Content Manager OnDemand for Multiplatforms

Introduction and Planning Guide

Version 7.1

GC27-0839-00

Content Manager OnDemand for Multiplatforms

Introduction and Planning Guide

Version 7.1

GC27-0839-00

Note

Before using this information and the product it supports, read the information in "Appendix. Notices" on page 139.

First Edition (March 2001)

This edition applies to IBM® Content Manager OnDemand for Multiplatforms, Version 7 Release 1 and to all subsequent releases and modifications until otherwise indicated in new editions.

Contents

About this publication

This book contains information about IBM Content Manager OnDemand for Multiplatforms Version 7.1 (OnDemand). OnDemand 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, OnDemand requires an archive storage manager, such as Tivoli® Storage Manager (TSM). If you plan to use the OnDemand server print or server FAX functions, then you must install IBM Infoprint® Manager (Infoprint) on a server that is on the same network as the OnDemand library server.

Note: The term *Windows® client* refers to the OnDemand client program that runs under Windows 95, Windows 98, Windows NT® 4.0 and Windows 2000. The term *Windows server* refers to the OnDemand server program that runs under Windows NT Server 4.0, Windows 2000 Server, and Windows 2000 Advanced Server.

Who should use this publication

This book is of primary interest to administrators planning to install, administer, and use 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 OnDemand. This part contains the following sections:

  "Chapter 1. About OnDemand" on page 3 provides an overview of the system

  "Chapter 2. Preparing for OnDemand" on page 19 contains information that can help you prepare your organization for 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. OnDemand and TSM" on page 27 illustrates how OnDemand works with TSM to maintain documents on archive media
- "Part 2. System requirements" on page 33 lists the hardware, software, network, and printing requirements for and other information about OnDemand servers and clients. This part contains the following sections:

"Chapter 4. Hardware and software" on page 35 lists the hardware, software, and licensing requirements for servers and clients, provides information about the different types of server configurations, contains information about downloading data to the server and printing reports from the server

"Chapter 5. Disk storage" on page 57 provides information that can help you plan the disk storage devices required to support the system

- "Part 3. Planning information" on page 65 contains information that can help you define reports[1] to OnDemand, index data, estimate storage requirements, and plan for backup and recovery of data on the system. This part contains the following sections:

  "Chapter 6. Reports and other data" on page 67 contains information that can help you plan for the reports that you will be storing on the system

  "Chapter 7. OnDemand objects" on page 77 contains information that can help you plan application groups, applications, and folders for your reports

  "Chapter 8. Storage requirements" on page 93 provides information that can help you estimate the amount of storage required to maintain reports on the system

  "Chapter 9. Backup and recovery" on page 125 provides information that you can use to develop a backup and recovery plan for the system

The "Glossary" on page 143 contains terms and definitions that you might find helpful as you and others in your organization learn about and use OnDemand.

Where to find more information

OnDemand includes a complete set of information to help you plan for, install, administer, and use the system. OnDemand product documentation is also available on the Web. OnDemand also includes product documentation for other products that you may use on your system.

Information included in your product package

The IBM Content Manager OnDemand for Multiplatforms Documentation CD-ROM contains each publication in the following formats:

1. In 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 an OS/390® application and a two-page Lotus® WordPro file are both reports to OnDemand.

Hypertext Markup Language (HTML)
A set of publications in HTML format are in the HTMLBOOKS directory. You can view the HTML files using a Web browser, such as Netscape Navigator.

Portable document format (PDF)
A set of publications in the Adobe Acrobat portable document format are in the PDFBOOKS directory. You can view the PDF files using the Adobe Acrobat Reader supplied by IBM.

PostScript (PS)
A set of publications in PostScript format are in the PSBOOKS directory. You can print the PostScript files on printers equipped with PostScript fonts.

Table 1 shows the OnDemand publications included on the IBM Content Manager OnDemand for Multiplatforms Documentation CD-ROM.

Table 1. OnDemand publications

File name	Title	Order Number
ARS1P071	Introduction and Planning Guide	GC27-0839
ARS1U071	Installation and Configuration Guide for UNIX® Servers	GC27-0834
ARS1W071	Installation and Configuration Guide for Windows Servers	GC27-0835
ARS5Q071	User's Guide	SC27-0836
ARS5U071	Windows Client Customization Guide and Reference	SC27-0837
ARS1B071	Administrator's Guide	SC27-0840
ARS1D071	Indexing Reference	SC27-0842
ARS5Y071	Web Enablement Kit Installation and Configuration Guide	SC27-1000

Note: You received a printed copy of *Installation and Configuration Guide for UNIX Servers*, *Installation and Configuration Guide for Windows Servers*, and *User's Guide* with the product.

Information available on the Web

You received the full set of OnDemand documentation in softcopy on the documentation CD-ROM (provided with the product). You can also obtain the OnDemand documentation from the Web.

Product documentation is available from the following Web sites:

IBM Content Manager OnDemand
Click **Library** from the product Web site at:
http://www.ibm.com/software/data/ondemand

IBM Publication Ordering System
http://www.elink.ibmlink.ibm.com/public/applications/publications/cgibin/pbi.cgi

Information about other products

The following products are included with the OnDemand product package. Each product includes a full set of documentation in softcopy on CD-ROM.

- DB2 UDB V7

  The DB2 Universal Database library consists of SmartGuides, online help, and books in HTML and PostScript format. See the *Quick Beginnings* book for your server operating system for details about how the DB2 library is structured, how to view information online, and how to print books.

  Note: If you plan to use some other database management product with OnDemand, then you must get the product documentation from the vendor.

- TSM V4.1

  The Tivoli Storage Manager Publications CD-ROM is included in the OnDemand product package. The CD-ROM contains a full set of TSM books in BookManager®, HTML, and PDF format. In addition, you can get the latest versions of TSM publications from the TSM home page on the Web at:

  http://www.tivoli.com/tsm

  Note: If you plan to use some other archive storage management product with OnDemand, then you must get the product documentation from the vendor.

If you plan to use the OnDemand server print or server FAX functions, then you must install and configure IBM Infoprint Manager (Infoprint) on a server that is on the same network as the OnDemand library server. You can get product documentation for Infoprint from the IBM Printing Systems Division Web site:

http://www.printers.ibm.com

If you plan to use OnDemand with IBM Content Manager or the IBM Enterprise Information Portal, you can get product documentation for them from the following Web sites:

IBM Content Manager
http://www.ibm.com/software/data/cm

IBM Enterprise Information Portal
http://www.ibm.com/software/data/eip

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. You can use either of the following methods to provide comments:

- Send your comments from the Web. Visit the IBM Data Management Online Reader's Comment Form (RCF) page at:

  `http://www.ibm.com/software/data/rcf`

- Send your comments by e-mail to:

  `ondemand@us.ibm.com`

  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).

Product support

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

`http://www.ibm.com/software/data/ondemand/`

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

`http://www.ibm.com/software/data/ondemand/`

The IBM support center maintains product updates (called Program Temporary Fix or PTF) for OnDemand. You can obtain the latest PTF from IBM service on the Web at:

`ftp://service.software.ibm.com/software/ondemand/fixes/v71`

If you encounter problems or errors running any of the OnDemand programs, you can call the IBM support center to obtain software problem and defect support. The phone number for the IBM support center is 1-800-237-5511. The OnDemand component ID is 5622-66200.

Part 1. Introduction

This section of the book provides an overview of the OnDemand system and contains information that can help you better understand how OnDemand works. This section describes how OnDemand manages reports and index data, includes important information about how OnDemand, the database manager, and the storage manager work to index, load, and retrieve documents[2], and contains a list of the tasks that OnDemand administrators typically perform to manage an OnDemand system.

2. In 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 OnDemand

Overview

OnDemand supports any organization that can benefit from hard copy or microfiche replacement and instant access to information. An OnDemand system can support small office environments and large enterprise installations with hundreds of system users. OnDemand can dramatically improve productivity and customer service in many businesses by providing fast access to information stored in the system.

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 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.

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

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

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

Some of the advantages that 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

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

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

System overview

An 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.

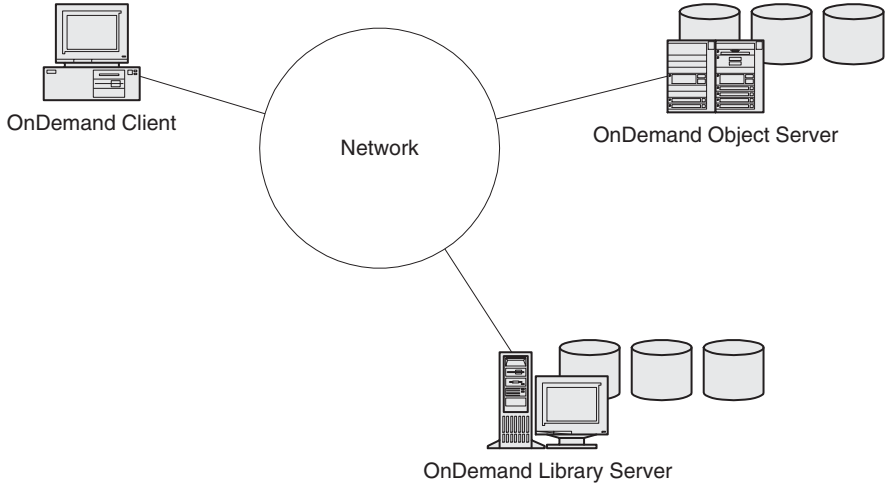

Figure 1. OnDemand system

OnDemand client programs run on PCs and terminals attached to the network and communicate with OnDemand servers. The OnDemand library server manages a database of information about the users of the system and the reports stored on the system. An OnDemand object server manages the reports on disk, optical, and tape storage devices. An 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.

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 OnDemand, view, print, and FAX copies or pages of documents, and attach electronic notes to pages of a document.

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

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 OnDemand library server. The library server returns the list of documents that match the query to the user. When the user selects a document for viewing, the client program retrieves a copy of the document from the object server where the document is stored, opens a viewing window, and displays the document.

