Content Manager OnDemand for Multiplatforms
Version 8 Release 5

Introduction and Planning Guide

GC19-2938-00

Content Manager OnDemand for Multiplatforms
Version 8 Release 5

Introduction and Planning Guide

GC19-2938-00

Note
Before using this information and the product it supports, read the information in "Notices" on page 105.

This edition replaces GC18-9236-02.

This edition applies to IBM Content Manager OnDemand for Multiplatforms, Version 8 Revision 5 (program number 5724-J33) and to all subsequent releases and modifications until otherwise indicated in new editions.

This product includes software developed by the Apache Software Foundation (http://www.apache.org/).

Contents

About this publication

This book contains information about IBM® Content Manager OnDemand for Multiplatforms Version 8 Release 5 (Content Manager OnDemand). The Content Manager OnDemand system requires a database management product, such as IBM DB2 Universal Database™ (DB2®). If you plan to copy or backup data to optical or tape storage volumes, Content Manager OnDemand requires an archive storage manager, such as IBM Tivoli® Storage Manager. If you plan to use the Content Manager OnDemand server print or server FAX functions, you must install IBM Infoprint Manager (Infoprint) on a server that is on the same network as the Content Manager OnDemand library server.

Note: In this publication, the term *Windows® client* refers to the Content Manager OnDemand client program that runs under Windows XP, Windows Vista, and Windows Server 2003; the term *Windows server* refers to the Content Manager OnDemand server program that runs under Windows 2000 Server and Windows Server 2003.

Who should use this publication

This book is of primary interest to administrators planning to install, administer, and use Content Manager OnDemand and other people in an organization who plan hardware, software, network, recovery, and applications for business systems.

How this publication is organized

This publication is organized into the following parts:

- Part 1, "Introduction," on page 1 provides an overview of Content Manager OnDemand. This part contains the following sections:

  Chapter 1, "About Content Manager Content Manager OnDemand," on page 3 provides an overview of the system

  Chapter 2, "Preparing for Content Manager Content Manager OnDemand," on page 21 contains information that can help you prepare your organization for Content Manager OnDemand, lists the types of administrative tasks required to maintain the system, and describes several ways that you can configure how the server and clients operate

  Chapter 3, "Content Manager OnDemand and Tivoli Storage Manager," on page 29 illustrates how Content Manager OnDemand works with Tivoli Storage Manager to maintain documents in archive storage

- Part 2, "System requirements," on page 35 includes Chapter 4, "Disk storage," on page 37 and provides information that can help you plan the disk storage devices required to support the system.

  For hardware and software requirements, see http://www.ibm.com/support/docview.wss?rs=129&uid=swg27016455 or search for 7016455 at http://www.ibm.com/.

- Part 3, “Planning information,” on page 43 contains information that can help you define reports[1] to Content Manager OnDemand, index data, estimate storage requirements, and plan for backup and recovery of data on the system. This part contains the following sections:

  Chapter 5, “Reports and other data,” on page 45 contains information that can help you plan for the reports that you will be storing on the system

  Chapter 6, “Content Manager OnDemand objects,” on page 55 contains information that can help you plan application groups, applications, and folders for your reports

  Chapter 7, “Storage requirements,” on page 69 provides information that can help you estimate the amount of storage required to maintain reports on the system

  Chapter 8, “Backup and recovery,” on page 95 provides information that you can use to develop a backup and recovery plan for the system

The “Glossary” on page 109 contains terms and definitions that you might find helpful as you and others in your organization learn about and use Content Manager OnDemand.

ibm.com and related resources

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

Support and assistance

Product support is available on the Web. Click Support from the product Web site at:

Content Manager OnDemand for Multiplatforms
www.ibm.com/software/data/ondemand/mp/support.html

Information Center

You can view the product documentation in an Eclipse-based information center that you can install when you install the product. By default, the information center runs in a Web server mode that other Web browsers can access. You can also run it locally on your workstation. See the information center at: www.ibm.com/software/data/ondemand/mp/support.html

PDF publications

You can view the PDF files online using the Adobe® Acrobat Reader for your operating system. If you do not have Acrobat Reader installed, you can download it from the Adobe Web site at www.adobe.com.

You can find PDF publications for IBM Content Manager OnDemand for Multiplatforms at: http://www.ibm.com/support/docview.wss?rs=129&uid=swg27012713

Accessibility information for OnDemand

For complete information about accessibility features that are supported by this product, see your Content Manager OnDemand *Administration Guide*.

1. In Content Manager OnDemand, the term *report* refers to any type of data that you want to store in the system. A 10,000-page general ledger generated by a z/OS application and a two-page Lotus® WordPro file are both reports to Content Manager OnDemand.

How to send your comments

Your feedback helps IBM to provide quality information. Please send any comments that you have about this publication or other OnDemand documentation. Visit the IBM Data Management Online Reader's Comment Form (RCF) page at www.ibm.com/software/data/rcf.

Be sure to include the name of the product, the version number of the product, and the name of the book. If you are commenting on specific text, please include the location of the text (for example, a chapter and section title, a table number, a page number, or a help topic title).

If you would like to help IBM make IBM Content Manager OnDemand for Multiplatforms easier to use, take the consumability survey at http://www.ibm.com/software/data/info/consumability-survey/.

What's new in Version 8.5

Lightweight Directory Access Protocol (LDAP) Secure Sockets Layer (SSL) support
You can now use LDAP with SSL.

Report Distribution enhancements
IBM Content Manager OnDemand Report Distribution for Multiplatforms has been enhanced to allow support for additional data type conversion engines, that is, from line data to PDF.

Enhanced reporting and analysis for managing the Content Manager OnDemand system
Content Manager OnDemand now provides ability to analyze system log data and generate daily activity reports.

Improved capability and productivity for indexing PDF data into Content Manager OnDemand
You can now index PDF documents using metadata values contained within the PDF document.

Expanded security capabilities to better comply with Federal Information Processing Standards (FIPS 140-2)
Content Manager OnDemand has added additional security to better comply with FIPS 140-2.

Additional language support
The Content Manager OnDemand server and ODWEK installation programs are translated into 9 additional languages. The Content Manager OnDemand client installation program is translated into 22 additional languages. The Content Manager OnDemand Configurator is translated into 9 additional languages.

HP-UX Itanium support for DB2 and Oracle databases
Content Manager OnDemand now supports HP-UX Itanium for DB2 and Oracle databases.

ARSXML system administration performance improvement
Performance improvements have been made to the ARSXML system administration command.

Install Anywhere replacing ISMP for servers
InstallAnywhere 2009 SP1 has replaced InstallShield Multiplatform (ISMP)

as the installer engine for Content Manager OnDemand server and ODWEK on Windows and Unix platforms.

Part 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[2]. This section also contains a list of the tasks that Content Manager OnDemand administrators typically perform to manage an Content Manager OnDemand system.

2. In Content Manager OnDemand, the term *document* refers to an indexed part of a report, such as a statement, policy, or other logical grouping of pages.

Chapter 1. About Content Manager Content Manager OnDemand

Overview

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. An 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

An 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. Figure 1 shows an example.

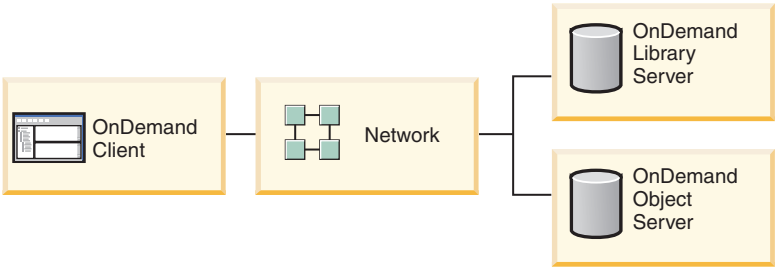

Figure 1. Content Manager OnDemand system

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 client programs operate on personal computers running Windows[3]. 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.

3. 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. See "Content Manager OnDemand Web Enablement Kit" on page 14 for more information.

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.

Concepts

The terms *application*, *application group*, *folder*, and *cabinet* describe how Content Manager OnDemand stores, manages, retrieves, views, and prints reports, and index data. 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. (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.

Application

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 will be 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 Group

An application group contains the storage management attributes of 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 may 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.

Figure 2 on page 6 illustrates the concepts described in this section.

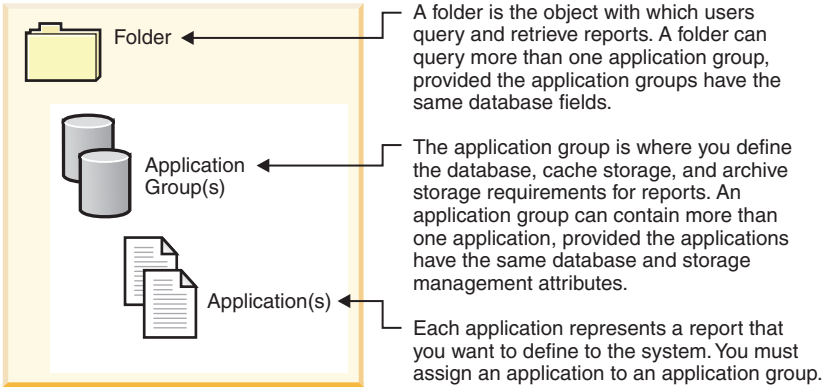

Figure 2. Folders, application groups, and applications (part 1 of 2)

Figure 3 shows an example.

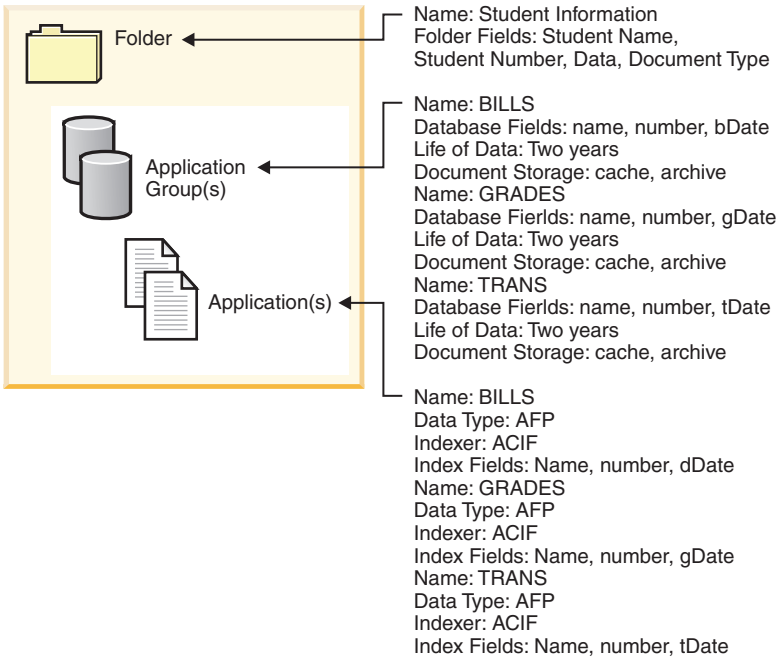

Figure 3. Folders, application groups, and applications (part 2 of 2)

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.

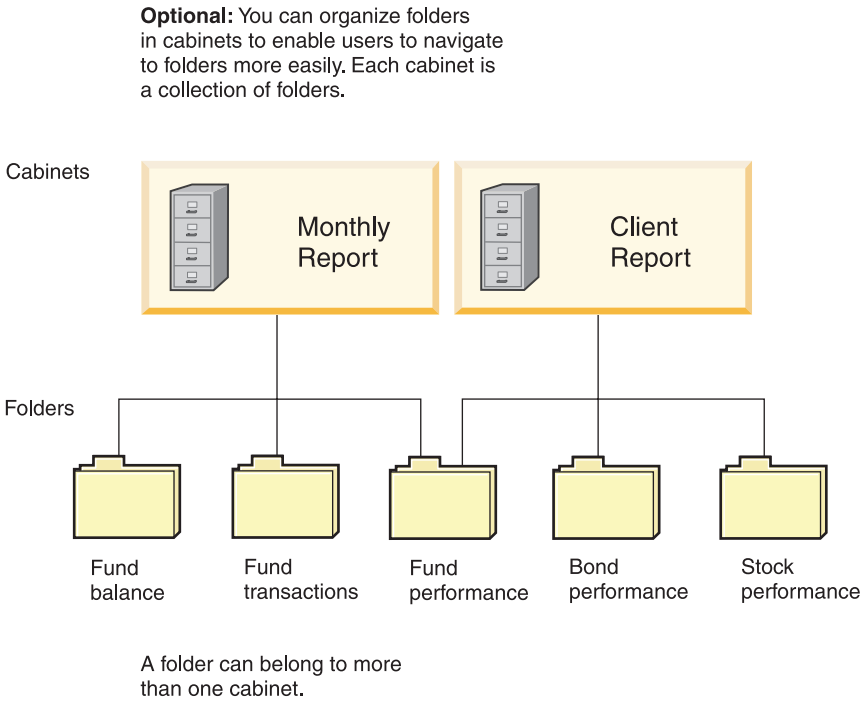

Figure 4. Cabinets are used to organize folders

Indexing methods

Content Manager OnDemand provides two basic ways to index data:

- Document indexing is used for reports that contain logical items such as policies, and statements. 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 32 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.
- Report indexing is used for reports that contain many pages of the same kind of data, such as a transaction log. 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.

Figure 5 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 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 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. OnDemand retrieves the group that contains the value entered by the user.

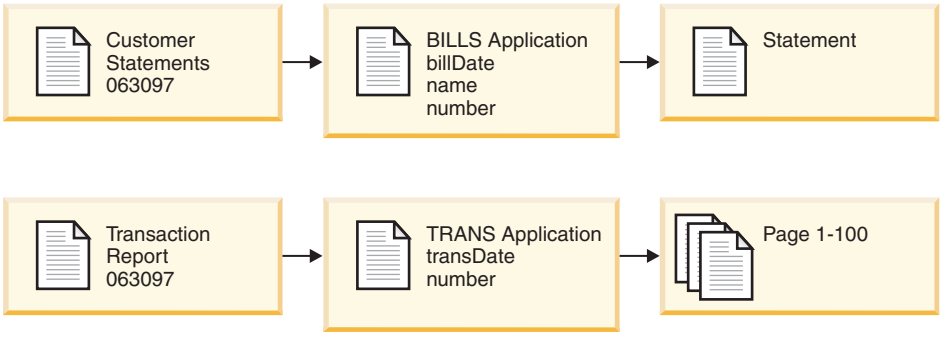

Figure 5. Applications and documents

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, 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, OnDemand data loading and maintenance programs, and optionally, the archive storage manager.

