affected instance:  
1. Delete the Hitachi Content Intelligence installation directory.  
2. Start the system installation or instance addition procedure over again.  
3. Use the `-i` option when running the setup script.  
Additionally, if you are installing a single instance system, you need to specify the `-m` option. For example:  

```
setup -i 172.1.2.3 -m 172.1.2.3
```

**Fix:** The setup script now correctly identifies the instance IP address. When installing a single instance system, you can continue to omit the `-m` option. |
Resolved Issues

Resolved in release 1.1.5

Note: Release 1.1.5 was not generally available. Releases 1.2.0 and later include all fixes from 1.1.5.

<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENS-6017</td>
<td>Text and Metadata Extraction stage</td>
<td><strong>Some fields fail to be extracted from journaled emails</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>The Text and Metadata Extraction stage fails to create fields from the information stored in the message/rfc822 portion of journaled emails. This includes information such as the list of BCC'd email addresses.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Fix:</strong> Email header fields are now correctly extracted from journaled emails.</td>
</tr>
</tbody>
</table>
## Resolved in release 1.1.4

<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
</table>
| ENS-5047         | Email Expansion stage | **Email Expansion stage fails to extract email attachments**  
Occasionally, the Email Expansion stage fails to extract attachments to email documents but does not report a document failure.  
**Fix:** The stage now correctly extracts attachments from email documents. |
| ENS-5544         | Workflows     | **SSL certificates fail to be retrieved over TLS 1.2**  
Hitachi Content Intelligence fails to download SSL certificates from data sources that use TLS version 1.2.  
**Fix:** Hitachi Content Intelligence now downloads SSL certificates from these data sources. |
| ENS-5545         | Workflows     | **Cannot connect to HCP data sources using a proxy server**  
The built-in HCP data connections cannot access HCP systems over a proxy server.  
**Fix:** Both the HCP rest and HCP MQE data connections can now access HCP data sources over a proxy server. |
| ENS-5546         | Text and Metadata Extraction stage | **BCC fields are not extracted from some email documents**  
The Text and Metadata Extraction stage fails to extract BCC fields from email documents that were exported from an email server by the server administrator, not by the user who originally sent the email.  
**Fix:** BCC fields are now correctly extracted for these emails. |
## Resolved Issues

<table>
<thead>
<tr>
<th>Reference number</th>
<th>Area affected</th>
<th>Description</th>
</tr>
</thead>
</table>
| ENS-5570, ENS-5583 | Workflow tasks | **Workflow task schedules encounter errors in time zones with positive offsets from UTC**  
If your browser uses a time zone with a positive offset from UTC, when you edit a workflow task schedule, the scheduled start and stop times for the task can be automatically shifted by several hours, even when you made no change to them.  
**Fix:** Workflow task schedules now retain their correct configurations in the affected time zones. |