Concepts

The terms *application*, *application group*, and *folder* represent how OnDemand stores, manages, retrieves, views, and prints reports and index data. When defining a new report or type of data to 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.

Application

An application describes the physical characteristics of a report to OnDemand. Typically you define an application for each program that produces output that will be stored in 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

3. OnDemand provides the capability to do most client functions from almost any operating system, by using a Web browser. See "OnDemand Web Enablement Kit" on page 16 for information about accessing data stored in OnDemand servers with a Web browser.

and processing instructions that 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 OnDemand. When you load a report into OnDemand, you must identify the application group where OnDemand will load the index data and store the documents. An application group is a collection of one or more 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 OnDemand. A folder provides users with a convenient way to find related information stored in 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 7 illustrates the concepts described in this section.

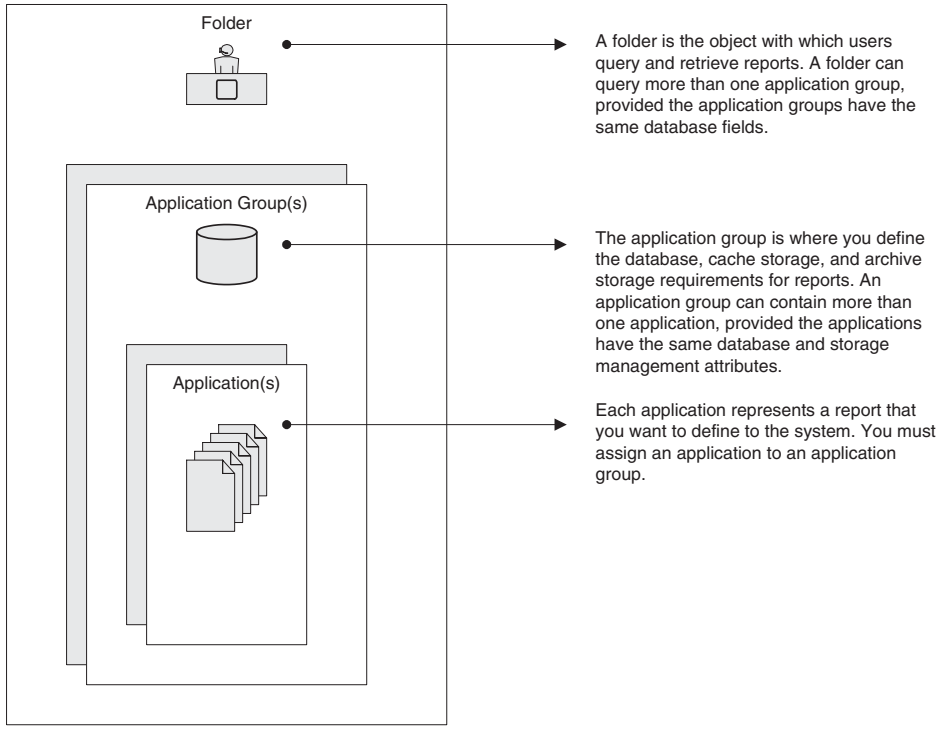

Figure 2. Folders, application groups, and applications (part 1 of 2)

Figure 3 on page 8 shows an example.

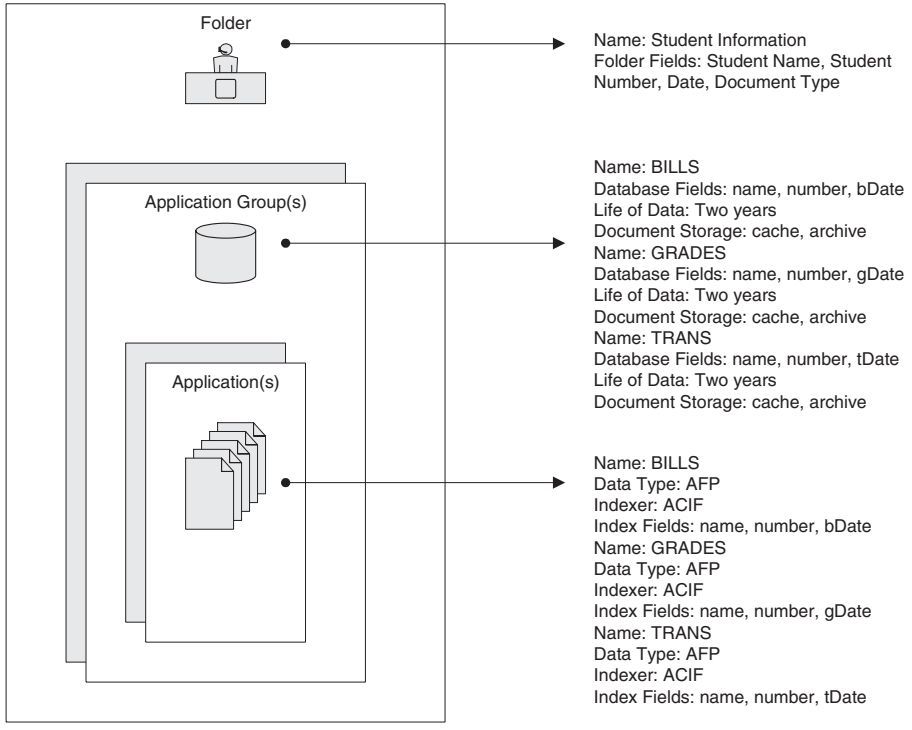

Figure 3. Folders, application groups, and applications (part 2 of 2)

Indexing methods

OnDemand provides two 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. 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 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. OnDemand stores the report as groups of pages and indexes each group. When reports include a sorted transaction value (for example, invoice number), 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

OnDemand documents represent indexed groups of pages. Typically an OnDemand document is a logical section of a larger report, such as an individual customer statement within a report of thousands of statements. An 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, 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 4 illustrates 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.

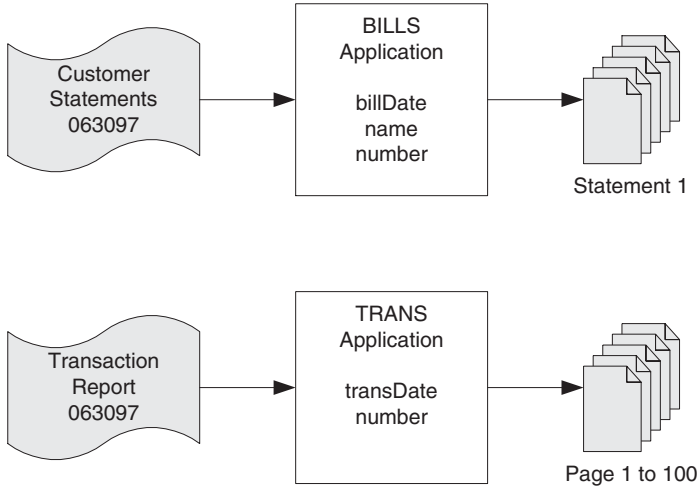

Figure 4. Applications and documents

Servers

The 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 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 on archive media, 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 on archive media. 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 on archive media attached to different servers.

The 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 (which is included with OnDemand). 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 on archive media, such as optical and tape storage libraries. TSM (which is included with OnDemand) is an example of an archive storage manager product. You can also use TSM to maintain DB2 archived log files and backup image files. This capability means that you do not have to manage these files on disk.
- A *download facility* that automatically transfers spool files to a server at high speed. We recommend that you use Download for OS/390, a licensed feature of Print Services Facility™ (PSF) for OS/390. Download provides the automatic, high-speed download of JES spool files from an OS/390 system to OnDemand servers.
- *Data indexing and conversion* programs. These programs create index data, collect required resources, and optionally convert line data reports to AFP data. OnDemand provides several indexing programs. The AFP Conversion and Indexing Facility (ACIF) can be used to index S/390® line data, ASCII data, and AFP files, collect resources required to view the reports, and convert line data files to AFP data. The OnDemand PDF Indexer can be used to create index data for Adobe Acrobat PDF files. The OnDemand Generic Indexer can be used to create index data for almost any other type of data that you want to store on the system, such as HTML documents, Lotus WordPro documents, TIFF files, and so forth. The indexing programs can run on any OnDemand server; ACIF can also run on an OS/390 system.
- *Data loading* programs that can be set up to automatically store report data into application groups and update the database. The data loading programs can run on any OnDemand server.
- Archived reports and resources.
- A *server print* facility that allows users to reprint a large volume of documents at high speed. OnDemand uses Infoprint, which must be purchased separately, to manage the server print devices.
- OnDemand *management programs* to maintain the 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 OnDemand request manager
- The OnDemand database manager
- The OnDemand storage manager
- Download
- Data indexing and loading
- OnDemand management programs

Request manager

The request manager processes search requests from 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 cache storage manager, if the document resides in cache storage, or to 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 OnDemand client program decompresses and displays the document.

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, 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 OnDemand programs in the System Log, for example, logon, query, print, and so forth. The messages contain the transaction number, userid, time stamp, and other information. Administrators can open the System Log folder and view the messages. 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

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

Storage manager

The 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 command to migrate documents from cache storage to archive media and to remove documents that have passed their life of data period.

OnDemand also supports an archive storage manager, such as TSM (which is provided with OnDemand). The archive storage manager maintains one or more copies of documents on archive media, such as optical or tape storage libraries. You decide which types of archive media that your 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 on archive media, 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 on archive media, TSM can also maintain DB2 archived log files and backup image files. This capability means that you do not have to maintain these files on disk. TSM can assist you with automating database backup and archival on a regular schedule. When you use the arsdb command to create a database or table space backup image, you can specify that you want TSM to manage the image. After completing the backup image, the arsdb command copies the archived log files to TSM-managed storage.

Download

Download is a licensed feature of PSF for OS/390. Download provides the automatic, high-speed download of JES spool files from an OS/390 system to an IBM Content Manager OnDemand for Multiplatforms server. Download can be used to transfer reports created on OS/390 systems to the server, where you can configure 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 an OS/390 system to the OnDemand server. See *PSF for OS/390: Download for OS/390* for more information about Download.