The basic 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 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, 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 Print Services Facility™ (PSF) for z/OS. Download provides the automatic, high-speed download of JES spool files from an z/OS system to Content Manager OnDemand servers.
- *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 Generic indexer can be used to create index data for almost any type of data that you want to store on the system, such as Hypertext Markup Language (HTML) files, Lotus WordPro files, compressed and uncompressed Tagged Image File Format (TIFF) images, and so on.
  - The Content Manager OnDemand IBM Content Manager OnDemand PDF Indexer for Multiplatforms can be used to create index data for Adobe Acrobat PDF input files.

  The indexing programs may be run on the library server or an object server. ACIF may 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 may 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.

OnDemand also supports an archive storage manager, such as Tivoli Storage Manager (which is provided with OnDemand). 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 an IBM Content Manager OnDemand for Multiplatforms 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 an 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. See the *IBM Content Manager OnDemand for Multiplatforms: Indexing Reference* for details about the indexing programs provided with Content Manager OnDemand.

Figure 6 shows an overview of the data preparation process.

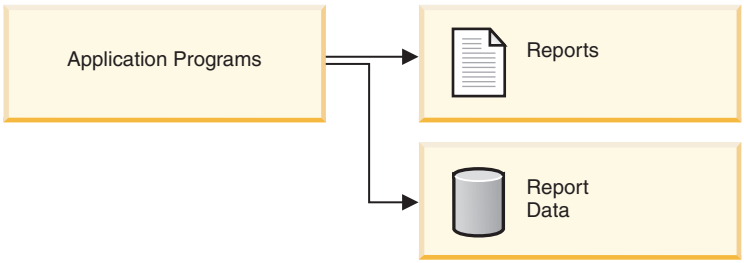

Figure 6. Data preparation, indexing, and loading (part 1 of 2)

In the picture, user-defined application programs generate printed reports and save report data to disk. The report data can be transmitted to an 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 an Content Manager OnDemand server.

Figure 7 on page 13 shows an overview of the data indexing and loading process.

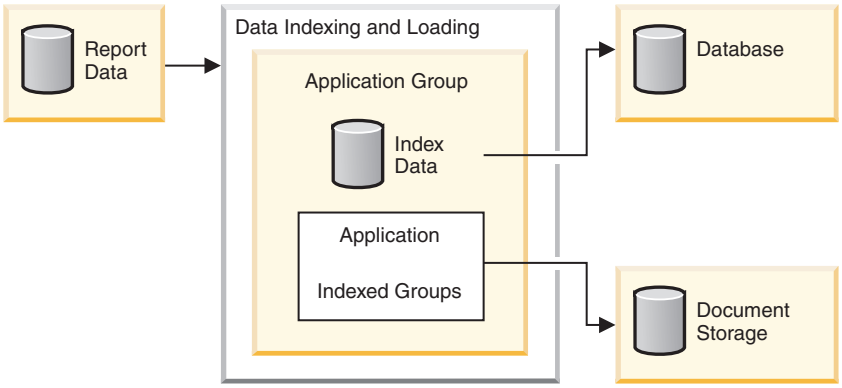

Figure 7. Data preparation, indexing, and loading (part 2 of 2)

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 may 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.

Remote library server

The Content Manager OnDemand Windows server configurator program can be used to add an instance of Content Manager OnDemand that is running on some other system. This feature supports using the IBM DB2 Content Manager CommonStore product on a Windows system to access an Content Manager OnDemand library server that is running on any other supported platforms. To configure the system to use a remote library server, you must install the Content Manager OnDemand software on the system that is running the CommonStore software and create a Content Manager OnDemand instance or archive. Do the following to configure a remote instance on the local OnDemand server:

1. Open the Content Manager OnDemand configurator on the local Windows server. Select **OnDemand for Windows Server**, **File** and **New Server**,
2. On the next screen, specify the **Server** and **Hostname** of the local Windows server, for example, **virtual1**, **virtual1**. Select **Local Windows Server**. This is an important step even though this is a remote server. Click **OK**.
3. From the list of **OnDemand for Windows servers**, click + next to the new local server you just created (for example, **virtual1**). Select **Instances**. On the right panel right-click on any instance name (you should already have one, for example, archive), then click **New** on the pull down menu.
4. On the next screen, type any name of your choice. This will be the name of a virtual instance, for example **vodinst**, pointing to the remote instance. Click **Next**.
5. On the new screen, select **Remote Library Server** and type the **Library Server Name** with the **hostname**, **alias**, or **IP address**. Click **Communications**.
6. On the next screen, type the port number of the remote instance on the remote library server. (This can be found in the ars.ini file on the remote server.) Click **OK** and **Finish**.

After a remote instance is configured on a Windows system, you can run the Content Manager OnDemand server programs (such as ARSDB, ARSDOC, and ARSLOAD) on the remote instance by issuing a command on Windows. For example: arsload -I vodinst -g application_group_name ... where **vodinst** is the name used in step 4). There is no need to specify the hostname, since Content Manager OnDemand only searches for the instance name under the local Content Manager OnDemand server, and the configuration process (steps 5 and 6 above) specifies the hostname and port number of the remote library server.

Content Manager OnDemand Web Enablement Kit

The Content Manager OnDemand Web Enablement Kit (ODWEK) allows users to access data that is stored in an Content Manager OnDemand server by using a Web browser or a user-written program. For example, you can provide some people with the Uniform Resource Locator (URL) of a Web page that permits them to log on to an Content Manager OnDemand server; you can provide other people

with the URL of a Web page that permits them to search a specific folder. 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, ODWEK displays a Web page that contains a list of the documents that match the query. The user selects a document to view and ODWEK sends the document to the browser.

Figure 8 shows a workstation with a Web browser that is being used to access data from an Content Manager OnDemand server.

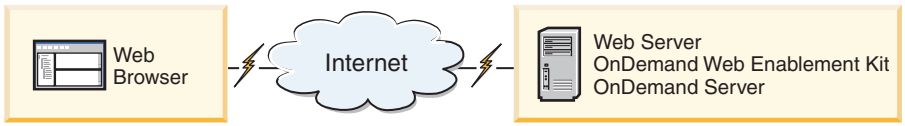

Figure 8. Accessing data stored in Content Manager OnDemand using ODWEK

ODWEK can search for and retrieve documents from Content Manager OnDemand servers that are running, IBM Content Manager OnDemand for Multiplatforms, Version 8.5, and IBM Content Manager OnDemand for z/OS, Version 8. **Note:** ODWEK can also search for and retrieve documents from an Content Manager OnDemand for OS/390® Version 2.1 server. However, see *IBM Content Manager OnDemand for Multiplatforms: Web Enablement Kit Implementation Guide* for a list of limitations in this environment.

ODWEK contains several components:

- The Content Manager OnDemand programming interface. The programming interface uses standard Content Manager OnDemand interfaces and protocols to access data stored in an Content Manager OnDemand server. No additional code is needed on the Content Manager OnDemand server to support ODWEK. You may use one of the following programming interfaces in your ODWEK application:
  - CGI program. The CGI program provides a way to access OnDemand data from a Web browser. The CGI program runs on a workstation that is running an HTTP server, such as the IBM HTTP Server.
  - Java™ servlet. The servlet provides a way to access OnDemand data from a Web browser. The servlet runs on a workstation that is running a Java-enabled HTTP server (running as a Java application server), such as the IBM WebSphere® Application Server. The servlet requires Java version 1.5 or later.
  - Java Application Programming Interface (Java API). The Java API provides a way to access OnDemand data from a user-written program. The Java API requires Java version 1.5 or later.
- The IBM OnDemand AFP Web Viewer. The AFP Web Viewer lets users search, retrieve, view, navigate, and print AFP documents from a Web browser.
- The IBM OnDemand Image Web Viewer. The Image Web Viewer lets users search, retrieve, view, navigate, and print BMP, GIF, JPEG, PCX, and TIFF documents from a Web browser.
- The Line Data Java applet. The Line Data applet lets users view line data documents from a Web browser.
- The AFP2HTML Java applet. The AFP2HTML applet lets users view the output generated by the ODWEK via the AFP2HTML Transform. For more information, see *IBM Content Manager OnDemand for Multiplatforms: Web Enablement Kit Implementation Guide*. The AFP2HTML Transform converts AFP documents and

resources into HTML files that can be displayed with the AFP2HTML applet. After installing and configuring the Content Manager OnDemand Advanced Function Presentation Transformations for Multiplatforms, an administrator enables the use of the AFP2HTML applet by configuring the ARSWWW.INI file.

Viewing and transforming documents

To view other types of documents that are stored in 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 OnDemand into PDF documents that you can view with the Adobe Acrobat viewer requires either the AFP2PDF transform or the Xenos d2e Platform. See your IBM representative for more information about these optional priced transforms.

To convert AFP, line data, or Metacode documents that are stored in OnDemand into HTML or Extensible Markup Language (XML) files that you can view with a Web browser requires the Xenos d2e Platform. See your IBM representative for information.

To convert line data documents that are stored in OnDemand into AFP documents that you can view with the AFP Web Viewer requires the Xenos d2e Platform. To convert line data documents that are stored in OnDemand into PDF documents that you can view with the Adobe Acrobat viewer requires the Xenos d2e Platform. See your IBM representative for information.

See *IBM Content Manager OnDemand for Multiplatforms: Web Enablement Kit Implementation Guide* for more information.

IBM Content Manager OnDemand Report Distribution for Multiplatforms

IBM Content Manager OnDemand Report Distribution for Multiplatforms is an optional priced feature of IBM Content Manager OnDemand for Multiplatforms. Report Distribution provides an easy way to automatically group reports and portions of related reports together, organize them, convert the report data into different formats, and send them through e-mail to multiple users or make them available for printing.

A *report* in report distribution is a document or a set of documents that is retrieved from the OnDemand system to be delivered to one or more users. A report can be e-mailed to the users or sent to their default server printers. Before you retrieve a report, the documents must have been loaded into OnDemand, and they might be in one of the following file formats:

- AFP
- Line data
- Unformatted ASCII data
- PDF

You can use one of the following methods to retrieve reports:

- Load. Building a list of documents based on the documents that are loaded during a specific time frame

- Named Query. Performing a database query using a public named query that was defined by the OnDemand client
- SQL. Performing a database query using an SQL query

After you build your report queries based on load or construct the database query, OnDemand retrieves the documents that meet your requirements. If a report query returns multiple documents from different application groups, the entire set of documents appear in the same report. You cannot specify or change the order in which the documents are retrieved, because it is optimized for performance.

Report Distribution introduces the following concepts:

Banner A banner contains information about the distribution, contents in the distribution, and the user ID that sends out the report distribution. Information about the user ID is defined when you added that user ID. Banners are optional in a report distribution.

Bundle A bundle contains at least one or more reports, and might optionally include banners and a manifest. Reports can be ordered within the bundle and converted to different data types. Multiple copies of the same report can exist in the same bundle.

Manifest A manifest is a list of the reports in a bundle. The manifest is always generated for inclusion into the OnDemand system log, however, it is an optional component of a bundle.

Schedule A schedule determines when and how often OnDemand sends out a distribution. A schedule can be time-based or load-based. You can set OnDemand to send out a distribution once, daily, weekly, or monthly. If you set the schedule to be load-based, OnDemand sends out the distribution as soon as reports that are required for the distribution is loaded into the system.

Recipient A recipient is a user or group of users assigned to receive reports using report distribution. When you create a distribution, you assign who should receive the reports contained in the bundle. Because one recipient in a recipient list can be a group of users, it is possible that a bundle can distribute reports to several users while identifying only one recipient. So, a recipient list can contain a combination of individual users and groups. Recipient and recipient list is associated only with distributions, not any other object in OnDemand.

Distribution A distribution consists of a set of reports that are contained in a bundle, one or more recipients to receive the reports, and a schedule that specifies when the distribution is delivered. All of the recipients receive all of the same reports in the same format. You can schedule a distribution by defining the schedules in report distribution. A distribution includes:

- Distribution name
- A bundle
- One or more recipients
- (Optional) Schedules

A bundle might include the following objects:

- A header banner
- Reports

- Separator banners between reports
- A manifest
- A trailer banner

See *IBM Content Manager OnDemand for Multiplatforms: Report Distribution Installation, Use, and Reference* for more information.

Why is Report Distribution useful?

OnDemand allows you to store several reports which you can create, open, modify, and access with the OnDemand administrative client. The Report Distribution feature automates distributing the most relevant reports to those who need them.

Report distribution allows you to:

- Group reports and portions of related reports into a bundle
- Convert the report data into different formats
- Distribute reports to multiple users
- Distribute single or multiple copies of each report
- Send reports by e-mail or make them available for printing
- Notify users of a delivered report

With Report distribution, you can conduct tasks such as print monthly sales statements, e-mail all documents related to an annual review of a customer's policy, or designate a schedule that distributes reports when something has been updated.

How does Report Distribution work with OnDemand?

You need to have the OnDemand client and administrative client installed on your machine and populated with documents. See *IBM Content Manager OnDemand for Multiplatforms: Report Distribution Installation, Use, and Reference* for the specific set up needs.

Report distribution has three main steps:

1. Report extraction
2. Bundle creation
3. Distribution delivery

Report extraction

Report distribution retrieves reports based on parameters that you supply. The parameters of the retrieval are defined in a report type query. The ARSRD program (Report Distribution program) provides three methods to extract documents from the OnDemand system. These methods are the public named query, the SQL query, and Load.

The public named query report type is created in the OnDemand client. A named query can retrieve one document, or, it can retrieve a group of documents from multiple application groups at once, depending on the public named query that you use.

