Content Manager OnDemand for Multiplatforms
Version 10 Release 5

Introduction and Planning Guide

SC19-3351-03

Note

Before using this information and the product it supports, read the information in “Notices” on page 89.

This edition applies to the following products and to all subsequent releases and modifications until otherwise indicated in new editions:

- Version 10 Release 5 of IBM® Content Manager OnDemand for Multiplatforms (product number 5724-J33)

Contents

Tables

ibm.com® and related resources

Product support and documentation are available from ibm.com®.

Support and assistance

From ibm.com, click **Support & downloads** and select the type of support that you need. From the Support Portal, you can search for product information, download fixes, open service requests, and access other tools and resources.

IBM Knowledge Center

See your online product information in IBM Knowledge Center at https://www.ibm.com/support/knowledgecenter/SSEPCD.

PDF publications

See the following PDF publications for your product at https://www.ibm.com/support/pages/node/1079037.

Contacting IBM

For general inquiries, call 800-IBM-4YOU (800-426-4968). To contact IBM customer service in the United States or Canada, call 1-800-IBM-SERV (1-800-426-7378).

For more information about how to contact IBM, including TTY service, see the Contact IBM website at http://www.ibm.com/contact/us/.

Accessibility

For complete information about accessibility features that are supported by this product, see your Content Manager OnDemand administration guide.

Chapter 1. Introduction

This section provides an overview of the Content Manager OnDemand system and contains information that can help you better understand how the system works.

This section describes how Content Manager OnDemand manages reports and index data, and includes important information about how Content Manager OnDemand, the database manager, and the storage manager work to index, load, and retrieve documents.

This section also contains a list of the tasks that Content Manager OnDemand administrators typically perform to manage an Content Manager OnDemand system.

Overview of Content Manager OnDemand

Content Manager OnDemand for Multiplatforms is a document archive solution. Content Manager OnDemand supports any organization that can benefit from hard copy or microfiche replacement and instant access to information.

A Content Manager OnDemand system can support small office environments and large enterprise installations with hundreds of system users. Content Manager OnDemand can dramatically improve productivity and customer service in many businesses by providing fast access to information stored in the system.

Content Manager OnDemand processes the print output of application programs, extracts index fields from the data, stores the index information in a relational database, and stores one or more copies of the data in the system. With Content Manager OnDemand, you can archive newly created and frequently accessed reports on high speed, disk storage volumes and automatically migrate them to other types of storage volumes as they age.

Content Manager OnDemand fully integrates the capabilities of Advanced Function Presentation (AFP), including management of resources, indexes, and annotations, and supports full fidelity reprinting of documents to devices attached to a workstation, Content Manager OnDemand server, or other server on the network.

Content Manager OnDemand provides administrators with tools to manage Content Manager OnDemand servers, to authorize users to access Content Manager OnDemand servers and data stored in the system, and to backup the database and data storage.

Content Manager OnDemand provides users the ability to view documents, print, and send copies of documents, and attach electronic notes to documents.

Some of the advantages that Content Manager OnDemand offers include:

- Easily locate data without specifying the exact report
- Retrieve the pages of the report that you need without processing the entire report
- View selected data from within a report

Content Manager OnDemand can provide you with an information management tool that can increase your effectiveness when working with customers.

Content Manager OnDemand does the following:

- Integrates data created by application programs into an online, electronic information archive and retrieval system
- Provides the controlled and reliable access to all of an organization's reports
- Retrieves the data that you need when you need it
- Provides a standard, intuitive client with features such as thumbnails, bookmarks, notes, and shortcuts

These features mean that Content Manager OnDemand can help you quickly retrieve the specific page of a report that you need to provide fast customer service.

System overview

A Content Manager OnDemand system consists of client programs and server programs that communicate over a network running the TCP/IP communications protocol, a database manager that maintains index data and server control information, and storage managers that maintain documents on various types of storage devices.

The diagram shows an example.

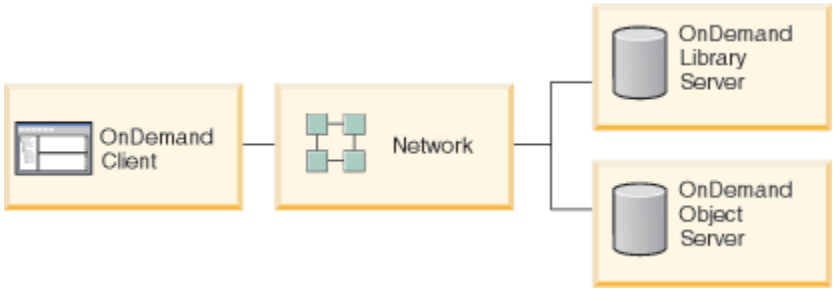

Figure 1. Content Manager OnDemand system

Content Manager OnDemand client programs run on workstations and terminals attached to the network and communicate with Content Manager OnDemand servers. The Content Manager OnDemand library server manages a database of information about the users of the system and the reports stored on the system. An Content Manager OnDemand object server manages the reports on disk, optical, and tape storage devices. An Content Manager OnDemand system has one library server and one or more object servers. An object server can operate on the same workstation or node as the library server or on a different workstation or node than the library server.

Content Manager OnDemand provides the capability to do many of the client functions from almost any operating system, by using a Web browser or a user-written program.

Content Manager OnDemand client programs operate on personal computers running Windows. The client program is the user's way to search for and retrieve reports stored on the system. Using the client program, users can construct queries and search for reports, retrieve documents from Content Manager OnDemand, view, print, and FAX copies or pages of documents, and attach electronic notes to pages of a document.

Content Manager OnDemand servers manage control information and index data, store and retrieve documents and resource group files, and process query requests from Content Manager OnDemand client programs. The documents can reside on disk, optical, and tape storage volumes. New reports can be loaded into Content Manager OnDemand every day. That way, Content Manager OnDemand can retrieve the latest information generated by application programs.

Content Manager OnDemand client programs and servers communicate over a computer network supported by TCP/IP. When a user submits a query, the client program sends a search request to the Content Manager OnDemand library server. The library server returns the list of documents that match the query to the user. When the user selects a document for viewing, the client program retrieves a copy of the document from the object server where the document is stored, opens a viewing window, and displays the document.

Indexers

You can use indexers to define the characteristics of your documents to IBM Content Manager OnDemand.

The indexers in IBM Content Manager OnDemand are programs that provide the following functions:

- A way to load and store reports
- Determining where one document ends and the next document begins
- Determining which index values are to be associated with each document
- Using index values to identify and retrieve documents for viewing or printing
- Creating a resource file that contains all the resources that are needed to view and print a document

Usually, the index values are extracted from the content of the documents; however, they can also be created manually or by a custom application (an exit). The choice of an indexer depends on the operating system, the format of the documents, and the function that is needed. Content Manager OnDemand provides the following indexers:

- ACIF indexer
- 390 indexer
- 400 indexer
- PDF indexer
- Generic indexer
- XML indexer

Enhanced Retention Management (ERM)

The Content Manager OnDemand Enhanced Retention Management feature enables you to manage and enforce retention of documents.

In a Content Manager OnDemand system, you retain documents for a specific amount of time. This practice is commonly referred to as retention management. Records management describes the process of retaining and deleting documents under a set of circumstances that are not necessarily bounded by time, for example, until the end of litigation.

Without the Enhanced Retention Management feature, Content Manager OnDemand cannot implement records management due to the following limitations:

- Records management requires that you have control over individual documents. Content Manager OnDemand manages application groups instead of individual documents and works with a storage manager to delete (expire) documents.
- Records management requires flexibility in defining the time when documents are deleted. However, Content Manager OnDemand defines only the time when application groups with fixed time ranges are deleted, for example, five years after Content Manager OnDemand loads documents.

To overcome these limitations, you can purchase and install the Enhanced Retention Management feature. With the Enhanced Retention Management feature, you can control individual documents by introducing holds, a mechanism that identifies the documents that you want to keep for some time. To expire a document with a hold, you need to remove the hold. Holds give you flexibility to choose when to delete documents because you control when to remove a hold. By using holds, you control when to delete a document. You can manage holds through any of the following interfaces:

- Content Manager OnDemand client or administrative client
- IBM Content Navigator
- ARSDOC command
- ODWEK Java APIs
- FileNet P8 when you integrate it with Content Federation Services for Content Manager OnDemand by enabling the Content Federation Services for Content Manager OnDemand. The Content Federation Services for Content Manager OnDemand also enables you to federate Content Manager OnDemand repositories. This feature connects your Content Manager OnDemand content to business process management (IBM BPM) and records management features of FileNet P8.

Content Manager OnDemand Distribution Facility (ODF)

Content Manager OnDemand Distribution Facility (ODF) is the report distribution feature for Content Manager OnDemand for Multiplatforms and Content Manager OnDemand for z/OS®.

ODF provides an easy way to automatically group reports and portions of reports and distribute the reports to multiple users. ODF distributions can be printed, created as an output file, or emailed as an attachment. ODF can retrieve reports that are stored in a Content Manager OnDemand server on z/OS or any of the operating systems that are supported by Content Manager OnDemand for Multiplatforms.

Content Manager OnDemand also provides an ODF monitor utility. The ODF monitor utility is an interactive workstation client program that enables you to check the status of distributions that are submitted for processing and to monitor distribution activity.

Content Manager OnDemand administrative client

The Content Manager OnDemand administrative client helps you maintain your Content Manager OnDemand system.

The administrative client provides tools that enable you to:

- Define reports to the system
- Add and maintain users and groups
- Add and maintain server printers
- Add and maintain storage sets and storage nodes
- Add and maintain users, groups, applications, application groups, storage sets, folders, printers, holds, and cabinets
- Add and maintain OnDemand Distribution Facility elements including recipients, recipient lists, report IDs, report bundles, and distributions
- Add and maintain servers
- Set system parameters for servers and client programs
- Copy items from one server to another
- Track changes to the system

When you use the administrative client to add or update an object in Content Manager OnDemand, information about the object is saved in the system log. The information includes the changes that you made.

The administrative client also provides a report wizard that enables you to:

- Activate Full Text Search in the application group and to create FTS folder fields.
- Create an application, application group, and folder using a wizard for reports that are loaded by using the generic indexer format.
- Provide the ability to define a generic indexer report.
- Enable the function to add Document Size and Page Count fields to the application group and folder definitions that are created.

Concepts

When defining a new report or type of data to Content Manager OnDemand, an administrator must create an application and assign the application to an application group.

The terms *application*, *application group*, *folder*, and *cabinet* describe how Content Manager OnDemand stores, manages, retrieves, views, and prints reports, and index data. (If an application group does not exist, the administrator must create one first.) Before users can search for and retrieve documents, an

administrator must create or update a folder to use the application group and application. To help users find folders quickly, administrators can create cabinets.

Applications

An application describes the physical characteristics of a report to Content Manager OnDemand.

Typically you define an application for each program that produces output that are stored in Content Manager OnDemand. The application includes information about the format of the data, the orientation of data on the page, the paper size, the record length, and the code page of the data. The application also includes parameters that the indexing program uses to locate and extract index data and processing instructions that Content Manager OnDemand uses to load index data in the database and documents on storage volumes.

Application Groups

An application group contains the storage management attributes and index fields for the data that you load into in Content Manager OnDemand.

When you load a report into Content Manager OnDemand, you must identify the application group where Content Manager OnDemand will load the index data and store the documents. An application group is a collection of one or more Content Manager OnDemand applications with common indexing and storage management attributes. You typically group several different reports in an application group so that users can access the information contained in the reports with a single query. All of the applications in the application group must be indexed on the same fields, for example, customer name, account number, and date.

Folder

A folder is the user's way to query and retrieve data stored in Content Manager OnDemand.

A folder provides users with a convenient way to find related information stored in Content Manager OnDemand, regardless of the source of the information or how the data was prepared. A folder allows an administrator to set up a common query screen for several application groups that might use different indexing schemes, so that a user can retrieve the data with a single query. For example, a folder called Student Information might contain transcripts, bills, and grades, which represents information stored in different application groups, defined in different applications, and created by different programs.

The diagram illustrates the concepts described in this section.

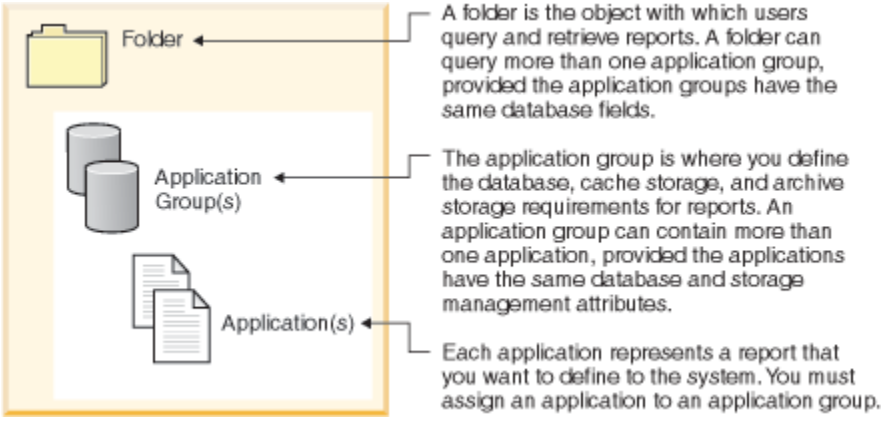

Figure 2. Folders, application groups, and applications (part 1 of 2)

The diagram shows an example.

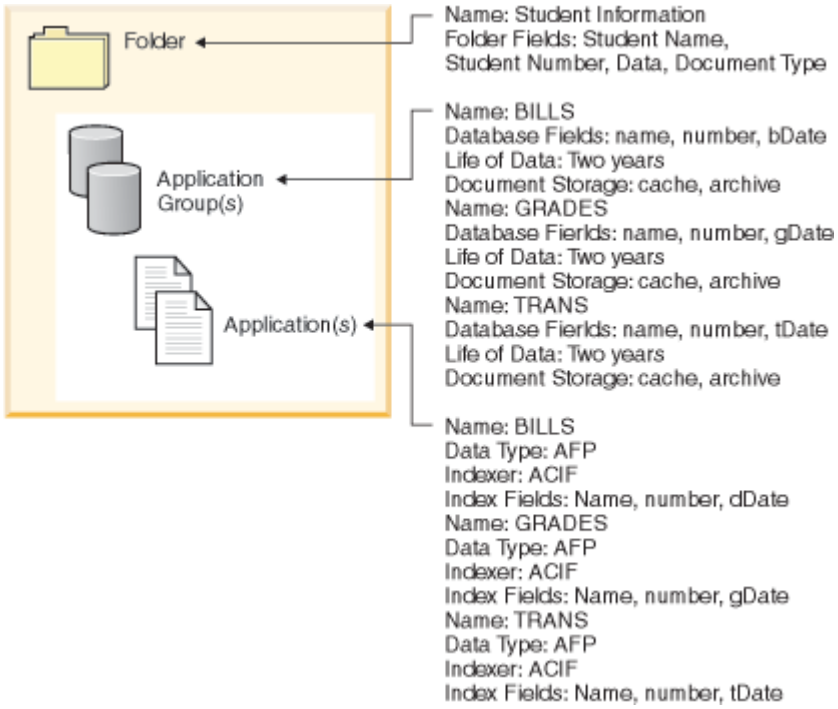

Figure 3. Folders, application groups, and applications (part 2 of 2)

Cabinet

If users have many folders, they might find it helpful to group their folders into cabinets. Cabinets are an optional feature that enable users to navigate to folders more easily.

Cabinets follow these rules:

- A cabinet can contain one or more folders.
- A folder can belong to zero or more cabinets.

Figure 4 on page 7 demonstrates how cabinets can be used to organize the folders that a user needs for generating different types of reports. In this example, the user needs to pull together information on fund performance, fund balance, and fund transactions for a monthly report. The "Monthly Report" cabinet contains folders for each type of information that the user needs to collect. The user also needs to generate investments performance reports for clients. Investment performance reports include information on stock performance, bond performance, and fund performance. The "Client Report" cabinet contains folders for stocks, bonds, and funds. Both cabinets contain the folder for fund performance because the user needs information on fund performance to generate both reports.

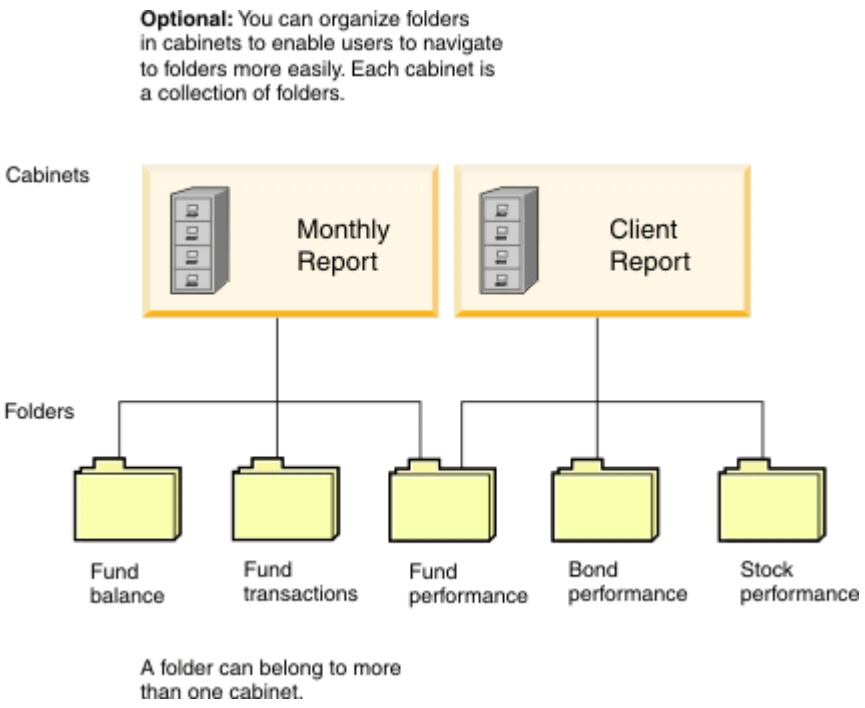

Figure 4. Cabinets are used to organize folders

Indexing methods

Document indexing is used for reports that contain logical items such as policies, and statements. The report indexing is used for reports that contain many pages of the same kind of data, such as a transaction log.

Content Manager OnDemand provides two basic ways to index data:

- Each of the items in a report can be individually indexed on values such as account number, customer name, and balance. Content Manager OnDemand supports up to 128 index values per item. With document indexing, the user does not necessarily need to know about reports or report cycles to retrieve a document from Content Manager OnDemand.
- Each line in the report usually identifies a specific transaction, and it would not be cost effective to index each line. Content Manager OnDemand stores the report as groups of pages and indexes each group. When reports include a sorted transaction value (for example, invoice number), Content Manager OnDemand can index the data on the transaction value. This is done by extracting the beginning and ending transaction values for each group of pages and storing the values in the database. This type of indexing lets users retrieve a specific transaction value directly.

Documents

Content Manager OnDemand documents represent indexed groups of pages.

Typically an Content Manager OnDemand document is a logical section of a larger report, such as an individual customer statement within a report of thousands of statements. An Content Manager OnDemand document can also represent a portion of a larger report. For reports that do not contain logical groups of pages, such as transaction logs, Content Manager OnDemand can divide the report into groups of pages. The groups of pages are individually indexed and can be retrieved to the client workstation much more efficiently than the entire report. Documents are always identified by date, and usually one or more other ways, such as customer name, customer number, or transaction number.

The applications and documents diagram illustrates Content Manager OnDemand applications and documents. An administrator could define the BILLS application for a report that contains logical items,

such as customer statements. The BILLS application uses the document indexing method to divide the report into documents. Each statement in the report becomes a document in Content Manager OnDemand. Users can retrieve a statement by specifying the date and any combination of name and number. An administrator could define the TRANS application for a report that contains lines of sorted transaction data. The TRANS application uses the report indexing method to divide the report into documents. Each group of 100 pages in the report becomes a document in Content Manager OnDemand. Each group is indexed using the first and last sorted transaction values that occur in the group. Users can retrieve the group of pages that contains a specific transaction number by specifying the date and the transaction number. Content Manager OnDemand retrieves the group that contains the value entered by the user.

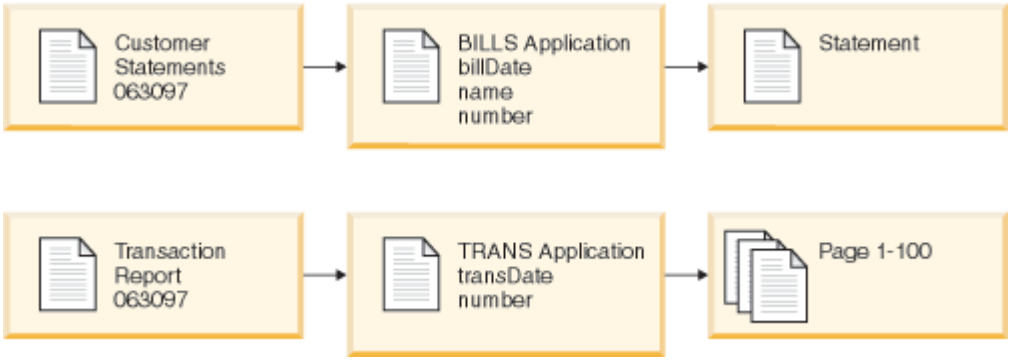

Figure 5. Applications and documents

Servers

The Content Manager OnDemand server environment includes a library server and one or more object servers residing on one or more workstations connected to a TCP/IP network.

The library server maintains a central database about the reports stored in Content Manager OnDemand. The database also contains information about the objects defined to the system, such as users, groups, printers, application groups, applications, folders, cabinets, holds and storage sets. The database manager provides the database engine and utilities to administer the database. The library server processes client logons, queries, and print requests and updates to the database. The major functions that run on the library server are the request manager, the database manager, and the server print manager.

An object server maintains documents on cache storage volumes and, optionally, works with an archive storage manager to maintain documents in archive storage, such as optical and tape storage libraries. An object server loads data, retrieves documents, and expires documents. The major functions that run on an object server are the cache storage manager, Content Manager OnDemand data loading and maintenance programs, and optionally, the archive storage manager.

The basic Content Manager OnDemand configuration is a library server and an object server on the same workstation or node. This single library/object server configuration supports the database functions and cache storage on one workstation. You can add an archive storage manager to the single library/object server configuration, to maintain documents in archive storage. You can also configure your Content Manager OnDemand system with the library server on one workstation and one or more object servers on different workstations. This configuration is known as a distributed library/object server system. The distributed library/object server configuration supports caching of documents on different servers. You can add an archive storage manager to one or more of the object servers to maintain documents in archive storage that is attached to different servers.

The Content Manager OnDemand server environment contains several components:

- A request manager that provides client, network, and operating system services, security, and accounting. The request manager resides on the library server.

- A database manager that maintains the index data for the reports that you store on the system. The database manager is a relational database management product, such as DB2® (included with your product package). The database manager resides on the library server.
- Database control information about the users, groups, application groups, applications, folders, storage sets, cabinets, holds and printers that you define on the system. The control information determines who can access the system, the folders that a user can open, and the application group data that a user can query and retrieve. The database resides on the library server.
- A cache storage manager that maintains documents in cache storage. Cache storage is for high-speed access to the most frequently used documents.
- An *archive storage manager*, which is an optional part of the system. The archive storage manager is for the long-term storage of one or more copies of documents in archive storage, such as optical and tape storage libraries. Tivoli® Storage Manager (included in your product package) is an example of an archive storage manager product. You can also use Tivoli Storage Manager to backup and restore DB2 databases. This capability means that you do not have to manage DB2 backup files on disk.
- A download facility that automatically transfers spool files to a server at high speed. It is recommended that you use Download for z/OS, a licensed feature of InfoPrint Manager for AIX. Download provides the automatic, high-speed download of JES spool files from a z/OS system to a Content Manager OnDemand server.
- Data indexing and conversion programs. These programs extract index data from input files or generate index data and, depending on the indexer, optionally collect resources and transform input data from one format to another. Content Manager OnDemand provides several indexing programs:
  - The AFP Conversion and Indexing Facility (ACIF) can be used to index line data, ASCII data, and AFP input files. ACIF can collect the resources that are required to view AFP documents and convert line data input into AFP data to be stored on the system.
  - The PDF indexer can be used to create index data for Adobe PDF input files. It is a separately priced feature.
  - The XML indexer can be used to index XML input files.
  - The 390 indexer can be used to index line data and AFP input files.
  - The 400 indexer can be used to index line data, AFP, SCS and SCS Extended input files. It is included with Content Manager OnDemand on IBM i only.