Data indexing and loading

The reports that you store in OnDemand must be indexed. 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. OnDemand data loading programs read the index data generated by ACIF and load it into the OnDemand database. The data loading programs obtain other processing parameters from the OnDemand database, such as parameters used to segment, compress, and store report data in cache storage and on archive media. If you plan to index reports on an 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. 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 files. See the *Indexing Reference* for details about the indexing programs provided with OnDemand.

Figure 5 shows an overview of the data preparation process.

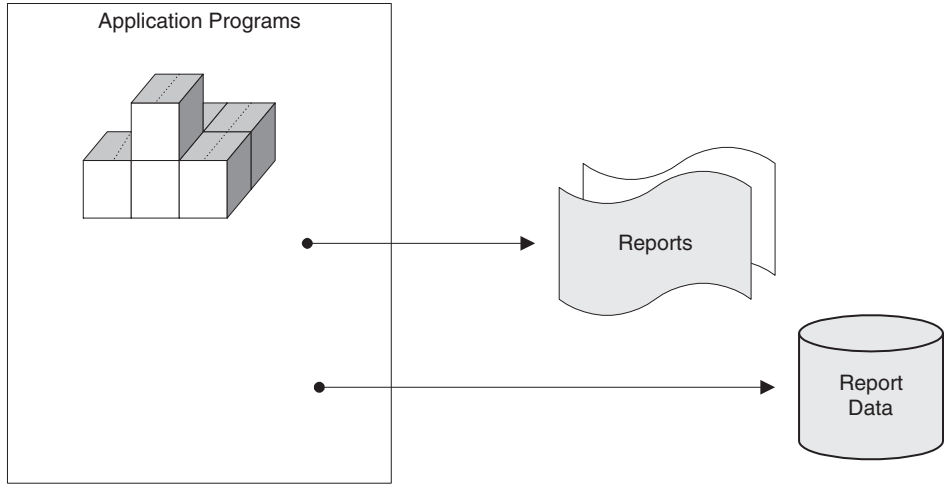

Figure 5. Data preparation, indexing, and loading (part 1 of 2)

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

Figure 6 shows an overview of the data indexing and loading process.

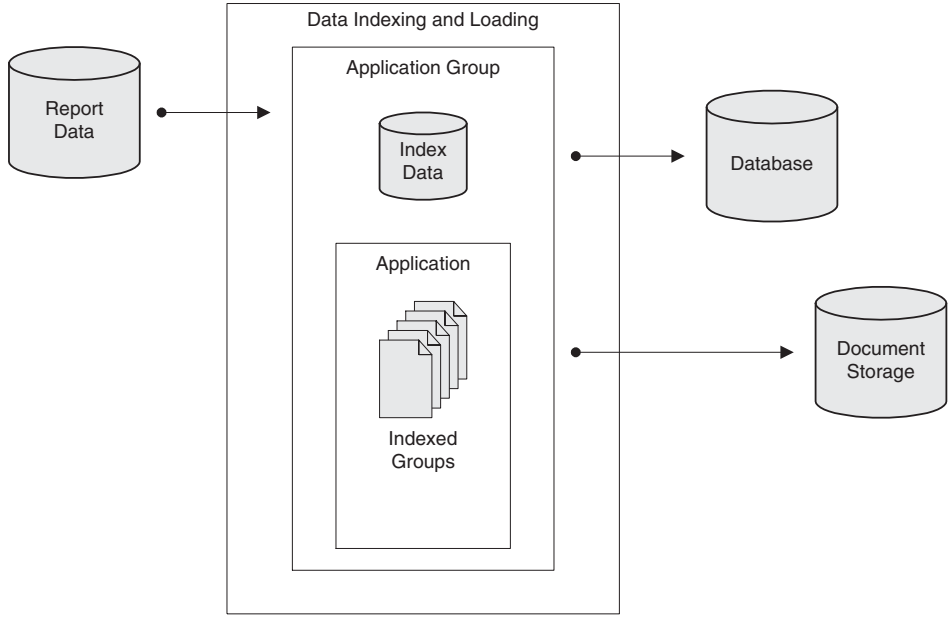

Figure 6. Data preparation, indexing, and loading (part 2 of 2)

The 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 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 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

OnDemand provides programs to maintain and optimize the database and maintain documents on 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 media. The programs use the information to migrate documents from cache storage to archive media, delete documents from cache storage, migrate index data from the database to archive media, and delete index data from the database. These functions are useful because OnDemand can reclaim the database and cache storage space released by expired and migrated data. We recommend that you configure your 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 media when it reaches its storage expiration date. An administrator defines management information to the archive storage manager to support the OnDemand data it manages. The management information includes the storage libraries and storage volumes that can contain OnDemand data, the number of copies of a report to maintain, and how long to keep data in the archive management system.

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 specify the same criteria to OnDemand and the archive storage manager. For example, The Life of Data, which is used by OnDemand, and the Retention Period, which is used by TSM, should specify the same value.

OnDemand Web Enablement Kit

The OnDemand Web Enablement Kit (ODWEK) is an optional feature of OnDemand that allows people in an organization to use a Web browser to access data stored in an OnDemand system[4]. 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 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 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.

4. The data must be stored in an OnDemand for Multiplatforms system running AIX®, HP-UX, Sun Solaris, or Windows server.

Figure 7 shows a workstation with a Web browser that is being used to access data from an OnDemand server.

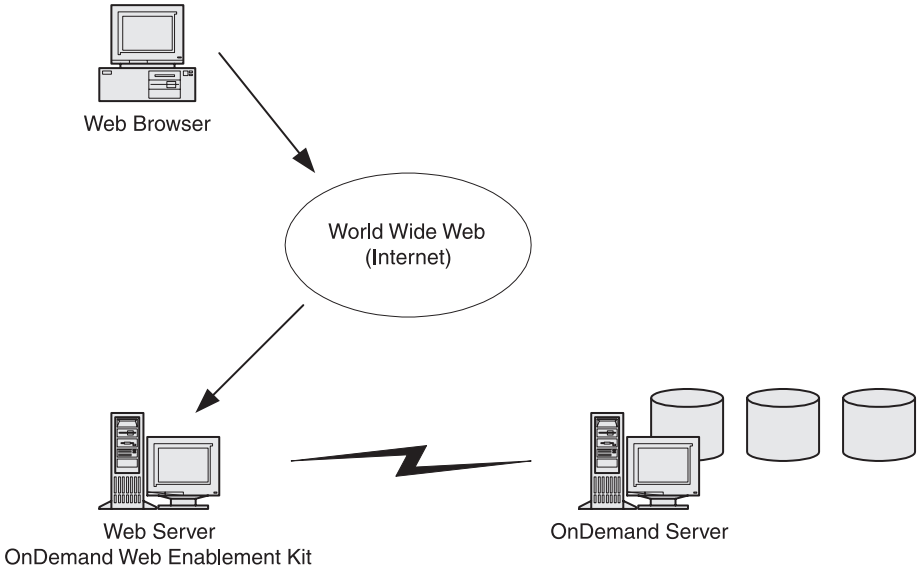

Figure 7. Accessing data stored in OnDemand using ODWEK

ODWEK contains several components:

- The Web server program. The server program uses standard OnDemand interfaces and protocols to access data stored in an OnDemand server. No additional code is needed on the OnDemand server to support ODWEK. You can use one of the following Web server programs to control ODWEK:
  - CGI program. The CGI program runs on a workstation running an HTTP server, such as the IBM HTTP Server.
  - Java™ servlet. The servlet runs on a Java-enabled HTTP server with a Java application server, such as the IBM WebSphere™ Application Server.
- The AFP plug-in. The AFP plug-in lets users search, retrieve, view, navigate, and print AFP documents from a Web browser. Each person in your organization that plans to use the AFP plug-in to view AFP documents must install it on their PC.
- The Image plug-in. The Image plug-in lets users search, retrieve, view, navigate, and print BMP, GIF, JPEG, PCX, and TIFF documents from a Web browser. Each person in your organization that plans to use the Image plug-in to view documents must install it on their PC.
- 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 IBM AFP2WEB Transform service offering. The AFP2WEB Transform converts AFP documents and resources into HTML files that can be displayed with the AFP2HTML applet. If you plan to use the AFP2HTML applet, then you must obtain the AFP2WEB Transform from IBM and install and configure it on the Web server. See your IBM representative for more information about the AFP2WEB Transform.

Note: To view other types of documents stored in OnDemand, you must obtain and install the appropriate plug-in. For example, to view Adobe Portable Data Format (PDF) documents, we recommend that you obtain the Adobe Acrobat Reader plug-in for the browsers used in your organization.

If you plan to convert AFP documents that are stored in OnDemand into PDF documents that you can view with the Acrobat Reader plug-in, then you must obtain the AFP2PDF Transform from IBM. See your IBM representative for more information about the AFP2PDF Transform.

To use the plug-ins and applets, the browser must be Netscape Navigator Version 4.06 or later or Internet Explorer Version 4.01 or later.

See *Web Enablement Kit Installation and Configuration Guide* for more information.

Chapter 2. Preparing for OnDemand

This section contains an outline that you may find helpful as you prepare your organization for the 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 OnDemand will affect their daily work.
- Develop an end-user training course or contract with IBM to provide training for 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 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 data to HTML data, we recommend that you use the AFP2WEB Transform, a service offering available from IBM.
- Identify the type of data contained in the report and determine how you will create index data.
- If you plan to use ACIF to index the report, 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 OnDemand. If you plan to index reports on an OS/390 system, then resources can be gathered into a resource group file. If you plan to index reports on an OnDemand server, then you must either transmit the resource group file to the server or provide access to the resource group file using the Network File System (NFS).
- If you plan to use ACIF to index the report, decide where to index reports: on an OS/390 system or an OnDemand server. Determine how to transmit

report and index data from the OS/390 system to the OnDemand server. We recommend that you use Download to transmit data from the JES spool to file systems on OnDemand servers. See *PSF for OS/390: Download for OS/390* for details about installing, configuring, and using Download.