An SQL query report type is created in the OnDemand administrative client. It retrieves one or more documents from one application group. The benefit for using this query over the named query is the fact that you can search all application group fields whereas the named query is limited to the defined folder fields.

The difference between the SQL query and Load is that SQL query retrieves documents on a time-based schedule while Load retrieves documents upon loading of the documents.

Load checks if documents have been loaded for the application group that is specified in the report definition since the last time the check was made. The amount of time to wait between checks is defined by the value called "Number of minutes between schedule searches" on the Report Distribution parameters window in the administrative client.

Bundle creation

After you define what reports to distribute, you create a bundle, which is a method to organize the reports. The bundle allows you to provide separators between each report, a header banner, and trailer banner. You even have the choice as to whether or not to include a manifest of what is included in the bundle.

Distribution delivery

Now that you have found and organized your reports, you must decide who these reports go to and when they should receive them. You select the recipients from the users already defined to the OnDemand system. You can select two types of schedules: one based on a fixed date or one that is load-driven. Report distribution runs quietly in the background and is activated by time. When an active distribution is scheduled for a fixed time and day, report distribution retrieves the reports from OnDemand and distributes the bundle. If new documents have been loaded since the last time report distribution checked for newly loaded documents, the documents will be extracted, bundled, and delivered.

You can e-mail distributions or elect to print them. The e-mail recipients must be defined to the OnDemand system. You do not need to have any specific print software to print reports, banners, or manifests.

See *IBM Content Manager OnDemand for Multiplatforms: Report Distribution Installation, Use, and Reference* for detailed information about the OnDemand Report Distribution feature.

Chapter 2. Preparing for Content Manager Content Manager OnDemand

This section contains an outline that you may find helpful as you prepare your organization for the Content Manager OnDemand environment and perform a pilot roll out of the system.

- Work with a single department or group of end-users. Send a memo to the users to explain how Content Manager OnDemand will affect 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.
- Obtain, install, and test any data transforms that you may need. For example, if you need to convert AFP documents to HTML files, it is recommended that you use either the IBM Content Manager OnDemand Advanced Function Presentation Transformations for Multiplatforms or the Xenos d2e Platform. See your IBM representative for information.
- Identify the type of data contained in the report and determine how you will create the index data.
- If you plan to store AFP or Metacode 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 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 OnDemand administrators are not the only people in an organization working on all these components.

Depending on the size of your organization, there may be one person or many people administering the system. If your organization is large, the administrative tasks may be divided among several people. For example, an Content Manager OnDemand system administrator could maintain Content Manager OnDemand storage sets, system printers, groups, and users; an 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 an Content Manager OnDemand system. Some of these tasks may be the responsibility of a person other than an 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 OnDemand's 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

Content Manager OnDemand provides information about the Object Linking and Embedding (OLE) control and how to customize the Windows client by specifying command line parameters, by invoking and manipulating Content Manager OnDemand from other Windows 32-bit applications with the Dynamic Data Exchange (DDE) interface, or by creating a Product Information File (PIF).

The *IBM Content Manager OnDemand: Windows Client Customization Guide* provides information about customizing the Windows clients.

For example, you can integrate Monarch[4] Version 5 with the Windows client so that users can load Content Manager OnDemand documents into Monarch. The user can then do complex data manipulation in Monarch, such as creating derived columns and generating charts and reports. See the *IBM Content Manager OnDemand: Windows Client Customization Guide* for more information.

Note that the only APIs that work in this instance are Object Linking and Embedding (OLE) on Win32, or Content Manager OnDemand Web Enablement Kit (ODWEK) on all platforms. You can also use arsdoc. For example, it is not possible to integrate OnDemand on AIX® to a non-Windows based system using another application programming interface.

Server programs

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

- The ARSXML administration utility enables you to perform any task that you can complete by using the administration client from the command line. You can use ARSXML administration utility to create, maintain, or delete applications, application groups, folders, groups, printers, storage sets, and users. You can also use the ARSXML administration utility to add, delete, and update 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 multi–purpose document processing program. You can use the ARSDOC program to query the library server and generate a list of

4. Monarch is a software program that is available from Datawatch Corporation.

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 OnDemand loads data into. During normal operation, 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 may 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 may 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 ARSLOG.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 *IBM 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 OnDemand. For example, you may want to deny access to the system after three incorrect logon attempts are made by a user; you may 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 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 OnDemand administrative functions
- Access to an OnDemand folder
- Access to an 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 *IBM Content Manager OnDemand for Multiplatforms: Installation and Configuration Guide* for more information about the security user exit.

Retrieval preview user exit

The 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 *IBM 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 may 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 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 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 OnDemand system tables.

For table and index creation, the installation can alter the SQL that will be used to create the table or index.

See the *IBM 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 may need a license for each concurrent connection to the database. Contact your Microsoft® representative for more information.

Chapter 3. Content Manager OnDemand and Tivoli Storage Manager

Introduction

You can configure an 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. Tivoli Storage Manager is the product that Content Manager OnDemand works with 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.

Content Manager OnDemand 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 shows the primary Content Manager OnDemand storage objects.

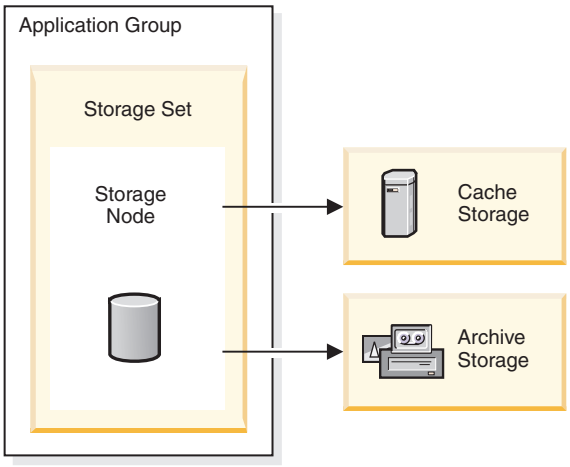

Figure 9. Content Manager OnDemand storage objects

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. 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. 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*.

Storage policy

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

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.

Figure 10 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.

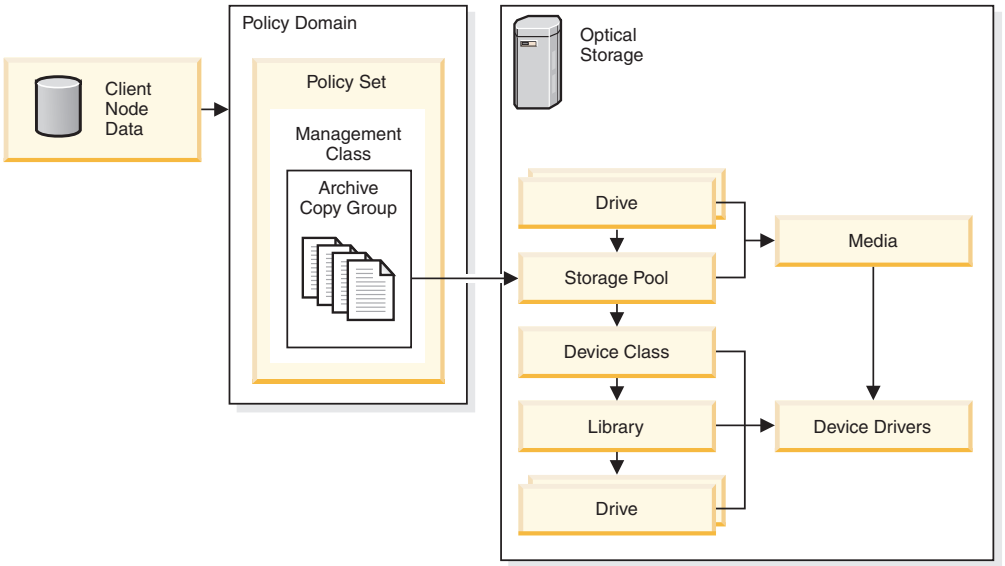

Figure 10. Tivoli Storage Manager storage objects

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.

For more information about how to determine your storage needs, see Chapter 7, "Storage requirements," on page 69.

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 an 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.

Operational considerations

Each storage set identifies the object server and storage node(s) 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. However, 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.

64-bit operating systems

OnDemand now supports UNIX 64-bit operating systems and databases. If you run OnDemand and use a DB2 Universal Database Version 8 32-bit server, you must migrate your data to 64-bit instances. See one of the following topics for information on how to migrate your instances:

- If you are running OnDemand on a Linux® or UNIX server, see "Migrating DB2 32-bit servers to 64-bit systems (Linux and UNIX)" in the IBM DB2 Database for Linux, UNIX, and Windows Information Center.
- If you are running OnDemand on an Oracle server, see "Migrating your data to 64-bit instances" in the *IBM Content Manager OnDemand for Multiplatforms: Installation and Configuration Guide*.

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 will delete the application group table(s) from the Content Manager OnDemand database but will 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.

Part 2. System requirements

This section includes the chapter on Chapter 4, “Disk storage,” on page 37 and provides information that can help you plan the disk storage devices required to support the system.

For hardware and software requirements, see http://www.ibm.com/support/docview.wss?rs=129&uid=swg27016455 or search for 7016455 at http://www.ibm.com/.

Chapter 4. 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. See Chapter 7, "Storage requirements," on page 69 contains information that can help you estimate the storage requirements for 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 may 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 may need to configure your disk storage devices into groups of volumes. For example, you may 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 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

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 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.

You should organize your disk storage devices into groups of storage volumes. 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 1 shows one way to configure disk storage into groups on a server that will support a large organization.

Table 1. 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
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 2 shows one way to configure disk storage devices on a server that will support a small workgroup.

Table 2. 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

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

You should plan to use at least ten disk drives. Use your sizing calculations to determine the size of these disk drives. You may need more or fewer 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). Table 3 shows one way to configure disk storage into groups on a server that will support a large organization.

Table 3. Disk storage groups for a large organization

Physical Disk	Partition	File System(s)	Purpose
Disk1	C:	\WIN32	Windows system files
Disk1	D:	\Program Files	OnDemand and other applications
Disk2	E:	\arsdb	OnDemand database
Disk3	F:	\arsdbpri, \arsdbarc	Database log files
Disk4	G:	\arscache1	Cache storage
Disk5	H:	\arscache2	Cache storage
Disk6	I:	\arscache3	Cache storage
Disk7	J:	\arscache4	Cache storage
Disk8	K:	\arsacif1	Data download and indexing
Disk9	L:	\arsload1	Data loading
Disk10	M:	\arstmp	Temporary storage

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

You should plan to use at least ten disk drives. Use your sizing calculations to determine the size of these disk drives. You may need more or fewer drives, depending on how much data that you need to maintain in cache storage (and how long you need to maintain the data).Table 4 on page 40 shows one way to configure disk storage devices on a server that will support a small workgroup.

Table 4. Disk storage groups for a small workgroup

Physical Disk	Partition	File System(s)	Purpose
Disk1	C:	\WIN32, \Program Files	Windows system files, OnDemand and other applications
Disk2	D:	\arsdb1	OnDemand database
Disk3	E:	\arscache1	Cache storage
Disk4	F:	\arsacif1	Data indexing and loading
Disk5	G:	\arsload1	Data downloading
Disk6	H:	\arsdbpri, \arsdbarc, arstmp	Database log files, database archive log files, 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 5 provides an overview of RAID implementations: http://publib16.boulder.ibm.com/doc_link/en_US/a_doc_lib/aixbman/prftungd/diskperf14.htm

Table 5. 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.

Table 5. RAID implementations (continued)

RAID Level	Description	Protection	Performance
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 OnDemand systems, use the RAID 1 or RAID 5 implementations.

RAID for cache storage

A typical use of RAID storage on an 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 may be difficult or impossible to recreate.

RAID for the Content Manager OnDemand database

You can store the Content Manager OnDemand database in one or more RAID storage subsystems. However, depending on the exact hardware and the implementation level of the RAID devices, the system may 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 may not attain the same level of performance as a system with the database on non-arrayed storage. (On the other hand, it may. 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.) However, the availability benefits provided by RAID storage subsystems typically outweigh any performance degradation that the users may experience. See your database and storage system specialists for help with configuring and using RAID storage devices.

Part 3. Planning information

This section is a planning source for Content Manager OnDemand administrators. Other people in an organization interested in this section may 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.

Chapter 5. 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 questions that you might ask users of the reports, provide information about the types of data that you can store in OnDemand, and provide information about indexing reports.

Collecting requirements

Planning for Content Manager OnDemand requires that you understand how the system will be deployed, who will use the system and how they will use it, 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:

- Will you operate a single Content Manager OnDemand server or a network of Content Manager OnDemand servers?
- What types of print data streams will the system support? Are transforms required to convert input data to other display formats (such as AFP to HTML)?
- 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 may not have a consistent format, such as reference materials or product literature.
- Will 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. 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.
- PCL (Printer Control Language). The term PCL refers to the compound data stream that is used by the Hewlett Packard (HP) printers. The input can be most PCL 4 or 5 that is designed for HP desktop printers, but not the HP PCL or HP Deskjet formats. OnDemand supports PCL input files by using the optional Xenos transforms.
- Xerox metacode. A print file designed to print on a Xerox printer can be converted to a fully-composed AFP datastream (MO:DCA-P). The input can be DJDE-Conditioned Line Data, mixed mode and fully composed Xerox metacode files, including FRMs, IMGs and LGOs. OnDemand supports Metacode input files by using the optional Xenos transforms. For example, you can use a Xenos transform to convert a Xerox metacode print file into AFP documents that can be indexed and loaded into the system. Font Correlation Tables give you full control of what AFP fonts are used in the output. Xerox font character bitmaps can selectively be put in the output as images for greater fidelity.

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 Netscape Navigator 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 are 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 may 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 OnDemand). Your users can search for and retrieve the logical items using identifiers such as account number, customer name, and date.

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 for Multiplatforms: Indexing Reference* for details about and examples on using the indexing programs that are provided with 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 32 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. Figure 11 shows an example of a report file and document indexing.