  The indexing programs might be run on the library server or an object server. ACIF might also run on a z/OS system, and the output can be transferred to the Content Manager OnDemand server for loading.
- Data loading programs that can be set up to automatically store report data into application groups and update the database. The data loading programs might run on the library server or on an object server.
- Archived reports and resources.
- A server print facility that allows users to reprint a large volume of documents at high speed. Content Manager OnDemand uses Infoprint, which must be purchased separately, to manage the server print devices.
- Content Manager OnDemand management programs to maintain the Content Manager OnDemand database and documents in cache storage.
- A system logging facility that provides administrators with tools to monitor server activity and respond to specific events as they occur. The interface to the system logging facility is through the System Log folder and the System Log user exit.

The following topics provide additional information:

- The Content Manager OnDemand request manager
- The Content Manager OnDemand database manager
- The Content Manager OnDemand storage manager
- Download
- Data indexing and loading

- Content Manager OnDemand management programs

Request manager

The request manager processes search requests from Content Manager OnDemand client programs. When a user enters a query, the client program sends a request over the network to the request manager.

The request manager works with the database manager to compile a list of the items that match the query and returns the list to the client program. When the user selects an item for viewing, the request manager sends a retrieval request to the storage manager: the cache storage manager, if the document resides in cache storage or the archive storage manager, if the document resides in archive storage. The storage manager retrieves the document and, optionally, the resources associated with the item. The Content Manager OnDemand client program decompresses and displays the document.

Content Manager OnDemand management programs include utilities that maintain the database and cache storage, including the ability to automatically migrate data from the database and cache storage to archive storage. These programs use the services of the request manager to manage index data, documents, and resource files.

When a user logs on to the system, Content Manager OnDemand assigns a unique transaction number to that instance of the client program. All activity associated with that instance of the client program contains the same transaction number. The request manager records messages generated by the various Content Manager OnDemand programs in the System Log, for example, logon, query, print, and so forth. These System Log messages contain the transaction number, user ID, time stamp, and other information. Administrators can open the System Log folder and view the messages. Content Manager OnDemand Also provides a System Log user exit so that you can run a user-defined program to process messages. For example, you could design a user-defined program to send an alert to an administrator when certain messages appear in the System Log. The messages in the System Log can also be used to generate usage and billing reports.

Database manager

Content Manager OnDemand uses a database management product, such as DB2 (provided with Content Manager OnDemand), to maintain the index data for the reports that you load into the system.

The database manager also maintains the Content Manager OnDemand system tables that describe the applications, application groups, storage sets, folders, groups, users, and printers that you define to the system. You should periodically collect statistics on the tables in the database to optimize the operation of the Content Manager OnDemand database.

Storage manager

The Content Manager OnDemand cache storage manager maintains a copy of documents, usually temporarily, on disk.

The cache storage manager uses a list of file systems to determine the devices available to store and maintain documents. You typically define a set of cache storage devices on each object server so that the data loaded on the server can be placed on the fastest devices to provide the most benefit to your users. The cache storage manager uses the ARSMAINT program to migrate documents from cache storage to archive storage and to remove documents that have passed their life of data period.

Content Manager OnDemand also supports an archive storage manager, such as Tivoli Storage Manager. The archive storage manager maintains one or more copies of documents in archive storage, such as optical or tape storage libraries. You decide which types of archive storage that your Content Manager OnDemand system must support, configure the storage devices on the system, and define the storage devices to the archive storage manager. To store application group data in archive storage, you must assign the application group to a storage set that identifies a storage node that is managed by the archive storage manager.

In addition to managing reports in archive storage, Tivoli Storage Manager can also maintain files used to backup and restore DB2 databases. This capability means that you do not have to maintain the DB2 backup files on disk. Tivoli Storage Manager can assist you with automating database backup on a regular schedule. When you use the ARSDB program to create a database or table space backup image, you can

specify that you want Tivoli Storage Manager to manage the image. After completing the backup image, the ARSDB program copies the archived log files to storage that is managed by Tivoli Storage Manager.

Download

Download is a licensed feature of Infoprint for z/OS. Download provides the automatic, high-speed download of JES spool files from a z/OS system to a Content Manager OnDemand server.

Download can be used to transfer reports created on z/OS systems to the server, where you can configure Content Manager OnDemand to automatically index the reports and store the report and index data on the system. Download operates as a JES Functional Subsystem Application (FSA) and can automatically route jobs based on a JES class or destination, reducing the need to modify JCL. Download uses TCP/IP protocols to stream data at high speed over a LAN or channel connection from a z/OS system to the Content Manager OnDemand server. See *PSF for z/OS: Download for z/OS* for more information about Download.

Data indexing and loading

The reports that you store in Content Manager OnDemand must be indexed. Content Manager OnDemand supports several types of index data and indexing programs. For example, you can use ACIF to extract index data from the reports that you want to store on the system.

An administrator defines the index fields and other processing parameters that ACIF uses to locate and extract index information from reports. Content Manager OnDemand data loading programs read the index data generated by ACIF and load it into the Content Manager OnDemand database. The data loading programs obtain other processing parameters from the Content Manager OnDemand database, such as parameters used to segment, compress, and store report data in cache storage and in archive storage. If you plan to index reports on a Content Manager OnDemand server, you can define the parameters with the administrative client. The administrative client includes a report wizard that lets you create ACIF indexing parameters by visually marking up sample report data. Content Manager OnDemand Also provides indexing programs that can be used to generate index data for Adobe PDF files and other types of source data, such as TIFF images.

The diagram shows an overview of the data preparation process.

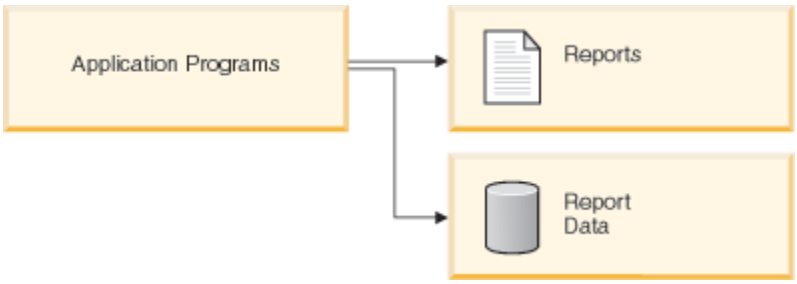

Figure 6. Data preparation, indexing, and loading (part 1 of 2)

In the picture, user-defined application programs generate printed reports and save report data to disk. The report data can be transmitted to a Content Manager OnDemand server for indexing and loading. There are a number of methods that you can use to transmit the report data to the server. For example, you can use Download to transmit data from the JES spool to a Content Manager OnDemand server.

The diagram shows an overview of the data indexing and loading process.

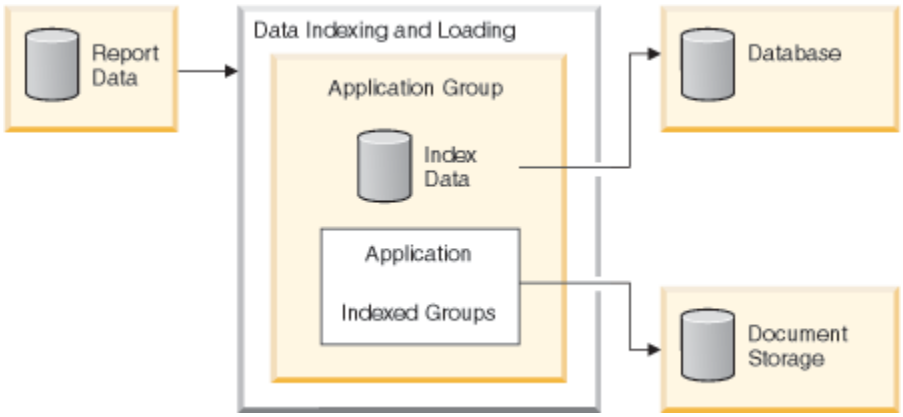

Figure 7. Data preparation, indexing, and loading (part 2 of 2)

The Content Manager OnDemand data loading program first determines whether the report needs to be indexed. If the report needs to be indexed, the data loading program calls the appropriate indexing program. The indexing program uses the indexing parameters from the Content Manager OnDemand application to process the report data. The indexing program can extract and generate index data, divide the report into indexed groups, and collect the resources required to view and reprint the report. After indexing the report, the data loading program processes the index data, the indexed groups, and the resources using other parameters from the application and application group. The data loading program works with the database manager to update the Content Manager OnDemand database with index data extracted from the report. Depending on the storage management attributes of the application group, the data loading program might work with the cache storage manager to segment, compress, and copy report data to cache storage and the archive storage manager to copy report data to archive storage.

Management programs

Content Manager OnDemand provides programs to maintain and optimize the database and maintain documents in cache storage.

An administrator usually determines the processing parameters for these programs, including the frequency with which the programs should run. When someone in your organization creates an application group, they specify other parameters that these programs use to maintain the report data stored in the application group. For example, when creating an application group, the administrator specifies how long documents should be maintained on the system and whether index data should be migrated from the database to archive storage. The programs use the information to migrate documents from cache storage to archive storage, delete documents from cache storage, migrate index data from the database to archive storage, and delete index data from the database. These functions are useful because Content Manager OnDemand can reclaim the database and cache storage space released by expired and migrated data. It is recommended that you configure your Content Manager OnDemand system to automatically start these management programs on a regular schedule, usually once every night or week.

The archive storage manager deletes data from archive storage when it reaches its storage expiration date. An administrator defines management information to the archive storage manager to support the Content Manager OnDemand data it manages. The management information includes the storage libraries and storage volumes that can contain Content Manager OnDemand data, the number of copies of a report to maintain, and how long to keep data in the archive management system.

Content Manager OnDemand and the archive storage manager delete data independently of each other. Each uses its own criteria to determine when to remove documents. Each uses its own utilities and schedules to remove documents. However, for final removal of documents from the system, you should always specify the same criteria to Content Manager OnDemand and the archive storage manager. For example, The Life of Data, which is used by Content Manager OnDemand, and the Retention Period, which is used by Tivoli Storage Manager, should specify the same value.

Configuring a remote instance on a Windows server

You can configure a remote instance of Content Manager OnDemand on a Windows server.

Procedure

To configure a remote instance on the local Content Manager OnDemand server:

1. Open the Content Manager OnDemand Configurator V10.5 on the local Windows server. Select **OnDemand Configurator V10.5** > **File** > **New Remote Instance**.
2. On the **Instance** page, type a remote instance name, for example `rodinst`. The name should point to the remote library server. Type the Library Server Name with the hostname, alias, or IP address. Accept or change the path name for the temporary file path. Click **Next**.
3. On the **Server Communications** page, type the port number of the remote instance on the remote library server. The port number can be found in the `ars.ini` file on the remote server. Click **Next**.
4. On the **Load** page, specify the Content Manager OnDemand login ID and password for the Load service. Specify one or more directories to be used by the load service. Click **Next**.
5. On the **Distribution** page, specify the Content Manager OnDemand login ID and password for the Content Manager OnDemand Distribution (ODF) service. Specify the number of print tasks, a working directory, output file location, the SMTP server name or IP address and sleep time intervals. Click **Finish**.

Results

After a remote instance is configured on a Windows system, you can run the Content Manager OnDemand server programs (such ARSDOC and ARSLOAD) on the remote instance by issuing a command on Windows.

Content Manager OnDemand Web Enablement Kit

The Content Manager OnDemand Web Enablement Kit (ODWEK) provides a set of programming interfaces that can search and retrieve documents from Content Manager OnDemand servers.

ODWEK allows users to access data that is stored in Content Manager OnDemand with IBM Content Navigator or a user-written program. An application that uses the Java API can verify permissions, manage hit lists and return data. For example, ODWEK verifies that the user information is valid on the Content Manager OnDemand server, such as permission to access the server and data stored in an application group. After the user submits a search, the ODWEK Java API returns a list of the documents that match the query. The user selects a document to view and IBM Content Navigator or the user-written program sends the document to the browser.

The diagram shows a workstation with a web browser that is being used to access data from a Content Manager OnDemand server.

Figure 8. Accessing data stored in Content Manager OnDemand using ODWEK

ODWEK can search for and retrieve documents from the Content Manager OnDemand server.

ODWEK contains several components:

- The ODWEK Java Application Programming Interface (Java API). The Java API uses standard Content Manager OnDemand interfaces and protocols to access data stored in a Content Manager OnDemand server. No additional code is needed on the Content Manager OnDemand server to support ODWEK. The Java API provides a way to access Content Manager OnDemand data from a user-written program. The

Java API also provides a way to access Content Manager OnDemand data from software such as IBM Content Navigator.

- The IBM Content Manager OnDemand AFP Web Viewer. The AFP Web Viewer lets users search, retrieve, view, navigate, and print AFP documents from a web browser.
- The HTML5 Line Data Viewer. The HTML5 Line Data Viewer lets users view line data documents from a web browser.

Viewing and transforming documents

To view other types of documents that are stored in Content Manager OnDemand, you must obtain and install the appropriate viewer.

For example, to view Adobe Portable Data Format (PDF) documents, it is recommended that you obtain the Adobe Acrobat viewer for the browsers that are used in your organization.

To convert AFP documents that are stored in Content Manager OnDemand into PDF documents, a transform is required.

Chapter 2. Preparing the server

This section contains an outline that can help to prepare your organization for the Content Manager OnDemand environment and perform a pilot roll out of the system.

To prepare you system:

- Use the Content Manager OnDemand administrative client to create a storage set to set up cache and archive storage nodes.
- Work with a single department or group of end-users. Send a memo to the users to explain how Content Manager OnDemand affects their daily work.
- Develop an end-user training course or ask your IBM representative about training for Content Manager OnDemand.
- Establish a support plan for the users. The plan should include the names and phone numbers of persons to contact for assistance and a list of troubleshooting tips.
- Develop a set of evaluation and completion criteria that you can use to compare against the actual performance of the system.
- Choose a report or set of reports for an initial migration to Content Manager OnDemand. Obtain hardcopy of the reports.
- Review the reports and determine the type of indexing required. Then select the fields from the reports for index, filter, and display fields.
- Review the selections with the users. Verify that the index, search, and display fields allow the users to retrieve the data that they need.
- Determine the viewing requirements of your users.
- Identify the type of data contained in the report and determine how you create the index data.
- If you plan to store AFP documents in the system, identify the resources used by the report. Resources are reusable objects found on pages of a report, such as overlays and page segments. Overlays contain constant data that is merged with variable report data during printing and viewing. Page segments are graphics and images that appear on pages of a report file, such as a company logo. Resources can be used by different applications in Content Manager OnDemand. If you plan to index the input data on a z/OS system, the resources can be gathered into a resource group file. If you plan to index the input data on an Content Manager OnDemand server, you must either transmit the resource group file to the server or provide access to the resource group file by using some other method such as the Network File System (NFS).
- If you plan to use ACIF or the PDF Indexer to index the report, decide where to index reports: on System z or a Content Manager OnDemand server. Determine how to transmit report and index data from the System z to the Content Manager OnDemand server. It is recommended that you use Download for z/OS to transmit data from the JES spool to file systems on Content Manager OnDemand servers. See *PSF for z/OS: Download for z/OS* for details about installing, configuring, and using Download for z/OS.
- Configure cache storage (magnetic storage devices) and archive storage (optical and tape storage devices) on the Content Manager OnDemand servers. Define and configure archive storage devices to Tivoli Storage Manager. Define storage management policies to Tivoli Storage Manager to support the reports that you plan to store on the system.
- Use the Content Manager OnDemand administrative client to create the application groups and applications required to support your reports.
- Use the administrative client to define the folders that users open to access data stored on the system.
- Use the administrative client to define users and groups to Content Manager OnDemand.
- Index the reports.
- Load the report, resources, and index data into the application group.

- Begin end-user testing. Survey the users about initial testing and index, search, and display fields.
- Collect additional information from users, report suppliers, production scheduling, and capacity planning. For example:
  - The frequency with which a report is generated and must be loaded into the system
  - The number of pages in a report
  - The number of indexed items, such as statements, contained in a report
  - The access frequency and patterns of your users
  - The length of time until a version of a report is out of date; the length of time that you need to maintain a report on the system
  - The number of copies of a report that must be maintained on the system
- Use the administrative client to update Content Manager OnDemand with the information that you collect.
- Survey users about their satisfaction with Content Manager OnDemand. Compare the performance of the system with the evaluation and completion criteria that you established. Prepare a list of issues to resolve.
- Update your company's vital records list to include the hardware and software required by the Content Manager OnDemand system. Update your company's operations and recovery manuals with information required to operate, support, and backup the Content Manager OnDemand system.

Administrative roles and responsibilities

Content Manager OnDemand administrators assume responsibility for and take care of the Content Manager OnDemand system.

The Content Manager OnDemand system includes all sorts of things, including hardware, application and system software, reports, and users.

- Hardware includes library and object server workstations, backup devices, archive storage devices, client workstations, terminals, printers, and the networking equipment.
- Software includes the base operating system, prerequisite software, and client and server programs, configuration files and shell scripts.
- Administrators define Content Manager OnDemand applications and decide how Content Manager OnDemand will manage data on the servers.
- Administrators define Content Manager OnDemand groups and users to the system and make sure that the client software is installed and operating properly.

While Content Manager OnDemand administrators are responsible for this collective environment from the viewpoint of Content Manager OnDemand users, it is likely the Content Manager OnDemand administrators are not the only people in an organization working on all these components.

Depending on the size of your organization, there can be one person or many people administering the system. If your organization is large, the administrative tasks can be divided among several people. For example, a Content Manager OnDemand system administrator could maintain Content Manager OnDemand storage sets, system printers, groups, and users; a Content Manager OnDemand application administrator could maintain application groups, applications, and folders; an operating system administrator could apply base operating system upgrades and perform problem determination; and a service administrator could maintain records of system and network hardware and software and make equipment changes.

The following list of items is typical of the tasks required to administer and maintain a Content Manager OnDemand system. Some of these tasks might be the responsibility of a person other than a Content Manager OnDemand administrator.

- Installing and upgrading equipment
- Installing and maintaining Content Manager OnDemand programs and other software

- Defining and labeling storage volumes
- Monitoring the space used by the database and the space available on the system
- Monitoring the space used for cache storage and the space available on the system
- Monitoring the space used for archive storage and the space available on the system
- Scheduling jobs to maintain the database, cache storage, and archive storage
- Working with users to determine report indexing and retrieval requirements
- Defining storage sets and storage nodes
- Defining Content Manager OnDemand system printers
- Defining reports to the system
- Defining Content Manager OnDemand groups and users
- Loading reports on the system
- Managing the backup and recovery process for the database and other areas that contain data critical to the operation of the system
- Monitoring server activity and tuning system parameters
- Solving server, network, and application problems
- Answering end-user questions
- Establishing security and audit policies, for example: set and maintain passwords and permissions; use the Content Manager OnDemand audit facilities to monitor application group and user activity; develop, document, and maintain change control procedures to prevent unauthorized changes to the system

Content Manager OnDemand provides an administrative client to allow administrators to maintain Content Manager OnDemand objects through an easy-to-use, graphical user interface. The administrative client runs on a Windows workstation. The administrative client allows administrators to define and maintain application groups, storage sets, storage nodes, folders, cabinets, system printers, applications, groups, and users. The administrative client includes features that allow administrators to process sample report data and create ACIF indexing parameters and logical views by visually marking up a sample of a report.

Content Manager OnDemand provides a set of administrative commands to help administrators maintain the system. For example, Content Manager OnDemand provides commands for loading and unloading reports, maintaining the database and cache storage, and querying and retrieving documents. Many of the administrative commands can be configured to run automatically, on a regular schedule.

Application programming interfaces

Content Manager OnDemand provides several kinds of application programming interfaces that you can use to customize Content Manager OnDemand clients and work with objects on the server.

Client customization

You can customize the Content Manager OnDemand Windows client to access data stored in Content Manager OnDemand.

The OnDemand client can be customized by specifying command line parameters, by invoking and manipulating the client from another application with the Dynamic Data Exchange (DDE) interface, by invoking another application's DLL file from the client, by retrieving related documents, by creating a Program Information File (PIF), and by auditing documents using the Document Audit Facility.

An example of invoking another application's DLL file is the integration of Monarch software with the OnDemand client. Users can view Content Manager OnDemand documents in Monarch and then perform complex data manipulation such as creating derived columns and generating charts and reports.

The OnDemand client (32-bit only) can also be customized by using the Object Linking and Embedding (OLE) APIs. The APIs are used by an application to make requests to the OnDemand client to perform functions such as logon to a server, open a folder, search for a document, view a document, and so on.

Another form of client customization is the use of the Content Manager OnDemand Web Enablement Kit (ODWEK) Java APIs. The Java APIs are used by a user-written program or IBM Content Navigator to search for, retrieve, and view documents.

Server programs

Content Manager OnDemand provides programs that you can use to work with objects on the system. For example:

- You can use ARSXML administration utility to create, maintain, or delete applications, application groups, folders, groups, printers, storage sets, and users. You can run the ARSXML administration utility from the command line or invoke it from a user–written program.
- The ARSDOC program is a multipurpose document processing program. You can use the ARSDOC program to do the following tasks:
  - Query the database and generate a list of items that match a query
  - Retrieve documents from the system
  - Add, delete, and update documents
  - Send documents to a server printer
  - Add and release holds on documents
  - Add or remove documents from full text indexing
  - Federate documents to IBM FileNet® P8 through Content Federation Services for Content Manager OnDemand

  query the library server and generate a list of items that match a query; retrieve documents from the system; add, delete, and update documents; and send documents to the server print facility. You can run the ARSDOC program from the command line or invoke it from a user–written program.
- The ARSTBLSP program can be run to change the table that Content Manager OnDemand loads data into. During normal operation, Content Manager OnDemand loads index rows into a table until the Maximum Rows value for the application group has been reached. Such a table is said to be open for loading. When the Maximum Rows value is reached, the table is closed and a new table and table space are created. Under certain circumstances, an installation might desire to close a table to loading before the Maximum Rows value is reached. For example, migration processing (by using `arsmaint -e`) will not process a table that is open for loading, and the installation might desire to migrate the table earlier than initially anticipated.

The *IBM Content Manager OnDemand for Multiplatforms: Administration Guide* provides details about these and other server programs.

Server logging

System logging facility

Content Manager OnDemand provides the system logging facility to help an administrator track activity and monitor the system.

Content Manager OnDemand can log messages that are generated by the various client and server programs. For example, you can configure the system to save a message in the system log every time a user logs on the system; you can configure the system to save a message every time an unsuccessful log on attempt occurs; and so on. When you use the administrative client to add objects to the system and update the database, Content Manager OnDemand saves information about the actions in the system log. You can use one of the Content Manager OnDemand client programs to search for and view messages from the system log by time stamp, severity, message number, user ID, and other search criteria.

System log user exit

Content Manager OnDemand provides a user exit that can be used to process the messages that are written to the Content Manager OnDemand system log.

A common use of the system log user exit is to watch for error conditions or certain messages and take the appropriate action, such as notifying an administrator or operator or running some other program.

The system log user exit runs the ARSLOG program (on UNIX servers; the ARSL0G.BAT file on Windows servers) after writing a record to the system log. However, the ARSLOG program that is provided with Content Manager OnDemand does not perform any functions. You must replace the one that is provided by IBM with your own program that performs the functions that you require. For example, you could create a program to check the message number and severity of each message written to the system log and, when appropriate, send an alert to the Tivoli system management console.

Content Manager OnDemand sends parameters to the system log user exit, such as the name of the Content Manager OnDemand instance, a time stamp, a log record identifier, the user ID that is associated with the action, accounting information for the user ID, a message severity, a message number, and the text of the message. The information that appears in the accounting information part of the message can be specified for each user defined to the system by using the add or update a user command. You can customize the text of the messages by selecting the application group fields (and values) to include in the message. You can further configure Content Manager OnDemand to provide specific information to the system log user exit by setting system and application group parameters with the administrative client.

See the *Content Manager OnDemand for Multiplatforms: Installation and Configuration Guide* for more information about the system log user exit.