- Configure cache storage and archive storage (optical and tape storage devices) on the OnDemand servers. Define and configure archive media devices to TSM. Define storage management policies to TSM to support the reports that you plan to store on the system.
- Use the 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 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 want OnDemand 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 OnDemand with the information that you collect.
- Survey users about their satisfaction with 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 OnDemand system. Update your company's operations and recovery manuals with information required to operate, support, and backup the OnDemand system.

Administrative roles and responsibilities

OnDemand administrators assume responsibility for and take care of the OnDemand system. The 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 media devices, client PCs, 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 OnDemand applications and decide how OnDemand will manage data on the servers.
- Administrators define OnDemand groups and users to the system and make sure that the client software is installed and operating properly.

While OnDemand administrators are responsible for this collective environment from the viewpoint of 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 OnDemand system administrator could maintain OnDemand storage sets, system printers, groups, and users; an 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 OnDemand system. Some of these tasks may be the responsibility of a person other than an OnDemand administrator.

- Installing and upgrading equipment
- Installing and maintaining 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 OnDemand system printers

- Defining reports to the system
- Defining 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

OnDemand provides an administrative client to allow administrators to maintain OnDemand objects through an easy-to-use, graphical user interface. The administrative client is a 32–bit Windows application. The administrative client allows administrators to define and maintain application groups, storage sets, storage nodes, folders, 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.

OnDemand provides a set of administrative commands to help administrators maintain the system. For example, 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

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

Client customization

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 OnDemand from other Windows 32-bit applications with the Dynamic Data Exchange (DDE) interface, or by creating a Product Information File (PIF).

The *Windows Client Customization Guide and Reference* provides information about customizing the Windows clients.

For example, you can integrate Monarch[5] Version 5 with the Windows client so that users can load 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 client *Customization Guide* for more information.

Server programs

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

- The ARSADM program is a utility for maintaining users, groups, printers, and storage sets. You can use the ARSADM program to add, delete, and update users, groups, printers, and storage sets. You can run the ARSADM program from the command line or invoke it from a user–defined 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 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–defined program.

The *Administrator's Guide* provides details about these and other server programs.

Server logging

System logging facility

OnDemand provides the system logging facility to help an administrator track activity and monitor the system. OnDemand can record the messages that are generated by the various client and server programs. For example, you can configure the system to record a message in the system log every time a user logs on the system; you can configure the system to record a message in the system log every time an unsuccessful log on attempt occurs; and so forth. When you use the administrative client to add objects to the system and update the database, OnDemand records information about your actions in the system log. You can use one of the OnDemand client programs to search for and view messages from the system log by time stamp, severity, message number, userid, and other search criteria.

System log user exit

OnDemand provides a user exit that can be used to process the messages that are written to the OnDemand system log. A common use of the system log

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

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 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.

OnDemand sends parameters to the system log user exit, such as the name of the OnDemand instance, a time stamp, a log record identifier, the userid that is associated with the action, accounting information for the userid, 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 OnDemand to provide specific information to the system log user exit by setting system and application group parameters with the administrative client.

See the *Installation and Configuration Guide* for your server operating system for more information about the system log user exit.

Logon user exit

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 logon 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 logon user exit to allow users that are not already in the OnDemand user database to access the system.

The logon 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 logon 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 logon user exit program.

Important: When you implement your own logon user exit program, you bypass the logon verification processing built into the base OnDemand product. We advise caution when you bypass the 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.

Please contact the IBM support center for more information about implementing a logon user exit.

License information

The License Information to use OnDemand, DB2, and TSM is included in the product package. You should read the License Information carefully before using the software that is provided with OnDemand.

If you plan to use Oracle instead of DB2, then 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, then you may need a Client Access License (CALs) for each concurrent connection to the database. Contact your Microsoft® representative for more information.

Chapter 3. OnDemand and TSM

Introduction

You can configure an OnDemand system to maintain copies of reports in cache storage and in archive storage. The copies in archive storage are for long-term storage. TSM is the product that OnDemand works with to maintain reports in archive storage. TSM supports a variety of optical and tape storage devices. TSM 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 TSM.
- 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 TSM maintains the data.
- An API that OnDemand uses to work with TSM. The TSM API is required on the library server and all object servers that use TSM.
- Device support modules which provide support for storage devices and storage libraries.

OnDemand storage objects

The storage management criteria that you specify on the library server determines where and when OnDemand stores reports and how it maintains them. Figure 8 on page 28 shows the primary OnDemand storage objects.

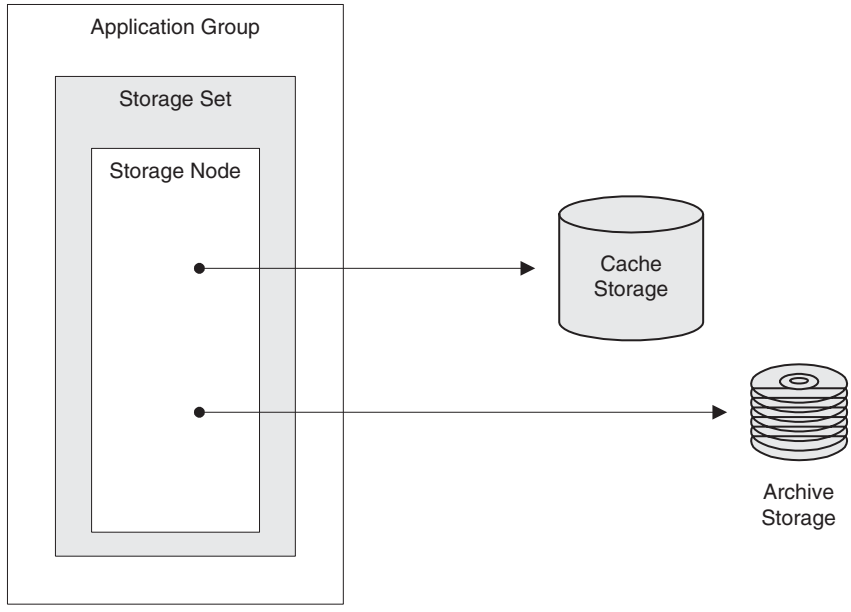

Figure 8. OnDemand storage objects

An administrator creates an 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 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. 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 TSM. If a storage node identifies a client node in TSM, then OnDemand automatically stores a copy of the report in archive storage, which is managed by TSM.

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

Overview of TSM

This section describes the basic TSM concepts. For more information about TSM and details about storage policies and devices and managing TSM storage, please see the TSM *Administrator's Guide*.

Storage policy

Client node
Represents an object server on which the TSM 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 TSM for long-term storage

Storage devices and media

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

Drive Each drive defined to TSM 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 9 on page 30 illustrates the concepts discussed in this section:

- A client node is registered in a policy domain. The other TSM 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.

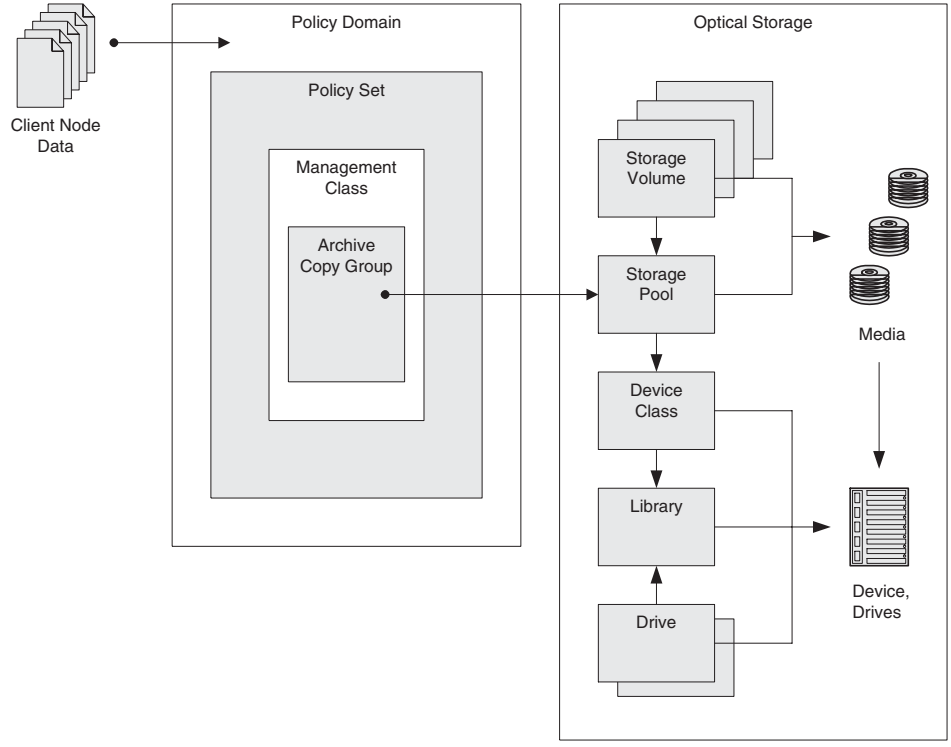

Figure 9. TSM storage objects

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 maintaing 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. The IBM 3995 Model C68 Optical Library is an example of a high-capacity optical storage library. The 3995-C68 library contains four drives that support optical platters with an unformatted capacity of up to 5.2 GB each. You can store up to 258 optical platters in the library at any one time, providing a total online (unformatted) capacity of 1.341 TB of data. Since OnDemand compresses data before storing it on storage volumes, you may obtain a much greater storage capacity. For example, if you achieve 5:1 compression on the data stored in a 3995-C68 library, you can store over six terabytes of data, online, in the library.
- 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 TSM can automatically maintain a backup copy of the reports.

After collecting the storage requirements, you typically work with a TSM administrator to configure storage devices on the system and define devices to TSM. For example, when you define an optical library to TSM, you specify the type of device, the number of drives in the library, and the capacity of the storage volumes. The TSM 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. TSM 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 TSM 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 TSM 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 OnDemand object server and, optionally, a client node in TSM. If the storage node identifies a client node in TSM, 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

OnDemand provides the ARS_ADSM program for use on UNIX object servers. You can use the ARS_ADSM program to create backup copies of the TSM database.

Each storage set identifies the object server and storage node(s) where data is to be written. OnDemand can write data to one storage node at a time. (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.