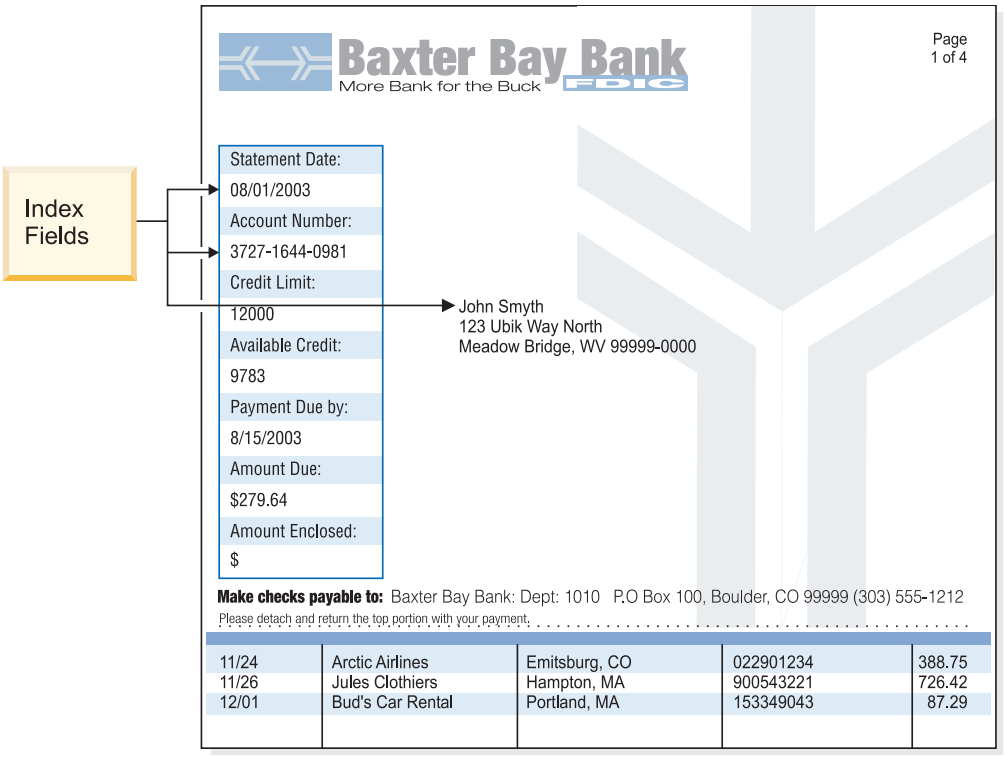

Figure 11. Document indexing method

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.Figure 12 on page 49 shows an example of a

report file and report indexing.

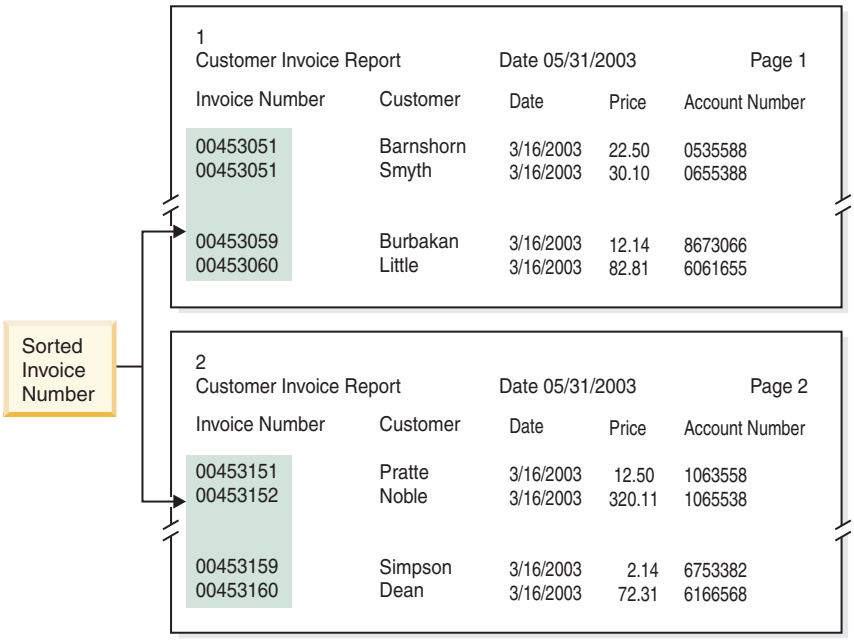

Figure 12. Report indexing method

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 32 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

Generating 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. However, you may find it necessary to modify your application program to add indexing functions.

Generating 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.

Inserting 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 IBM Content Manager OnDemand PDF Indexer for Multiplatforms

The Content Manager OnDemand IBM Content Manager OnDemand PDF Indexer for Multiplatforms 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 IBM Content Manager OnDemand PDF Indexer for Multiplatforms 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 IBM Content Manager OnDemand PDF Indexer for Multiplatforms supports PDF Version 1.6 and lower input and output files. See the documentation provided with Acrobat Distiller for more information about preparing input data for the Distiller.

The IBM Content Manager OnDemand PDF Indexer for Multiplatforms can logically divide reports into individual items, such as statements, policies, and bills. You can define up to 32 index fields for each item in a report.

The IBM Content Manager OnDemand PDF Indexer for Multiplatforms 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. OnDemand provides the Report Wizard in the Administrator to help you create indexing parameters for the IBM Content Manager OnDemand PDF Indexer for Multiplatforms. Content Manager OnDemand 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 *IBM Content Manager OnDemand for Multiplatforms: Indexing Reference* provides details about the IBM Content Manager OnDemand PDF Indexer for Multiplatforms and shows examples about how to use it to process PDF input files.

Indexing with the Xenos transforms

The Xenos transforms can be used to extract index data from input print files that contain AFP, Metacode/DJDE, or PCL data. You can use the Xenos transforms to convert Metacode to AFP and AFP, Metacode, and PCL to PDF. The Xenos transforms enhance your ability to view, archive, or retrieve individual pages or groups of pages from large print files.

The Xenos transforms accepts data from application programs in these formats:

- AFP data
- Metacode/DJDE print files
- PCL print files

The Xenos transforms can process application print data and resources to produce and index file and an output document file.

Note: When Xenos transforms a data stream with resources, it does not create new resources as separate entities in the output data stream. All of the resources that are required to view and print the document are contained within the generated document. Therefore, no separate resource file is produced.

With the files that the Xenos transforms create, you can store the data into Content Manager OnDemand and then do the following:

- Use the Windows client to search for and retrieve, view, and print the documents.
- Use the Content Manager OnDemand Web Enablement Kit feature to search for and retrieve, view, and print the documents.

To use the Xenos transforms on your system, you must obtain the transform program, license, and documentation from your IBM representative. Your IBM representative can also provide information about education that is available and other types of help and support for processing input files with the transform programs.

Using the Content Manager OnDemand 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. However, because the documents do not contain AFP data, line data, Metacode, PCL or PDF data, you cannot index them with ACIF, the IBM Content Manager OnDemand PDF Indexer for Multiplatforms, or Xenos. However, 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 data set that you want to store in and retrieve from Content Manager OnDemand. You specify the index data in a parameter data set. The parameter data set contains the index fields, index values, and information about the input data sets or documents that you want to process. The Generic indexer retrieves the indexing information from the parameter data set and generates the index data that is loaded into the database. Content Manager OnDemand creates one index record for each input data set (or document) that you specify in the parameter data set. The index record contains the index values that uniquely identify a data set 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 data set or document that you want to process with the Generic indexer. You can define up to 32 index fields for each data set or document.

See the *IBM Content Manager OnDemand for Multiplatforms: 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, 2069. OnDemand also supports a date/time field. A date/time field can contain date values from January 1, 1970 to December 31, 2038.

Running ACIF on z/OS Systems

Indexing reports on z/OS systems can provide certain benefits in a production archival process, including:

- The potential to balance processing and resources across the z/OS system and Content Manager OnDemand servers. If you index data as a step in the jobs that generate reports for Content Manager OnDemand, you can transmit report and index data directly to one or more Content Manager OnDemand servers. The Content Manager OnDemand servers can load the data into the appropriate application groups. Additional storage and processing cycles to do indexing on the Content Manager OnDemand server would not be required.
- The indexing program can collect resources for reports into a resource group file. You can transmit the resource group file to the Content Manager OnDemand server on which you plan to load the report or you can make the resource group file available to the OnDemand server using a directory service, such as NFS. If you index reports on an Content Manager OnDemand server, you must either transmit the individual resources to the Content Manager OnDemand server on which the report will be indexed or make the resource library available to the server that is indexing the report using a directory service such as NFS.

Running ACIF on Content Manager OnDemand servers

Indexing reports on an Content Manager OnDemand server can provide certain benefits in a production archival process, including:

- 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 an 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 6. Content Manager OnDemand 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.

An 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 an 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 the 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. In normal use, documents in OnDemand are retained for a specified time, after which they are removed. However, under some circumstances, you may need to keep certain documents beyond the normal retention period. For example, in some business sectors companies are required by law to maintain certain classes of documents for a specified length of time.

You can enable document retention by placing a 'hold' on documents to prevent them from being deleted. Documents on which holds are placed are never deleted, and if you no longer want to retain them you delete the hold. You can place holds on individual documents or all documents in a load. Documents can later be added to existing holds.

You create holds by using the Apply Holds to Document(s) menu option. For each hold you specify the name of the hold, its description, and access permissions for individual users and groups.

You grant permissions for holds to users, groups, and *PUBLIC. By default, no permission is granted to *PUBLIC. Using the Permissions tab on the Add a Hold dialog box, you select specific users or groups and apply either 'Access' or 'Administrator' permissions to them. Any granted permission can later be revoked. Access permissions enable users or groups to view, search, or summarize hold objects. Administrator permissions enable users or groups to update, delete, or view hold objects.

You can search for holds either by name or by description. You can also create hold reports, which summarize the details of the hold object.

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 32 *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 32 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 may be organized differently, and may 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 may 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

An 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 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 Table 6, a user assigned to both groups can access the Student Bills and Student Transcripts folders.

Table 6. 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, cabinet, and application group permissions when you add or update a folder, cabinet, or application group. You can also set folder, cabinet, 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.

Notes:

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.

When naming a user, the name that you specify:

- Can contain from one to 128 characters (bytes)
- Cannot include the ' (apostrophe), * (asterisk), % (percentage) + (plus), _ (underscore), [(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)

Note: 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)

  Notes:

  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 the user types more than eight characters, the client will display an error message unless the SRVR_FLAGS_IGNORE_PASSWD_MAX parameter in the ARS.INI file is set to 1 (one). The default value for the SRVR_FLAGS_IGNORE_PASSWD_MAX parameter is 0 (zero).
  3. 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.

4. 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)

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

The 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

In addition, for customers that are upgrading to Version 7.1.1 from Version 7.1.0.6 or earlier, any existing user who has never changed their password will be prompted to change the password the first time that they log on to the system after the upgrade is completed.

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

When naming a group, the name that you specify:

- Can contain from one to 128 characters (bytes)
- Cannot include the ' (apostrophe), * (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), _ (underscore) , [(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
- Cannot include the ' (apostrophe), % (percentage), _ (underscore) [(left bracket),] (right bracket), or " (double quote) characters
- 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

Table 7 lists the types of application group and folder fields supported by Content Manager OnDemand.

Table 7. Application group and folder field types

Field Type	Description
Small Integer	Contains whole numbers between –32,767 and 32,767
Integer	Contains whole numbers between –2147483648 and 2147483647
Big Integer	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. **Note:** DB2 and SQL Server support the Big Integer data/field type.
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, 1970 to December 31, 2069. If you need to index reports that contain dates that occur before January 1, 1970, you must define the date as a String field. Content Manager OnDemand checks a date value to make sure it is valid.
Time	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, 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, 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.

Table 8 on page 67 lists additional types of fields that are supported in folders.

Table 8. 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 text 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. 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.
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. **Note:** The (sequential) text search takes place on the server. A text search will delay the generation of the document list. Only documents that meet all of the criteria specified in the other folder fields will be 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 will 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 7. Storage requirements

Overview

Estimating storage requirements for an Content Manager OnDemand system begins with understanding and documenting end-user requirements for storing and accessing data. Chapter 5, “Reports and other data,” on page 45 provides information that can help you gather end-user requirements.

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 32 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 may 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 an 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 may 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 may 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 may 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 may be zero, for reports that do not require cache storage. However, a very large amount of disk space may 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.[5] 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 may 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.

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

5. For sorted transaction data, the examples and calculations that follow assume that Content Manager OnDemand will create 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 *IBM Content Manager OnDemand for Multiplatforms: Indexing Reference* provides move information.

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 may 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 may 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 may 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 will 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.