Security user exit

Content Manager OnDemand provides a user exit that allows you to implement your own user exit program to identify and authenticate users that log on to the system.

You can use the security user exit to authenticate a user's password by some means other than the way that is built in to Content Manager OnDemand. For example, you might want to deny access to the system after three incorrect logon attempts are made by a user; you might want to enforce some sort of password uniqueness; and so forth. You can also use the security user exit to allow users that are not already in the Content Manager OnDemand user database to access the system.

The security user exit allows you to augment the security related processing of the following activities or events:

- Logon
- Change Password
- Add User ID or Delete User ID by using the Content Manager OnDemand administrative functions
- Access to an Content Manager OnDemand folder
- Access to an Content Manager OnDemand application group

When driven for these activities, a user-written exit routine (or set of exit routines) can interact with some other security system to determine if the given activity is to be allowed or disallowed.

The security user exit runs the ARSUSEC program when a user attempts to logon to the system. A sample C program is provided in the EXITS directory. To implement your own security user exit program, you should add your specific code to the sample provided (for example, you could call another program from the ARSUSEC program). See the ARSCSXIT.H file for information about functions, parameters, and return codes. You then compile the ARSUSEC program (a Makefile is provided) and move or copy the executable program to the BIN directory. Then restart the library server to begin using the security user exit program.

Important: When you implement your own security user exit program, you bypass the logon verification processing that is built into the base Content Manager OnDemand product. IBM advises caution when you bypass the Content Manager OnDemand user and password restrictions. The security of the system could easily be subverted by malicious or defective code. Only use code that you trust.

See the *Content Manager OnDemand for Multiplatforms: Installation and Configuration Guide* for more information about the security user exit.

Retrieval preview user exit

The Content Manager OnDemand retrieval preview user exit allows you to process document data before the document is presented to the client.

The retrieval preview user exit can be used to add, remove, or reformat data before the document is presented to the client. For example:

- Remove pages from the document, such as banner pages, title pages, all pages but the summary page, and so on.
- Remove specific words, columns of data, or other information from the document. That is, omit ("white out") sensitive information such as salaries, social security numbers, and birth dates.
- Add information to the document, for example, a summary page, data analysis information, and Confidential or Copy statements.
- Reformat data contained in the document, for example, reorder the columns of data.

See the *Content Manager OnDemand for Multiplatforms: Installation and Configuration Guide* for more information about the retrieval preview user exit.

Download user exit

The ARSJESD program is the component of Download for z/OS that runs on the workstation.

The **-x** parameter of the ARSJESD program might be used to specify the name of a user-written program to process additional job information sent by PSF through the APSUX15 user exit.

The additional job information is installation dependent. See *PSF for z/OS: Download for z/OS* for details about the APSUX15 user exit and the content, format, and purpose of the additional job information. The processing done by the user-written program is also installation dependent. See your Infoprint Manager or PSF information for information about processing the additional job information with a user-written program.

If the ARSJESD program was invoked with the **-x** parameter, it calls the specified user-written program. The ARSJESD program passes the file name and the additional job information to the user-written program. Using this exit, it is possible to do functions such as parse the additional job information that is sent by PSF and rename the input file by using one of the PSF parameters.

See the *IBM Content Manager OnDemand for Multiplatforms: Installation and Configuration Guide* for more information about the Download user exit.

Report Specifications Archive Definition exit

The Report Specifications Archive Definition exit allows an installation to modify some of the parameters used by Content Manager OnDemand when document data is being captured (loaded) by the ARSLOAD program.

The following parameters can be modified:

- The Application Group name.
- The Application name.
- The name of the Object Server to be used for data storage.
- The name of the Storage Node to be used for data storage.
- The indexer parameters set.
- The input file control character type, logical record length and record format.

See the *IBM Content Manager OnDemand for Multiplatforms: Installation and Configuration Guide* for more information about the Report Specifications Archive Definition exit.

Table space creation exit

The table space creation exit allows an installation to take action when Content Manager OnDemand is going to create a table space, table, or index tables that will be used to store application index data.

The table space creation exit is not called for the Content Manager OnDemand system tables. For table and index creation, the installation can alter the SQL that is used to create the table or index.

See the *Content Manager OnDemand for Multiplatforms: Installation and Configuration Guide* for more information about the table space creation exit.

User exit programming

User exits provided by Content Manager OnDemand are specific points in the program where an experienced programmer can specify processing routines to enhance or replace the default Content Manager OnDemand functions.

For example, the logon user exit provides a point on the library server where you can identify and authenticate users that log on to the system. Programmers require a working knowledge of the tools needed to develop a user exit program. The following list identifies the main skills and tools that are needed:

- Skills
  - C and C++ programming
  - Operating system programming
  - Experience with relational database technology
  - Knowledge of compiling and linking programs in the C, C++, and operating system environment
  - DB2 UDB, Oracle, or SQL Server (if writing your own SQL code)
- Tools
  - IDE
  - C or C++ compiler

If you do not have these skills, see your IBM representative.

License information

The license information to use Content Manager OnDemand, DB2, and Tivoli Storage Manager is included in the product package.

You should read the License Information booklet carefully before using the software that is provided with Content Manager OnDemand.

If you plan to use Oracle instead of DB2, you must contact your Oracle sales representative for information about concurrent user licensing and upgrading your licenses.

If you plan to use SQL Server instead of DB2, you might need a license for each concurrent connection to the database. Contact your Microsoft representative for more information.

Chapter 3. System requirements

Review your system specifications to help plan the disk storage required to support your system.

For hardware and software requirements, see http://www.ibm.com/support/docview.wss?uid=swg27049168 or search for 7049168 at http://www.ibm.com/

Chapter 4. Storage

You can configure a Content Manager OnDemand system to maintain copies of reports in cache storage and in archive storage.

The copies in archive storage are for long-term storage. You can use Tivoli Storage Manager or other external storage managers to maintain reports in archive storage.

Tivoli Storage Manager supports a variety of optical and tape storage devices. Tivoli Storage Manager includes the following components:

- A server program that maintains a database of information about the devices and data that it manages. The server program also controls the storage media and devices that you define to Tivoli Storage Manager.
- An administrative client program that you can use to control and monitor the server program activities and define storage management policies. The activities include expiration processing, which is the process of deleting data that is eligible to be removed from the system, and reclamation processing, which is the process of reclaiming the space taken by expired data. Storage volumes that have been reclaimed can be reused. The storage management policies determine where data is stored and how long Tivoli Storage Manager maintains the data.
- An API that Content Manager OnDemand uses to work with Tivoli Storage Manager. The Tivoli Storage Manager API is required on the library server and all object servers that use Tivoli Storage Manager.
- Device support modules which provide support for storage devices and storage libraries.

Other external storage managers provide support for cloud and external file system storage.

Storage objects

The storage management criteria that you specify on the library server determines where and when Content Manager OnDemand stores reports and how it maintains them.

Figure 9 on page 25 shows the primary Content Manager OnDemand storage objects.

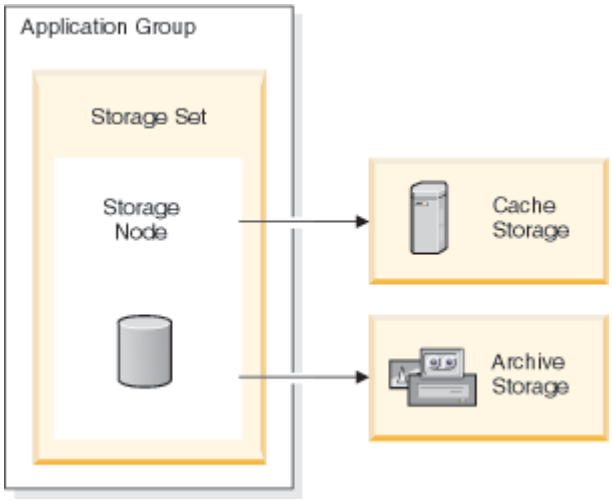

Figure 9. Content Manager OnDemand storage objects

An administrator creates an Content Manager OnDemand application for each report that is to be stored on the system. Applications with similar storage characteristics can be placed into a collection called an application group.

When you load a report into Content Manager OnDemand, you assign it to an application group. The application group identifies a storage set. The storage set contains one or more storage nodes. A storage

node identifies an object server on which data is stored. Content Manager OnDemand will automatically store a copy of the report in cache storage on the object server, unless you specify otherwise. A storage node can also identify a *client node* in Tivoli Storage Manager or another external storage manager for cloud or external file system storage. If a storage node identifies a client node in Tivoli Storage Manager, then Content Manager OnDemand automatically stores a copy of the report in archive storage, which is managed by Tivoli Storage Manager.

One or more application groups can specify the same storage set. However, a storage set can write to only one (archive) storage node at a time. This means that all of the data that is written to a storage node will be maintained using the same *policy*, for example, the type of media, the devices, the length of time to maintain data on the system, and so forth.

If you use Tivoli Storage Manager to maintain reports, you should specify the same storage management criteria to Content Manager OnDemand and Tivoli Storage Manager. For example, the Life of Data and Indexes, which is used by Content Manager OnDemand, and the Retention Period, which is used by Tivoli Storage Manager, should be the same value.

Overview of Tivoli Storage Manager

This section describes the basic Tivoli Storage Manager concepts. Tivoli Storage Manager is also known as IBM Spectrum Protect.

For more information about Tivoli Storage Manager and details about storage policies and devices and managing Tivoli Storage Manager storage, please see the Tivoli Storage Manager *Administrator's Guide*.

Figure 10 on page 27 illustrates the concepts discussed in this section:

- A client node is registered in a policy domain. The other Tivoli Storage Manager policy objects are within the policy domain.
- When a report is copied to archive storage, it is bound to a management class. The management class and the archive copy group within it specify where the report is stored and how it is managed.
- A storage pool is the destination for reports that are copied to archive storage. An archive copy group specifies the name of the storage pool. The storage pool is mapped to a device class, which represents a device. The storage pool contains volumes as indicated in the device type that is associated with the device class. For example, a storage pool that is mapped to a device class with a device type of OPTICAL contains only optical storage volumes. All devices require a device class that specifies a device type. Optical and tape devices also require a library and drive for management of media, including the mounting of that media.

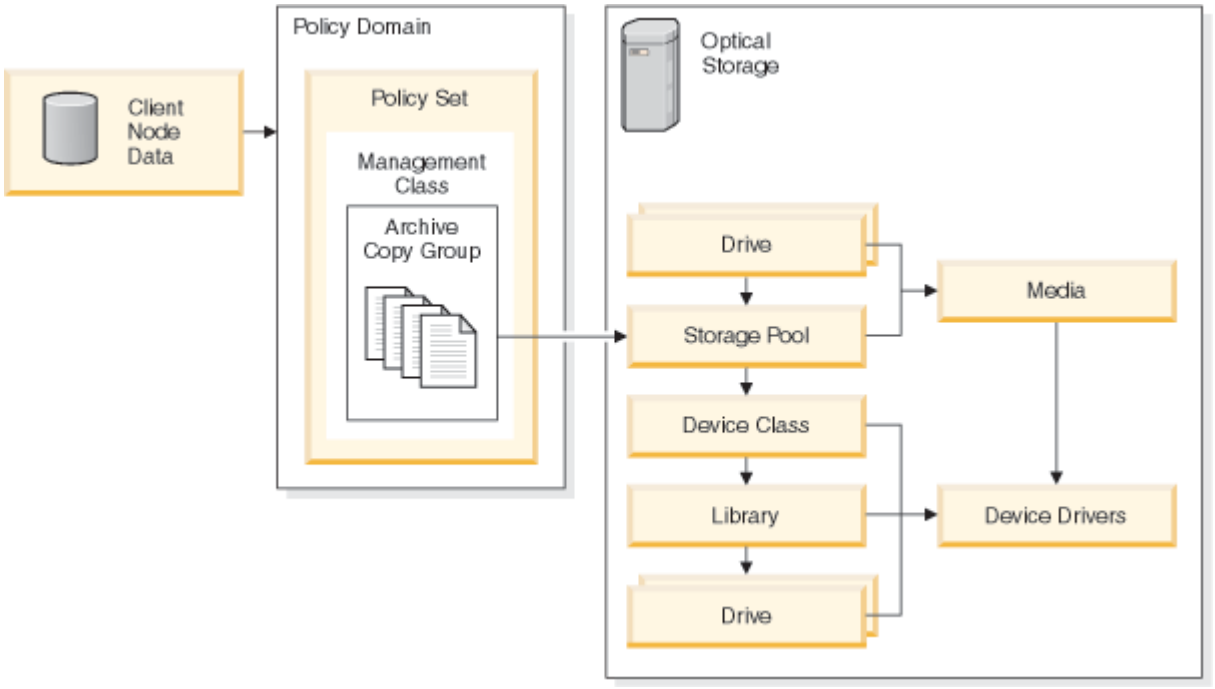

Figure 10. Tivoli Storage Manager storage objects

Storage policy

You can use different storage options to archive your system.

Client node
Represents an object server on which the Tivoli Storage Manager backup-archive client program has been installed, and has been assigned to a policy domain

Policy domain
Contains the policy set, management class, and archive copy group that is used by the client nodes that are assigned to the policy domain

Policy set
Contains the rules that are currently in use by all client nodes that are assigned to the policy domain

Management class
Determines where data is stored and how it is managed

Archive copy group
Used to copy data to Tivoli Storage Manager for long-term storage

Storage devices and media

You can use several options to store devices and media on your system.

Library
A Tivoli Storage Manager library is one or more drives (and possibly robotic devices) with similar media mounting requirements

Drive
Each drive defined to Tivoli Storage Manager represents a drive mechanism in a tape or optical device

Device Class
Each device is associated with a device class that specifies the device type and how the device manages its media.

Storage Pools and Volumes
A storage pool is a named collection of storage volumes of the same media type. A storage pool is associated with a device class. For example, an OPTICAL storage pool contains only optical storage volumes. A storage pool volume is associated with a specific storage pool.

Other external storage managers

Content Manager OnDemand supports external cloud storage managers such as Amazon Simple Storage Service (S3), Apache Hadoop Distributed File System (HDFS), Hitachi Content Platform, IBM Cloud Object Storage, Microsoft Azure, and OpenStack Swift. Content Manager OnDemand can also store data to external file systems.

Cloud storage options

The Content Manager OnDemand server can be configured to maintain copies of its stored data in both cache storage, managed by Content Manager OnDemand, and in archive storage, now referred to as external storage. Historically, IBM Tivoli Storage Manager (TSM) has been the only option used by Content Manager OnDemand to maintain data stored in external storage. The addition of Amazon S3, Apache HDFS, Hitachi Content Platform, IBM Cloud Object Storage, Microsoft Azure, and OpenStack Swift support augments the storage capabilities of Content Manager OnDemand by providing multiple external cloud storage solution options. Cloud storage solutions allow Content Manager OnDemand users to leverage the advantages that such storage provides such as cost savings, data replication, and disaster recovery. This functionality is configured in Content Manager OnDemand and behaves much in the same way that communicating with TSM does. This means that data in Content Manager OnDemand can be stored in cache as well as stored in Amazon S3, Apache HDFS, Hitachi Content Platform, IBM Cloud Object Storage, Microsoft Azure, or OpenStack Swift. The storing of data to any external cloud storage manager can take place at the same time that data is written to the Content Manager OnDemand cache or can be scheduled to migrate at a later date.

Amazon S3, Apache HDFS, Hitachi Content Platform, IBM Cloud Object Storage, Microsoft Azure, and OpenStack Swift storage options complement the functionality provided by TSM. Content Manager OnDemand servers can be configured to use any combination of Amazon S3, Apache HDFS, Hitachi Content Platform, IBM Cloud Object Storage, Microsoft Azure, OpenStack Swift, and TSM.

Additional information for each cloud storage solution can be found on the web at the following locations:

Amazon S3
https://aws.amazon.com/s3/

Apache HDFS
https://hadoop.apache.org/

Hitachi Content Platform
https://www.hitachivantara.com/en-us/products/cloud-object-platform/content-platform.html

IBM Cloud Object Storage
https://www.ibm.com/cloud-computing/infrastructure/object-storage/

Microsoft Azure
https://azure.microsoft.com/

OpenStack Swift
http://docs.openstack.org/developer/swift/

Using a file system for external storage

The Content Manager OnDemand server can be configured to maintain copies of its stored data in a file system accessible to Content Manager OnDemand. As with the cloud storage options, the storing of data to an external file system can take place at the same time that data is written to the Content Manager OnDemand cache or can be scheduled to migrate at a later date.

Detailed configuration information can be found in the *Content Manager OnDemand for Multiplatforms Installation and Configuration Guide*.

Chapter 5. Defining the storage configuration

Before you begin loading reports on the system, you need to determine the amount of storage required to hold the report data.

You should also determine how long the system should maintain a version of a report, how many copies of a report the system should maintain, on what type of media a report should be stored, and any other business, legal or operational requirements for storing and maintaining data. For example:

- You determine that many of the reports generated by your organization have the same basic space, media, and retention requirements. You can dedicate a high-capacity optical library to these reports. For more information about high-capacity optical libraries available from IBM, see www.ibm.com/systems/storage/optical/.
- You determine that several of the reports contain critical data and you need the system to maintain a backup copy of the data. You can install, configure, and define a second storage library to the system so that Tivoli Storage Manager can automatically maintain a backup copy of the reports. You can also use other external storage managers for cloud or external file system storage.
- You determine that several of the reports contain critical data and you need to keep those reports beyond the time you keep other reports. You install and configure the Enhanced Retention Management feature, then place a hold on those reports. Content Manager OnDemand does not delete those reports when you run the expiration policy.

If you are going to use Tivoli Storage Manager for your archive storage, then after collecting the storage requirements, you typically work with a Tivoli Storage Manager administrator to configure storage devices on the system and define devices to Tivoli Storage Manager. For example, when you define an optical library to Tivoli Storage Manager, you specify the type of device, the number of drives in the library, and the capacity of the storage volumes. The Tivoli Storage Manager administrator defines a device class and a storage pool for each storage library. A storage pool contains a set of storage volumes that belong to the same library. Tivoli Storage Manager keeps track of the storage volumes that belong to each storage pool, including the utilization percentage and the date a storage volume was last written to or read from.

After defining the storage devices, the Tivoli Storage Manager administrator defines the storage management policies, using the information that you collected about the reports that you plan to maintain on the system. For example, the policy information includes the length of time that Tivoli Storage Manager should keep the data that it manages.

When you load a report into the system, you identify an application group. The application group identifies a storage set. The storage nodes in a storage set determine how many copies of a report are maintained and where the copies are maintained. A storage node identifies a Content Manager OnDemand object server and, optionally, a client node in Tivoli Storage Manager. If the storage node identifies a client node in Tivoli Storage Manager, a copy of the report is stored in the library that is associated with the domain in which the client node is registered.

Chapter 6. Operational considerations

Each storage set identifies the object server and storage nodes where data is to be written. Content Manager OnDemand can write data to one storage node at a time. (Content Manager OnDemand can read data from several storage nodes.)

If the storage set contains more than one storage node, an administrator must identify the specific storage node where data is to be written. An administrator can update the storage set to change the storage node where data is to be written and add storage nodes to the storage set.

Content Manager OnDemand segments and compresses report data into storage objects. A storage object is a container of compressed documents that is maintained by the storage manager. Content Manager OnDemand does not require Tivoli Storage Manager to compress the storage objects. Content Manager OnDemand extracts and decompresses a small portion of the storage object, as required, when users retrieve the report.

Tivoli Storage Manager places storage objects on the storage volumes that it manages. The data on these storage volumes can be copied to a copy storage pool, providing a backup copy of the reports that are stored in archive storage.

An administrator specifies migration and expiration criteria for each application group:

- Migration is the process by which data is copied to archive storage. In general, it is recommended that most customers plan to migrate (copy) data to archive storage when a report is loaded on the system.
- Expiration is the process of deleting data that is eligible to be removed from the system. Content Manager OnDemand and Tivoli Storage Manager delete data independently of each other. Each uses their own criteria to determine when data expires and should be removed from the system. Each uses their own utilities to remove data. For final removal of data from the system, you should specify the same criteria to Content Manager OnDemand and to Tivoli Storage Manager. The Life of Data and Indexes, which is used by Content Manager OnDemand, and the Retention Period, which is used by Tivoli Storage Manager, should be the same value.

  If any documents have *holds* on them when it is time to expire data, Content Manager OnDemand does not delete any of those documents. The Enhanced Retention Management feature, which provides the holds functionality, requires you to disable Tivoli Storage Manager; Tivoli Storage Manager is not able to delete these documents.

Deleting application groups

On UNIX servers, if the owner of the Content Manager OnDemand instance is not root, when a user deletes an application group, Content Manager OnDemand deletes the application group tables from the Content Manager OnDemand database but does not delete the application group data from Tivoli Storage Manager.

In this case, Content Manager OnDemand issues message number ARS0022, which states that the user must manually delete the application group data from Tivoli Storage Manager. To delete the data from Tivoli Storage Manager, log on to Tivoli Storage Manager and use Tivoli Storage Manager utilities to delete the filespace in Tivoli Storage Manager that is associated with the application group. The name of the filespace is specified in message number ARS0022. See your Tivoli Storage Manager information for details about logging on to Tivoli Storage Manager and using Tivoli Storage Manager utilities to delete data.

Chapter 7. Storage requirements

Estimating storage requirements for a Content Manager OnDemand system begins with understanding and documenting end-user requirements for storing and accessing data.

Before you turn requirements into a storage subsystem to support your system, you must also review the various operational and performance issues. For example, Content Manager OnDemand supports up to 128 index fields for each report. However, users should not need a lot of indexes to locate a specific version of a report or a document within a report. The number of index fields that you define has a direct impact on the amount of disk space that you will need for your database. In addition, the more indexes that you define for a report, the longer it will take to load the report into the system. It is important to work with users and understand their data retrieval requirements. Define only the number of index fields that they need. You might have to balance end-user requirements with disk space, the amount of time required to load a report, and other performance issues.

Maintaining a copy of reports in cache storage can have a significant impact on the amount of disk storage that you need on your system. Most customers store the latest versions or most frequently accessed reports in cache storage. You should review how users search for and retrieve information from the reports that you plan to store in Content Manager OnDemand. For example, if most retrievals occur in the first 90 days after a report is generated, then you probably want to store the report in cache storage for at least that length of time. You should choose a time frame to cache each report which meets the requirements of your users and also makes the best use of available cache storage space.

There are several components that you need to measure to determine the amount of disk, optical, and tape storage required to support a Content Manager OnDemand system. For example, the following components of the system require disk storage:

- Storage space for application programs and system software, including the base operating system, the Content Manager OnDemand server software, and the database manager and optional components such as the archive storage manager and the server print manager.
- Storage space for configuration files and control files.
- Storage space for the Content Manager OnDemand system logging facility.
- Temporary storage space for reports received from other systems. In general, you should plan for enough disk space to hold either the largest single report that you will be loading on the system or the total of several reports that might be staged for loading at the same time, which ever requires the most storage space. In many organizations, most versions of a report are similar in size. However, there might be times when a report is much larger than average. For example, a report generated at the end of the month or the end of the quarter might greatly exceed the average report size.
- Temporary storage space for indexing a report on the Content Manager OnDemand server.
- Temporary storage space for loading a report on the Content Manager OnDemand server.
- Cache storage. This might be zero, for reports that do not require cache storage. However, a very large amount of disk space might be required for reports that must remain in cache storage for several months or longer.

  Content Manager OnDemand compresses report data before storing it on storage volumes. The compression ratio can have a significant impact on the amount of disk space that you need to store a report in cache storage. Content Manager OnDemand can achieve up to 30:1 compression on line data reports. However, for reports that contain AFP data or image data that is already compressed, the compression achieved will be much lower.
- Storage space for the database, which includes Content Manager OnDemand system tables (control information and objects that you define to Content Manager OnDemand) and application group tables (index data extracted from reports). The amount of database space that you should plan for a report is a factor of the number of items contained in the report, the number of index fields that you define for the report, the number of versions of a report (or the frequency with which you load a report on the system), and how long you need to maintain a report on the system.

For reports that contain sorted transaction data, Content Manager OnDemand can divide the report into groups of a fixed number of pages and create one index row for each group of pages. For reports that contain logical items, such as statements, and policies, Content Manager OnDemand can create one index row for each logical item in the report. Typically the database space required for indexing sorted transaction data is much less than the database space required for indexing reports that contain logical items. Also, index fields provide fast lookup, but require a significant amount of database space.