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. OnDemand does not require TSM to compress the storage objects. OnDemand extracts and decompresses a small portion of the storage object, as required, when users retrieve the report.

TSM 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, we recommend 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. OnDemand and TSM 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 OnDemand and to TSM. The Life of Data and Indexes, which is used by OnDemand, and the Retention Period, which is used by TSM, should be the same value.

Part 2. System requirements

This section of the book contains information about the hardware, software, networking, and printing requirements for OnDemand servers and clients.

Chapter 4. Hardware and software

The exact hardware and software configuration that you need for OnDemand to support your organization depends on the volume of data that you plan to maintain on the system, the number of concurrent users that the system must support, the backup and recovery requirements of your organization, and the performance levels that the system must meet. At a minimum, you need one processor for a standard OnDemand library/object server.

AIX server requirements

The minimum configuration for an OnDemand for AIX server requires the following:

- AIX Version 4.3.3 or later
- DB2 Universal Database Enterprise Edition Version 7 or later (included with OnDemand) or Oracle 8i (8.1.5.0) on the library server
- RS/6000® Model F50
- 128 megabytes of memory
- TCP/IP
- Magnetic and optical storage for the database, report data, and temporary work space. A minimum of eight gigabytes of space and two separate disks are required, in addition to sufficient disk space for report storage.
- A CD-ROM drive for program installation
- A 4mm or 8mm tape drive, automated tape library, or optical library for data backup and recovery. A tape drive can be used with a small system configuration. An automated tape library or dedicated optical library is recommended for medium and large system configurations. (See "Typical server hardware configurations" on page 44 for details about system configurations.)
- TSM Version 4.1 or later (included with OnDemand), if you plan to maintain report data on archive media or maintain DB2 archived log files and backup image files on archive media. You must install TSM on at least one object server.
- Infoprint, if you plan to use the OnDemand server print or server FAX functions. You must install Infoprint on a system that belongs to the same network as the OnDemand library server. See "Printing" on page 42 for more information.

You can use the following general guidelines to help determine the type of RS/6000 required to support OnDemand.

Processor

Most customers will need at least one RS/6000 Model J40 server.

A system that supports a small workgroup of users or a system on which you plan to load or maintain very little data on disk can usually run on a single RS/6000 Workgroup Server Model F50.

A system that supports a large organization with a large number of concurrently logged on users or a system on which you plan to load and maintain a large amount of data on disk and archive media may require one or more RS/6000 Enterprise Server Model R50 systems or a Scalable POWERparallel® (SP) large scale server.

Memory

You should plan to configure each server with at least 128 MB of memory to support the basic OnDemand programs. You should plan to add 2 MB of memory for each concurrently logged on OnDemand user.

HP-UX server requirements

The minimum configuration for an OnDemand for HP-UX server requires the following:

- HP-UX Version 11.00 or later
- DB2 Universal Database Enterprise Edition Version 7 or later (included with OnDemand) or Oracle 8i (8.1.5.0) on the library server
- HP 9000 server
- 128 megabytes of memory
- TCP/IP
- Magnetic and tape storage for the database, report data, and temporary work space. A minimum of eight gigabytes of space and two separate disks are required, in addition to sufficient disk space for report storage.
- A CD-ROM drive for program installation
- A 4mm or 8mm tape drive or automated tape library for data backup and recovery. A tape drive can be used with a small system configuration. An automated tape library is recommended for medium and large system configurations. (See “Typical server hardware configurations” on page 44 for details about system configurations.)
- TSM Version 4.1 or later (included with OnDemand), if you plan to maintain report data on archive media or maintain DB2 archived log files and backup image files on archive media. You must install TSM on at least one object server.

- Infoprint, if you plan to use the OnDemand server print or server FAX functions. You must install Infoprint on a system that belongs to the same network as the OnDemand library server. See “Printing” on page 42 for more information.

You can use the following general guidelines to help determine the type of HP-UX system required to support OnDemand.

Processor

Most customers will configure a system with one or more HP-UX 9000 servers.

A system that supports a small workgroup of users or a system on which you plan to load or maintain very little data on disk can usually run on a single HP-UX 9000 server.

A system that supports a large organization with a large number of concurrently logged on users or a system on which you plan to load and maintain a large amount of data on disk and archive media may require one or more HP-UX 9000 Series 800 Business servers.

Memory

You should plan to configure each server with at least 128 MB of memory to support the basic OnDemand programs. You should plan to add 2 MB of memory for each concurrently logged on OnDemand user.

Solaris server requirements

The minimum configuration for an OnDemand for Sun Solaris server requires the following:

- Solaris Version 2.6 or later
- DB2 Universal Database Enterprise Edition Version 7 or later (included with OnDemand) or Oracle 8i (8.1.5.0) on the library server
- Sun Enterprise Ultra 10S server
- 128 megabytes of memory
- TCP/IP
- Magnetic and optical storage for the database, report data, and temporary work space. A minimum of eight gigabytes of space and two separate disks are required, in addition to sufficient disk space for report storage.
- A CD-ROM drive for program installation
- A 4mm or 8mm tape drive, automated tape library, or optical library for data backup and recovery. A tape drive can be used with a small system configuration. An automated tape library or dedicated optical library is

recommended for medium and large system configurations. (See "Typical server hardware configurations" on page 44 for details about system configurations.)

- TSM Version 4.1 or later (included with OnDemand), if you plan to maintain report data on archive media or maintain DB2 archived log files and backup image files on archive media. You must install TSM on at least one object server.
- Infoprint, if you plan to use the OnDemand server print or server FAX functions. You must install Infoprint on a system that belongs to the same network as the OnDemand library server. See "Printing" on page 42 for more information.

You can use the following general guidelines to help determine the type of Sun Solaris system required to support OnDemand.

Processor

Most customers will need at least one Sun Enterprise 3500 server.

A system that supports a small workgroup of users or a system on which you plan to load or maintain very little data on disk can usually run on a single Sun Enterprise Ultra 10S server.

A system that supports a large organization with a large number of concurrently logged on users or a system on which you plan to load and maintain a large amount of data on disk and archive media may require one or more Sun Enterprise 5500 servers or a Sun Enterprise 10000 server.

Memory

You should plan to configure each Solaris workstation with at least 128 MB of memory to support the basic OnDemand programs. You should plan to add 2 MB of memory for each concurrently logged on OnDemand user.

Windows server requirements

The minimum configuration for an OnDemand Windows server requires the following:

- Microsoft Windows 2000 Server, Microsoft Windows 2000 Advanced Server, or Microsoft Windows NT Version 4.0 SP3 or later
- DB2 Universal Database Enterprise Edition Version 7 or later (included with OnDemand), Microsoft SQL Server 2000, or Oracle 8i (8.1.5.0) on the library server
- An IBM-compatible PC with an Intel Pentium® 166 MHz or faster processor
- 128 MB of RAM
- TCP/IP

- A minimum of 8 GB of disk storage on at least two physical disks plus sufficient storage space for documents
- A CD-ROM drive for program installation (required)
- A 4mm or 8mm tape drive, automated tape library, or optical library for data backup and recovery. A tape drive can be used with a small system configuration. An automated tape library or dedicated optical library is recommended for medium and large system configurations. (See "Typical server hardware configurations" on page 44 for details about system configurations.)
- TSM Version 4.1 or later (included with OnDemand), if you plan to maintain report data on archive media or maintain DB2 archived log files and backup image files on archive media. You must install TSM on at least one object server.
- Infoprint, if you plan to use the OnDemand server print or server FAX functions. You must install Infoprint on a system that belongs to the same network as the OnDemand library server. See "Printing" on page 42 for more information.

You can use the following general guidelines to help determine the type of Windows server system required to support OnDemand.

Processor

Most customers will need at least one IBM compatible PC with two or more Intel Pentium II 266 MHz or faster processors.

A system that supports a small workgroup of users or a system on which you plan to load or maintain very little data on disk can usually run on a single IBM compatible PC with an Intel Pentium 166 MHz or faster processor.

A system that supports a large organization with a large number of concurrently logged on users or a system on which you plan to load and maintain a large amount of data on disk and archive media may require several IBM compatible PCs, each with multiple Intel Pentium III 550 MHz or faster processors.

Memory

You should plan to configure each Windows server with at least 128 MB of memory to support the basic OnDemand programs. You should plan to add 2 MB of memory for each concurrently logged on OnDemand user. You should also consider using ECC (Error Correction Code) memory for increased system reliability.

Optical and tape storage

The recommended option to support maintaining data on archive media (optical and tape storage) requires:

- TSM Version 4.1 or later Server with the appropriate optical or tape device support module
- TSM Version 4.1 or later Administrative Client with a minimum of a 10 concurrent-user TSM license. (More users may be required depending on the number of client nodes and storage pools that you define to TSM to support your OnDemand system.)
- TSM Version 4.1 or later TSM API

Note: TSM Version 4.1 is provided with OnDemand. Your OnDemand license permits you to install one TSM server with Extended Device Support and as many TSM clients as necessary to support your licensed use of OnDemand. If you wish to install additional TSM servers, then you must purchase additional Proofs of Entitlements.

If you plan to use TSM to maintain files on optical and tape storage volumes, consult the documentation provided with TSM for a list of the devices supported by the TSM server. Table 2 lists IBM optical libraries supported by TSM on AIX, Sun Solaris, and Windows servers. The IBM 3995 C-Series Optical Library family features the IBM industry-standard 5.25 inch, 5.2 GB (8x) Extended Multifunction Optical Drive. The 5.2 GB (8x) drives, when used with 5.2 GB optical cartridges, provide library capacities from 104 GB to 1.341 TB of online optical storage, depending on model configuration. The 3995 library supports magneto-optical (MO) rewritable, permanent write once/read many (WORM), and Continuous Composite WORM (CCW, also known as non-permanent WORM) optical media technologies.