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 may 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 may 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 13. 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 14. 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

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 OnDemand can archive
- The amount of time that you need to store a report in 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. 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 may 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. 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 15. 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
```

Content Manager OnDemand 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

Table 9 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.

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

Table 9. Index field types and sizes

Field Type	Field Size (DB2)	Field Size (Oracle)	Field Size (SQL Server)
Small Integer	2 bytes	7 bytes	2 bytes
Integer	4 bytes	11 bytes	4 bytes
Big Integer	8 bytes	Not Supported	8 bytes
DecFloat (16)	8 bytes	not available	not available

6. 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 32 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 9. Index field types and sizes (continued)

Field Type	Field Size (DB2)	Field Size (Oracle)	Field Size (SQL Server)
DecFloat (34)	16 bytes	not available	not available
String (Fixed)	1 – 254; 1 byte per character for single-byte character sets, and up to 4 bytes per character for multi-byte character sets	1 – 254; 1 byte per character for single-byte character sets, and up to 4 bytes per character for multi-byte character sets	1 – 254; 1 byte per character for single-byte character sets, and up to 4 bytes per character for multi-byte character sets
String (Variable)	1 – 2000 • Single-byte character set: 1 byte per character, plus 4 bytes overhead • Multi-byte character set: up to 4 bytes per character, plus 4 bytes overhead Unused characters do not consume storage	1 – 2000 • Single-byte character set: 1 byte per character, plus 4 bytes overhead • Multi-byte character set: up to 4 bytes per character, plus 4 bytes overhead Unused characters do not consume storage	1 – 2000 • Single-byte character set: 1 byte per character, plus 4 bytes overhead • Multi-byte character set: up to 4 bytes per character, plus 4 bytes overhead Unused characters do not consume storage
Date	4 bytes	7 bytes	4 bytes
Time	4 bytes	11 bytes	4 bytes
Date/Time	4 bytes	11 bytes	4 bytes
Date/Time (TZ)	4 bytes	11 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.

Note: 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:

- Table 9 on page 75 lists the sizes of the various data types supported by Content Manager OnDemand.
- 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.

Figure 16 shows the calculation that you can use to estimate database space requirements when the database manager is DB2. The formula 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 16. Calculating database storage space for DB2

Figure 17 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 17. Calculating database storage space for Oracle

Figure 18 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 18. Calculating database storage space for SQL Server

Examples

Note: The examples that follow assume that the database manager is DB2. If Oracle is the database manager, use the calculations from Figure 17 on page 77; if SQL Server is the database manager, use the calculations from Figure 18 on page 77.

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. Table 10 lists information about the database fields.

Table 10. 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. Table 11 lists information about the database fields.

Table 11. 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
Begin Transaction Value	Fixed String	10 bytes	Index
End Transaction Value	Fixed String	10 bytes	Index

Table 11. Database storage for a report that contains a sorted transaction value (continued)

Field Name	Field Type	Field Size	Index or Filter
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

Note: This section contains information for Oracle users.

The database storage space requirements for your OnDemand system includes a rollback segment, an area on your disk subsystem that is used when OnDemand makes changes to the database. 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 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 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 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 `IndexSize` in "Estimating database storage space" on page 76.
3. The 40 bytes of system information that OnDemand adds to each database row.

Once you know these values, you can use the formula shown in Figure 19 on page 80 to estimate the size of the rollback segment that is required for your system.

Note: 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 OnDemand overhead bytes ) )
        * 2 = estimated size of rollback segment
```

Figure 19. Formula for Estimating the Size of the Rollback Segment

For example, suppose that your largest OnDemand load is for a statement application that loads 150,000 statements in a single load file. The OnDemand application group database fields require approximately 50 bytes of database storage per document. Figure 20 shows the example calculation, which requires approximately 27 MB of rollback segment space.

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

Figure 20. 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 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 21. Calculating primary log storage space

- You can find the calculations for `TableSize` and `IndexSize` in "Content Manager OnDemand database storage" on page 75.
- 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

Important: 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, using the calculation in "Primary log storage space for a 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, using the calculation in "Primary log storage space for a report" on page 81. 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 22. 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.

Archived log storage

If you are using DB2, the storage space required for archived logs is a factor of the number and size of the logs created between full backups of the database.

Note: If you use Tivoli Storage Manager to maintain DB2 archived log files, do not allocate disk space for the archived logs. Allocate optical storage space for the archived logs instead (see "Database archived log storage" on page 86).

It is recommended that you backup the database on a regular schedule. For example, you could backup the database each time that you load a report into the system or you could backup the database once a day or once a week. When you backup the database with the arsdb command, Content Manager OnDemand automatically removes the archived logs from the archived database log path, releasing the space taken by logs that are no longer needed to recover the database. Taking regular backups can also reduce the time required to rebuild the database, in the event that you need to do so.

It is critical that you allocate adequate disk space for the archived database log path. If the database manager does not have enough disk space to copy log files to the archived database log path, then it leaves the files in the primary database log path. (You then run the risk of running out of primary log space.)

As a guideline, it is recommended that you allocate two times the space that you estimate for the primary logs. However, you must allocate enough space to hold all of the logs created between full backups of the database.

Use the following calculation to estimate the amount of archived log space that you require:

```
ArchiveLogSpace = ( 2 * PrimaryLogSpace )
```

Figure 23. Calculating archived log storage space

The following example illustrates the archive log space required, when the space allocated for the primary log is 172 MB.

```
ArchiveLogSpace =  ( 2 * 172 MB )  = 344 MB
```

In the example, Content Manager OnDemand requires approximately 344 MB of archive log file storage.

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 24. 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 to 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

Note: 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 may need to allocate 1 GB or more free space to this file system. For example, based on the estimate for the sample reports in "Content Manager OnDemand database storage" on page 75 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 25. 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 calculated in "Report storage space" on page 84.

Storage for database backup images

Note: 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 may 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 may 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.

Note: Depending on the variables that listed above, you may 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 may 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