- Storage space for database log files. You should plan for disk space for active or primary log files and for log files that are not active but might still be needed for recovery (sometimes known as archived log files). If you use Tivoli Storage Manager facilities to backup and restore DB2 databases, you should plan for additional disk space for the primary log files, but you will not need disk space for the archived log files.
- Storage space for the database and logs used by the archive storage manager.
- Temporary storage space for server print and FAX.
- Temporary storage space for importing migrated indexes from archive storage to the database.

For sorted transaction data, the examples and calculations that follow assume that Content Manager OnDemand creates one indexed item for each group of 100 pages in a report. The number of pages in a group is a parameter that you can configure when you index a report with ACIF. The *Content Manager OnDemand Indexing Reference* provides move information. The following components of the system require archive storage (optical and tape storage):

- Reports that you plan to maintain in archive storage.
- Backup copies of reports stored in archive storage. (For critical applications, some customers require that the system maintain two copies of a report in archive storage.)
- Database archived log files and backup image files, if you use Tivoli Storage Manager facilities to backup and restore DB2 databases.

When you calculate archive storage requirements, you should also determine the number of storage volumes and libraries that you need to hold the data that will be stored on your system. Optical libraries are capable of holding a large amount of data, with the storage capacity usually expressed in amounts of uncompressed data. Depending on the compression ratio achieved for your reports, an optical library might be able to hold more than the stated amount. For example, if Content Manager OnDemand can achieve a 6:1 compression ratio on the reports to be stored in Tivoli Storage Manager, then the library could hold multiple terabytes of report data, depending on the exact hardware configuration.

You can replace full optical storage volumes as needed, if the availability requirements of your system allow you to do so. For example, you might decide to remove full storage volumes from a library one year after the last time that Content Manager OnDemand wrote report data to the storage volume. You could replace the full storage volumes with newly initialized storage volumes to hold the latest reports stored on the system. That way, the latest versions of a report are always available in the library. However, if you need to keep many years of report data online in the library or you store massive amounts of data in your application groups, then you might need to plan on having several optical libraries for your system.

Storage hierarchy

There are several different storage management strategies that you can use with Content Manager OnDemand and most archive storage managers.

For example, Tivoli Storage Manager is a hierarchical storage management system that manages storage pools of disk devices, optical devices, and tape devices. Tivoli Storage Manager allows data to be migrated from one storage pool to another using criteria defined by an administrator. For most customers, Content Manager OnDemand does not use the hierarchical storage management capabilities of Tivoli Storage Manager, because of the time required to migrate data from one storage medium to another. However, if you need to, you can configure Tivoli Storage Manager to migrate the data that it maintains from one storage medium to another.

Content Manager OnDemand maintains a cache storage system independently of the archive storage manager. The cache storage system should contain the fastest storage devices, for high-speed access to

reports. When you load a report on the system, Content Manager OnDemand can automatically store one copy of the report in cache storage and another copy of the report in archive storage. Content Manager OnDemand also supports the option of storing reports in cache storage and then later migrating them to archive storage. However, it is recommended that you always plan to copy reports to cache storage and archive storage at the same time (when you load the report). Doing so usually eliminates the need for you to periodically backup cache storage, because a backup copy of your reports already exists on archive storage. Copying reports to cache storage and archive storage at the same time also eliminates the need for you to migrate reports to archive storage.

Reports expire (are eligible to be removed) from cache storage when they reach their cache storage expiration date. You specify the cache storage expiration date for a report when you create an application group. For example, you can specify that a report should expire from cache storage after it has been stored there for ninety days. Content Manager OnDemand provides a utility that you can use to automatically remove expired reports from cache storage on a regular schedule. After you run expiration processing, Content Manager OnDemand reclaims the space taken by expired documents. If any reports have *holds* on them when you run expiration processing, Content Manager OnDemand does not remove any of those reports. The Enhanced Retention Management feature, which provides the holds functionality, requires you to disable the archive storage manager; therefore, the archive storage manager will not be able to delete these reports.

Content Manager OnDemand and the archive storage manager maintain documents independently of each other. For example, each use their own criteria to determine when data expires and should be removed from the system; each use their own utilities to remove documents. However, for removal of documents from the system, you should specify the same criteria to Content Manager OnDemand and the archive storage manager. For example, the Life of Data and Indexes, which is used by Content Manager OnDemand, should specify the same length of time as the Retention Period, which is used by Tivoli Storage Manager.

Data compression

Content Manager OnDemand can compress report data using several different data compression algorithms, before storing the data in cache storage and archive storage.

The compression ratio that Content Manager OnDemand can achieve has a significant impact on the amount of space required to store reports.

The compression ratios that Content Manager OnDemand can achieve vary widely depending on the type of data and the format of the data. You cannot always accurately estimate the compression ratio by simply examining the data. On average, you can expect to achieve between 2:1 and 15:1 compression for AFP documents and up to 30:1 compression for line data reports. Compression for AFP documents is based on the output data file produced by ACIF, and not the input file, which could have been line data. When ACIF formats line data with a page definition, it might increase the size of the data by adding AFP controls for positioning text.

To properly estimate the amount of storage space required by a report, it is recommended that you measure the compression ratio achieved on a sample of the report. You can measure the compression ratio by using the ARSADMIN program. For example:

- For reports that contain logical items, such as statements and policies, use the following example:

```
arsadmin compress -l 200000 -s inputFile -o outputFile
```

  Where `inputFile` is the report that you want to measure and `outputFile` is the compressed output.

  To determine the compression ratio, divide the size of `outputFile` by the length (`-l 200000`). For example, if the size of `outputFile` is 66,000 bytes, then the compression ratio is 66000/200000 or 0.33 (3:1 compression).
- For reports that contain line data and include a sorted transaction value, such as a general ledger, first determine the size of an indexed group of pages, for example, 100 pages. Then extract a group of pages from a larger report and process them with the ARSADMIN program. For example:

```
arsadmin compress -s groupPages -o outputFile
```

Where `groupPages` is a file that contains a representative group of pages from a larger report and `outputFile` is the compressed output.

To determine the compression ratio, divide the size of `outputFile` by `groupPages`. For example, if the size of `outputFile` is 40,000 bytes and the size of the group of pages is 200,000 bytes, then the compression ratio is 40000/200000 or 0.20 (5:1 compression).

See the *IBM Content Manager OnDemand for Multiplatforms: Administration Guide* for more information about the ARSADMIN program.

Estimating disk storage requirements

System software

Most Content Manager OnDemand servers require at least 4 GB of disk storage space for software products.

This includes the operating system software, swap space, temporary work space, user space, the database manager software, the archive storage manager software, the server print manager software, and the Content Manager OnDemand server software.

Download

Content Manager OnDemand requires temporary storage space to hold reports that are transmitted (downloaded) from other systems.

For example, you might need to transmit reports from an z/OS system to an Content Manager OnDemand object server. Many customers download report data during the day when their application programs generate the reports, but do not load the data into Content Manager OnDemand until the evening or other periods of little or no other activity on the system. This method requires enough disk space to hold all of the data generated in one day. (Or, if your organization defers the loading of data for several days, enough disk space to hold all of the data that accumulates before you begin loading the data.) It is recommended that you dedicate one or more disk storage volumes to data download storage.

Use the following calculation to determine the amount of disk space required to hold data downloaded from other systems:

```
 Report
Download  =  Total data for  *  1.20
 space       largest cycle
```

Figure 11. Calculating report download disk storage space

Where `Total data for largest cycle` is the size in bytes of the largest version of a report or the total size of all of the reports that the server must hold before you begin loading the data (if you defer the loading of reports).

For example, if you download 400 MB of data in a single day, then the download space required on the Content Manager OnDemand server is:

```
 Report
Download  =  400 MB  *  1.20  =  480 MB
 space
```

Temporary space for indexing

Content Manager OnDemand requires temporary storage space on disk to index reports.

The temporary space required by Content Manager OnDemand is a factor of the largest version of a report and the number of reports that you plan to index at the same time. You also need to know where you will index the reports: on a z/OS system or on the Content Manager OnDemand server.

Use the following calculation to determine the amount of temporary space required to index reports:

- If you plan to index reports on a z/OS system, then no temporary space is required.
- If you plan to index reports on the Content Manager OnDemand server:

```
Temporary   =  Largest report  *  1.5
space          file size
```

Figure 12. Calculating temporary space for indexing

Where `Largest report file size` is the size in bytes of the largest version of a report to be indexed or the total size of all of the reports that the server must index at the same time (if you index more than one report at a time).

For example, if the largest report is 400 MB and the report is indexed on the Content Manager OnDemand server, then the temporary space required to index the report is:

```
Temporary   =  400 MB  *  1.5  =  600 MB
space
```

Cache storage

Most customers store reports in cache storage for a short period of time, to provide the fastest retrieval for the most frequently used reports.

The amount of disk space that you should dedicate to cache storage varies greatly based on your storage requirements including:

- Number of reports that you want to store on the system
- The compression ratio that Content Manager OnDemand can archive
- The amount of time that you need to store a report in cache storage

As reports age, and retrieval requests for them are much less frequent, the reports can be retrieved from archive storage. Another reason to keep reports in cache storage is if lots of users access them at the same time. Because the archive storage manager might require from six and sixty seconds to mount an optical or tape storage volume and retrieve a report, it is usually not possible to support a high transaction rate for reports stored on archive storage.

The number of resources that need to be stored in the system is also a factor in deciding the amount of disk space for cache storage. Content Manager OnDemand always stores resources in cache storage to provide fast retrieval when a user selects an item for viewing. The ARSLOAD program saves only one copy of a resource on the system even if several reports use the same resource. When the ARSLOAD program processes a resource group file, it checks the resource identifier to determine whether the resource already exists on the system.

Another use of cache storage is for reports that have a short life, such as one week or one month. You can store these types of reports in cache storage and the system can be configured to automatically delete them when they reach their expiration date. Cache storage can also be used to hold reports for which you do not need a backup copy.

Use the following calculation to determine the amount of disk space required for cache storage:

```
Cache Storage = Size of Data per week
                  * Number of Weeks to cache
                      * Data Compression ratio
                          * 1.1
```

Figure 13. Calculating cache storage

For example, if you plan to load 2 GB of report data on the system each week, the reports must be maintained in cache storage for 12 weeks, and the compression ratio is 3:1 (0.33), then the disk space required for cache storage space can be calculated as follows:

```
Cache      =  2 GB  *  12  *  .33  * 1.1  =  8.71 GB
Storage
```

Database storage

When you load a report into the system, Content Manager OnDemand extracts index data from the report and stores it in an application group table in the database. For reports that contain logical items, such as statements and policies, Content Manager OnDemand can create one database row for every item found in the report. For reports that contain sorted transaction data, Content Manager OnDemand can creates one database row for every indexed group of pages (by default, 100 pages in a group).

A database row contains a fixed amount of information that Content Manager OnDemand uses to maintain reports (approximately 40 bytes) and any additional index and filter fields that you define for the application group. Index fields, which allow users to locate documents quickly, require significantly more disk storage space than filter fields. (Index fields also require more time to load into Content Manager OnDemand.)

There are four major factors that determine the amount of disk space required for the Content Manager OnDemand database:

1. The number of index and filter fields
2. The size of the index and filter fields
3. The number of indexed items per month
4. The number of months that Content Manager OnDemand maintains the index data in the database

Note for Oracles users: For decimal fields, the actual field size varies, depending on the average precision for each column. It is recommended that you read your Oracle information about the NUMBER data type.

Table 1. Index field data types. The table lists the types of database fields supported by Content Manager OnDemand and the number of bytes required to hold a value in each type of index field.

Field Type	Description	Field Size (DB2)	Field Size (Oracle)	Field Size (SQL Server)
Small Int	Contains whole numbers between -32768 and 32,767.	2 bytes	21 bytes	2 bytes
Integer	Contains whole numbers between -2147483648 and 2147483647.	4 bytes	21 bytes	4 bytes

[1] Oracle stores decimal values in a variable length format. Each value is stored in scientific notation, with one byte used to store the exponent and up to twenty bytes to store the mantissa. The resulting value is limited to 128 digits of precision. Oracle does not store leading and trailing zeros. For example, the number `412.50` is stored in a format similar to `4.125 * 100`, with one byte used to store the exponent (2) and two bytes used to store the four significant digits of the mantissa (`4, 1, 2, 5`). Taking this into account, the column data size for a particular decimal value `NUMBER (p)`, where p is the precision of a given value (scale has no effect), can be calculated using the formula: `1 + FLOOR( p/2 ) + 2`. Therefore, the system requires a minimum of four bytes to hold a decimal value.

Table 1. Index field data types. The table lists the types of database fields supported by Content Manager OnDemand and the number of bytes required to hold a value in each type of index field. *(continued)*

Field Type	**Description**	**Field Size (DB2)**	**Field Size (Oracle)**	**Field Size (SQL Server)**
Big Int	Contains whole numbers between -9223372036854775808 and 9223372036854775807.	8 bytes	21 bytes	8 bytes
Decimal	Contains numbers between 10(307) and 10(308) with up to 15 significant digits.	8 bytes	21 bytes	8 bytes
DecFloat (16)	Decimal floating-point number with 16 digits of precision.	8 bytes	Not available	Not available
DecFloat (34)	Decimal floating-point number with 34 digits of precision.	16 bytes	Not available	Not available
String (Fixed)	Contains letters, numbers, special symbols, such as the % and #, and any other printable character.	1 - 254 bytes. • **For SBCS string:** 1 byte per character • **For MBCS string:** Up to 4 bytes per character	1 - 254 bytes. • **For SBCS string:** 1 byte per character • **For MBCS string:** Up to 4 bytes per character	1 - 254 bytes. • **For SBCS string:** 1 byte per character • **For MBCS string:** Up to 4 bytes per character
String (Variable)	Contains letters, numbers, special symbols, such as the % and #, and any other printable character.	1 - 2000 bytes. • **For SBCS string:** 1 byte per character plus 2 bytes of overhead • **For MBCS string:** Up to 4 bytes per character plus 2 bytes of overhead Unused bytes do not use storage. See your database documentation for more information on variable string type.	1 - 2000 bytes. • **For SBCS string:** 1 byte per character plus 2 bytes of overhead • **For MBCS string:** Up to 4 bytes per character plus 2 bytes of overhead Unused bytes do not use storage. See your database documentation for more information on variable string type.	1 - 2000 bytes. • **For SBCS string:** 1 byte per character plus 2 bytes of overhead • **For MBCS string:** Up to 4 bytes per character plus 2 bytes of overhead Unused bytes do not use storage. See your database documentation for more information on variable string type.

Table 1. Index field data types. The table lists the types of database fields supported by Content Manager OnDemand and the number of bytes required to hold a value in each type of index field. *(continued)*

Field Type	Description	Field Size (DB2)	Field Size (Oracle)	Field Size (SQL Server)
Date	Contains a valid date from January 1, 0001 to December 31, 9999.	4 bytes	7 bytes	3 bytes
Date (old style)	Contains a valid date from January 1, 1970 to December 31, 2058.	2 bytes	2 bytes	2 bytes
Time (old style)	Contains times of day, stored in three second increments, since midnight, and limited to 24 hours.	2 bytes	2 bytes	2 bytes
Date/Time	Contains a valid date from January 1, 0001 to December 31, 9999.	13 bytes	11 bytes	8 bytes
Date/Time (old style)	Contains both a date and time value. The date can be from January 1, 1970 to January 18, 2038. The time is stored in one second increments.	4 bytes	4 bytes	4 bytes
Date/Time (TZ)	Contains a valid date from January 1, 0001 to December 31, 9999.	13 bytes	11 bytes	8 bytes
Date/Time (TZ) (old style)	Contains both a date and time value. A Date/Time (TZ) field is exactly like a Date/Time field, but uses the time zone from the client.	4 bytes	4 bytes	4 bytes

Estimating database storage space

You can use the following calculations to estimate the space required in the Content Manager OnDemand database to hold the index data for a report.

In general, it is recommended that you add a 10 or 20 percent buffer to the value returned by these formulas. These calculations can be used for reports that contain logical items and reports that contain a sorted transaction value.

Estimating the size of database objects is an imprecise undertaking. Overhead caused by disk fragmentation, free space, and the use of variable length fields (including numbers) make size estimation difficult, because there is such a wide range of possibilities for field types and row lengths. After initially estimating your database size, you should create a test database and populate it with representative data.

These examples uses the following assumptions:

- The system adds a 19-byte index to each table.
- `Index n length` is the size of a database field for which you want Content Manager OnDemand to build an index. For example, a date field requires 4 bytes to hold the date value. DB2 requires an additional eight bytes for each index that you define.
- The system adds approximately 40 bytes of control information to each row in a table.
- When the report contains logical items, the `Number of indexed items per month` is the number of statements, policies, and so forth.
- When the report contains a sorted transaction value, the `Number of indexed items per month` is the number of groups of indexed pages (by default, the system indexes a report in groups of 100 pages). You can specify the size of an indexed group of pages when you index a report with ACIF.

The examples shows the formula that was derived from information provided with DB2. See the DB2 product information for details.

```
TableSize = ( Sum of column lengths )

IndexSize = 19 + ( Index 1 length + 8 ) + ( Index 2 length + 8 ) + ...

DatabaseSize = ( ( TableSize + 40 ) * 1.5 ) + ( IndexSize * 2 )
                 * Number of indexed items per month
                      * Number of months to keep index in database
```

Figure 14. Calculating database storage space for DB2

The example shows the calculation that you can use to estimate database space requirements when the database manager is Oracle. The formula was derived in part from information provided by Oracle. For more information, or if you have special requirements or if you need to do more, see the Oracle product information. Also, the formula does not include space requirements related to file management overhead required by the operating system, including file block size and directory control space.

```
TableSize = ( Sum of column lengths ) + 3 + ( Number of columns * 2 )

IndexSize = 19 + ( Index 1 length + 8 ) + ( Index 2 length + 8 ) + ...

DatabaseSize = ( ( TableSize + 40 ) * 1.2 ) + ( IndexSize * 1.2 )
                 * Number of indexed items per month
                      * Number of months to keep index in database
```

Figure 15. Calculating database storage space for Oracle

The example shows the calculations that you can use to estimate database space requirements when the database manager is SQL Server. The formula was derived in part from information provided by Microsoft. For more information, or if you have special requirements or if you need to do more, see the SQL Server product information.

```
TableSize = ( Sum of column lengths ) + 6 + Number of VARCHAR columns

IndexSize = 19 + ( Index 1 length + 11 + ( 1 if VARCHAR ) ) +
            ( Index 2 length + 11 + ( 1 if VARCHAR ) ) + ...

DatabaseSize = ( ( TableSize + 40 ) * 1.2 ) + ( IndexSize * 1.2 )
                    * Number of indexed items per month
                         * Number of months to keep index in database
```

Figure 16. Calculating database storage space for SQL Server

Examples

The examples that follow assume that the database manager is DB2.

If Oracle is the database manager, use the calculations from Oracle; if SQL Server is the database manager, use the SQL Server calculation.

1. The following example illustrates how to calculate the database storage space required for a report that contains logical items, such as statements. The example plans for indexing one million items per month and keeping the index data in the database for 24 months. The table lists information about the database fields.

   Table 2. Database storage for a report that contains logical items

Field Name	Field Type	Field Size	Index or Filter
Report Date	Date	4 bytes	Index
Account Number	Fixed String	12 bytes	Index
Invoice Balance	Decimal	8 bytes	Filter
Customer Name	Variable String	20+4 bytes	Filter

   ```
   TableSize = ( 4 + 12 + 8 + ( 20 + 4 ) ) = 48

   IndexSize = 19 + ( 4 + 8 ) + ( 12 + 8 ) ) = 51

   DatabaseSize = ( ( 48 + 40 ) * 1.5 ) + ( 51 * 2 ) ) = 234
                   * 1,000,000 = 234000000
                     * 24  =  5616000000
   ```

   Content Manager OnDemand requires 5.6 GB of magnetic disk space to store 24 months of report index data in the database.

2. The following example illustrates how to calculate the database storage space required for a report that contains line data with a sorted transaction value. Because only one database row is generated for each indexed group of pages in the report, in general, significantly less database storage space is required than for reports that contain logical items.

   Reports that contain line data with a sorted transaction value use a fixed type of indexing, where each database row contains the beginning value, the ending value, and the beginning page number for the group of pages. Content Manager OnDemand maintains the beginning and ending values as indexes and the page number as a filter. The main parameters for the calculation are the length, in bytes, of the sorted transaction value, the number of pages generated in a month, the size of a group of indexed pages, and the number of months that Content Manager OnDemand maintains the index data in the database.

   The example plans for indexing one million pages per month, in groups of 100 pages, and keeping the index in the database for 24 months. The table lists information about the database fields.

   Table 3. Database storage for a report that contains a sorted transaction value

Field Name	Field Type	Field Size	Index or Filter
Report Date	Date	4 bytes	Index

Table 3. Database storage for a report that contains a sorted transaction value (continued)

Field Name	Field Type	Field Size	Index or Filter
Begin Transaction Value	Fixed String	10 bytes	Index
End Transaction Value	Fixed String	10 bytes	Index
Page Number	Integer	4 bytes	Filter

```
TableSize = ( 4 + 10 + 10 + 4 )  = 28

IndexSize = ( 19 + ( 4 + 8 ) + ( 10 + 8 ) + ( 10 + 8 ) ) = 67

DatabaseSize = ( ( 28 + 40 ) * 1.5 ) + ( 67 * 2 ) ) = 236
                 * ( 1,000,000/100 ) = 2360000
                    * 24  =  56640000
```

Content Manager OnDemand requires 56.6 MB of magnetic disk to store 24 months of report index data in the database.

Estimating the size of rollback segments

The database storage space requirements for your Content Manager OnDemand system includes a rollback segment, an area on your disk subsystem that is used when Content Manager OnDemand makes changes to the database.

This section contains information for Oracle users. The size of the rollback segment should be based on the types of transactions that run against the database.

In general, there are two types of transactions in Content Manager OnDemand: loading data into the system and deleting data from the system. Loading data into the system is a batch job of (usually) long running transactions. The Content Manager OnDemand actions that load data into the system include the ARSLOAD program and the ARSADMIN **LOAD** command. Deleting data from the system can be either a long running transaction (for example, when deleting an entire report from the system) or a short transaction (for example, when deleting a document from the system). The Content Manager OnDemand actions that delete data from the system include the ARSADMIN **UNLOAD** command, the ARSDOC **DELETE** command, and the ARSMAINT program. Data is deleted by the ARSMAINT program when you set an expiration type in your application groups.

Because of this mix of transaction sizes, most customers should plan a large rollback segment that can handle transactions of any size. Most customers should plan to allocate enough storage space for a rollback segment that can hold the transactions for the largest input file that will be loaded into the system.

To estimate the size of the rollback segment, you need to consider three factors:

1. The number of documents that are loaded into the system during the single largest load process.
2. The number of bytes that are allocated to the user-defined database fields for the application group that is associated with the single largest load process. This is the value of the index size.
3. The 40 bytes of system information that Content Manager OnDemand adds to each database row.

Once you know these values, you can use a formula to estimate the size of the rollback segment that is required for your system. The formula was derived in part from information provided by Oracle. For more information, or if you have special requirements or if you need to do more, see the Oracle product information. Also, the formula does not include space requirements related to file management overhead required by the operating system, including file block size and directory control space.

```
     ( Number of documents in largest load *
          ( Number of bytes in application group + 40 Content Manager OnDemand overhead
bytes ) )
               * 2 = estimated size of rollback segment
```

Figure 17. Formula for Estimating the Size of the Rollback Segment

For example, suppose that your largest Content Manager OnDemand load is for a statement application that loads 150,000 statements in a single load file. The Content Manager OnDemand application group database fields require approximately 50 bytes of database storage per document. The following example shows the calculation, which requires approximately 27 MB of rollback segment space.

```
( 150,000 statements *
       ( 50 bytes in application group + 40 Content Manager OnDemand overhead bytes ) )
            * 2 = 27,000,000
```

Figure 18. Example of Estimating the Size of the Rollback Segment

If you expect to delete data from the system by using the ARSDOC **DELETE** command (this is a somewhat unusual requirement), then instead of using the size of the single largest load file, you should substitute the largest number of records that you expect to delete during a delete process.