Table 2. Optical Libraries

Device	Drives	Storage Cells	Uncompressed Capacity	SCSI IDs Required
IBM 3995-C60	1, 5.25 inch	20	104 GB	2
IBM 3995-C62	2, 5.25 inch	52	270 GB	3
IBM 3995-C64	2, 5.25 inch	104	540 GB	3
IBM 3995-C66	4 or 6, 5.25 inch	156	810 GB	5 or 7
IBM 3995-C68	4, 5.25 inch	258	1.341 TB	5

TSM also supports optical libraries from Hewlett-Packard, ATG Cygnet, Kodak, DISC, and Phillips/LMS, tape libraries from IBM and StorageTek, and

many other storage devices. Visit the Tivoli home page at http://www.tivoli.com/tsm for the latest information about supported devices.

Download

There are three methods that most customers use to transfer reports from an OS/390 system to an OnDemand server:

- SNA protocols
- TCP/IP protocols
- Download

There are advantages and disadvantages to each approach.

SNA protocols

You can use third-party RJE packages to provide a seamless download process from the JES spool to queues on an OnDemand server. This option works well for low volume applications or line of business applications with less critical availability requirements.

TCP/IP

TCP/IP provides a wide variety of options for data download. TCP/IP is a powerful and flexible industry-standard way of connecting multiple systems. There are many ways you can customize TCP/IP to fit the specific needs of your system.

TCP/IP provides very fast data download. Up to one gigabyte of data can be transferred per hour over a 16 megabyte Token Ring gateway (IBM 3172 controller). Even faster speeds can be obtained by directly attaching the host system and the OnDemand server over an ESCON® channel.

Several TCP/IP protocols can be used to move print data between the host system and an OnDemand server:

- FTP

  FTP permits data transfer between dissimilar host systems. FTP is designed to be used by applications, creating and removing datasets and directories and transferring multiple files in a single request. FTP provides security by passing user and account passwords between hosts.
- LPR/LPD

  LPR is the TCP/IP queueing facility. LPR places print data streams on print queues by copying files into spooling datasets. LPD is the printer daemon that controls remote printing between print servers and other hosts on the network. With OnDemand, LPD can be used to transfer files from a host or LAN system to an OnDemand server.
- NFS

NFS allows host systems on the TCP/IP network to share datasets and disks. The shared datasets and disks appear to each remote host system as just another local dataset or disk.

Download

Download Version 3 Release 2.0 for OS/390 is a licensed feature of Print Services Facility (PSF) for OS/390. Download automatically transmits the output of OS/390 application programs to OnDemand servers for indexing and loading. Download transmits the output from the JES spool to file systems on the server using a JES class or printer destination. You can configure your OnDemand system to automatically process the data that is transmitted by Download. See *PSF for OS/390: Download for OS/390* for detailed information about installing, configuring, and using Download.

To use Download, the OS/390 system requires:

- OS/390 Version 2 Release 7 or later
- OS/390 e-Network Communications Server with the latest PTFs
- PSF for OS/390 Version 3 Release 2 or later
- Download for OS/390 feature of PSF for OS/390

Indexing with ACIF

ACIF is an optional feature of PSF for OS/390 that lets you convert a S/390 line data print file to an AFP document, retrieve the resources used by the document, and index the file for archiving and viewing with OnDemand.

To use ACIF on an OS/390 system requires:

- OS/390 Version 2 Release 7 or later
- PSF for OS/390 Version 3 Release 2 or later
- ACIF feature of PSF for OS/390

You can index reports with either the standard ACIF (free with PSF) or an enhanced version of ACIF that can be ordered without charge by customers who are entitled to OnDemand. See the README file provided with the OnDemand product package for ordering information.

Printing

The OnDemand clients let users print directly from the document that they are viewing or print selected items from the document list. There are three ways that users can print from the client:

- Print to a local printer. The OnDemand client can reprint all types of documents (AFP, line data, and image file) under Windows. This method is most commonly used to print to PostScript and PCL printers that are already defined and used by other programs on the user's PC.

When printing AFP documents to local printers, the fidelity of the output can vary. The best fidelity is achieved when printing AFP documents that have been formatted with the IBM Core Interchange Fonts. If an exact reprint of the document is required, then you must use one of the other printing methods, which requires Infoprint.

- Print to devices that are attached to the network and managed by Infoprint. Some users will redirect a local printer port on their PC to an Infoprint print queue. This method can allow reprinting of high volumes of documents at high speed, depending on the capabilities of the network printer, because the printing is off-loaded to a print server.
- Print through the OnDemand server print function. This is the highest performance print option, because the documents are not retrieved to the user's PC before printing. The server print function is also designed to allow many documents to be selected for reprint from the document list. When the OnDemand server print function is used, the client sends a print request to the OnDemand server. The server sorts the documents by storage volume before retrieving them. After the documents are retrieved, a print job is submitted to a server print queue, which is then processed by Infoprint. When a user prints an AFP document, the resource group that was archived at the time that the document was loaded into the system is put into the print data stream, to make sure that the document prints with the correct resources.

  An OnDemand server printer is an interface between the user and a server print device. OnDemand supports two types of server print devices: a FAX machine and a physical printer. A server print device can be physically connected to the library server or it can be attached to some other workstation on the network. Server print devices are managed by Infoprint. A server printer identifies an Infoprint print queue.

  To configure the system to support the server print function, you must do the following:

  – Define a server printer to OnDemand with the administrative client; the server printer identifies a print queue in Infoprint
  – Identify the TCP/IP host name or IP address of the workstation on which the Infoprint server is running
  – Install and configure Infoprint on a workstation that is attached to the same network as the OnDemand library server
  – Define Infoprint printers, print devices, and print queues

Printing jobs with PSF for OS/390

The AFP Upload feature of PSF for OS/390 lets you submit a print job to Infoprint Manager for AIX for printing on any printer that is supported by PSF for OS/390. AFP Upload receives the print data from AIX and places it on the JES spool for printing by PSF. Jobs submitted to PSF can contain any type of data stream that PSF can transform to MO:DCA-P.

Printing documents with PSF for OS/390 is desirable for customers that want to reprint archived bills, statements, and so forth through the same process that was used to print the original output data stream.

Server printing requirements

The OnDemand server print function requires:

- Infoprint Manager for AIX Version 3 Release 2 or later or Infoprint Manager for Windows NT and Windows 2000
- PSF Version 3.2 or later for OS/390 with the AFP Upload feature, if you plan to submit print jobs from Infoprint Manager for AIX to PSF for OS/390

Typical server hardware configurations

You can use the following general guidelines to help determine the server hardware required to support OnDemand. The guidelines are based on the maximum number of users logged on to the server at any one time and the average number of pages stored on the system each month. The configurations include enough disk storage to cache reports for three months. Each configuration includes a data backup device.

When you plan your OnDemand system configuration, the business, legal, and operational requirements of your organization may result in a system configuration that is much different than ones we recommend here. For example, if your systems need to be available to users at all times, then you will likely require more than one physical server and significantly more disk storage and archive storage. If you plan to use TSM to maintain DB2 database log files and backup image files, we recommend that you include an automated tape library or a dedicated optical library in your configuration. In addition, the configurations shown here **do not** include archive media storage for mirroring of databases or maintaining a backup (second) copy of reports in archive storage. In addition to understanding your organization's business and operational needs, we recommend that you use other information in this publication to help determine the hardware required for your OnDemand system.

Small server configuration

The following OnDemand server hardware can usually support up to 15 concurrently logged on users. The amount of data loaded into the system in one year, at a rate of 200,000 pages a month, requires approximately 6 GB of archive storage. The configuration includes enough disk storage to cache reports for three months.

Table 3. Small Server Configuration

Memory	Cache Storage	Archive Media Storage	Backup Device
128 MB	8 GB	One ATL or IBM 3995-C60	One 8 mm tape drive

Medium server configuration

The following OnDemand server hardware can usually support up to 100 concurrently logged on users. The amount of data stored in the system in one year, at a rate of 4,000,000 pages a month, requires approximately 120 GB of archive storage. The configuration includes enough disk storage to cache reports for three months.

Table 4. Medium Server Configuration

Memory	Cache Storage	Archive Media Storage	Backup Device
256 MB	54 GB	One ATL or IBM 3995–C64	Second ATL or optical library

Large server configuration

The following OnDemand server hardware can usually support up to 500 concurrently logged on users. The amount of data loaded into the system in one year, at a rate of 20,000,000 pages a month, requires approximately 600 GB of archive storage. The configuration includes enough disk storage to cache reports for three months.

Table 5. Large Server Configuration

Memory	Cache Storage	Archive Media Storage	Backup Device
1024 MB	234 GB	One ATL or IBM 3995–C68	Second ATL or optical library

Typical server software configurations

Library/object server

Figure 10 on page 46 shows the standard OnDemand library/object server.

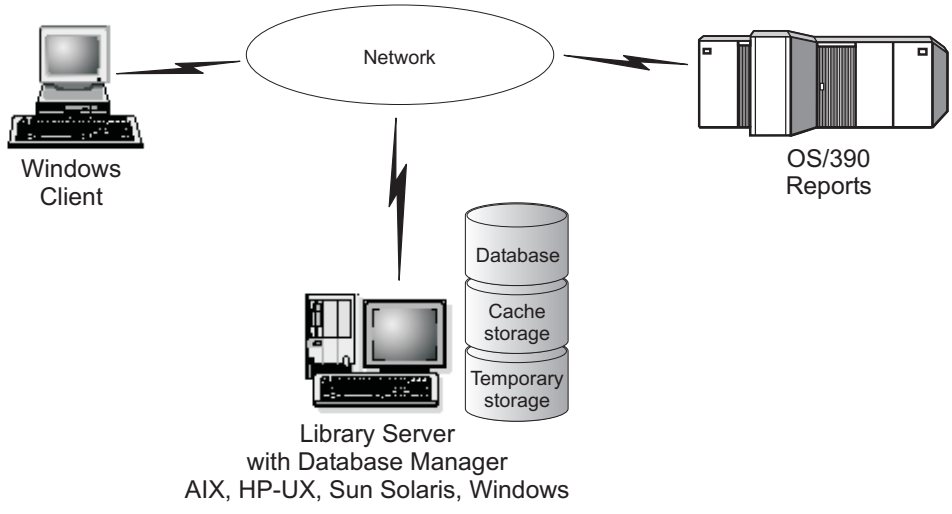

Figure 10. Standard library/object server

The standard library/object server includes the cache storage manager and the programs required to index reports and load data on the system. You can stage reports on temporary storage volumes for the data indexing and loading programs. This environment is ideal for customers who do not require backup copies of reports on archive media or need to run OnDemand on a single processor.