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

```
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

Note: 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 (see "Archived log storage" on page 82).

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 26. 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 may 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 may 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 27. 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 the previous example (see “Report storage space” on page 84), 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.

However, depending on the operational and management requirements of your organization, you may 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 may 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. However, 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. However, 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

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

Table 12. 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 13. 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 DB columns from Table 12	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)
Index on z/OS or Content Manager OnDemand server	Content Manager OnDemand server

Table 14. 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 migration index data	500,000,000
Total Disk Storage Required (Bytes)	18,971,040,000

Table 15. 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

Note: 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

Table 16. 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 17. 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

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

Report Characteristic	Report Estimate
Largest cycle (number of pages)	50,000
Largest single report file size (bytes)	200,000,000
Group of indexed pages	100
Number DB columns from Table 16	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)
Index on z/OS or Content Manager OnDemand server	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	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 19. 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

Notes:

1. Report data will not be stored in cache storage.
2. 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.
3. Index data will not be migrated to archive storage.
4. DB2 archived log files and backup image files are maintained on disk (or tape) independently of the archive storage manager.

Storage sizing worksheets

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. See "Estimating disk storage requirements" on page 73 and "Estimating archive storage requirements" on page 84 for the formulas that you can use to calculate the storage requirements.

Report that contains logical items

Table 20. Report that contains logical items. Part 1 of 4. Database Columns.

Column Number	Name	Index or Filter	Bytes
1			
2			
3			
4			

Table 21. 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 DB columns from Table 20	
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)	
Index on z/OS or Content Manager OnDemand server	

Table 22. 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	
Database logs	
Archive storage manager database, logs	
Server print	
Imported migrated indexes	
Total Disk Storage Required (Bytes)	

Table 23. 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 24. Report that contains transaction data. Part 1 of 4. Database Columns.

Column Number	Name	Index or Filter	Bytes
1			
2			
3			
4			

Table 25. 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)	

Table 25. Report that contains transaction data. Part 2 of 4. Report Profile. (continued)

Report Characteristic	Report Estimate
Number DB columns from Table 24	
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)	
Index on z/OS or Content Manager OnDemand server	

Table 26. 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 27. 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)	

Chapter 8. Backup and recovery

Overview

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

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 may 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.

Server information

When you installed and configured Content Manager OnDemand, you specified information that customized Content Manager OnDemand to operate in your environment. For UNIX servers, this information is stored in various control files. For Windows servers, this information is stored in the Registry. It is recommended that you backup the control files or Registry immediately after you have verified the installation of Content Manager OnDemand. In addition, if you periodically make changes to the Content Manager OnDemand server information, It is recommended that you backup the control files or Registry on a regular basis, perhaps once a day.

Also, if you periodically make changes to the system, including the database, archive storage manager, and server print manager, you may find it helpful to backup the control files or Registry on a regular basis, perhaps once a week. An administrator should schedule regular backup copies of the file systems on the server(s). Verify that all of the control files and other information required to operate the Content Manager OnDemand system are included in the backup.

You can use several different commands to make backup copies of files and file systems and to schedule regular backups of selected file systems, directories, and files. For example, in UNIX, you can schedule the TAR and MKSYSB commands with the CRON facility; in Windows, you can use the backup application or the REGBACK and RDISK commands to backup the Registry and schedule the backup with the AT command. In addition, you can use an archive storage manager, such as Tivoli Storage Manager, to backup files, including the Registry, maintain the backups, and assist with recovery. Tivoli Storage Manager is available on all of the server operating systems supported by OnDemand.

See the operating system and device publications for your server for details about backup and restore concepts and commands.

Content Manager OnDemand database

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 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 may 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:

- 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 OnDemand for Windows product allows for the scheduling of database backups. 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 OnDemand services and the rest of your organization cannot connect to the database while the backup task is running.

To schedule an offline backup with the Configurator program, you must do the following:

1. Manually disconnect all other processes from the database before the backup task is scheduled to begin. This includes stopping the 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 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 OnDemand LibSrvr service and the OnDemand MVSD Server and OnDemand Load Data services on the library server. If you stopped OnDemand services on an object server, manually restart the services.

Using Tivoli Storage Manager to maintain backup image files

If you use DB2, you can configure Tivoli Storage Manager to maintain the DB2 backup image files. This eliminates the need for you to manage DB2 backup image files on disk. When you use the ARSDB program to create table space backup images or backup the database, you can specify that you want to store the database or table space backup image files in storage that is managed by Tivoli Storage Manager.

Before you can use Tivoli Storage Manager to maintain the DB2 backup image files, you must define a Tivoli Storage Manager storage hierarchy to manage the files. The storage hierarchy includes definitions that identify the type of media and storage devices on which Tivoli Storage Manager stores the files, the length of time that Tivoli Storage Manager maintains the files, and the number of versions of a file that Tivoli Storage Manager maintains. Before you schedule a backup command, you must make sure that Tivoli Storage Manager storage volumes exist with sufficient free space to hold the backup image.

See the *IBM Content Manager OnDemand for Multiplatforms: Installation and Configuration Guide* for details about how to configure the system to use Tivoli Storage Manager to maintain the backup image files. See the *IBM Content Manager OnDemand for Multiplatforms: Administration Guide* for details about how to use the ARSDB program to backup table spaces to storage that is managed by Tivoli Storage Manager.

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.

Using Tivoli Storage Manager to maintain DB2 archived log files

If you use DB2, you can use Tivoli Storage Manager to maintain the DB2 archived log files. This eliminates the need for you to manage the log files on disk.

Before you can use Tivoli Storage Manager to maintain the DB2 archived log files, you must define a Tivoli Storage Manager storage hierarchy to manage the files.

The storage hierarchy includes definitions that identify the type of media and storage devices on which Tivoli Storage Manager stores the files, the length of time that Tivoli Storage Manager maintains the files, and the number of versions of a file that Tivoli Storage Manager maintains. Before DB2 creates archived log files, you must make sure that Tivoli Storage Manager storage volumes exist with sufficient free space to hold the archived log files.

See the *IBM Content Manager OnDemand for Multiplatforms: Installation and Configuration Guide* for details about how to configure the system to use Tivoli Storage Manager to maintain the DB2 log files.

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 may 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 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.

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

Recovering 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 may 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.

Reports

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 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.

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.

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 OnDemand to use archive storage. You configure 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.

Note: 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. However, 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 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 OnDemand storage node identifies the object server and the client node in archive storage where the primary copy of a report is maintained. 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 may 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.

Notices

This information was developed for products and services offered in the U.S.A.

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

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

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

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

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
1623-14, Shimotsuruma, Yamato-shi
Kanagawa 242-8502 Japan

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

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

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

IBM may use or distribute any of the information you supply 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 Corporation
J46A/G4
555 Bailey Avenue
San Jose, CA 95141-1003
U.S.A.

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.

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

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

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

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

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which 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.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at `www.ibm.com/legal/copytrade.shtml`

Adobe, Acrobat, Portable Document Format (PDF), and PostScript are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, other countries, or both.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

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

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

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

Portions of the OnDemand Windows client program contain licensed software from Pixel Translations Incorporated, © Pixel Translations Incorporated 1990, 2003. All rights reserved.

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

Glossary

A

access. The ability to read, update, or otherwise use a resource. Access to protected resources is usually controlled by system software.

ACIF. Advanced Function Presentation Conversion and Indexing Facility

active log file. The subset of files consisting of primary log files and secondary log files that are currently needed by the database manager for rollback and recovery.

active policy set. In Tivoli Storage Manager, the activated policy set that contains the policy rules currently in use by all client nodes assigned to the policy domain. See policy set.

active storage node. In a storage set, the storage node that is currently being used to load data.

adapter. A part that electrically or physically connects a device to a computer or to another device.

addressable point. Any point in a presentation surface that can be identified by a coordinate from the coordinate system of the presentation medium.

administrative client. (1) In OnDemand, the program that provides administrators with functions to maintain Content Manager OnDemand groups, users, printers, applications, application groups, storage sets, and folders. (2) In Tivoli Storage Manager, the program that allows administrators to control and monitor the server through administrator commands.

administrator. In OnDemand, a person authorized to maintain the system. For example, an OnDemand administrator can add, update, and delete users and folders.

Advanced Function Presentation (AFP). A set of licensed programs that use the all-points-addressable concept to print data on a wide variety of printers or display data on a variety of display devices.

Advanced Function Presentation Application Programming Interface (AFP API). An AFP program that creates the AFP data stream from the COBOL and PL/1 high-level programming languages. This program is shipped with certain releases of PSF.

Advanced Function Presentation Conversion and Indexing Facility (ACIF). A program shipped with Content Manager OnDemand that you can use to convert a print file into a MO:DCA-P document, to retrieve resources used by the document, and to index the file for later retrieval and viewing.

Advanced Function Presentation data stream (AFPDS). A presentation data stream that is processed in the AFP environment. MO:DCA-P is the strategic AFP interchange data stream. IPDS is the strategic AFP printer data stream.

AFP. Advanced Function Presentation

AFP API. Advanced Function Presentation Application Programming Interface

AFPDS. A term formerly used to identify the composed page, MO:DCA-P-based data stream interchanged in AFP environments.

AIX Acrobat Libraries. A subset of the Acrobat Libraries ported to AIX for use by OnDemand.

AIX Print Services Facility/6000. An IBM licensed program that produces printer commands from the data sent to it.

all-points-addressable (APA). The capability to address, reference, and position data elements at any addressable position in a presentation space or a physical medium.

alphabetic character. A letter or other symbol, excluding numerals, used in a language. Usually the uppercase and lowercase letters A through Z plus other special symbols (such as $ and _) allowed by a particular language. See also alphanumeric character.

alphanumeric character. Consisting of letters, numerals, and often other symbols, such as punctuation marks and mathematical symbols. See also alphabetic character.

alphanumeric string. A sequence of characters consisting solely of the letters a through z and the numerals 0 through 9.

American National Standards Institute (ANSI). An organization consisting of producers, consumers, and general interest groups, that establishes the procedures by which accredited organizations create and maintain voluntary industry standards in the United States.

anchor point. The point in a document that signals to ACIF the beginning of a group of pages, after which it adds indexing structured fields to delineate this group.

annotations. Comments, clarifications, and reminders that can be attached to a document.

ANSI. American National Standards Institute.

ANSI carriage control character. A character that specifies that a write, space, or skip operation should be performed before printing the line that contains the carriage control. ANSI carriage control characters are encoded in ASCII or EBCDIC.

APA. all points addressable.

API. application programming interface.

application. In OnDemand, an object that describes the physical attributes of a report or input file, such as the type of data found in the input file, the code page, and whether the input data contains carriage control characters. An application also contains instructions that the data indexing and loading programs use to process the input data. Most customers define an application for each different output print data stream or source of data that they plan to store in Content Manager OnDemand.

application group. A collection of one or more Content Manager OnDemand applications that have similar indexing and storage management requirements. For example, two reports that can be retrieved using the same index fields and that are to be maintained by the system in the same storage locations for the same length of time might be placed in the same application group.

archive copy group. In Tivoli Storage Manager, a policy object containing attributes that control the generation, destination, and expiration of archive files. An archive copy group belongs to a management class.

archive log file. The subject of files consisting of primary log files and secondary log files that are no longer needed for normal database processing.

archive media. Devices and volumes on which the long-term or backup copy of a report is stored. For example, an optical storage library is one type of archive media supported by Content Manager OnDemand.

archive storage. The storage in which the long-term or backup copy of a report is maintained. Includes the devices and volumes on which the files are stored and the management policies that determine how long data is maintained in archive storage.

archive storage manager. The software product that manages archive media and maintains files in archive storage. See Tivoli Storage Manager.

ASCII (American Standard Code for Information Interchange). The standard code, using a coded character set consisting of 7-bit coded characters (8-bits including parity check), that is used for information interchange among data processing systems, data communication systems, and associated equipment. The ASCII set consists of control characters and graphic characters.

attachment. A device or feature attached to a processing unit, including required adapters.

authentication. The process of checking a user's password before allowing the user access to resources or the server.

authorize. (1) To grant to a user the right to communicate with or make use of a computer system or display station. (2) To give a user either complete or restricted access to an object, resource, or function.

B

back-end program. In the AIX operating system, the program that sends output to a particular device.

banner. An optional report page that contains information from a recipient's user ID for the purpose of distribution. The three different banner types are: header, separator, and trailer.

Bar Code Object Content Architecture (BCOCA). An architected collection of control structures used to interchange and present bar code data.

BCOCA. Bar Code Object Content Architecture.

bitmap. A file that contains a bit-mapped graphic.

bundle. A bundle is an IBM Content Manager OnDemand Report Distribution for Multiplatforms object that allows you to package, organize, and optionally provide additional information about the reports that you want to send to the recipients. A bundle contains at least one or more reports, and might optionally include banners and a manifest. A distribution contains only one bundle, but a bundle can belong to more than one distribution.

byte. The amount of storage required to represent 1 character; a byte is 8 bits.

C

cabinet. A cabinet is a container for folders. You can use cabinets to manage folders and enable users to navigate to folders more easily. A folder can belong to one or more cabinets.

cache storage. The storage in which the primary or short-term copy of a report is stored. Usually disk storage. Most customers configure the system to maintain the most recent and frequently used versions of reports in cache storage.

carriage control character. The first character of an output record (line) that is to be printed; it determines how many lines should be skipped before the next line is printed.

case-sensitive. Able to distinguish between uppercase and lowercase letters.

CCITT. Consultative Committee on International Telegraphy and Telephone

channel. A device connecting the processor to input and output devices.

channel adapter. A communication controller hardware unit used to attach the controller to a System/370 data channel.

channel-attached. Pertaining to devices attached to a controlling unit by cables, rather than by telecommunication lines.

character rotation. The alignment of a character with respect to its character baseline, measured in degrees in a clockwise direction. Examples are 0°, 90°, 180°, and 270°. Zero-degree character rotation exists when a character is in its customary alignment with the baseline.

character set. A group of characters used for a specific reason; for example, the set of characters a printer can print or a keyboard can support.

client. (1) In a distributed file system environment, a system that is dependent on a server to provide it with programs or access to programs. (2) A workstation connected to a network running Content Manager OnDemand software that can log on and query the library server, retrieve documents from Content Manager OnDemand, and view and print documents.

client domain. The set of optical drives and storage volumes used by Tivoli Storage Manager to store report files and resources belonging to an application group.

client node. An application group that has been registered to the Tivoli Storage Manager server.

COBOL. Common Business Oriented Language. A high-level programming language, based on English, that is used primarily for applications.

code page. An ordered set of up to 256 predefined display symbols. The first 32 code points of each code page are reserved for control codes and are the same for all code pages, leaving up to 224 distinct display symbols per page.

Code Page Global Identifier (CPGID). A unique code page identifier that can be expressed as either a two-byte binary or a five-digit decimal value.

coded font. An AFP font that associates a code page and a font character set.

command. A request to perform an operation or run a program. When parameters, values, flags, or other operands are associated with a command, the resulting character string is a single command.

command line. The area of the screen where commands are displayed as they are typed.

communication method. The method used by Content Manager OnDemand and Tivoli Storage Manager to exchange information.

communication protocol. A set of defined interfaces that allow computers to communicate with each other.

composed page. In Advanced Function Presentation, a page that can be printed only on an all-points-addressable output medium. It may contain composed text and raster images.

composed-text data file. A file containing text data and text control information that dictates the format, placement, and appearance of the data to be printed.

compression. A technique for removing strings of duplicate characters, gaps, empty fields, and trailing blanks before transmitting data.

concatenate. (1) To link together. (2) To join two character strings.

concatenated field. Two or more fields from a physical file record format that have been combined to make one field in a logical file record format.

conditional processing. A page definition function that allows input data records to partially control their own formatting.

console. A terminal connected directly to the computer and used for communication between the operator and the computer.