Database log file storage

The Content Manager OnDemand database includes recovery logs which are used to recover from application or system errors. In combination with database backups, they are used to recover the consistency of the database right up to the point in time when an error occurs. Some logs, called *active* or *primary* logs, contain transactions which have not been committed to the database. These logs are stored in the primary database log path. Other logs, called *archived* or *secondary* logs, contain transactions which have been committed to the database. These logs are stored in the secondary database log path. Both types of logs can be used with database backups to enable forward recovery of the database to any point in time before a failure.

When you load a report into Content Manager OnDemand, the database manager records changes made to the database in a recovery log:

- If you are using DB2, when a log fills, the database manager closes the full log and opens a new log. When all changes to the database have been made, the database manager closes the last log. After the load process disconnects from the database, Content Manager OnDemand copies the closed logs from the primary database log path to the secondary database log path. When you create a full backup image of the database with the arsdb command, Content Manager OnDemand deletes all of the logs from the secondary database log path. (When you create a full backup image of the database, it invalidates the secondary logs that were created before the time that the backup was taken.)
- If you are using SQL Server, when a log file fills, the database manager closes the full log file and opens a new log file, provided that you have configured the transaction log to use multiple log files. SQL Server also uses the *auto grow* feature to reduce the potential of running out of transaction log space. The log files are truncated after a successful backup of the transaction log and can be reused.

The amount of disk space that you need to store log files is a factor of the number and size of the log files and the length of time between full backups of the database.

When you install and configure Content Manager OnDemand, you set parameters that determine the number and size of the log files and where Content Manager OnDemand should store the log files.

Primary log storage space for a report

You can use calculations to estimate the amount of primary log space required for a report:

You can use the following calculation to estimate the amount of primary log space required for a report:

```
PrimaryLogSpace = ( ( ( TableSize + 40 ) * 1.5 ) + ( IndexSize * 2 ) )
                    * Number of indexed items
                    * 4
```

Figure 19. Calculating primary log storage space

- You can find the calculations for `TableSize` and `IndexSize`.
- Content Manager OnDemand adds approximately 40 bytes of control information to each indexed item.

- The `Number of indexed items` is the number of logical items or indexed groups of pages contained in the report. The number of indexed items depends on the organization of data in the report and how you index the report.

The following example illustrates the amount of primary log space required for a report, where the `TableSize` is 48, the `IndexSize` is 51, and the `Number of indexed items` added to the database is 50,000.

```
PrimaryLogSpace =  ( ( ( 48 + 40 ) * 1.5 ) + ( 51 * 2 ) ) = 234
                      * 50000 = 11700000
                      * 4   =   46800000
```

Content Manager OnDemand requires approximately 47 MB of primary log space for the sample report.

Primary log storage space for the system
It is critical that you allocate enough disk space for the primary logs. Content Manager OnDemand cannot load reports if the database manager runs out of space for the primary logs.

The amount of primary log space required on the system is a factor of the largest report that you plan to load into the system (in terms of the number of indexed items, the number of indexes, and the size of the indexes), the maximum number of reports that you plan to load into the system at any one time, and some buffer space. In addition, for DB2, if you use Tivoli Storage Manager to maintain the DB2 archived log files, it is recommended that you triple the amount of space that you estimate for the primary logs. To estimate the total amount of primary log space required for your system:

- Determine the primary log space required for the largest report.
- Estimate the maximum number of reports that Content Manager OnDemand must process at any one time. Determine the storage space required for each report. Total the values.
- Double the sum of the previous two values. The result is the storage space required for the primary logs.
- Using the previous value, verify that the database manager will allocate enough primary log space. Use the following calculation and database configuration parameters:

```
PrimaryLogSpace = ( ( ( logprimary + logsecond )
                    * ( logfilsiz + 2 ) * 4096 ) + 8192 )
```

Figure 20. Calculating primary log storage space

  - The `logprimary` is 40
  - The `logsecond` is 2
  - The `logfilsiz` is 1000

By default, Content Manager OnDemand allocates approximately 172 MB of primary log space.

Archive storage manager database and recovery log

The archive storage manager maintains a database of information about the storage devices that it manages, the storage objects that it maintains, and the management policies that it uses to maintain the storage objects.

The archive storage manager uses the information to store, retrieve, and expire report data that is stored on optical and tape storage volumes.

You can use the following calculation to determine the size of the archive storage manager database:

```
ASMDatabase = ( Data per Month / Object Size )
                * 700
                  *  Life of data in months
```

Figure 21. Calculating archive storage manager database storage space

For example, if you plan to store 8 GB of data per month into an application group, the size of a storage object in Content Manager OnDemand is 10 MB (the default), and the archive storage manager needs to

maintain the data for seven years (84 months), you should plan to allocate approximately 47 MB of disk storage space for the database:

```
ASMDatabase = ( 8,000,000,000 / 10,000,000 )
                * 700
                  * 84   =   47.04 MB
```

If you plan to maintain a backup copy of data stored in an application group, that is, you need two copies of the data on archive storage, double the space required for the archive storage manager database. If you plan to mirror the database, double the space required for the database. If you need a backup copy of the data and plan to mirror the database, quadruple the space required for the database.

Server print storage space

Content Manager OnDemand requires temporary work space to process requests for the server print manager.

You must allocate enough disk space to support the maximum number of concurrent print requests that the server must manage. It is recommended that you define a dedicated file system on which Content Manager OnDemand can store the temporary print files. It is recommended that at least 500 MB of free space be available in this file system at all times. If your storage configuration permits, it is recommended that you allocate 1 GB or more of free space to this file system.

Temporary space for importing index data

If you do not plan to migrate index data from the database to archive storage, then you do not need to allocate temporary storage space for importing the migrated index data.

Content Manager OnDemand requires temporary work space to import migrated index data from archive storage into the database. You must allocate enough disk space to support the maximum number of concurrent import requests that the server must manage. It is recommended that you define a dedicated file system where Content Manager OnDemand can store temporary data created by the programs that import migrated index data. The amount of space that you allocate to this file system is based on the size of your application group tables and the number of tables that you must import to satisfy a query for migrated data.

It is recommended that you have at least 500 MB of free space available in this file system at all times. If your storage configuration permits or the size of your database tables dictates, you might need to allocate 1 GB or more free space to this file system. For example, based on the estimate for the sample reports and making some assumptions about how the data is stored in Content Manager OnDemand, you need approximately 500 MB of space to import one application group table. If you need to import two application group tables to satisfy a query, then the import program requires at least 1 GB of temporary disk space.

Estimating archive storage requirements

Report storage space

When you estimate the amount of space required to store a report in archive storage, you must consider the size of the report, the compression ratio achieved, and the length of time that the archive storage manager maintains the report. archive storage can be optical storage or magnetic tape. Use the following calculation to estimate that amount of space required:

```
ArchiveStorageSpace = ( Data per month * life of data in months )
                        * compression ratio
                            * 1.1
```

Figure 22. Calculating archive storage space

For example, if you plan to store 8 GB of report data per month, the archive storage manager must maintain the data for seven years, and Content Manager OnDemand can achieve a compression ratio of 3:1 (0.33), you would require approximately 244 GB of archive storage space:

```
OpticalSpace  =  ( 8 GB  *  84 )
                   *  0.33
                     *  1.1  =  244 GB
```

Backup report storage space

The Content Manager OnDemand system can maintain a backup (second) copy of reports that you store on archive storage.

You typically maintain multiple copies of reports that are critical to the operation of your company or difficult or impossible to recreate.

The method that Content Manager OnDemand uses to maintain the backup copy depends on the archive storage manager that you use. For example, with Tivoli Storage Manager, you can configure a copy storage pool. With this method, Tivoli Storage Manager maintains a backup copy of data stored in a primary storage pool independently and transparently to Content Manager OnDemand. Tivoli Storage Manager automatically retrieves the backup copy if the primary copy becomes damaged, lost, or unusable.

If you need Content Manager OnDemand to maintain a backup copy of your reports, double the archive storage space that you previously calculated.

Storage for database backup images

If you do not plan to use Tivoli Storage Manager to maintain DB2 backup image files, you do not need to allocate space for the backup image files on archive storage.

The storage pool where Tivoli Storage Manager maintains DB2 backup image files must contain enough storage to hold all of the backup image files needed to recover your database:

- The number of backup image files that Tivoli Storage Manager maintains depends on the type of database backups taken and how often that you take backup images.
- The storage required to hold the backup image files also depends on the size of the database and the table spaces contained in the database.
- If you migrate application group data to table spaces, then Tivoli Storage Manager must maintain a backup image for each table that you migrate.
- Tivoli Storage Manager can maintain multiple copies of each backup image. For example, for added protection, you might want Tivoli Storage Manager to maintain two copies of each backup image.

Content Manager OnDemand supports full database backups and incremental table space backups. To recover a database using incremental table space backups, you must create and maintain at least one full database backup image (taken before any changes are made to the database and prior to the first incremental table space backup).

You must configure Tivoli Storage Manager to maintain a backup image as long as it is needed. For example, if you plan to create a full backup image of the database every week, it is recommended that you configure Tivoli Storage Manager to maintain two versions of the backup image and two copies of each version. If you need to recover the database, you would always start with the latest version of the backup image. The latest version should be no more than one week old. If, for some reason, either copy of the latest version could not be used, you could use the prior version, which should be no more than two weeks old.

You might need to regularly initialize and load scratch storage volumes into the storage library where Tivoli Storage Manager maintains the database backup images. If Tivoli Storage Manager determines that there is not enough space available in the storage pool, it can request a mount for a scratch storage volume. However, the backup command cannot complete until the mount request is satisfied. If you operate in an unattended environment, this could have an adverse affect on system availability, especially when running an offline backup.

Depending on the variables listed above, you might need to maintain hundreds of backup image files in storage that is managed by Tivoli Storage Manager. It is recommended that you contact Tivoli Storage Manager and database specialists to help plan your storage requirements. The following storage calculations might not accurately estimate the amount of storage that you need for the backup image files required by your system.

Storage space for a full database backup image
Use the following calculation to estimate the archive storage space required to maintain full backup images of the database.

The calculation uses the maximum size of the database, to allocate enough storage space to hold the largest backup image file required to recover the database. The compression ratio is the compression that Tivoli Storage Manager can achieve on the backup image files.

```
DB2BackupImageSpace = ( MaxDBSize * compression ratio )
                        * CopiesMaintained
                          * VersionsMaintained
```

For example, if the maximum size of the database is 5.6 GB and you need Tivoli Storage Manager to maintain two versions of the backup image and two copies of each version, then the archive storage required to hold the backup image files is:

```
DB2BackupImageSpace =  (5.6 GB  *  .33)
                         *  2
                           *  2   =  7392000000
```

Content Manager OnDemand requires approximately 7.4 GB of archive storage space to hold the backup image files.

Storage space for table space backup images
The calculation uses the maximum size of the table space, to allocate enough storage space to hold the largest backup image file required to recover the table space.

Use the following calculation to estimate the archive storage space required to maintain backup images of a table space.

```
TSBackupImageSpace = ( MaxTSSize * compression ratio )
                       * CopiesMaintained
                         * VersionsMaintained
```

For example, if the maximum size of the table space is 560 MB and you want Tivoli Storage Manager to maintain two versions of the backup image and two copies of each version, then the optical storage required to hold the backup image files is:

```
TSBackupImageSpace =  ( 560 MB  *  .33 )
                         *  2
                            *  2  =  740 MB
```

Content Manager OnDemand requires approximately 740 MB of optical storage to hold the backup image files.

Database archived log storage

If you do not plan to use Tivoli Storage Manager to maintain DB2 archived log files, you do not need to allocate space for the archived log files on archive storage. Allocate disk storage space for the archived log files instead.

The storage pool where Tivoli Storage Manager maintains the DB2 archived log files must contain enough storage to hold the log files that are needed to recover the database. There are many factors that you should consider when estimating the storage space needed to hold them:

- How often do you load reports into the system?
- How often do you add to or update the Content Manager OnDemand system tables (for users, groups, system printers, storage sets, application groups, applications, and folders? In addition, the System Log

tables usually gets updated every time someone logs on or off the system, data is stored, queried, retrieved, and printed, and so forth.

- Do you store application group data in table spaces?
- What is the size of the database; the table spaces?
- What is the frequency and type of database backups taken?
- How long do you need to keep archived log files?
- What is the compression ratio that the archive storage manager can achieve on archived log files?

If you take full backup images of the database on a regular schedule, such as once a day or once a week, It is recommended that you allocate two times the space that you have estimated for the active log files. However, you must allocate enough space to hold all of the archived log files created between full backups of the database. After a full database backup image is created, archived log files created prior to the backup are no longer needed and can be deleted. The following calculation can be used to estimate the amount of archive storage needed to hold archived log files:

```
DB2ArchiveLogSpace = ( 2 * ActiveLogSpace )
                           * compression ratio
```

Figure 23. Calculating database archived log file storage

The following example illustrates the archive storage required to hold archived log files, when the space allocated for the active log files is 516 MB:

```
DB2ArchiveLogSpace =  ( 2 * 516 MB )
                            *  .33  =  340 MB
```

Content Manager OnDemand requires approximately 340 MB of archive storage to hold the archived log files.

If you do not take full backup images of the database, it is recommended that you keep the archived log files indefinitely. Accordingly, you must carefully estimate the amount of archive storage that you will need. For example, a single archived log file requires approximately 1.3 MB of (uncompressed) storage space. Depending on the variables that listed above, you might need to maintain hundreds of archived log files in storage that is managed by Tivoli Storage Manager. It is recommended that you contact Tivoli Storage Manager and database specialists to help plan your storage requirements.

You might need to regularly initialize and load scratch storage volumes into the storage library where Tivoli Storage Manager maintains the archived log files. If Tivoli Storage Manager determines that there is not enough space available in the storage pool, it can request a mount for a scratch storage volume. However, the backup command cannot complete until the mount request is satisfied. If you operate in an unattended environment, this could have an adverse affect on system availability, especially when running an offline backup.

Migrated index storage space

Content Manager OnDemand supports automatic migration of indexes from the database to archive storage so that you can maintain seldom used indexes for long periods of time. However, migration of indexes should be done only after there is no longer a need to retrieve the reports to which they point. For example, suppose that all of the queries for a report occur in the first 24 months after the report is loaded into the system. After that time, there are almost no queries for the report. The indexes could be eligible to be migrated from the database to archive storage. Migration of index data is optional; you can choose to migrate indexes for all, some, or none of the application groups on your system. In addition, you determine the length of time that indexes stay in the database before Content Manager OnDemand migrates them to archive storage.

You can use the following calculation to determine the archive storage space required to hold migrated indexes:

```
ArchiveMediaDBSpace = ( Database size per month * compression ratio )
                    * ( life of data - months before migrating data )
```

Figure 24. Calculating migrated index storage space

For example, if the index data requires 234 MB of space in the database per month, you need to maintain the indexes for 84 months, and the indexes remain in the database for 24 months before being migrated, then the archive storage required to hold the migrated indexes is:

```
ArchiveMediaDBSpace  =  ( 234 MB  * .33 )
                     *  ( 84   -  24 )  =  4633200000
```

Content Manager OnDemand requires approximately 4.6 GB of archive storage to hold the migrated indexes.

Storage volumes and libraries

Estimating the amount of archive storage required to hold your reports also helps you determine the number of archive storage volumes that you need to plan for.

In previous examples, approximately 244 GB of archive storage space is required. Assuming that the formatted capacity of a 5.25 inch optical storage volume is about 30 GB , you would need approximately 8 storage volumes to hold the data for 10 years.

Depending on the operational and management requirements of your organization, you might need to plan for additional storage volumes and storage libraries. For example:

- If you use Tivoli Storage Manager to maintain DB2 backup image files and archived log files, it is recommended that you dedicate a library for that purpose.
- If you need to maintain a backup copy of the reports that you store on archive storage, and the archive storage manager that you are using supports it, it is recommended that you store the backup copy in a different library than the primary copy.

It is possible to reduce the number of storage libraries by removing storage volumes from a library and placing them in offline storage. For example, you might find that you can remove a storage volume from a library one year after the last time that Content Manager OnDemand stored data on or retrieved data from the storage volume. The archive storage manager should provide commands that you can use to determine when a storage volume was last written to or read from and to dismount a storage volume from a library. Before report data can be retrieved from an offline storage volume, an operator must usually locate the storage volume and mount it in the library.

You can also reduce the number of storage libraries by storing different types of reports in the same library. The archive storage manager usually uses one management policy to maintain all data stored in a library. The management policy determines the length of time that the archive storage manager maintains data in the library.

Storage sizing examples

You can use examples to estimate storage requirements for two types of reports.

The following examples illustrate how to estimate storage requirements for two types of reports:

- Report that contains logical items, such as statements or policies
- Report that contains sorted transaction data

Each example contains four parts:

- Database Columns
- Report Profile
- Disk Storage Space
- Archive Storage Space

Report that contains logical items

You can create reports that contain logical items.

Table 4. Report that contains logical items. Part 1 of 4. Database Columns.

Column Number	Name	Index or Filter	Bytes
1	Account Number	Index	12
2	Report Date	Filter	4
3	Customer Name	Filter	24
4	Balance	Filter	8

Table 5. Report that contains logical items. Part 2 of 4. Report Profile.

Report Characteristic	Report Estimate
Volume of data per month (bytes)	8,000,000,000
Average statement size (bytes)	8,000
Number of statements per month	1,000,000
Number of cycles per month	20
Largest cycle data size (bytes)	400,000,000
Largest single report file size (bytes)	400,000,000
Largest cycle (number of statements)	50,000
Number of database columns from Table Table 4 on page 51	4
Life of Data (days)	2555
Number of days to cache data	90
Number of days to keep index in database	730
Compression percentage (ratio)	0.33 (3:1)

Table 6. Report that contains logical items. Part 3 of 4. Disk Storage Requirements in Bytes.

Storage Component	Storage Requirement
Base system storage	2,000,000,000
Data download	480,000,000
Indexing	600,000,000
Cache storage	8,712,000,000
Content Manager OnDemand database	5,616,000,000
Database logs	516,000,000
Archive storage manager database logs	47,040,000
Server print	500,000,000
Imported migrated index data	500,000,000
Total Disk Storage Required (Bytes)	18,971,040,000

Table 7. Report that contains logical items. Part 4 of 4. Archive Storage Requirements in Bytes

Storage Component	Storage Space Requirement
Report data	244,000,000,000
Migrated index data	4,633,200,000
DB2 log files	0
DB2 backup image files	0
Total Archive Storage Required (Bytes)	248,633,200,000

In this example, the DB2 archived log files and backup image files are maintained on disk (or tape) independently of the archive storage manager.

Report that contains transaction data

Various reports can contain transaction data.

Table 8. Report that contains transaction data. Part 1 of 4. Database Columns.

Column Number	Name	Index or Filter	Bytes
1	Beginning Invoice Number	Index	10
2	Ending Invoice Number	Index	10
3	Report Date	Filter	4
4	Page Number	Filter	4

Table 9. Report that contains transaction data. Part 2 of 4. Report Profile.

Report Characteristic	Report Estimate
Volume of data per month (bytes)	5,000,000,000
Average page size (bytes)	5,000
Number of pages per month	1,000,000
Number of cycles per month	20
Largest cycle (data size in bytes)	200,000,000
Largest cycle (number of pages)	50,000
Largest single report file size (bytes)	200,000,000
Group of indexed pages	100
Number database columns from Table Table 8 on page 52	4
Life of data (days)	730
Number of days to cache data	0
Number of days to keep indexes in database	730
Compression percentage (ratio)	0.25 (4:1)

Table 10. Report that contains transaction data. Part 3 of 4. Disk Storage Requirements in Bytes.

Storage Component	Storage Space Requirement
Base system storage	2,000,000,000
Data download	240,000,000
Indexing	300,000,000
Cache storage	0
Content Manager OnDemand database	56,640,000
Database logs	172,000,000
Archive storage manager database logs	8,400,000
Server print	500,000,000
Imported migrated indexes	0
Total Disk Storage Required (Bytes)	3,277,040,000

Table 11. Report that contains transaction data. Part 4 of 4. Archive Storage Requirements in Bytes.

Storage Component	Storage Space Requirement
Report data	33,000,000,000
Migrated index data	0
DB2 log files	0
DB2 backup image files	0
Total Archive Storage Required (Bytes)	33,000,000,000

- Report data is not stored in cache storage.
- Database log space for the largest report requires approximately 512 KB. The example uses the default value provided for active log space (172 MB), which should be more than enough to hold not only the active logs, but also the archived logs.
- Index data is not migrated to archive storage.
- DB2 archived log files and backup image files are maintained on disk (or tape) independently of the archive storage manager.

Storage sizing worksheets

You can estimate the storage requirements for two types of reports.

The following worksheets can help you estimate the storage requirements for two types of reports:

- Report that contains logical items, such as statements or policies
- Report that contains sorted transaction data

Each worksheets contains four parts:

- Database Columns
- Report Profile
- Disk Storage Space
- Archive Storage Space

Make a copy of the worksheets on the following pages for each report that you want to store in Content Manager OnDemand. Complete the worksheets to calculate the storage requirements for the report.

Report that contains logical items

You can create a report that contains logical items such as database columns, report profiles, and disk and archive storage requirements.

Table 12. Report that contains logical items. Part 1 of 4. Database Columns.

Column Number	Name	Index or Filter	Bytes
1			
2			
3			
4			

Table 13. Report that contains logical items. Part 2 of 4. Report Profile.

Report Characteristic	Report Estimate
Volume of data per month (bytes)	
Average item size (bytes)	
Number of items per month	
Number of cycles per month	
Largest cycle data size (bytes)	
Largest single report file size (bytes)	
Largest cycle (number of items)	
Number of database columns from Table Table 12 on page 54	
Life of data in days	
Number of days to cache data	
Number of days to keep indexes in database (default is Life of Data)	
Compression ratio: image, PDF 1; AFP 8:1 (0.13); Line data 20:1 (0:05)	

Table 14. Report that contains logical items. Part 3 of 4. Disk Storage Requirements in Bytes.

Storage Component	Storage Space Requirement
Base system storage **Note:** The base system software requirement of 2 GB is per server.	2,000,000,000
Data download	
Indexing	
Cache storage	
Content Manager OnDemand database	

Table 14. Report that contains logical items. Part 3 of 4. Disk Storage Requirements in Bytes. (continued)

Storage Component	Storage Space Requirement
Database logs	
Archive storage manager database logs	
Server print	
Imported migrated indexes	
Total Disk Storage Required (Bytes)	

Table 15. Report that contains logical items. Part 4 of 4. Archive Storage Requirements in Bytes

Storage Component	Storage Space Requirement
Report data	
Migrated index data	
DB2 log files	
DB2 backup image files	
Total Archive Storage Required (Bytes)	

Report that contains transaction data

Table 16. Report that contains transaction data. Part 1 of 4. Database Columns.

Column Number	Name	Index or Filter	Bytes
1			
2			
3			
4			

Table 17. Report that contains transaction data. Part 2 of 4. Report Profile.

Report Characteristic	Report Estimate
Volume of data per month (bytes)	
Average page size (bytes)	
Number of pages per month	
Number of cycles per month	
Size of largest cycle (bytes)	
Size of largest cycle in pages	
Size of largest single report file (bytes)	
Pages in an indexed group (default is 100)	
Number of database columns from Table Table 16 on page 55	
Life of data in days	
Number of days to cache data	

Table 17. Report that contains transaction data. Part 2 of 4. Report Profile. (continued)	
Report Characteristic	**Report Estimate**
Number of days to keep indexes in database (default is Life of Data)	
Compression ratio: image, PDF 1; AFP 8:1 (0.13); Line data 20:1 (0:05)	
Index on OS/390® or Content Manager OnDemand server	

Table 18. Report that contains transaction data. Part 3 of 4. Disk Storage Requirements in Bytes.	
Storage Component	**Storage Space Requirement**
Base system storage **Note:** The base system software requirement of 2 GB is per server	2,000,000,000
Data download	
Indexing	
Cache storage	
Content Manager OnDemand database	
Database logs	
Archive storage manager database, logs	
Server print	
Imported migrated indexes	
Total Disk Storage Required (Bytes)	

Table 19. Report that contains transaction data. Part 4 of 4. Archive Storage Requirements in Bytes.	
Storage Component	**Storage Space Requirement**
Report data	
Migrated index data	
DB2 log files	
DB2 backup image files	
Total Archive Storage Required (Bytes)	

Disk storage

Overview

Before you begin defining reports to Content Manager OnDemand and loading them on the system, it is important that you estimate the storage required to hold your reports.