Table 6 lists the software requirements for the standard library/object server.

Table 6. Software for standard library/object server

Function	Req/Opt	Notes
OnDemand base	Required	Base OnDemand functions, such as enhanced ACIF. The cache storage manager maintain reports on disk storage.
Database manager	Required	Database engine and administration
OnDemand client	Required	Windows client program
Download	Optional	Transmit files from OS/390 systems
Infoprint	Optional	Server print and server FAX

Library/object server with TSM

Figure 11 shows the standard library/object server with TSM.

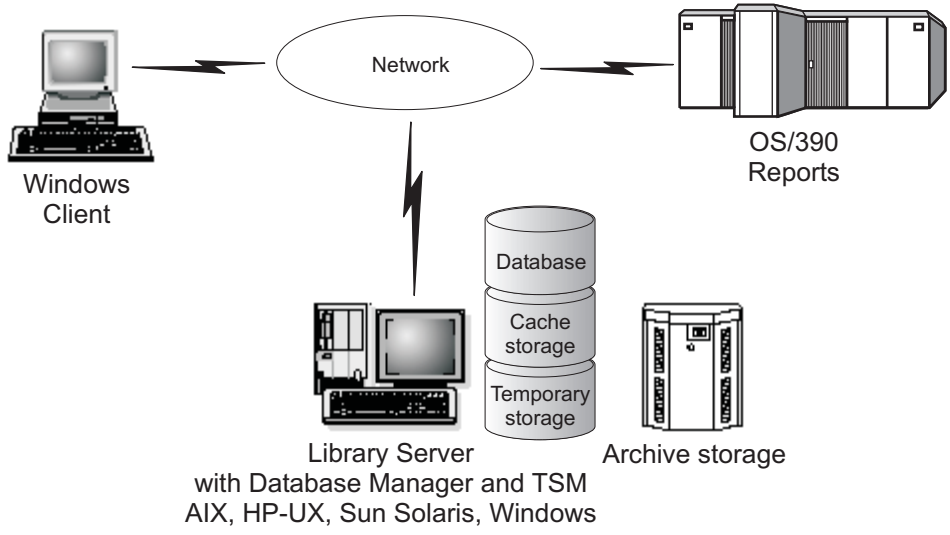

Figure 11. Standard library/object server with TSM

TSM is included with OnDemand. You can use TSM to maintain report data on archive media (optical and tape). The library/object server includes the cache storage manager and the programs required to index reports and load data on the system. You can stage reports on temporary storage volumes for the data indexing and loading programs. This environment is ideal for customers who require backup copies of data on archive media and need to run OnDemand on a single processor.

Table 7 lists the software requirements for the standard library/object server with TSM.

Table 7. Software for standard library/object server with TSM

Function	Req/Opt	Notes
OnDemand base	Required	Base OnDemand functions, such as enhanced ACIF. The cache storage manager maintains reports on disk storage. The TSM API is used to store data on and retrieve data from archive media.
Database manager	Required	Database engine and administration
OnDemand client	Required	Windows client program
TSM	Required	Maintain data on optical and tape storage volumes
Download	Optional	Transmit files from OS/390 systems
Infoprint	Optional	Server print and server FAX

Distributed library/object server

Figure 12 shows the distributed OnDemand library/object server system.

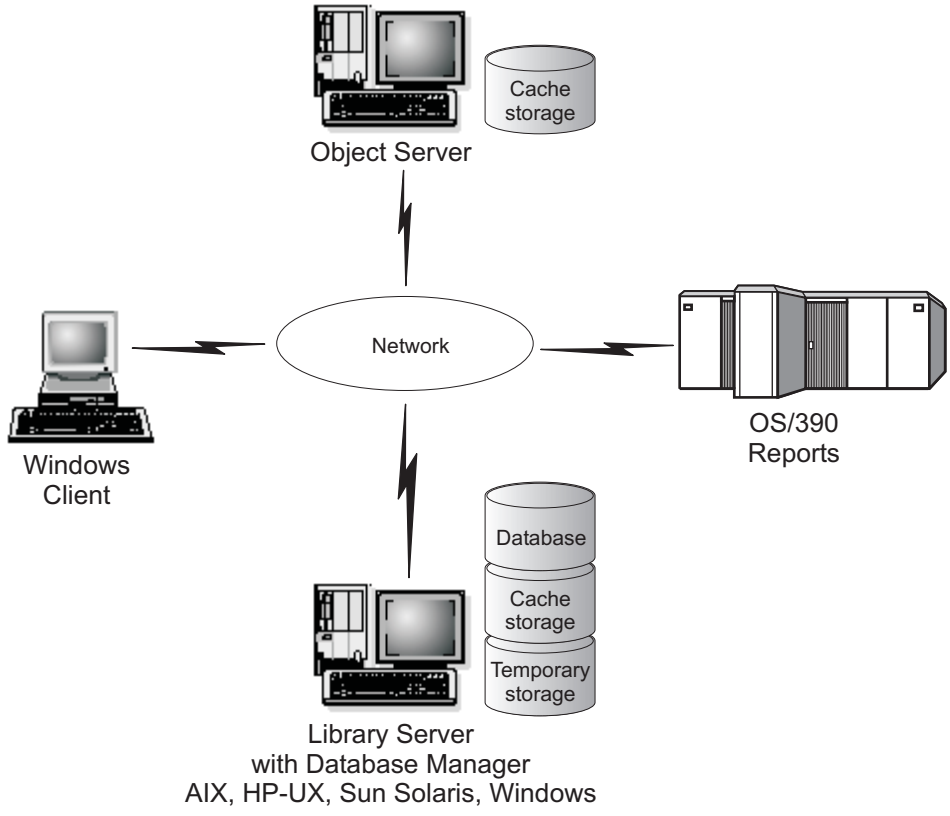

Figure 12. Distributed library/object server system

OnDemand supports storing data on and retrieving data from more than one server. In the distributed environment, users submit queries to the library server and OnDemand retrieves documents from the object server on which the data is stored. You can load reports on any of the object servers that are part of the system. The index data is always stored on the library server. You can stage reports on temporary storage volumes for the data indexing and loading programs. This environment is ideal for customers who need to distribute the loading and accessing of reports over more than one processor. The servers can reside on nodes in one physical machine, such as an SP™ processor, or on separate systems in different physical locations.

Table 8 lists the software requirements for the OnDemand library server.

Table 8. Software for distributed library/object server part 1 of 2

Function	Req/Opt	Notes
OnDemand base	Required	Base OnDemand functions, such as enhanced ACIF. The cache storage manager maintains reports on disk storage volumes attached to the library server.
Database manager	Required	Database engine and administration
OnDemand client	Required	Windows client program
Download	Optional	Transmit files from OS/390 systems
Infoprint	Optional	Server print and server FAX

Table 9 lists the software requirements for the OnDemand object servers.

Table 9. Software for distributed library/object server part 2 of 2

Function	Req/Opt	Notes
OnDemand base	Required	Base OnDemand functions, such as enhanced ACIF. The cache storage manager maintains reports on disk storage volumes attached to the object server.

Distributed library/object server with TSM

Figure 13 shows the distributed library/object server with TSM.

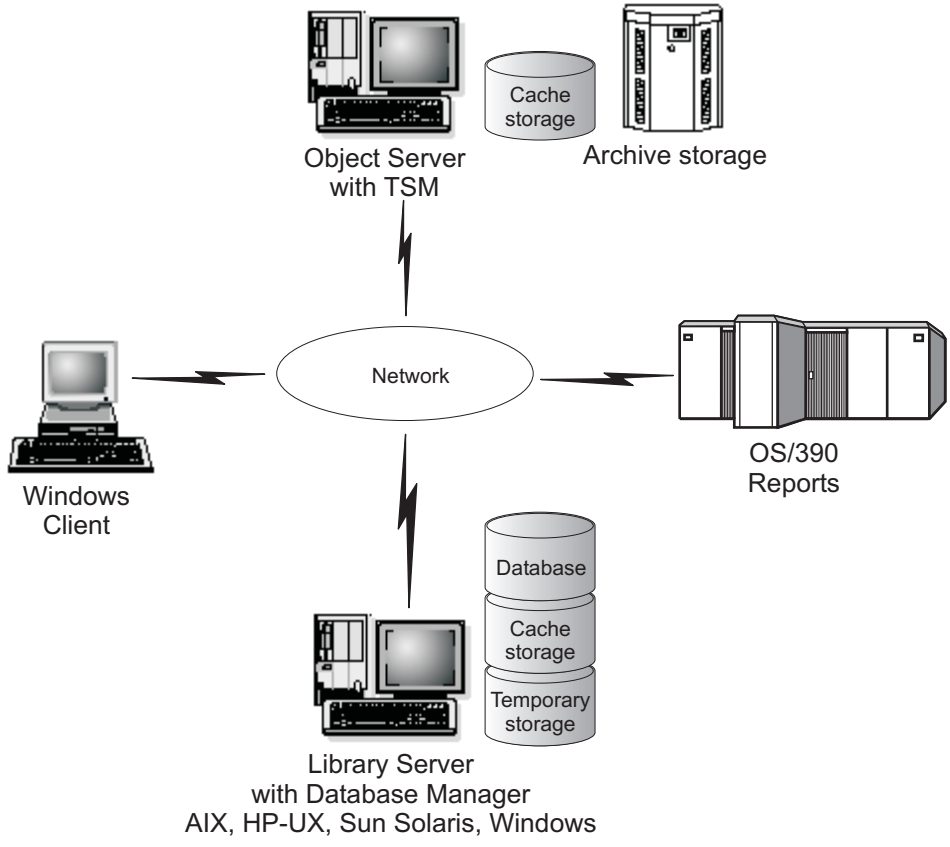

Figure 13. Distributed library/object server with TSM

TSM can be installed on one or more object servers to maintain report data on archive media (optical and tape storage devices). TSM is included with the OnDemand program package. Installing TSM on an object server is ideal for customers who want to move the archive storage part of the system off of the library server and for customers who need to distribute the loading and accessing of reports over more than one processor. The servers can reside on nodes in one physical machine, such as an SP processor, or on separate systems in different physical locations.