Consultative Committee on International Telegraphy and Telephone (CCITT). A United Nations Specialized Standards group whose membership includes common carriers concerned with devising and proposing recommendations for international telecommunications representing alphabets, graphics, control information, and other fundamental information interchange issues.

control character. A character that is not a graphic character such as a letter, number, or punctuation mark. Such characters are called control characters because they frequently act to control a peripheral device.

control interval. In a key-sequenced data set or file, the set of records pointed to by an entry in the sequence-set index record.

control interval definition field (CIDF). In VSAM, a field located in the 4 bytes at the end of each control interval; it describes the free space, if any, in the control interval.

controller. A device that coordinates and controls the operation of one or more input/output devices, such as workstations, and synchronizes the operation of the system as a whole.

conversion. In programming languages, the transformation between values that represent the same data item but belong to different data types.

copy group. In Tivoli Storage Manager, a policy object that contains attributes that control the generation, destination, and expiration of backup and archive files. There are two kinds of copy groups: backup and archive. Copy groups belong to management classes.

copy storage pool. A named collection of storage volumes that contains copies of files that reside in primary storage pools. Copy storage pools are used to back up the data stored in primary storage pools.

CPGID. code page global identifier.

D

database. A collection of interrelated or independent data items stored together to serve one or more applications.

data set. See file.

data stream. A continuous stream of data elements being transmitted, or intended for transmission, in character or binary-digit form using a defined format.

data transfer. The movement, or copying, of data from one location and the storage of the data at another location.

data type. The type, format, or classification of a data object.

date/time. Describe various ODF fields that carry date/time stamps from the CPU to record events in the ODF and base systems.

DCF. Document Composition Facility

decimal. Pertaining to a system of numbers to the base 10. The decimal digits range from 0 through 9.

decompression. A function that expands data to the length that preceded data compression. See also compression.

default. A value, attribute, or option that is assumed when no alternative is specified by the user.

default directory. The directory name supplied by the operating system if none is specified.

default printer. A printer that accepts all of the printed output from a display station that is assigned to it.

default value. A predetermined value, attribute, or option that is assumed when no other is explicitly specified.

device class. A named group of Tivoli Storage Manager storage devices. Each device class has a unique name and represents a device type of disk, tape, or optical disk.

device driver. A program that operates a specific device, such as a printer, disk drive, or display.

device-independent. Pertaining to a function that can be accomplished without regard for the characteristics of particular types of devices.

device type. A type of Tivoli Storage Manager storage device. Each device class must be categorized with one of the following devices types: disk, tape, or optical disk.

directory. (1) A type of file containing the names and controlling information for other files or directories. (2) A listing of related files arranged in a useful hierarchy.

Distiller. A batch utility that converts PostScript files to Adobe PDF files. The distiller runs on AIX, HP-UX Itanium, Solaris, and Windows servers.

distribution. A set of reports contained in a bundle and have the same recipient list.

document. (1) In Content Manager OnDemand, a logical section of a larger file, such as an individual invoice within a report of thousands of invoices. A document can also represent an indexed group of pages from a report. (2) A file containing an AFP data stream document. An AFP data stream document is bounded by Begin Document and End Document structured fields and can be created using a text formatter such as Document Composition Facility (DCF).

Document Composition Facility (DCF). An IBM licensed program used to prepare printed documents.

domain. See Policy Domain or Client Domain.

download. To transfer data from one computer for use on another one. Typically, users download from a larger computer to a diskette or fixed disk on a smaller computer or from a system unit to an adapter.

driver. The end of a stream closest to an external interface. The principal functions of the driver are handling any associated device, and transforming data and information between the external device and stream.

E

EBCDIC (Extended Binary-Coded Decimal Interchange Code). A coded character set of 256 8-bit characters developed for the representation of textual data.

enqueue. To place items in a queue.

environment variable. A variable that is included in the current software environment and is therefore available to any called program that requests it.

error condition. The state that results from an attempt to run instructions in a program that are not valid or that operate on data that is not valid.

error log. A file in a product or system where error information is stored for later access.

error log entry. In AIX, a record in the system error log describing a hardware or software failure and containing data that was captured at the time of the failure.

error message. An indication that an error has been detected. (A)

error recovery. The process of correcting or bypassing the effects of a fault to restore a computer system to a prescribed condition. (T)

error type. Identifies whether an error log entry is for a permanent failure, temporary failure, performance degradation, impending loss of availability, or undetermined failure.

Ethernet. A 10-megabit baseband local area network that allows multiple stations to access the medium at will without prior coordination, avoids contention by using carrier sense and deference, and resolves contention by using collision detection and transmission.

exit program. A user-written program that is given control during operation of a system function.

exit routine. A routine that receives control when a specified event occurs, such as an error.

expiration. The process of deleting index data and reports based on storage management information. The Content Manager OnDemand database manager and the storage managers run expiration processing to remove data that is no longer needed from storage volumes and reclaim the space.

Extended Binary-Coded Decimal Interchange Code (EBCDIC). A coded character set consisting of eight-bit coded characters.

external library resource (member). See external object.

external object. Objects that can be used by other program products while running print jobs; for example, coded fonts, code pages, font character sets, form definitions, page definitions, and page segments.

F

FCB. forms control buffer

field. A specified area in a record used for a particular type of data; for example, a group of characters that represent a customer's name.

file. (1) A named set of records stored or processed as a unit. (T) (2) The major unit of data storage and retrieval. A file consists of a collection of data in one of several prescribed arrangements and described by control information to which the operating system has access.

file system. The collection of files and file management structures on a physical or logical mass storage device, such as a diskette or a minidisk.

file transfer. In remote communications, the transfer of a file or files from one system to another over a communications link.

File Transfer Protocol (FTP). In TCP/IP, the protocol that makes it possible to transfer data among hosts and to use foreign hosts indirectly.

fixed disk. A flat, circular, nonremovable plate with a magnetizable surface layer on which data can be stored by magnetic recording. A rigid magnetic disk.

fixed-disk drive. The mechanism used to read and write information on a fixed disk.

folder. A container for related information, such as statements, invoices, or correspondence, regardless of the source of the information or where the data is stored. When you open a folder, you have access to all of the information that it contains. For example, a billing folder might contain all of the reports for customer transactions over the past two years.

font. (1) A family of characters of a given size and style, for example 9-point Helvetica. (2) A set of characters in a particular style. See Raster Font.

font character set. Part of an AFP font that contains the raster patterns, identifiers, and descriptions of characters.

forms control buffer (FCB). A buffer for controlling the vertical format of printed output. The FCB is a line-printer control that is similar to the punched-paper, carriage-control tape used on the IBM 1403 Printer. For AFP page printers, the forms control buffer is replaced by the page definition.

form definition (FORMDEF). Specifies the number of copies to be printed, whether the sheet should be printed on both sides, the position of a page of data on the sheet, text suppression, and overlays to be used (if any).

FORMDEF. form definition

FSA (functional subsystem application). A collection of programs residing in the FSS address space that control a device.

FSI (functional subsystem interface). A set of services that allow communication between the JES address space or direct printer services subsystem (DPSS) and the Print Services Facility functional subsystem.

FSS (functional subsystem). The PSF address space created by the job entry subsystem (JES).

FTP. File Transfer Protocol.

G

GB. gigabyte

GIF. graphical interchange format

gigabyte. A unit of memory or space measurement equal to approximately one billion bytes. One gigabyte equals 1,000 megabytes.

GOCA. Graphic Object Content Architecture

graphic. A symbol produced by a process such as handwriting, drawing, or printing. (I) (A)

graphical interchange format (GIF). A digital format that is used to compress and transfer graphical information over computer networks. For example, GIF is a common format for graphical information on the Internet.

graphic character. A character that can be displayed or printed.

Graphic Object Content Architecture. An architecture that provides a collection of graphical values and control structures used to interchange and present graphical data.

graphics. A type of data created from such fundamental drawing units such as lines, curves, polygons, and so forth.

group. (1) A named collection of sequential pages that form a logical subset of a document. (2) A named collection of users assigned a specific role on the system or belonging to a specific department.

GUI. graphical user interface

H

header banner. An optional report delimiter containing recipient address information.

hexadecimal (hex). Pertaining to a system of numbers in the base sixteen; hexadecimal digits range from 0 (zero) through 9 (nine) and A (ten) through F (fifteen).

hold. In Content Manager OnDemand, documents are stored and retained for a specified period of time. After that specified period of time, the documents are removed from Content Manager OnDemand. In certain situations, a hold object enables you to keep one or more documents beyond the expiration date.

host-based computer. (1) In a computer network a computer that provides end users with services such as computation and databases and that usually performs network control functions. (T) (2) The primary or controlling computer in a multiple-computer installation.

host system. (1) The controlling or highest level system in a data communication configuration, for example, an OS/390 system is the host system for the terminals connected to it. (2) In TCP/IP, a computer that is a peer system in a network.

I

icon. A 32 by 32 pixel bitmap used by the windows manager to represent an application or other window.

image. (1) An electronic representation of a picture produced by means of sensing light, sound, electron radiation, or other emanations coming from the picture or reflected by the picture. An image can also be generated directly by software without reference to an existing picture. (2) An electronic representation of an original document recorded by a scanning device.

Image Object Content Architecture. An architected collection of constructs used to interchange and present images.

index. (1) A process of segmenting a print file into uniquely identifiable groups of pages (a named collection of sequential pages) for later retrieval. (2) A process of matching reference points within a file and creating structured field tags within the MO:DCA-P document and the separate index object file.

index object file. An index-information file created by ACIF that contains the Index Element (IEL) structured fields, which identify the location of tagged groups in the AFP file. The indexing tags are contained in the Tagged Logical Element (TLE) structured fields.

Infoprint Manager. A sophisticated IBM print subsystem that drives AFP printers, PostScript printers, and PCL printers. Infoprint Manager is supported

under AIX, OS/390, and Windows. Infoprint Manager manages printer resources such as fonts, images, electronic forms, form definitions, and page definitions, and provides error recovery for print jobs.

When printing line data, Infoprint Manager supports external formatting using page definitions and form definitions. This external formatting extends page printer functions such as electronic forms and use of typographic fonts without any change to applications that generate the data.

informational message. (1) A message that provides information to the end-user or system administrator but does not require a response. (2) A message that is not the result of an error condition.

input file. A file opened in order to allow records to be read.

intelligent printer data stream (IPDS). An all-points-addressable data stream that allows users to position text, images, and graphics at any defined point on a printed page.

interface. Hardware, software, or both, that links systems, programs, or devices.

Internet. A wide area network connecting thousands of disparate networks in industry, education, government, and research. The Internet network uses TCP/IP as the protocol for transmitting information.

Internet Protocol (IP). In TCP/IP, a protocol that routes data from its source to its destination in an Internet environment.

IOCA. Image Object Content Architecture

IP. Internet Protocol

IPDS. intelligent printer data stream

J

job. One or more related procedures or programs grouped into a procedure, identified by appropriate job control statements.

job queue. A list of jobs waiting to be processed by the system.

Joint Photographic Experts Group (JPEG). An image compression standard developed to handle larger images with many colors. JPEG uses a lossy algorithm, which means there is some loss of detail when saving and viewing images in this format. However, JPEG files can offer as much as 35% improvement in file size and compression.

JPEG. See Joint Photographic Experts Group

K

KB. kilobyte

kernel. The part of an operating system that performs basic functions such as allocating hardware resources.

kernel extension. A program that modifies parts of the kernel that can be customized to provide additional services and calls. See kernel.

keyword. Part of a command operand that consists of a specific character string.

kilobyte (KB). 1024 bytes in decimal notation when referring to memory capacity; in all other cases, it is defined as 1000.

L

LAN. local area network

LAN server. A data station that provides services to other data stations on a local area network; for example, file server, print server, mail server.

laser printer. A nonimpact printer that creates, by means of a laser beam directed on a photosensitive surface, a latent image which is then made visible by toner and transferred and fixed on paper. (T)

Lempel Ziv Welsh (LZW). A data compression algorithm. OnDemand uses the 16-bit version of LZW to compress data.

library. System storage for generated form definitions and page definitions.

library resource (member). A named collection of records or statements in a library.

library resource name. A name by which an object can be called from a library by AFP as part of a print job. Includes the 2-character prefix for the type of object, such as P1 for page definitions, F1 for form definitions, or O1 for overlays.

library server. In Content Manager OnDemand, the workstation or node that users must go through to access the system. The library server controls the Content Manager OnDemand database.

licensed program. A separately priced program and its associated materials that bear a copyright and are offered to customers under the terms and conditions of a licensing agreement.

Lightweight Directory Access Protocol (LDAP). The Lightweight Directory Access Protocol (LDAP) is an open industry standard that has evolved to share information between distributed applications on the same network, organize information clearly and

consistently, and prevent unauthorized modification or disclosure of private information.

line data. Data prepared for printing on a line printer, such as an IBM 3800 Model 1 Printing Subsystem. Line data is usually characterized by carriage-control characters and table reference characters.

line-data print file. A file that consists of line data, optionally supplemented by a limited set of structured fields.

line printer. A device that prints a line of characters as a unit. (I) (A).

line printer daemon (LPD). In TCP/IP, the command responsible for sending data from the spooling directory to a printer.

line printer requestor (LPR). In TCP/IP, a client command that allows the local host to submit a file to be printed on a remote print server.

literal. (1) A symbol or a quantity in a source program that is itself data, rather than a reference to data. (2) A character string whose value is given by the characters themselves; for example, the numeric literal 7 has the value 7, and the character literal CHARACTERS has the value CHARACTERS.

loading. The logical process of archiving reports in OnDemand. During the loading process, Content Manager OnDemand processes reports, creates index data, and copies report data and resources to cache storage and archive storage.

local. Pertaining to a device accessed directly without use of a telecommunication line.

local area network (LAN). (1) A computer network located on a user's premises within a limited geographical area. Communication within a local area network is not subject to external regulations; however, communication across the LAN boundary might be subject to some form of regulation. (2) A network in which a set of devices is connected to one another for communication and that can be connected to a larger network. See also Token-Ring Network.

log file. A fixed-length file used to record changes to a database.

logical volume. The combined space from all volumes defined to either the Tivoli Storage Manager database or recovery log. The database resides on one logical volume and the recovery log resides on a different logical volume.

LPD. line printer daemon.

LPR. line printer requestor.

LZW. See Lempel Ziv Welsh

M

MB. megabyte

machine carriage control character. A character that specifies that a write, space, or skip operation should be performed either immediately or after printing the line containing the carriage control.

mainframe. A large computer, particularly one to which other computers can be connected so that they can share facilities the mainframe provides. The term usually refers to hardware only.

management class. A logical area of storage that is managed by Tivoli Storage Manager. A management class is a policy object that is a named collection of copy groups. A management class can contain one backup copy group, one archive copy group, a backup and archive copy group, or zero copy groups. Users can bind each file to a management class to specify how the server should manage backup versions or archive copies of files. See copy group.

manifest. An optional list of the reports in a bundle.

mapping. (1) A list that establishes a correspondence between items in two groups. (2) The process of linking database fields in an application group to folder search and display fields.

megabyte (MB). When used with hard drive, diskette, or removable media storage capacity, 1,000,000 bytes. When referring to system memory capacity, 1,048,576 bytes.

memory. Program-addressable memory from which instructions and other data can be loaded directly into registers for subsequent running or processing. Memory is sometimes referred to as "storage".

menu bar. The area at the top of a window that contains choices that give a user access to actions available in that window.

message. Information from the system that informs the user of a condition that might affect further processing of a current program.

migration. (1) The process of moving data from one computer system to another without converting the data. (2) The process of moving report files, resources, and index data from cache storage to long-term (optical or tape) storage.

mirroring. In Tivoli Storage Manager, a feature that protects against data loss with the database or recovery log by writing the same data to multiple disks at the same time. Mirroring supports up to three exact copies of each database or recovery log.

Mixed Object Document Content Architecture - Presentation (MO:DCA-P). A subset of MO:DCA that defines presentation documents.

MO:DCA-P. Mixed Object Document Content Architecture - Presentation.

mount. To make a file system accessible.

N

named query. A set of entry field values on the Search Criteria and Document List window that a user has named and saved for selection as search criteria at a future time.

network. A collection of data processing products that are connected by communication lines for information exchange between locations.

Network File System (NFS). A protocol developed by Sun Microsystems that uses Internet Protocol to allow a set of cooperating computers to access each other's file system as if they were local.

NFS. Network File System

node. A workstation that operates as an Content Manager OnDemand library server or object server and is connected to a TCP/IP network.

non-IPDS printer. In this publication, a printer that is not channel-attached and which does not accept the Intelligent Printer Data Stream.

notes. Electronic comments, clarifications, and reminders that can be attached to an Content Manager OnDemand document.

numeric. Pertaining to any of the digits 0 through 9.

O

object. (1) A collection of structured fields. The first structured field provides a begin-object function and the last structured field provides an end-object function. The object can contain one or more other structured fields whose content consists of one or more data elements of a particular data type. An object can be assigned a name, which can be used to reference the object. Examples of objects are text, graphics, and image objects. (2) A resource or a sequence of structured fields contained within a larger entity, such as a page segment or a composed page. (3) A collection of data referred to by a single name.

object server. In OnDemand, a workstation or node controlled by a storage manager to maintain reports in cache storage, and optionally, archive storage.

offset. The number of measuring units from an arbitrary starting point in a record, area, or control block to some other point.

online. Being controlled directly by or directly communicating with the computer.

operating environment. (1) The physical environment; for example, temperature, humidity, and layout. (2) All of the basic functions and the user programs that can be executed by a store controller to enable the devices in the system to perform specific operations. (3) The collection of store controller data, user programs, lists, tables, control blocks, and files that reside in a subsystem store controller and control its operation.

operating system. Software that controls the running of programs and that also can provide such services as resource allocation, scheduling, input and output control, and data management.

optical library. A storage device that houses optical disk drives and optical disks, and contains a mechanism for moving optical disks between a storage area and optical disk drives.

optimize. To improve the speed of a program or to reduce the use of storage during processing.