Depending on the types of reports that you plan to store in the system, and their number, size, and other storage requirements, you might need to add many disk storage devices to the system.

After you know how much storage is needed for your reports, you can begin to plan the number and size of the disk storage devices that you need. If your server needs to hold lots of report data in the database and on cache storage, you might need to configure your disk storage devices into groups of volumes. For example, you might have a group of storage volumes dedicated to the database, a group dedicated to cache storage, and so forth. Configuring storage devices in this way allows you to manage them as your storage needs grow and configure them for high availability and maximum performance.

The type of disk storage you choose is less important than the speed of the device. A fast disk storage device enables you to maintain a healthy Content Manager OnDemand system. For information about disk storage devices available from IBM, see www.storage.ibm.com.

For details about configuring and managing storage devices, see your operating system documentation. If you are not familiar with configuring and managing storage devices, review the information in your operating system documentation before you continue.

Disk storage devices on a UNIX® server

You should organize your disk storage devices into groups of storage volumes.

The examples that follow show one way to configure disk storage devices on a server. The examples assume that disk storage is needed for the various Content Manager OnDemand software programs, for data transmitted from z/OS systems, to index data on the server, for the database and database log files, for the archive storage manager database, and for cache storage.

The number of storage devices that you can put in a group and the number of groups that you can define to the system will depend on the operating system. For example, in AIX 5L Version 5.3, you can have up to 1024 volume groups on the system.

Regardless of the number of disk storage devices that you have on the server or the capacity of the devices, organize the available disk storage as described in the examples that follow. You should adopt the suggested convention for naming the groups, file systems, directories, and files.

If you are configuring a server for a large organization or a server on which you plan to load and maintain a large amount of data on disk, organize your disk storage devices into the following groups:

- Software programs, control files, resources, and temporary storage
- Data transmitted from other systems and indexing data on the server
- Content Manager OnDemand database and database log files
- Archive storage manager database and recovery logs
- Cache storage

Table 20. Disk storage groups for a large organization

Volume Group	Logical Volume	File System	Physical Volumes
acifvg	aciflv1	/arsacif/acif1	hdisk1
acifvg	aciflv2	/arsacif/acif2	hdisk2
cachevg	cachelv1	/arscache/cache1	hdisk3, hdisk4
cachevg	cachelv2	/arscache/cache2	hdisk5, hdisk6
cachevg	cachelv3	/arscache/cache3	hdisk7, hdisk8
db2vg	dblv	/arsdb	hdisk9, hdisk10
db2vg	primloglv	/arsdb_primarylog	hdisk11
db2vg	archloglv	/arsdb_archivelog	hdisk12
rootvg	arstmplv	/arstmp	hdisk13
adsmvg	dsmdblv	none	hdisk14

Table 20. Disk storage groups for a large organization (continued)

Volume Group	Logical Volume	File System	Physical Volumes
adsmvg	dsmloglv	none	hdisk15

If you are configuring a server for a small workgroup or a server on which you plan to load and maintain very little data on disk, organize your disk storage devices into two groups:

- Software programs, control files, resources, and temporary storage
- Data transmitted from other systems, indexing data on the server, the Content Manager OnDemand database and database log files, and cache storage

Table 21. Disk storage groups for a small workgroup

Volume Group	Logical Volume	File System	Physical Volumes
arsvg	aciflv1	/arsacif/acif1	hdisk1
arsvg	cachelv1	/arscache/cache1	hdisk2, hdisk3
arsvg	dblv	/arsdb	hdisk4
arsvg	primloglv	/arsdb_primarylog	hdisk1
arsvg	archloglv	/arsdb_archivelog	hdisk1

Disk storage devices on a Windows® server

You should plan to use at least ten disk drives. Use your sizing calculations to determine the size of these disk drives.

The examples that follow show one way to configure disk storage devices on a server. The examples assume that disk storage is needed for the various Content Manager OnDemand software programs, for data transmitted from z/OS systems, to index data on the server, for the database and database log files, for the archive storage manager database, and for cache storage.

If you are configuring a server for a large organization or a server on which you plan to load and maintain a large amount of data on disk, organize your disk storage devices into the following groups:

- System software and files
- Content Manager OnDemand software
- Data transmitted from other systems and indexing data on the server
- Content Manager OnDemand database and database log files
- Cache storage

You might need more drives depending on the size of the database and how much data you need to maintain in cache storage (and how long you need to maintain the data).

If you are configuring a server for a small workgroup or a server on which you plan to load and maintain very little data on disk, organize your disk storage devices into the following groups:

- Software programs
- Content Manager OnDemand database and database log files
- Cache storage
- Data transmitted from other systems, indexing data on the server, temporary storage

Data storage and protection

Overview

This section provides information about RAID storage subsystems and the IBM Enterprise Storage Server®.

RAID stands for Redundant Array of Inexpensive Disks and provides a method of classifying the different methods of using multiple disks to increase availability. With RAID, multiple physical disks appear to the Content Manager OnDemand server as one logical disk. RAID carries out the concept of data striping by spreading data over multiple disks; a single file is segmented and stored on multiple disks. RAID carries out the concept of data mirroring by duplicating data from one disk to a second disk; a single file is stored twice, on two different disks. A failed disk still allows users to access data on the array, and a replacement disk or online spare can be recreated while the array is in use.

Table 22. RAID implementations

RAID Level	Description	Protection	Performance
RAID 0	Data striping on multiple disk drives.	Poor; single disk failure.	Best; read and write requests can be met by any disk.
RAID 1	Disk mirroring.	Good; any disk can fail and data is still accessible.	Good; read request can be met by any disk.
RAID 3	Disk striping with parity disk, using interleaved bytes.	Good; if any disk fails, data can be accessed by using information from other disks and parity disk.	Good for large data transfers.
RAID 4	Disk striping with parity disk, using interleaved sectors.	Good; if any disk fails, data can be accessed by using information from other disks and parity disk.	Good for large data transfers.
RAID 5	Disk striping with distributed parity data.	Good; if any disk fails, data can be accessed by using information from other disks and parity information.	Good for small block sizes.
RAID 5 Orthog – onal	Disk striping with distributed parity data, using dual controllers.	Best; if any disk fails, data can be accessed by using information from other disks and parity information, with additional protection from any single disk controller failure.	Good for small block sizes; improved performance because of use of dual controllers to read and write data.
RAID 6	An extension of RAID 5 with a second portion of parity data to ensure disk recovery in the event of a second disk failure before the first drive can be fixed.	Provides independent access to the disks with floating parity to avoid the separate parity disk bottleneck of RAID 4. Multiple concurrent accesses to the array devices are supported to satisfy multiple concurrent I/O requests.	Data is striped in blocks across disk drives. Single records or tracks can be read from one disk drive without accessing other drives in the array.

Important: For most Content Manager OnDemand systems, use the RAID 1 or RAID 5 implementations.

RAID for cache storage

A typical use of RAID storage on a Content Manager OnDemand server is for cache storage.

Orthogonal RAID 5 (redundant disk controllers) provides excellent protection from a single disk or controller failure. Disk striping with distributed parity data allows the Content Manager OnDemand server to remain available if a single disk fails. Redundant disk controllers provide excellent availability, enabling users to continue to access data if a controller fails. Customers who do not maintain copies of reports in archive storage should use RAID storage. Without a backup copy of a report in archive storage or an up-

to-date backup image of the cache storage file systems, the system is exposed to loss of data that are difficult or impossible to recreate.

RAID for the database

You can store the Content Manager OnDemand database in one or more RAID storage subsystems.

Depending on the exact hardware and the implementation level of the RAID devices, the system might not achieve the same level of database performance as a library server using non-arrayed disk storage. That is, when you load reports on the system or when lots of users query the database at the same time, a system with the database on arrayed storage might not attain the same level of performance as a system with the database on non-arrayed storage. (For example, RAID 1 can improve read performance by using both copies; RAID 5 can improve multiple, short data transfers by distributing them to multiple disks.)

The availability benefits provided by RAID storage subsystems typically outweigh any performance degradation that the users might experience. See your database and storage system specialists for help with configuring and using RAID storage devices.

Reports

Most customers configure the system to copy reports to cache storage and archive storage at the same time, when they load a report into the system.

Content Manager OnDemand can store copies of reports in cache storage and archive storage:

- The primary purpose of cache storage is short-term, high-speed storage and retrieval of reports. Cache storage consists of disk storage volumes maintained by Content Manager OnDemand on one or more object servers.
- The primary purpose of archive storage is long-term storage and retrieval of reports. Reports in archive storage can also be used as backup copies, in the event that cache storage becomes corrupted or unavailable. Archive storage consists of optical or tape storage volumes managed by the archive storage manager, such as Tivoli Storage Manager.

Content Manager OnDemand can retrieve a copy of a report from archive storage after the report has been removed from cache storage or if the copy on cache storage is unavailable. However, you must configure the system to support multiple copies of reports. You must install and configure an archive storage manager, define devices to the archive storage manager, and configure Content Manager OnDemand to use archive storage. You configure Content Manager OnDemand to use archive storage by defining storage sets with storage nodes that are registered with the archive storage manager, assigning application groups to the storage sets, and configuring data migration and caching information in application groups.

If you do not plan to copy reports to archive storage, then it is recommended that you take regular backups of the file systems that comprise cache storage. If a media failure occurs or cache storage becomes corrupted, users cannot retrieve reports until the file systems are restored.

Cache storage

Cache storage is the primary, short-term storage location for reports.

If you do not copy reports to archive storage when you store them in Content Manager OnDemand, then you need to consider how you can recover the reports in the event that you need to do so (for example, if a device fails).

Cache storage can be protected by maintaining it on high-availability storage devices. If no high-availability storage is available, it is recommended that backups of reports in cache storage (the file systems) be taken on a regular schedule.

Archive storage

The Content Manager OnDemand storage node identifies the object server and the client node in archive storage where the primary copy of a report is maintained.

Content Manager OnDemand retrieves the primary copy of the report from archive storage after the report has been removed from cache storage. Customers with special business, legal, or performance reasons

might want the system to maintain a backup copy of their reports in archive storage. The backup copy can be used if the primary copy becomes corrupted or unavailable.

You must configure the archive storage manager to maintain a backup copy of reports in archive storage. For example, with Tivoli Storage Manager, you would define a copy storage pool. With a copy storage pool, Tivoli Storage Manager manages a backup copy of files that are stored in a primary storage pool independently and transparently to Content Manager OnDemand. The backup copy is stored in a copy storage pool that can be used to restore the original files if they become damaged, lost, or unusable. The copy storage pool can be assigned to the same library as the primary storage pool. However, you would typically assign the copy storage pool to a different library. You can copy data from one or more primary storage pools to the same copy storage pool. Copy storage pools require additional space in the Tivoli Storage Manager database. A copy storage pool must reside on the object server where the primary storage pool resides. Tivoli Storage Manager includes a central scheduling component that allows the automatic processing of administrative commands, such as copying data from a primary storage pool to a copy storage pool. Each scheduled event is tracked by the server and recorded in the database. You can set up an administrative command schedule by defining schedule parameters, such as the start day, date, and time, specifying the command to be executed, such as the **BACKUP STGPOOL** command, and activating the schedule.

See your archive storage manager information for details about defining and managing multiple copies of reports, backup and recovery of data, and scheduling operations.

Chapter 8. Planning information

This section is a planning source for Content Manager OnDemand administrators. Other people in an organization interested in this section might include technical and service support personnel, database administrators, network administrators, application administrators, and anyone else who has responsibility for making decisions about business systems, such as people responsible for physical site planning, operations, and backup and recovery.

This part describes activities that it is recommended that Content Manager OnDemand administrators perform to plan for the installation of Content Manager OnDemand and prepare for the operation of Content Manager OnDemand.

Reports and other data

This section contains information that can help you plan for the reports that you will be storing into Content Manager OnDemand. You can use the information to help determine the hardware configuration that you need to support your Content Manager OnDemand system. This section contains a list of questions that you might ask users of the reports, provide information about the types of data that you can store in Content Manager OnDemand, and provide information about indexing reports.

Collecting requirements

Planning for Content Manager OnDemand requires that you understand how the system is deployed, who uses the system and how, and other end-user requirements.

Answers to these questions provide information that allows you to properly configure your Content Manager OnDemand system, including the storage and network configuration, to support your applications and users:

- Do you operate a single Content Manager OnDemand server or a network of Content Manager OnDemand servers?
- What is the logical organization of the print data streams?
  - Page organization: a consistent stream of pages of transaction or ledger data.
  - Logical groups of information, such as statements or policies.
  - Data that might not have a consistent format, such as reference materials or product literature.
- Does Content Manager OnDemand support short-term report management, long-term archival storage, or both?
- What is the volume of input to process? How large are your reports (in pages and bytes); how many reports; how many versions of reports?
- What index values do the users of a report need to retrieve a specific version of a report (or a document)?
- How much time is available to load reports into Content Manager OnDemand? Daily? Weekly?
- How long do you plan to maintain report data on the system?
- How many concurrent, logged-on users do you anticipate on average; at peak times?
- How many active users do you anticipate?
- What is the transaction rate of the active users?

Input data formats

Content Manager OnDemand supports several types of input data:

- AFP print data streams (AFP or MO:DCA-P), including line data mixed with AFP structured fields and line data formatted with a page definition.
- Line data, also known as IBM S/390® line data with ANSI or machine carriage control characters.
- Unformatted ASCII data that is typically generated in the workstation environment.
- Adobe Portable Data Format (PDF) files (**Note:** An input file cannot exceed 4 GB in size).
- Image files in the following formats:
  - BMP (Bitmap). A file that contains a bit-mapped graphic.
  - GIF (Graphic Interchange Format). A bit-mapped color graphics file format for IBM-compatible computers. GIF uses an efficient compression technique for high resolution graphics.
  - JFIF (JPEG Format Image File). A file that contains image data compressed using the JPEG (Joint Photographic Experts Group) standard.
  - PCX (Picture Exchange Format). A file that contains a graphic in the PCX file format, widely used by PC applications, such as the PC Paintbrush program. Compressed using PackBytes compression.
  - TIFF (Tagged Image File Format). A bit-mapped graphics image format for scanned images with resolutions up to 300 DPI. TIFF simulates gray-scale shading. Content Manager OnDemand supports single and multipage TIFF files that are uncompressed or are compressed using JPEG, CCITT Group 3, CCITT Group 3 / 2D, and CCITT Group 4 compression.
- Extensible Markup Language (XML) files.

In addition to the types of data listed above, Content Manager OnDemand allows you to store almost any other type of data on the system. For example, you can define an application for HTML documents. When you define the application, you must identify the file type of the data. The file type determines the program that the client starts when the user retrieves a document. For example, if the file type is HTM, then the client could start a web browser to view the document.

In the z/OS environment, Content Manager OnDemand allows application programs that produce 1403 or 3211 data stream formats to take advantage of overlays, page segments, and typographic fonts. This is done using a page definition that specifies how data is mapped on the page. The definition allows text to be moved to different positions on the page, fonts to be changed, and conditional processing. When combined with a form definition, the page definition allows sophisticated pages to be produced by existing line data applications without changing the application that generates the data.

You can use ACIF to convert line data to AFP data before loading it into the system. The resulting AFP data could add color or an electronic form to line data, making presentation of the information more effective. However, archiving line data without conversion usually results in much higher compression ratios.

AFP supports graphics, presentation text, image, and bar code objects. Storing AFP data on the system allows full-fidelity viewing of presentation text and image objects. For example, users can retrieve and view customer statements that Content Manager OnDemand presents using an electronic form, fonts, and images. The user views a copy of the statement that appears the same as the statement the customer received in the mail. AFP also supports navigation within a report file, using a table of contents.

When you load reports that contain AFP data, you must also load the resources into Content Manager OnDemand. The resources include overlays, page segments, form definitions, and fonts. The resources must be resident on the processor where the data is to be indexed. If data will be indexed on the z/OS system, then the indexing program must gather the resources into a resource group so that the resource group can be transferred to the Content Manager OnDemand server on which you plan to load the data. If data will be indexed on an Content Manager OnDemand server, then the resources must be resident on the Content Manager OnDemand server (or be accessible from the Content Manager OnDemand server) on which you plan to index and load the data.

Indexing data

One of the main operations that you do with Content Manager OnDemand is to index reports.

When you index a report, Content Manager OnDemand extracts index values from the report and stores them in the database. The database fields that you define for your application groups hold the index

values. When a user opens a folder, Content Manager OnDemand displays a list of search fields, which represent the database fields. To perform a query, the user enters values in the search fields. Content Manager OnDemand compares the values from the search values with the values in the database fields and retrieves the items that match the query.

Index information can be added to reports at the same time that the application program generates the print data or, more typically, the output print data can be processed by one of the indexing programs that are supported by Content Manager OnDemand.

When you index a report, you can divide a large report into smaller, uniquely identifiable units of information. For example, when an application program generates customer bills, it might produce a large print stream made up of thousands of individual customer bills. With Content Manager OnDemand, you can identify the individual customer bills within the report as smaller, separate information units, or logical items (known as documents in Content Manager OnDemand). Your users can search for and retrieve the logical items using identifiers such as account number, customer name, and date.

Content Manager OnDemand supports two basic methods of indexing:

- Document Indexing. For reports made up of logical items, such as statements, bills, policies, and invoices.
- Report Indexing. For reports that (typically) contain line data, with sorted values on each page, such as a transaction log or general ledger.

If a report does not contain logical items or sorted line data, it can usually be indexed by using the report indexing method.

See the *IBM Content Manager OnDemand Indexing Reference* for details about and examples on using the indexing programs that are provided with Content Manager OnDemand.

Document indexing

Document indexing can be used to index reports that are made up of logical items or to index reports that contain unique values such as an account number or a customer name.

When searching and retrieving these types of reports, Content Manager OnDemand returns a list of the items that match the user's query and transfers the individual items to the Content Manager OnDemand client program for viewing and printing. Content Manager OnDemand supports up to 128 fields as indexes or filters for document-type data. The fields do not have to be sorted and can contain numeric or text information. The fields are stored in the database as indexes or filters. The following diagram shows an example of a report file and document indexing.

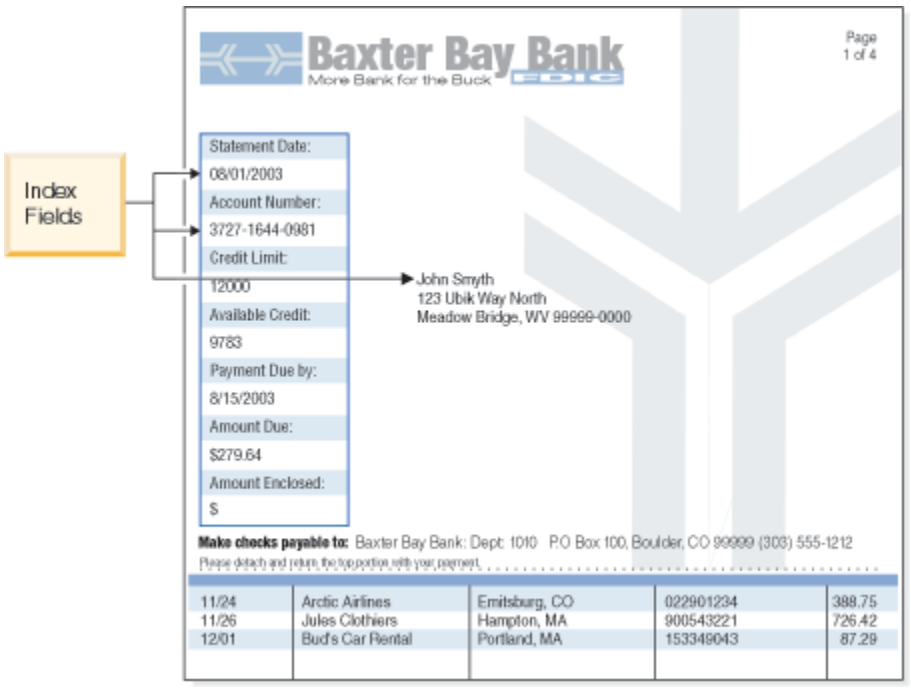

Figure 25. Document indexing method

Report indexing

Report indexing allows users to search sorted report data and retrieve the first occurrence of the value that they specified in the query.

Content Manager OnDemand divides the report data into groups of pages and stores the first and last index values contained in each group of pages in the database. When the user enters a query, Content Manager OnDemand returns a list of the items that match the query. When the user selects an item for viewing, Content Manager OnDemand performs a text search within the item for the value specified by the user. The Content Manager OnDemand client program displays the first page that contains the value specified by the user. Content Manager OnDemand uses a single, unique sorted index value for the retrieval of the report data, for example, an invoice number or a transaction identifier. The following diagram is an example of a report file and report indexing.

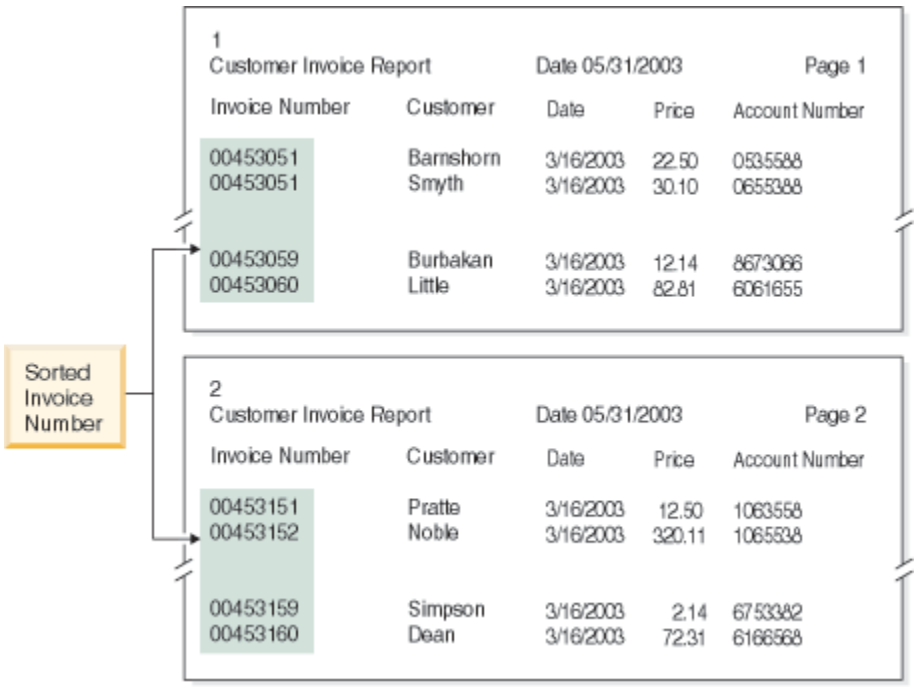

Figure 26. Report indexing method

Indexing data with ACIF

ACIF is a powerful tool for indexing the print data streams of z/OS application programs. ACIF indexes reports based on the organization of the data in the report.

You can optionally convert line data print streams into AFP data. ACIF processes three input sources:

- Indexing parameters that specify how the data should be indexed. You can create the indexing parameters when you define an Content Manager OnDemand application.
- AFP resources required to view and print the data, if the data was created by an AFP application.
- The print data stream.

The output of ACIF is either a fully composed AFP data stream or the original line data input. ACIF can convert line data input to AFP data, can produce an index file that Content Manager OnDemand uses to create index data for the database, and optionally, can collect resources into a resource group file.

ACIF produces a resource group file for AFP data. To create a resource group file, ACIF must have access to the resources required by the input data stream. Content Manager OnDemand usually stores the resources in cache storage and retrieves the resources associated with a specific document when a user selects the document for viewing.

ACIF can logically divide reports into individual items, such as statements, policies, and bills. You can define up to 128 index fields for each item in a report.

ACIF is designed to index reports that contain line data with a consistent structure and format. You can also use ACIF to index AFP input files that contain indexing controls and information. The following topics provide additional information:

- Generating index data in application programs
- Generating index data with AFP application programs
- Generating index data with the AFP API
- Inserting AFP records into a data stream

Generate index data in application programs
As an alternative to using ACIF to index reports, you can create index information in the application program that generates the report.

Some application programs already provide support to add indexing information. You might find it necessary to modify your application program to add indexing functions.

Generate index data with AFP™ application programs
The IBM Document Composition Facility (DCF) is a product that can be used to create indexed AFP data.

The primary function of DCF is to prepare and format documents for printing. Along with its many other features, DCF provides the ability to add both group-level and page-level indexing tags. DCF allows specific indexing information to be included in the output print data stream. You can process the output file created by DCF with ACIF to create an index file that can be processed with the Content Manager OnDemand data loading program.

In addition to DCF, there are several popular third-party programs that can produce indexed AFP data.

Generating index data with the AFP™ API
The AFP Application Programming Interface (AFP API) is a product that can be used to index print data.

Using the AFP API, a programmer who knows COBOL or PL/1 can format complex output without knowing the syntax and semantics of MO:DCA-P. Using the AFP API, you can index AFP files with both group-level and page-level indexing tags, which allows more specific information to be included in the output file. The indexing information is added at the same time that the application program generates the print data. You then process the output file with an indexing program, such as ACIF, to create the index data that the Content Manager OnDemand data loading program stores in the database.

Insert AFP™ records in the data stream
A common way of indexing unstructured, mixed-mode data is to add NOP structured fields to the data stream. ACIF can then be used to process the data stream and locate the NOP fields and extract the index values.

Using the PDF indexer
The Content Manager OnDemand PDF indexer is a program that you can use to extract index data from or generate index data about Adobe PDF files.

The index data can enhance your ability to store, retrieve, and view PDF documents with Content Manager OnDemand. The PDF indexer processes PDF input files. A PDF file is a distilled version of a PostScript file, adding structure and efficiency. A PDF file can be created by Acrobat Distiller or a special printer driver program called a PDFWriter. The PDF indexer supports PDF Version 1.9 and lower input and output files. See the documentation provided with Acrobat Distiller for more information about preparing input data for the Distiller.

The PDF indexer can logically divide reports into individual items, such as statements, policies, and bills. You can define up to 128 index fields for each item in a report.

The PDF indexer uses a coordinate system to locate the text strings that determine the beginning of a group and the index values. The coordinate system uses x and y pairs imposed on a page. For each text string, you identify its upper left and lower right position on the page. The upper left corner and lower right corner form a string box. The string box is the smallest rectangle that completely encloses the text string. The origin is in the upper left hand corner of the page. The x coordinate increases to the right and y increases down the page. You also identify the page on which the text string appears. Content Manager OnDemand provides the Report Wizard in the Administrator to help you create indexing parameters for the IBM Content Manager OnDemand PDF Indexer for Multiplatforms. Also provides the **ARSPDUMP** program to help you identify the locations of text strings on the page.

The *IBM Content Manager OnDemand for Multiplatforms: Administration Guide* provides details about the Report Wizard and gives examples of how to use the Report Wizard to process line data input files. Using the Report Wizard to process PDF input files is similar to processing line data input files.

The *Content Manager OnDemand Indexing Reference* provides details about the PDF Indexer and shows examples about how to use it to process PDF input files.

Generic indexer
Content Manager OnDemand provides the Generic indexer so that you can specify indexing information for input data that you cannot or do not want to index with ACIF or the other indexing programs.

For example, suppose that you want to load word processing documents into the system. The documents can be stored in Content Manager OnDemand in the same format in which they were created. The documents can be retrieved from Content Manager OnDemand and viewed with the word processor. The documents do not contain AFP data, line data, Metacode, PCL or PDF data, so you cannot index them with ACIF, the Content Manager OnDemand PDF Indexer. You can specify index information about the documents to the Generic indexer and load the documents into the system. Users can then search for and retrieve the documents using one of the Content Manager OnDemand client programs.

To use the Generic indexer, you must specify all of the index data for each input files that you want to store in and retrieve from Content Manager OnDemand. You specify the index data in a parameter file. The parameter file contains the index fields, index values, and information about the input files or documents that you want to process. The Generic indexer retrieves the indexing information from the parameter file and generates the index data that is loaded into the database. Content Manager OnDemand creates one index record for each input file (or document) that you specify in the parameter file. The index record contains the index values that uniquely identify a file or document in Content Manager OnDemand.

The Generic indexer supports group-level indexes. Group indexes are stored in the database and used to search for documents. You must specify one set of group indexes for each file or document that you want to process with the Generic indexer. You can define up to 128 index fields for each file or document.

See the *Content Manager OnDemand Indexing Reference* for more information about the Content Manager OnDemand Generic indexer.

Indexing reports using date fields
A date value should never be used as the first index value of a report or document.

The Content Manager OnDemand indexer design uses the first index value to break a larger report into smaller more manageable pieces. In most cases, a date value will not change within the report.

To store data in the system, each report must be indexed with a date field. When querying the database, Content Manager OnDemand uses the date in a report to determine one report's data from another. Content Manager OnDemand also uses the report date to determine when to remove reports from cache storage and how long to maintain report data (index data and documents) on the system.

You can use the date that appears in the report, such as the run date, a transaction date, or the statement date. If the data that you want to store in Content Manager OnDemand does not contain a date, you can use the date that the report was loaded into the system.

Content Manager OnDemand supports date values in the range of January 1, 1970 to December 31, 9999. Content Manager OnDemand also supports a date/time field. A date/time field can contain date values from January 1, 1970 to December 31, 9999.

Running ACIF on servers
Indexing reports on a Content Manager OnDemand server can provide certain benefits in a production archival process.

Indexing reports benefits include:

- You can schedule and run the indexing program as part of the data loading process. The *IBM Content Manager OnDemand for Multiplatforms: Installation and Configuration Guide* describes how to configure the Content Manager OnDemand data indexing and loading programs to run on a regular schedule, automatically processing input data.
- You can offload part of the z/OS processing to a Content Manager OnDemand server, by running the indexing program on the Content Manager OnDemand server. If your Content Manager OnDemand system consists of several Content Manager OnDemand servers, you can distribute the indexing workload among the servers, increasing the resources available to process reports and store them on the system.

Chapter 9. Objects

This section contains information that can help you plan application groups, applications, folders, and cabinets for your reports.

Overview

When you install and configure the Content Manager OnDemand software, you create and initialize a set of database tables that form the internal framework of the system. When you define reports to the system, Content Manager OnDemand adds an application group table structure and other control information to the database.

Content Manager OnDemand uses a set of objects to describe the database tables, fields, and data that make up the system. The standard objects are: user, group, application, application group, folder, cabinet, printer, and storage set. When you define an object to Content Manager OnDemand, such as an application group, Content Manager OnDemand stores the choices that you make and the information that you enter about the application group into the database. Every time that you load a report into an application group, Content Manager OnDemand updates the database with control information, inserts rows of index data into an application group table, and stores report data and resource files on storage volumes.

Users of the Content Manager OnDemand system open a folder to query and access reports that are stored on the system. A folder is the primary Content Manager OnDemand object that users deal with. A folder provides users the means to search for and retrieve data stored in Content Manager OnDemand. Users open a folder to construct queries and retrieve the reports that are stored in the application groups referenced by the folder. A folder can reference one or more application groups.

A cabinet can be used to organize folders. Cabinets are an optional feature that enable users to navigate to folders more easily. For example, a cabinet can be used to group together folders that a user needs to perform a certain task.

An application group represents the index and report data that you load into Content Manager OnDemand. The Content Manager OnDemand database contains tables of application group data. Records in an application group table contain index values extracted from reports and pointers to report data (documents) located on storage volumes. An application group can contain one or more applications that have the same storage characteristics and index fields.

A Content Manager OnDemand application includes a description of the physical characteristics of a report, such as the type of data contained in the report and the record format of the input file, instructions to the indexing and loading programs that process the report, and information that Content Manager OnDemand uses to display and print pages of the report. Typically, you define an application for each type of report that you plan to store in Content Manager OnDemand. You can group applications that have the same storage characteristics and index fields into an application group.

You assign a unique name to each object that you define to Content Manager OnDemand, such as application groups, applications, folders, and cabinets.

Content Manager OnDemand uses properties to describe the appearance, behavior, and internal structure of the objects that make up a Content Manager OnDemand system. For example, Display Format is a property of a folder field that determines how Content Manager OnDemand client programs display the values of the field in the document list. The properties are grouped in categories. For example, the General category under folders contains properties that describe general information about a folder, such as the name and description of the folder and the application groups contained in the folder.

Folders

A folder provides users the means to search for and retrieve related reports stored on the system.

Users open folders, construct queries, and retrieve reports from application groups. (However, it is not necessary that users know about or understand application groups.) When you create a folder, you define the search and display fields that appear when the user opens the folder. You map the folder fields to database fields in the application groups referenced by the folder. The database fields contain index values extracted from the reports that are loaded into the application groups. For example, the folder search field Customer Account Number could be mapped to the `acct#` application group database field. Content Manager OnDemand creates database records that include the index values for the `acct#` field when you load a report into the application group. When the user enters a query, Content Manager OnDemand retrieves records from the database if the values of the `acct#` database field match the value that the user typed in the Customer Account Number search field.

When you define a folder to Content Manager OnDemand, you add one or more application groups to the folder, select index fields from the application groups to appear as search and display fields when the user opens the folder, and specify the properties of the search and display fields. For example, you can determine the layout of the search fields on the screen and specify values that will automatically appear in the search fields when the user opens the folder.

Content Manager OnDemand maintains information about the name of the folder and its structure in the Content Manager OnDemand database. For example, the database contains information that describes the search and display fields that you defined and the database fields that you selected from application groups referenced by the folder.

You define a folder to Content Manager OnDemand through properties and values grouped in categories. A category is a set of related properties. Content Manager OnDemand provides folder categories for general information, permissions, field definitions, field information, and field mapping. The general category is where you specify general properties about the folder, such as the name of the folder and the application groups contained in the folder. The permissions category is where you determine the groups and users that can open the folder. You can assign other types of folder authorities in the permissions category, such as specifying someone to administer the folder. The field definitions category is where you define the search and display fields for the folder. The field information category is where you specify the attributes of the search and display fields. For example, you can specify the search operators available for each field and determine the order that the search fields appear on the screen. The field mapping category is where you map the folder search and display fields to database fields in application groups referenced by the folder.

Cabinets

Cabinets are used to organize folders into useful groups.

For example, if users need to retrieve reports from multiple folders to complete a task, you can create a cabinet to help them find task-related folders more easily. When you create a cabinet, you select which folders are grouped together and which users have access to the cabinet.

Cabinets are useful when users work with a large number of folders. Additionally, if users need the same folder to complete multiple tasks, the folder can be added to multiple cabinets.

Holds

This requires you to install the Enhanced Retention Management Feature option.

To learn more about holds, see the *Enhanced Retention Management Guide*.

Application groups

An application group is a collection of one or more applications that have the same index fields and storage characteristics.

The application group is the object that Content Manager OnDemand uses to maintain the reports that you load into the system. The application group holds index data for reports, documents, management information, permissions for the groups and users authorized to access application group, and so forth.

When you define an application group, you specify the name and type of the database fields that will hold the index data extracted from the reports that are loaded into the application group. You specify whether a database field is used to index or filter data, and specify other characteristics of the fields. When you define an application group, Content Manager OnDemand creates an application group table structure in the database, with a column for each database field that you defined. When you load a report into the application group, Content Manager OnDemand inserts rows into an application group table for each indexed item found in the report. An indexed item can be a logical item, such as a policy or statement, or a group of pages, depending on how the report is organized and how you decide to index the report. Users search for reports using one or more of the fields that you defined for the application group.

Content Manager OnDemand supports up to 128 index and filter fields for each application group:

- Index fields allow fast access to a specific record using a key, but generally require a large amount of disk storage to implement and require longer to load data into the application group. Content Manager OnDemand uses index fields to locate the records in the database that meet the search criteria entered by the user. The index record contains the physical location of an item on a storage volume.
- Filter fields are used to refine queries, retrieving only a subset of the records found with an index field. Filter fields are generally used with an index field to identify a specific item in the application group. Filter fields can also be used to display additional information in the document list, for example, an address.

Content Manager OnDemand requires a segment field for each application group that you define. Content Manager OnDemand uses the segment field to organize and maintain application group data and to locate items that match a query. The segment field must be one of the following date field or a date/time fields:

- Report Date. The date that the application program created the report file. Typically the date found on pages of the report.
- Load Date. The date that you loaded the report into the application group. Use the load date if the report does not contain a date.

Storage requirements and index fields are the primary considerations when you define an application group and identify the applications that you can place in an application group. A third factor is the organization of the information contained in the report. Content Manager OnDemand can index, store, and retrieve data contained in a report based on the structure of the data that it contains:

- Some reports are made up of logical groups of information, such as statements, invoices, and policies. These groups, or logical items, can contain one or more pages of information. Content Manager OnDemand can index, store, and retrieve the logical items contained in a report. Each logical item can be indexed on up to 128 values, for example, account number, customer name, and balance. Content Manager OnDemand creates a row in the database for each logical item it finds in the report.
- Other reports might be organized differently, and might not necessarily contain logical items. For example, a report could contain thousands of pages of transaction or general ledger data. Content Manager OnDemand can index, store, and retrieve information from these types of reports using index values such as date, page number, and a sorted value, such as transaction number. Content Manager OnDemand divides these types of reports into groups of pages and indexes each group of pages. While these types of reports might contain logical items, it probably would not be cost effective to index every item in the report. That is, indexing every item in these types of reports would probably result in thousands of index records being added to the database each time that a report is loaded into the application group.

When you create an application group, you specify how Content Manager OnDemand should store the index data for the reports that you load into the application group. Content Manager OnDemand provides the following method to determine how index records are loaded into the database and how users can query the application group:

- Multiple Loads per Database Table

  Each time that you load a report into the application group, Content Manager OnDemand inserts the index records into an existing database table. Index records for every report loaded into the application group are stored in the same logical database table. Content Manager OnDemand maintains the application group data so that, as far as a user querying the application group knows, they appear to reside in one database table. Content Manager OnDemand automatically segments the application group data when it grows beyond a certain size. Content Manager OnDemand maintains a segment table for each application group. The segment table provides faster query performance by limiting searches to a specific table of application group data, using a date value to construct the query. You can use this method to organize your database when the users that search for data stored in the application group do not necessarily know or care what version of a report generated the information that they need.

When you create an application group, you specify the storage characteristics of the report, such as the length of time that Content Manager OnDemand maintains data stored in the application group and the data caching and migration values. The storage characteristics also determine whether Content Manager OnDemand stores a copy of the report on archive media, whether Content Manager OnDemand should create a backup copy of the report, and when Content Manager OnDemand removes report data when it is no longer needed.

Content Manager OnDemand can perform three types of processing on application group data:

- Database expiration processing

  Index data expires (is eligible for removal from the system) when it reaches its Life of Data and Indexes period. (You specify the Life of Data and Indexes period when you create an application group.) Content Manager OnDemand provides a utility that you can use to remove index data. You typically set up the utility to run automatically on a regular schedule. Database expiration processing also reclaims the disk space taken by deleted index data.

- Cache migration processing

  Cache migration is the process of copying reports from cache storage to archive storage. You specify when a report should be copied from cache storage to archive storage when you create an application group. Content Manager OnDemand provides a utility that you can use to copy reports to archive storage. You typically set up the utility to run automatically on a regular schedule. Cache migration optimizes the use of cache storage, while providing excellent performance for short-term retrievals of reports. As a report ages, and in all likelihood accesses becomes less frequent, Content Manager OnDemand can automatically copy the report to long-term (archive) storage. You can also use cache migration to defer the loading of reports to archive storage to a time when there is little or no other system activity.

- Cache expiration processing

  Cache expiration is the process of deleting reports from cache storage. You specify how long a report should remain in cache storage when you create an application group. Content Manager OnDemand provides a utility that you can use to delete reports from cache storage. You typically set up the utility to run automatically on a regular schedule. Cache expiration reclaims cache storage space taken by expired reports, so that the system has space for newer versions of reports.

Applications

A Content Manager OnDemand application describes the physical characteristics of a report, processing instructions for the indexing and data loading programs, and information about how Content Manager OnDemand displays and prints pages of a report.

You can specify default settings for viewing and printing pages of a report at the Content Manager OnDemand application level. For example, if you select a default printer for the application, when a user prints a document that is associated with the application, Content Manager OnDemand sends the document to the printer that you specified. Typically you define an application for each different report that you plan to load into the system.

When you create an application, you specify properties of the input data, such as whether the data contains carriage control characters or table reference characters, and the record format of the input data. Content Manager OnDemand uses the information that you specify to properly interpret the data for viewing.

The Content Manager OnDemand application is where you specify information to the indexing and data loading programs, such as the technique that Content Manager OnDemand uses to compress the report file, the parameter used to index the data, and information that Content Manager OnDemand uses to process index data before loading index records into the database. Content Manager OnDemand uses the indexing parameters, options, and data values that you specify to locate index data in and extract index data from the report.

You can set up one or more *logical views* of a report. A logical view determines how Content Manager OnDemand displays line data reports and governs other viewing characteristics. For example, you can set up a logical view so that when a user selects a document for viewing, the Content Manager OnDemand client program automatically locks the heading of the report in place when the user moves up or down lines on a page.

Users and groups

Each user logs on to Content Manager OnDemand with a user ID.

Content Manager OnDemand authenticates user IDs and determines the usage and administrative authority available to the user based on the log on user ID. An Content Manager OnDemand user ID does not necessarily have to identify an individual user. However, for accounting purposes, you probably want to assign an Content Manager OnDemand user ID to each person that uses the system.

Content Manager OnDemand automatically creates the ADMIN user ID when you initialize the system. The ADMIN user ID has system administrator authority. A system administrator can perform the basic user functions, such as logging on the system and opening folders, and administrative functions, such as defining users and groups and creating, updating, and deleting application groups, applications, folders, storage sets, and printers. **Note:** For a local server, the default administrative user ID is `admin`. The system does not set an initial password for the `admin` user ID on a local server. See *IBM Content Manager OnDemand for Multiplatforms: Installation and Configuration Guide* for more information on local servers.

The Content Manager OnDemand security function requires that a user's password be changed:

- When a user accesses the system for the first time
- When a System Administrator or a User Administrator changes or resets a user's password

Content Manager OnDemand groups are a means to organize users by function, authorization, or any other purpose you might require. When you define an Content Manager OnDemand group, you can organize users by department or function and set folder and application group permissions that are common to all of the users assigned to the group. The permissions determine the types of actions that users assigned to the group can perform. You do not have to assign a user to a group, but doing so can simplify administration of users with similar requirements and capabilities.

Permissions

Overview

As both a convenience and security measure, you can assign a user to a group.

When you assign a user to a group, the user obtains the permissions of the group. For example, suppose you create a group and authorize the group to open the Student Information folder. Any user that you assign to the group automatically obtains permission to open the Student Information folder.

If you assign a user to more than one group, the user normally obtains the permissions of all of the groups. For example, using the group settings listed in the table, a user assigned to both groups can access the Student Bills and Student Transcripts folders.

Table 23. Group permissions

Group	Folders
Accounting	Student Bills
Admissions	Student Transcripts

However, there are exceptions to this rule. See information about permissions in the *IBM Content Manager OnDemand for Multiplatforms: Administration Guide* for details.

You can set folder, cabinet, and application group permissions for every user and group defined to Content Manager OnDemand. If you set permissions for a specific group, the group permissions take precedence over the permissions set at the folder level or the application group level. If you set permissions for a specific user, the user permissions take precedence, regardless of any group that includes the user or the permissions set at the folder level or the application group level.

You can set folder and application group permissions when you add or update a folder, cabinet, or application group. You can also set folder and application group permissions when you add or update a user or a group.

Folder permissions

You can set folder permissions at the folder, group, and user levels. Setting permissions at the folder level provides all Content Manager OnDemand users and groups that are not otherwise given permissions with the permissions that you define.

Setting permissions at the group level provides all of the users that you assign to the group with the permissions that you define. Group level permissions override folder level permissions. Setting permissions at the user level provides a specific user with the permissions that you define. User level permissions override group level permissions and folder level permissions.

By default, only the user that created the folder, users with administrator permission for the folder, application group/folder administrators, and system administrators can access the folder.

You can set the following types of folder permissions:

- Access. Users can open the folder with Content Manager OnDemand client programs and search for and retrieve data from the application groups referenced in the folder.

  To search for and retrieve items, users must have access permission for the folder, and access permission to one or more of the application groups referenced in the folder.
- Fields. Users can open the folder with Content Manager OnDemand client programs and can modify the folder field information with the administrator interface. Content Manager OnDemand maintains a set of folder fields for each user given fields permission for the folder.
- Named Queries. A named query is a set of search criteria, saved by name, that can be selected and restored into folder search fields. Content Manager OnDemand supports two types of named queries: public, that is, a named query that is available to all users that can open the folder, and private, that is, a

named query available only to the user that created the named query. Users can be given authority to view, create, modify, and delete named queries.

- Administrator. A folder administrator can modify and delete the folder. A folder administrator can change user and group permissions, add and remove users and groups from the folder, and make changes to the folder field information.

Application group permissions

You can set application group permissions at the application group, group, and user levels.

Setting permissions at the application group level provides all Content Manager OnDemand users and groups that are not otherwise given permissions with the permissions that you define. Setting permissions at the group level provides all of the users that you add to the group with the permissions that you define. Group level permissions override application group level permissions. Setting permissions at the user level provides a specific user with the permissions that you define. User level permissions override group level permissions and application group level permissions.

By default, only the user that created the application group, users with administrator permission for the application group, application group/folder administrators, and system administrators can access the application group.

You can set the following types of application group permissions:

- Access. Users can search for and retrieve data stored in the application group using Content Manager OnDemand client programs.
- Document. Determines the types of document functions users can perform. The default document permissions are view, print, FAX, and copy.
- Annotation. Determines the types of annotation functions users can perform. The default annotation permissions are view and add.
- Logical Views. Logical views determine how Content Manager OnDemand displays report file pages. Users can define their own logical views with Content Manager OnDemand client programs.
- Administrator. An application group administrator can modify and delete the application group. An application group administrator can change user and group permissions, add and remove users and groups from the application group, change message logging options, update the storage management settings for the application group, and make changes to the application group field information.
- Query restriction. Limits access to application group data. You typically set up a query restriction to limit the data that a specific user or group of users can access.

Naming rules

When you create objects in Content Manager OnDemand, you assign names to the various objects.

Remember:

1. If you install Content Manager OnDemand with a language that requires multiple bytes per character (for example, Kanji), the number of characters permitted for a name is less than the number listed in the sections that follow.
2. Trailing blank characters are trimmed from the name when adding or updating an object. As a reminder, for existing objects, the name specified to a command line program must include trailing blank characters that are part of the name.

Remember: When naming a user, the name that you specify:

- Can contain from one to 128 characters (bytes)
- Cannot include the * (asterisk), % (percentage) + (plus), [(left bracket),] (right bracket), “ (double quote), or blank characters
- Must be unique to the library server

- By default, Content Manager OnDemand converts lowercase letters in a user name to uppercase (for example, laguarde is stored as LAGUARDE)

If your organization implements the logon user exit, then you can determine the characteristics of user IDs on your system.

When creating a password, the value that you specify:

- Can contain from one to 128 characters (bytes)

  Remember:

  1. Content Manager OnDemand security verifies only the first eight characters that are entered by the user. The additional characters are provided for customers who choose to implement their own password security by enabling the security user exit.
  2. If your organization enables the security user exit, you should set the Minimum Password Length option to Permit Blank Password so that Content Manager OnDemand security does not validate passwords that are entered by your users (when they set or change a password). Also, Content Manager OnDemand security ignores the Maximum Password Age option when you enable the security user exit.
  3. Unless your organization enables the security user exit, It is recommended that you specify a value of eight or less for the Minimum Password Length option.
- By default, Content Manager OnDemand converts lowercase letters in a password to uppercase (for example, laguarde is stored as LAGUARDE)

If your organization implements the security user exit, then you can determine the characteristics of passwords on your system.

The Content Manager OnDemand security function requires that a user's password be changed:

- When a user accesses the system for the first time
- When a System Administrator or a User Administrator changes or resets a user's password

This change does not affect customers that use the security user exit to implement their own user / password security (and bypass the Content Manager OnDemand security function).