outline fonts. (1) Fonts whose graphic character shapes are defined as mathematical equations rather than by raster patterns. (2) Fonts created in the format described in *Adobe Type 1 Font Format*, a publication available from Adobe Systems, Inc.

overlay. A collection of predefined, constant data such as lines, shading, text, boxes, or logos, that is electronically composed and stored as an AFP resource file that can be merged with variable data on a page while printing or viewing.

P

page. Part of an AFP document bracketed by a pair of Begin Page and End Page structured fields.

PAGEDEF. page definition

page definition. A resource used by Content Manager OnDemand that defines the rules of transforming line data into composed pages and text controls.

page printer. A device that prints one page as a unit. (I) (A).

page segment. In Advanced Function Presentation, a resource that can contain text and images and can be positioned on any addressable point on a page or an electronic overlay.

parallel device. A device that can perform two or more concurrent activities.

parameter. (1) Information that the user supplies to a panel, command, or function. (2) In the AIX operating system, a keyword-value pair.

partitioned data set. A data set in direct access storage that is divided into partitions, called members, each of which can contain a program, part of a program, or data.

path. In a network, any route between any two nodes.

path name. A name that specifies the location of a directory within a file system. Path names are used to locate and reference directories and their contents.

PCL. printer control language

PCX. picture exchange files

PDF. Portable Document Format

permissions. Codes that determine the users who can access a system, that determine how data can be used by any users who can access the system, and that determine other types of tasks users of the system can perform.

picture exchange files (PCX). A file that contains a graphic in the PCX graphics file format, which was originally developed for the PC Paintbrush program, but is now widely used by other programs.

piobe. The printer input/output back end program used by AIX for printing tasks.

pipe. To direct the data so that the output from one process becomes the input to another process. The standard output of one command can be connected to the standard input of another with the pipe operator (|). Two commands connected in this way constitute a pipeline.

point. A unit of typesetting measure equal to 0.01384 inch (0.35054 mm), or about 1/72 of an inch. There are 12 points per pica.

point size. The height of a font in points. See also point.

policy domain. In Tivoli Storage Manager, a policy object that contains policy sets, management classes, and copy groups that is used by a group of client nodes. See policy set, management class, copy group, and client node.

policy set. In Tivoli Storage Manager, a policy object that contains a group of management class definitions that exist for a policy domain. At any one time, there can be many policy sets within a policy domain but only one policy set can be active. See management class and active policy set.

port. (1) A part of the system unit or remote controller to which cables for external devices (display stations, terminals, or printers) are attached. The port is an access point for data entry or exit. (2) A specific communications end point within a host. A port is identified by a port number.

Portable Document Format. A distilled version of PostScript data that adds structure and efficiency. PDF data has the same imaging model as PostScript but does not have its programmability. PDF also provides direct access to pages and allows hypertext links, bookmarks, and other navigational aids required for viewing. The text in a PDF file is usually compressed using LZW methods. The images is a PDF file are usually compressed using CCITT or JPEG methods.

PostScript. Adobe's page description language used for printing. PostScript is a very flexible programming language and imaging model but is not as structured as AFP. PostScript cannot be parsed to determine page boundaries, it must be interpreted. Because of this limitation, PostScript is not practical for archiving and viewing. Adobe created PDF for archiving and viewing.

primary log file. A set of one or more log files used to record changes to a database. Storage for these files is allocated in advance.

primary storage pool. A named collection of storage volumes in which Tivoli Storage Manager stores archive copies of files.

printer control language (PCL). The data stream used by Hewlett-Packard LaserJet II and III and other compatible printers.

print file. (1) The output of a user-defined program that is to be indexed and loaded into the system. (2) A file that a user wants to print.

print job. A series of print files scheduled for printing. At print submission time, the user can request one or more files to be printed; therefore, a print job consists of one or more print files.

print queue. A file containing a list of the names of files waiting to be printed.

print spooler. The print spooler directs the printing of data from different applications. It temporarily stores information in separate files until they are printed.

process. An activity within the system that is started, such as a command, a shell program, or another process.

profile. (1) A file containing customized settings for a system or user. (2) Data describing the significant features of a user, program, or device.

program level. The version, release, modification, and fix levels of a program.

prompt. A displayed symbol or message that requests information or operator action.

protocol. A set of semantic and syntactic rules that determines the behavior of functional units in achieving communication.

PSF for AIX. Print Services Facility for AIX.

Q

qdaemon. The daemon process that maintains a list of outstanding jobs and sends them to the specified device at the appropriate time.

qualified name. (1) A data name explicitly accompanied by a specification of the class to which it belongs in a specified classification system. (I) (A) (2) A name that has been made unique by the addition of one or more qualifiers.

queue. (1) A line or list formed by items waiting to be processed. (2) To form or arrange in a queue.

queue device. A logical device defining characteristics of a physical device attached to a queue.

R

RAM. random access memory. Specifically, the memory used for system memory. Sometimes this memory is referred to as main storage.

raster. In Advanced Function Presentation, an on/off pattern of electrostatic images produced by the laser print head under control of the character generator.

raster font. A font in which the characters are defined directly by the raster bit map. See font.

raster graphics. Computer graphics in which a display image is composed of an array of pixels arranged in rows and columns.

read access. In computer security, permission to read information.

recipient. An OnDemand user, or group of users, who receive distributions from the report distribution system.

recipient list. Consists of one or more recipients who receive the same distribution.

record. (1) In programming languages, an aggregate that consists of data objects, possibly with different attributes, that usually have identifiers attached to them. (2) A set of data treated as a unit. (3) A collection of fields treated as a unit.

recovery log. In Tivoli Storage Manager, a log of updates that are about to be written to the database. The log can be used to recover from system and media failures.

recovery procedure. (1) An action performed by the operator when an error message displays. This action usually permits the program to run the next job. (2) The method of returning the system to the point where a major system error occurred and running the recent critical jobs again.

register. To define a client node to Tivoli Storage Manager.

remote. Pertaining to a system or device that is accessed through a communications line.

remote print. Issuing print jobs to one machine (client) to print on another machine (server) on a network.

remote system. A system that is connected to your system through a communication line.

report. A document, or set of documents, that have been loaded into the OnDemand system and are defined by a named query, SQL statement, or supplied load parameters. You use only one query or SQL statement for each report.

report. A print data stream produced by a user-defined program or other software program that can contains hundreds or thousands of pages of related information. Most reports can be logically divided and indexed into single and multiple page objects called documents.

resolution. (1) In computer graphics, a measure of the sharpness of an image, expressed as the number of lines and columns on the display screen. (2) The number of pels per unit of linear measure.

resource. A collection of printing instructions, and sometimes data to be printed, that consists entirely of structured fields. A resource can be stored as a member of a directory and can be called for by the Print Services Facility when needed. The different resources are: coded font, character set, code page, page segment, overlay, and form definition.

resource directory. The location where resource files are stored.

resource management. The function that protects serially accessed resources from concurrent access by computing tasks.

retention. The amount of time, in days, that archived files will be retained in Tivoli Storage Manager before they are deleted.

retry. To try the operation that caused the device error message again.

return code. (1) A value that is returned to a program to indicate the results of an operation issued by that program. (2) A code used to influence the running of succeeding instructions.

root. On UNIX servers, the user name for the system user with the most authority.

root file system. In UNIX environments, the file system that contains all of the default installation and program directories in the system.

root user. In UNIX environments, an expert user who can log in and execute restricted commands, shut down the system, and edit or delete protected files.

root volume group. In UNIX environments, the volume group, identified with a single / (forward slash) that contains all the directories in the root file system.

rotation. (1) The alignment of a character with respect to its character baseline, measured in degrees in a clockwise rotation. Examples are 0°, 90°, 180°, and 270°. Zero-degree character rotation exists when a character is in its customary alignment with the baseline. (2) The number of degrees a character is turned relative to the page coordinates.

routing. The assignment of the path by which a message will reach its destination.

S

schedule. Specifies the time to start a distribution. Each distribution that is created is assigned to a schedule.

secondary log file. A set of one or more log files used to record changes to a database. Storage for these files is allocated as needed when the primary log fills up.

segment. (1) A collection of composed text and images, prepared before formatting and included in a document when it is printed. See Page Segment. (2) The resource that contains the structured-field definition of a page segment. (3) A 100 page portion of a report file. Content Manager OnDemand divides report files into segments to provide enhanced performance and maintenance.

segment table. A high-level index to data stored in an application group. Each row in the segment table identifies a table of application group index data. OnDemand uses the segment table to limit a query to a specific table of application group index data.

separator banner. An optional report page that contains information from a recipient's user ID for the purpose of distribution. This particular banner is used to separate individual reports from each other.

serial device. A device that performs functions sequentially, such as a serial printer that prints one byte at a time.

server. (1) On a network, the computer that contains the data or provides the facilities to be accessed by other computers on the network. (2) A program that handles protocol, queuing, routing, and other tasks necessary for data transfer between devices in a computer system. (3) A workstation connected to a TCP/IP network that runs the Content Manager OnDemand programs that store, retrieve, and maintain report files. Content Manager OnDemand supports two types of servers: a library server an object server.

server options file. The Tivoli Storage Manager file that specifies processing options for communication methods, tape handling, pool sizes, language, and date, time, and number formats.

server printer. A printer that is attached to a network server and is managed by a server print manager, such as IBM Infoprint Manager.

shell. In UNIX environments, a software interface between a user and the operating system of a computer. Shell programs interpret commands and user interactions on devices such as keyboards and pointing devices and communicate them to the operating system.

skip-to-channel control. A line printer control appearing in line data. Allows space to be left between print lines. Compatible with page printers when the data is formatted by page definitions.

SMIT. System Management Interface Tool

SMS. System-managed storage

software. Programs, procedures, rules, and any associated documentation pertaining to the operating of a system.

spool file. (1) A disk file that contains output that has been saved for later printing. (2) Files used in the transmission of data among devices.

spooling (simultaneous peripheral operation online). Performing a peripheral operation such as printing while the computer is busy with other work.

spooling subsystem. A part of the system that provides the operating environment for the programs that read jobs onto job queues to wait for processing and write files from an output queue to an output device. IBM supplies one spooling subsystem: QSPL.

stand-alone workstation. A workstation that can perform tasks without being connected to other resources such as servers or host systems.

standard input. The primary source of data going into a command. Standard input comes from the keyboard unless redirection or piping is used, in which case standard input can be from a file or the output from another command.

standard output. The primary destination of data coming from a command. Standard output goes to the display unless redirection or piping is used, in which case standard output can be to a file or another command.

status. The current condition or state of a program or device.

storage. (1) The location of saved information. (2) In contrast to memory, the saving of information on physical devices such as disk or tape.

storage device. A functional unit for storing and retrieving data.

storage hierarchy. A logical ordering of storage devices. Generally, the ordering is based on the speed and capacity of the devices.

storage node. A named object that identifies the locations used to hold report data. A storage node can identify cache storage and a Tivoli Storage Manager domain on an Content Manager OnDemand object server.

storage object. A portion of a storage volume managed as a single entity. A storage object can contain many segments of report data.

storage pool. In Tivoli Storage Manager, a named collection of storage volumes that is the destination for archived files.

storage pool volume. In Tivoli Storage Manager, a volume that has been assigned to a storage pool to store archived files.

storage set. A named collection of storage nodes that determines the locations that can hold report data.

storage volume. A volume that has been assigned to hold report data on an Content Manager OnDemand server.

string. A series or set of alphabetic or numeric characters.

structure. A variable that contains an ordered group of data objects. Unlike an array, the data objects within a structure can have varied data types.

structured field. (1) A self-identifying, variable-length, bounded record that can have a content portion that provides control information, data, or both. (2) A mechanism that permits variable length data to be encoded for transmission in the data stream. See field.

subdirectory. In the file system hierarchy, a directory contained within another directory.

subroutine. (1) A sequenced set of statements or coded instructions that can be used in one or more computer programs and at one or more points in a computer program. (2) A routine that can be part of another routine.

summary file. A file that contains all of the values for a particular object.

syntax. The grammatical rules for constructing a command, statement, or program.

syntax diagram. A diagram for a command that displays how to enter the command on the command line.

system console. A console, usually equipped with a keyboard and display screen, that is used by an operator to control and communicate with a system.

system integrity. In computer security, the quality of a system that can perform its intended function in an unimpaired manner, free from deliberate or inadvertent unauthorized manipulation of the system.

System Managed Space (SMS). A type of DB2 table space. An SMS table space is managed by the filesystem manager.

system management. The tasks involved in maintaining the system in good working order and modifying the system to meet changing requirements.

System Management Interface Tool (SMIT). In the AIX operating system, a series of panels that allow you to perform system functions without directly issuing any commands.

system memory. The part of internal storage into which instructions and other data must be loaded for running or processing.

system prompt. The system prompt is the symbol that appears at the command line of an operating system. The system prompt indicates that the operating system is ready for the user to enter a command. See command line.

T

table. A named collection of data consisting of rows and columns.

table reference character (TRC). (1) Usually, the second byte on a line in the user's data. This byte contains a value (0–126) that is used to select a font to

be used to print that line. (2) In the 3800 Printing Subsystem, a numeric character (0, 1, 2, or 3) corresponding to the order in which the character arrangement table names have been specified with the CHARS keyword. It is used for selection of a character arrangement table during printing.

table space. An abstraction of a collection of containers into which database objects are stored. A table space provides a level of indirection between a database and the tables stored within the database. A table space:

- Has space on media storage devices assigned to it.
- Has tables created within it.

tag. (1) A type of structured field used for indexing in an AFP document. Tags associate an index attribute-value pair with a specific page or group of pages in a document. (2) In text formatting markup language, a name for a type of document element that is entered in the source document to identify it.

Tagged Image File Format (TIFF). A bit-mapped graphics format for scanned images with resolutions of up to 300 dpi. TIFF simulates gray scale shading.

TB. terabyte

TCP. Transmission Control Protocol

TCP/IP. Transmission Control Protocol/Internet Protocol

terabyte. A unit of memory or space measurement capacity equal to approximately one trillion bytes. One terabyte is equal to 1,000 gigabytes, or one million megabytes.

text. (1) A type of data consisting of a set of linguistic characters (letters, numbers, and symbols) and formatting controls. (2) In word processing, information intended for human viewing that is presented in a two-dimensional form, such as data printed on paper or displayed on a screen.

throughput. A measure of the amount of work performed by a computer system over a period of time, for example, the number of jobs per day. (I)

TIFF. Tagged Image File Format.

token name. An eight-byte name that can be given to all data stream objects.

token-ring network. A ring network that allows unidirectional data transmission between data stations, by a token passing procedure, such that the transmitted data return to the transmitting station. (T)

toolbar. The region directly beneath the menu bar of the main window in Content Manager OnDemand client programs that support a graphical user interface.

toolbar button. A small bitmap on the toolbar that represents a command in Content Manager OnDemand client programs that support a graphical user interface. Click a toolbar button to quickly access a command.

trailer banner. An optional report delimiter that separates the last report of a bundle from the manifest.

transfer. To send data to one place and to receive data at another place.

transform. To change the form of data according to specified rules without significantly changing the meaning of the data. (I) (A)

Transmission Control Protocol (TCP). A communications protocol used in Internet and in any network that follows the U.S. Department of Defense standards for inter-network protocol. TCP provides a host-to-host protocol between hosts in packet-switched communications networks and in interconnected systems of such networks. It assumes that the Internet protocol is the underlying protocol.

Transmission Control Protocol/Internet Protocol (TCP/IP). A set of communications protocols that support peer-to-peer connectivity functions for both local and wide area networks.

TRC. table reference character

trigger. Data values that ACIF searches for in the input data stream, to delineate the beginning of a new group of pages. The first trigger is then the anchor point that ACIF uses to locate index values.

type. To enter specific information using the keyboard, typing characters exactly as given.

U

UNC. Universal Naming Convention.

unformatted print data. Data that is not formatted for printing. A page definition can contain controls that map unformatted print data to its output format.

upload. To transfer data from one computer to another. Typically, users load data from a small computer to a large one.

user. A person authorized to log on to an Content Manager OnDemand server.

user exit. (1) A point in an IBM-supplied program at which a user-defined program can be given control. (2) A programming service provided by an IBM software product that can be requested during the execution of an application program for the service of transferring control back to the application program upon the later occurrence of a user-specified event.

user interface. The hardware, software, or both that implements a user interface, allowing the user to interact with and perform operations on a system, program, or device. Examples are a keyboard, mouse, command language, or windowing subsystem.

V

value. (1) A set of characters or a quantity associated with a parameter or name. (2) A quantity assigned to a constant, variable, parameter, or symbol.

variable. (1) A name used to represent a data item whose value can change while the program is running. (2) In programming languages, a language object that can have different values at different times. (3) A quantity that can assume any of a given set of values.

version number. The version level of a program, which is an indicator of the hardware and basic operating system upon which the program operates. The version, release, modification, and fix levels together comprise the program level or version of a program.

viewing window. The window in OnDemand where you browse documents.

virtual printer. A view of a printer that refers only to the high-level data stream, such as ASCII or PostScript, that the printer understands. It does not include any information about how the printer hardware is attached to the host computer or the protocol used for transferring data to and from the printer.

volume. The basic unit of storage for a database, log file, or a storage pool. A volume can be an LVM logical volume, a standard file system file, a tape cartridge, or an optical platter. Each volume is identified by a unique volume identifier.

W

Web administrative client. A Web-based administrative client that allows you to add, view, update, and delete users, groups, applications, application groups, folders, printers, and storage sets.

wildcard character. Search characters that represent other letters, numbers, or special characters. In Content Manager OnDemand, the % (percentage) and the _ (underscore) are wildcard characters.

window. A part of a display screen with visible boundaries in which information is presented.

workstation. A terminal or microcomputer, usually one that is connected to a mainframe or to a network, at which a user can perform applications.

write access. In computer security, permission to write to an object.

writer. A JES function that processes print output.

Index

A

B

C

D

E

F

G

H

I

J

L

M

N

O

P

Q

R

S

T

U

W

X

Program Number: 5724-J33

GC19-2938-00