When naming a group, the name that you specify:

- Can contain from one to 128 characters (bytes)
- Cannot include the * (asterisk), % (percentage) + (plus), _ (underscore), [(left bracket),] (right bracket), “ (double quote), or blank characters
- Must be unique to the library server
- Can be mixed case; however, Content Manager OnDemand ignores the case (for example, LaGuarde is the same as laguarde)

When naming an application group, application, or folder, the name that you specify:

- Can contain from one to sixty characters (bytes), including embedded blanks
- Cannot include the ' (apostrophe), % (percentage), [(left bracket),] (right bracket), or “ (double quote) characters
- Can be mixed case; however, Content Manager OnDemand ignores the case (for example, LaGuarde is the same as laguarde)
- An application name must be unique to the application group where you assign the application
- An application group or folder name must be unique to the library server

When naming a database field, the name that you specify:

- Can contain from one to eighteen characters (bytes)
- Must begin with the letter A through Z
- Can include the letters A through Z, the numbers 0 through 9, and the @ (at sign), $ (dollar), _ (underscore), and # (number sign)

- Can be mixed case; however, Content Manager OnDemand does not create a unique name (for example, rDate is the same as rdate)
- Must be unique to the application group
- Cannot be any of the Content Manager OnDemand reserved words:

```
annot        doc_off
comp_len     doc_type
comp_off     prt_nid
comp_type    resource
doc_len      res_comp_type
doc_name     sec_nid
```

- Cannot be any of the words reserved by the database manager. (For a list of reserved words, see the documentation provided with your database manager product.)

When naming a logical view, the name that you specify:

- Can contain from one to thirty characters (bytes)
- Can be mixed case
- A public view must be unique to the application
- A private view must be unique to the user

When naming a folder field, the name that you specify:

- Can contain from one to sixty characters (bytes), including embedded blanks
- Can be mixed case
- Must be unique to the folder

When naming a storage set, the name that you specify:

- Can contain from one to sixty characters (bytes)
- Can be mixed case; however, Content Manager OnDemand ignores the case (for example, LaGuarde is the same as laguarde)
- Must be unique to the library server

When naming a primary storage node, the name that you specify:

- Can contain from one to 128 characters (bytes)
- Can be mixed case; however, Content Manager OnDemand ignores the case (for example, LaGuarde is the same as laguarde)
- Must be unique to the storage set

When naming a server printer, the name that you specify:

- Can contain from one to sixty characters (bytes)
- Can be mixed case; however, Content Manager OnDemand ignores the case (for example, LaGuarde is the same as laguarde)
- Must be unique to the library server

When naming a server printer queue, the name that you specify:

- Can contain from one to sixty characters (bytes)
- Must be a valid printer queue name on the library server

Data types and field types

When you define an application group, Content Manager OnDemand creates a structure for a database table with the index and filter fields that you define.

When you store a report in the application group, Content Manager OnDemand extracts index data from the report, places the index data into the database fields, and inserts rows into the application group table. The database fields that you define for the application group can contain different types of data. When you define the database fields, you select a data type for each field. The data type tells Content Manager OnDemand what kind of data can be stored in the field.

When you define a folder to Content Manager OnDemand, the fields that you define can be used in two ways:

- For search fields, in which users enter values to construct queries
- For display fields, to identify the items in the document list

The table lists the types of application group and folder fields supported by Content Manager OnDemand.

Table 24. Application group and folder field types

Field Type	Description
Small Int (2)	Contains whole numbers between -32,767 and 32,767
Integer	Contains whole numbers between -2147483648 and 2147483647
Big Int	Contains whole numbers between -922337036854775807 and 922337036854775807. Big integer fields hold a 64-bit integer representation of a number or a character string in the form of an integer constant.
DecFloat (16)	Decimal floating-point number with 16 digits of precision.
DecFloat (34)	Decimal floating-point number with 34 digits of precision.
String (Fixed)	Contains letters, numbers, special symbols, such as the % and #, and any other printable character. A fixed length string field requires one byte per character declared; unused characters do consume storage.
String (Variable)	Contains letters, numbers, special symbols, such as the % and #, and any other printable character. A variable length string field requires one byte per character plus four bytes for overhead; unused characters do not consume storage.
Date	Contains a valid date from January 1, 0001 to December 31, 9999.
Date (old style)	Contains a valid date from January 1, 1970 to December 31, 2069. Content Manager OnDemand checks a date value to make sure it is valid.
Time (old style)	Contains times of day, stored in three-second increments, since midnight, and limited to 24 hours
Date/Time	Contains both a date and time value. The date can be from January 1, 0001 to December 31, 9999. The time is stored in one-second increments.
Date/Time (old style)	Contains both a date and time value. The date can be from January 1, 1970 to December 31, 2038. The time is stored in one-second increments.
Date/Time (TZ)	Contains both a date and time value. The date can be from January 1, 0001 to December 31, 9999. A Date/Time (TZ) field is exactly like a Date/Time field, but uses the time zone set on the client PC.

Table 24. Application group and folder field types (continued)

Field Type	Description
Date/Time (TZ) (old style)	Contains both a date and time value. The date can be from January 1, 1970 to December 31, 2038. A Date/Time (TZ) field is exactly like a Date/Time field, but uses the time zone set on the client PC.

The table lists additional types of fields that are supported in folders.

Table 25. Additional folder field types

Field Type	Description
Annotation Color Search	Use to search annotations to a document by specifying a color. A match occurs and an item is added to the document list if the color of the text in one or more of the annotations to a document is the same as the color that is specified in the search field. A folder can have one annotation color search field.
Annotation Text Search	Use to search annotations to a document for the specified string. A match occurs and an item is added to the document list if one or more of the annotations to a document contain the text that is specified in the search field. A folder can have one annotation text search field.
Application Group	For a search field, contains a list of the application groups that can be searched from the folder. When you create a folder that contains more than one application group, you can define an application group field. If enabled for queries, users can select the name of the application group that Content Manager OnDemand searches, rather than searching all of the application groups contained in the folder (the default). For a display field, lists the name of the application group in which the document was found. A folder can have one application group field.
Segment	Contains a list of the tables of index data that are stored in the application groups that can be searched from the folder. Each item in the list represents a segment of application group data. Content Manager OnDemand segments application group data by date. If enabled for queries, users can select a specific segment of application group data to search. A folder can have one segment field. Specifying a search value in the segment field can improve the performance of queries.

Table 25. Additional folder field types (continued)

Field Type	Description
Text Search	Used to find documents that contain a non-indexed word or phrase. A match occurs and an item is added to the document list when one or more lines in a document contain the word or phrase exactly as specified. The search string can contain letters, numbers, special symbols, such as the % and #, and any other printable character. A folder can have one text search field. The (sequential) text search takes place on the server. A text search delays the generation of the document list. Only documents that meet all of the criteria specified in the other folder fields are searched for the specified word or phrase. A typical use of a text search field is to provide users an additional search field without incurring database overhead. For example, assume that a report is indexed on date and transaction number. A text search field would allow users to optionally enter a customer's name, phone number, or any other information contained in the document or documents that the user needs to retrieve (the information is not contained in the database). However, a text search field has a direct impact on the generation of the document list and the performance of the server. A large number of users performing text searches at the same time can usually drain the resources of even the most powerful library server.
Hold	Uses a padlock icon to indicate that a document is held in a hold. When a document is held in a hold, its retention policy is suspended indefinitely and the document is not be deleted until the hold is released, and the retention policy is resumed.
Document	Indicates what format the document is (PDF, ACIF, user-defined, etc).

Chapter 10. Backup and recovery

Backup and recovery overview

Backup and recovery for Content Manager OnDemand is provided with recommendations about methods and procedures that an administrator can use to ensure that critical Content Manager OnDemand components can be recovered.

This section of the book describes backup and recovery for Content Manager OnDemand and provides recommendations about methods and procedures that an administrator can use to make sure that the following critical Content Manager OnDemand components can be recovered when needed:

- Content Manager OnDemand software
- Content Manager OnDemand server information, created or modified during installation, configuration, and ongoing operation of Content Manager OnDemand
- The Content Manager OnDemand database
- The Tivoli Storage Manager database
- Archived reports

Content Manager OnDemand supports storing index data in table spaces and incremental backup of table spaces. Table spaces enhance the management of index data and provide improved performance, especially for database backups. An incremental table space backup completes much quicker than a full database backup, providing you with increased flexibility in scheduling report loads. Incremental backup images also require less storage space than full database backups.

If you use DB2, you can use Tivoli Storage Manager facilities to backup and restore DB2 databases. This capability means that you do not have to manage the DB2 files on disk.

Server software

If a media failure or some other unforeseen event occurs, you might be required to restore the Content Manager OnDemand software programs, database software, archive manager software, server print manager software, and other application and user-defined software that you use on the system.

You can usually use the original product media to restore the software programs.

It is important that you store the original product media in a safe location. It is recommended that you register Content Manager OnDemand as part of your business recovery plan and store the original product media in the same place that you store the other programs and files that are vital to the operation of your systems.

Database table spaces

Database table space support provides enhanced flexibility and improved performance for your application group data.

For example, after you store a report in Content Manager OnDemand, you can create a backup image of the table that changed during the load process, rather than creating a backup image of the entire database. You can also create an incremental backup image of the database, which contains only those tables that changed since the last backup image. Because the backup image only contains the changes made to the database, the backup process typically runs much faster than a full backup.

Content Manager OnDemand creates one table space for each segment of application group data. After Content Manager OnDemand closes the segment and you back up the table space, you do not need to back up the table space again, unless it is recovered or restored.

When you use the incremental table space backup capability, It is recommended that you backup the Content Manager OnDemand database after each report file load. If your schedule does not permit you to run the backup command after each load, It is recommended that you backup the database once a day (assuming that you load multiple reports each day). While incremental backup images can be used to recover the database, It is recommended that you periodically create a full backup image of the database. A full backup image of the database is the quickest way to recover the database in the event that you need to do so. However, if your Content Manager OnDemand database is very large, and it cannot be backed up in a reasonable amount of time or requires a prohibitive number of storage volumes to hold, you might find that maintaining full backup images of the database is not possible.

The *IBM Content Manager OnDemand for Multiplatforms: Installation and Configuration Guide* for your server provides details about how to configure the system to support table spaces.

Database backup

Content Manager OnDemand provides the ARSDB program so that you can create backup images of the Content Manager OnDemand database.

The ARSDB program supports table space and full database backups:

- Content Manager OnDemand provides support for incremental table space backups and full database backups.
- An online backup can be taken when other applications or processes are connected to the database. That is, other applications and processes can continue to read or modify data while the backup is in process.
- During an offline backup, only the backup task is connected to the database. Before starting an offline backup, It is recommended that you stop the Content Manager OnDemand system to make sure that no other applications or processes are connected to the database.
- When you back up the database with the ARSDB program, Content Manager OnDemand removes the log files from the archived log file directory, releasing the space taken by files that are no longer needed. However, if you use Tivoli Storage Manager to manage DB2 log files, then the policy domain determines when archived log files are eligible to be removed.

If your production schedule allows, It is recommended that you create offline backups on a regular schedule, perhaps once a week. Regularly scheduled offline backups can reduce the time required to rebuild table spaces or the database, if you need to do so. It is recommended that you write offline backup images to removable media or storage that is managed by Tivoli Storage Manager. Keep backup images in a safe place, until the next time that you create an offline backup image of the table space or database.

If your schedule does not provide time to take offline backups (that is, your system must always remain available to users), you should take online backups on a regular schedule. The *IBM Content Manager OnDemand for Multiplatforms: Installation and Configuration Guide* shows how to use the cron facility to create online backups of the database with the ARSDB program automatically on a regular schedule (for UNIX servers) and shows how to use the Windows server configuration program to schedule online backups of the database.

The *IBM Content Manager OnDemand for Multiplatforms: Administration Guide* provides details about the ARSDB program parameters and options.

See your database manager product information for details about backing up a database.

Database backup in Windows®

The configurator program that is provided with the Content Manager OnDemand for Windows product allows for the scheduling of database backups.

About this task

You can perform a backup while the database is either online or offline.

- If the backup is to be performed online, other applications or processes can continue to connect to the database, as well as read and modify data while the backup operation is running.
- If the backup is to be performed offline, only the backup operation can be connected to the database; other Content Manager OnDemand services and the rest of your organization cannot connect to the database while the backup task is running.

Procedure

To schedule an offline backup with the Configurator program:

1. Manually disconnect all other processes from the database before the backup task is scheduled to begin.

   This includes stopping the Content Manager OnDemand LibSrvr, MVSD Server, and Load Data services on the library server. In addition, if you load data to the library server from another object server, then you should manually stop the Content Manager OnDemand ObjSrvr and Load Data services on the object server.
2. Run the offline backup.
3. Verify that the offline backup completed successfully.
4. Manually restart the Content Manager OnDemand LibSrvr service and the Content Manager OnDemand MVSD Server and Content Manager OnDemand Load Data services on the library server.

   If you stopped Content Manager OnDemand services on an object server, manually restart the services.

Database logging

The database manager uses transaction logging to record information about changes to the Content Manager OnDemand database.

The information in the log file is used to recover from corruption of data in the database. Logging ensures that no data is lost. By combining the information in the log files with a backup copy of the database, the Content Manager OnDemand database can be recovered to any point in time.

Database recovery

There are two types of database recovery. The first type recovers from failures that occur while update transactions are taking place.

The log helps correct this type of failure by allowing the transactions received before the failure to either be reapplied to the database or to be rolled-out. Rolling-out transactions is a way to return the database to the state it was in before the transaction that caused the failure.

The second type of recovery deals with corruption of the Content Manager OnDemand database and is usually caused by media failure. The combination of log files and a backup copy of the database can be used to recreate an image of the Content Manager OnDemand database at a particular point in time.

If a catastrophic failure occurs, the system administrator will need to intervene to recover the database. Recovery from catastrophic failure starts with restoration of the latest full backup copy of the database.

Next, the system administrator reapplies the transactions recorded in the log files. These steps will recreate a mirror image of the Content Manager OnDemand database before the catastrophic failure.

The Content Manager OnDemand database and database log files should reside on different physical volumes. The database backup image should be written to removable media. Unless multiple disk and tape volumes are damaged or lost at the same time, there is no possibility of losing the information contained in the Content Manager OnDemand database.

Tivoli® Storage Manager database

Tivoli Storage Manager maintains a database that contains information about the devices and files that it manages.

When you store a copy of a report into archive storage, Tivoli Storage Manager updates its database and stores a copy of the report on a storage volume. When you define archive storage devices and register client nodes, Tivoli Storage Manager updates its database. When Tivoli Storage Manager maintains the storage that it manages, it updates the database with status information about files and storage volumes. The database is critical to proper operation of Tivoli Storage Manager in storing objects on and retrieving objects from the optical and tape storage volumes that it manages.

It is recommended that you mirror the Tivoli Storage Manager database. When you mirror the database, Tivoli Storage Manager replicates the database onto different physical storage. Tivoli Storage Manager automatically keeps track of and refreshes both copies of the database. When you configure physical storage so that Tivoli Storage Manager can mirror the database on different physical devices and adapters, you can provide protection for the database because of a failure of a single device. With mirroring, Tivoli Storage Manager can continue operation without interruption if a database volume fails by using a mirrored copy of the failed volume. Mirroring requires additional storage space for the mirrored volumes. See the Tivoli Storage Manager information for details about mirroring the database.

To protect the information in the database, and make sure that it can be restored if a disaster occurs, you must periodically create a backup copy of the database. You can recover the database to its most current state or to a specific point in time with the backup copy.

- You should take a full backup image of the database after you perform initial installation and configuration of Tivoli Storage Manager with Content Manager OnDemand. In addition, It is recommended that you periodically create a full backup of the database. A full backup copy of the database should be written to removable media and stored in a safe place.
- You should take an incremental backup image of the database more frequently, perhaps one or more times a day, depending on the amount of activity on the system. An incremental backup will record changes that have occurred since the last backup of the database (full or incremental). If you write incremental backup images of the database to disk, make sure that the disk is on a different controller and disk than any of the database or recovery log volumes.

There are several factors to consider when you decide the type and frequency of backups.

- A full backup takes longer to run than an incremental backup.
- Recovery time is faster with a full backup. Incremental backups increase the time it takes to recover the database because a full backup must be loaded first, followed by some or all of the incremental backups.
- A full backup is required under specific conditions. For example, you should create a full backup after you complete initial installation and configuration of Tivoli Storage Manager with Content Manager OnDemand. In addition, there might be restrictions on the number of incremental backups that can be taken between each full backup. See the Tivoli Storage Manager documentation for details.

It is recommended that you plan to backup the database after you load reports into the system and after Tivoli Storage Manager maintains its storage volumes (for example, you should backup the database after expiration and reclamation processing). Most customers, under typical conditions, should plan to backup the database every day.

Tivoli Storage Manager includes a central scheduling component that allows the automatic processing of administrative commands, such as database backup. The scheduled commands should be tracked by the server and recorded in the database. You usually set up an administrative command schedule by defining schedule parameters, such as the start day, date, and time, specifying the command to be executed, and activating the schedule. See the Tivoli Storage Manager information for details about scheduling administrative commands and automating database backups.

In addition to Content Manager OnDemand automatically migrating data from cache storage to storage managed by Tivoli Storage Manager, you can also use the standard Tivoli Storage Manager backup commands to do file system backups from the server. See the Tivoli Storage Manager Administrator's Reference for detailed information about the commands that you can use and to schedule backups.

Important: If you use the standard Tivoli Storage Manager backup commands to backup file systems on the server, you probably want to exclude the database file systems and the cache file systems from the backup. The database file systems should be backed up using the facilities provided by Content Manager OnDemand, such as the ARSDB program. (Although the database backup images created by the ARSDB program can be maintained by Tivoli Storage Manager.) You should use the data loading and migration facilities of Content Manager OnDemand to make sure that data in cache storage is copied to Tivoli Storage Manager-managed storage as needed.

Some other points to note:

- While you can store the file system backups in the same optical library as the Content Manager OnDemand data, you should define different client nodes to hold the file system backups. The client nodes should be assigned to a different domain than the Content Manager OnDemand data. The retention period of the file system backups will likely be different than the retention period for the Content Manager OnDemand data.
- File system backups generally require rewritable media. Many customers use write-once optical disks for your Content Manager OnDemand data
- You should plan a library with at least two drives: One for the file system backup. One for the Content Manager OnDemand data. If you plan to run the file system backups at the same time as you load data into Content Manager OnDemand, then you should plan a library with four drives, or you might store the file system backups in a different library than the Content Manager OnDemand data.

Recovery log

The recovery log is critical to the operation of Tivoli Storage Manager.

If the recovery log is unusable, then Tivoli Storage Manager is usually unavailable to store and retrieve data. With the recovery log available, and a restored backup image of the database, you can recover the database to its most current state.

To ensure fast recovery time and high availability of the database, you should always mirror the recovery log. Mirroring the recovery log requires much less space than mirroring the database. If you do not mirror the recovery log, then you should allocate the recovery log on a disk other than the one on which the database resides. See the Tivoli Storage Manager information for details about mirroring the recovery log.

When a database backup is completed, the recovery log records preceding the backup are deleted, freeing up recovery log storage for reuse. Taking frequent database backups reduces recovery log storage requirements, and reduces the time required to recover the database.

Storage volume history

Up-to-date storage volume history is vital for recovery of a lost or damaged database.

The storage volume history contains information that Tivoli Storage Manager needs about the storage volumes to use for database backups. The storage volume history also contains information that you will need to audit storage volumes after a recovery.

Tivoli Storage Manager cannot obtain storage volume history from the database during a restore of the database. Therefore, you should store at least one backup copy of the storage volume history on a disk other than the one on which the database resides.

See the Tivoli Storage Manager information for details about backing up the storage volume history.

Device configuration history

When you define, update, or delete storage objects such as devices, drives, and libraries, Tivoli Storage Manager updates the database and makes an entry in a device configuration history file.

To restore the database, Tivoli Storage Manager requires a definition for the device from which backup data is to be read. This definition is maintained in the device configuration history.

When the database is being restored, no definitions can be read from the database. Therefore, you should have at least one backup copy of the device configuration history on a disk other than the one on which the database resides.

See the Tivoli Storage Manager information for details about backing up the device configuration history.

Database recovery

Recovering using mirrored copies of the database

If a database volume fails because of media failure and you have enabled mirroring, then you can recover the database by using mirrored copies of the database.

After fixing the failing device, you can allocate space for the new mirrored copy and define the volume to Tivoli Storage Manager. After you define the volume to Tivoli Storage Manager, the server synchronizes the volume with the database.

Recoveries using backup copies of the database

Tivoli Storage Manager provides programs to recover the database, should a catastrophic failure occur.

These programs restore the database from the latest available full backup copy, apply all incremental backups that apply, and use the recovery log to apply any changes made to the database since the last backup was created.

If you restore the database to its most current state, Tivoli Storage Manager should automatically synchronizes the database and storage volumes.

If you restore the database to a specific point in time, you must audit all storage volumes to check for and resolve any inconsistencies between the information in the database and the actual information on the storage volumes. Depending on the number of storage volumes and the amount of activity that occurred after the database backup that you restored, the audit might require a significant amount of time.

To perform a database recovery, you should have the following information, preferably stored at a different location:

- Back up volumes of the database
- Copy storage pool volumes
- Server options file
- Storage volume history
- Device configuration history
- Output from commands that provide details of the database and recovery log setup

See the Tivoli Storage Manager information for details about recovering data.

Notices

This information was developed for products and services that are offered in the US. This material might be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

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

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
US

For license inquiries regarding double-byte character set (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japan

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

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

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you provide in any way it believes appropriate without incurring any obligation to you.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
US

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

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

The client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions.

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

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

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

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

Each copy or any portion of these sample programs or any derivative work must include a copyright notice as follows:

© (your company name) (year).
Portions of this code are derived from IBM Corp. Sample Programs.
© Copyright IBM Corp. _enter the year or years_.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at "Copyright and trademark information" at www.ibm.com/legal/copytrade.shtml.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

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

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

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

Other company, product, and service names may be trademarks or service marks of others.

Terms and conditions for product documentation

Permissions for the use of these publications are granted subject to the following terms and conditions.

Applicability

These terms and conditions are in addition to any terms of use for the IBM website.

Personal use

You may reproduce these publications for your personal, noncommercial use provided that all proprietary notices are preserved. You may not distribute, display or make derivative work of these publications, or any portion thereof, without the express consent of IBM.

Commercial use

You may reproduce, distribute and display these publications solely within your enterprise provided that all proprietary notices are preserved. You may not make derivative works of these publications, or reproduce, distribute or display these publications or any portion thereof outside your enterprise, without the express consent of IBM.

Rights

Except as expressly granted in this permission, no other permissions, licenses or rights are granted, either express or implied, to the publications or any information, data, software or other intellectual property contained therein.

IBM reserves the right to withdraw the permissions granted herein whenever, in its discretion, the use of the publications is detrimental to its interest or, as determined by IBM, the above instructions are not being properly followed.

You may not download, export or re-export this information except in full compliance with all applicable laws and regulations, including all United States export laws and regulations.

IBM MAKES NO GUARANTEE ABOUT THE CONTENT OF THESE PUBLICATIONS. THE PUBLICATIONS ARE PROVIDED "AS-IS" AND WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, NON-INFRINGEMENT, AND FITNESS FOR A PARTICULAR PURPOSE.

IBM Online Privacy Statement

IBM Software products, including software as a service solutions, (“Software Offerings”) may use cookies or other technologies to collect product usage information, to help improve the end user experience, to tailor interactions with the end user or for other purposes. In many cases no personally identifiable information is collected by the Software Offerings. Some of our Software Offerings can help enable you to collect personally identifiable information. If this Software Offering uses cookies to collect personally identifiable information, specific information about this offering’s use of cookies is set forth below.

This Software Offering does not use cookies or other technologies to collect personally identifiable information.

If the configurations deployed for this Software Offering provide you as customer the ability to collect personally identifiable information from end users via cookies and other technologies, you should seek your own legal advice about any laws applicable to such data collection, including any requirements for notice and consent.

For more information about the use of various technologies, including cookies, for these purposes, see IBM’s Privacy Policy at www.ibm.com/privacy and IBM’s Online Privacy Statement at www.ibm.com/privacy/details the section entitled “Cookies, Web Beacons and Other Technologies” and the “IBM

Software Products and Software-as-a-Service Privacy Statement" at www.ibm.com/software/info/product-privacy.

Index

D

E

F

G

H

I

J

L

M

N

O

P

Q

R

S

T

U

V

W

Product Number: 5724-J33

SC19-3351-03