Table 10 lists the software requirements for the OnDemand library server.

Table 10. Software for distributed library/object server with TSM part 1 of 2

Function	Req/Opt	Notes
OnDemand base	Required	Base OnDemand functions, such as enhanced ACIF. The cache storage manager maintains reports on disk storage volumes attached to the library server.
Database manager	Required	Database engine; database administration
OnDemand client	Required	Windows client program
Download	Optional	Transmit files from OS/390 system
Infoprint	Optional	Server print and server FAX

Table 11 lists the software requirements for the object server with TSM.

Table 11. Software for distributed library/object server with TSM part 2 of 2

Function	Req/Opt	Notes
OnDemand base	Required	Base OnDemand functions, such as enhanced ACIF. The cache storage manager maintains reports on disk storage volumes attached to the object server. The TSM API is used to store data on and retrieve data from archive media.
TSM	Required	Maintain report data on optical and tape storage volumes attached to the object server

Distributing the TSM workload

You can distribute the TSM workload across your OnDemand servers. For example, you can use TSM on your object servers to maintain copies of reports; you can use TSM on the library server to maintain DB2 archived log files and backup image files; you can use TSM on one object server to maintain a backup copy of the data that is stored on other object servers. The servers can reside on nodes in one physical machine, such as an SP processor, or on separate systems in different physical locations. Figure 14 shows an example.

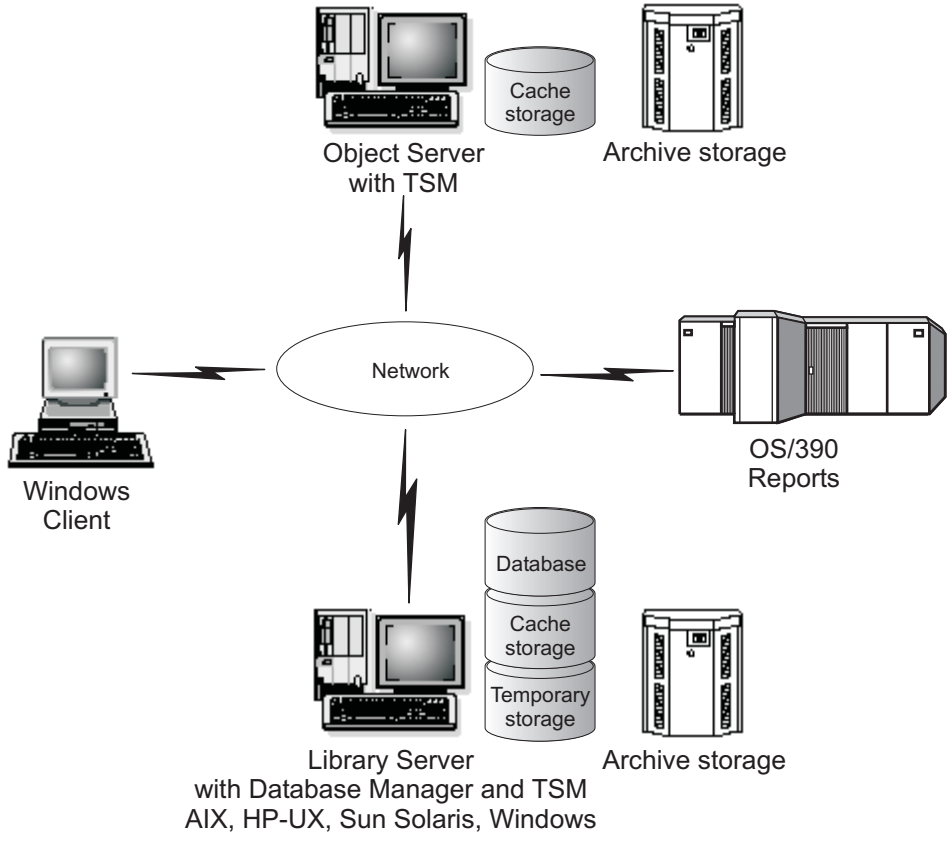

Figure 14. Distributing the TSM workload

Table 12 lists the software requirements for the OnDemand library server with TSM.

Table 12. Software for library server with TSM

Function	Req/Opt	Notes
OnDemand base	Required	Base OnDemand functions, such as enhanced ACIF. The cache storage manager maintains reports on disk storage volumes attached to the library server. The TSM API is used to store data on and retrieve data from archive media attached to the library server.
Database manager	Required	Database engine; database administration
OnDemand client	Required	Windows client program
TSM	Required	Maintain report data on archive media that is attached to the library server
Download	Optional	Transmit files from OS/390 systems
Infoprint	Optional	Server print and server FAX

Table 13 lists the software requirements for the OnDemand object server with TSM.

Table 13. Software for object server with TSM

Function	Req/Opt	Notes
OnDemand base	Required	Base OnDemand functions, such as enhanced ACIF. The cache storage manager maintains reports on disk storage volumes attached to the object server. The TSM API is used to store data on and retrieve data from archive media attached to the object server.
TSM	Required	Maintain report data on archive media that is attached to the object server

Windows client

The OnDemand Windows client runs under Windows 2000, Windows 95, Windows 98, and Windows NT 4.0 and requires the following hardware and software:

- Physical connection to the network, such as a Token Ring or Ethernet network adapter
- A minimum of 32 megabytes of memory

- An Intel Pentium 166 MHz or faster processor
- A super-VGA display and adapter with a minimum resolution of 800x600
- A minimum of 100 megabytes of free hard disk space
- The Windows TCP/IP networking protocol

Administrative software

OnDemand provides the following software to help you administer the system:

- Configurator. The client used to configure and maintain OnDemand Windows servers, services, and scheduled tasks. The configurator typically runs on an OnDemand Windows server.
- Administrative Client. The client used to maintain users, groups, application groups, applications, folders, storage sets, and printers. The administrative client runs on a Windows 32–bit operating system.

The administrative software requires the following:

- A minimum of 32 MB of memory
- An Intel Pentium 166 MHz or faster processor
- A super-VGA display and adapter with a minimum resolution of 800x600
- A minimum of 10 MB of free hard disk space
- The Windows TCP/IP networking protocol

Installing and running client software on a network

A personal computer attached to the network can share a single copy of the OnDemand client software. In general, you install client software on a network file server to save hard disk space on PCs and make it easier to upgrade software later. You typically install OnDemand client software on a PC for better performance and to reduce network traffic.

OnDemand provides a network installation procedure to copy OnDemand control files from a network file server to a user's PC (no program files are copied to the user's PC). As part of a network installation, the setup program builds a directory structure on the user's PC and creates folders and icons on the desktop. The OnDemand program icons point to a copy of the software installed on the network file server.

When the user runs an OnDemand client program from a network file server, the operating system loads the OnDemand programs from the network file server into memory on the user's PC. When the user selects items for viewing, OnDemand allocates temporary work space on the user's PC for documents and resources.

Before you install a client program using the network installation option, you must install the client software on a network file server. See the *User's Guide* for information about the installation options and an overview of the network installation procedure.

Client start up parameters

OnDemand provides command line parameters that you can use to customize the operation of the client program. For example, you can automate the logon process so that the user does not have to specify a server, userid, or password. See the *Windows Client Customization Guide and Reference* for information about the parameters that you can specify for the Windows client.

Adobe software provided with OnDemand

IBM provides Adobe Acrobat for the Windows client so that users can view PDF documents that are retrieved from OnDemand. The Adobe software license agreement provided with OnDemand allows you to use Acrobat to view PDF documents that are retrieved from an OnDemand server. If you wish to view PDF documents that are not stored in OnDemand, or to use Acrobat for other purposes, then you must purchase a full Acrobat license from Adobe.

To view documents that require Adobe Type 1 fonts, Adobe Type Manager (ATM) must be installed on the client PC. On Windows 2000 systems, ATM is part of the base operating system. If you plan to use the OnDemand Windows client with other operating systems, then we recommend that you obtain the ATM software from Adobe and install it before you install the OnDemand client software on the PC. You should plan to use ATM if your users will view AFP documents that use the IBM Core Outline Fonts or the Sonoran Metric Outline Fonts (that are provided with OnDemand). You can find out more about ATM from the Adobe Web site at:

```
http://www.adobe.com
```

Chapter 5. Disk storage

Overview

Before you begin defining reports to OnDemand and loading them on the system, it is important that you estimate the storage required to hold your reports. "Chapter 8. Storage requirements" on page 93 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.

For details about configuring and managing storage devices, see your operating system documentation. If you are not familiar with configuring and managing storage devices, we recommend that you 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 OS/390 systems, to index data on the server, for the database and database log files, for the archive storage manager database, and for cache storage.

We recommend that you 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, a group can support up to 32 storage devices of varying sizes and types and you can have up to 255 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, we recommend that you organize the available

disk storage as described in the examples that follow. We strongly encourage you to 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, we suggest that you 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
- OnDemand database and database log files
- Archive storage manager database and recovery logs
- Cache storage

Table 14 shows one way to configure disk storage into groups on a server that will support a large organization.

Table 14. 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, we suggest that you 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 OnDemand database and database log files, and cache storage

Table 15 on page 59 shows one way to configure disk storage devices on a server that will support a small workgroup.

Table 15. 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 OS/390 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, we suggest that you 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
- OnDemand database and database log files
- Cache storage

We suggest that you plan to use at least ten, 4.5 GB 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 16 shows one way to configure disk storage into groups on a server that will support a large organization.

Table 16. 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

Table 16. Disk storage groups for a large organization (continued)

Physical Disk	Partition	File System(s)	Purpose
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, we suggest that you 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

We recommend that you plan to use at least four 4.5 GB 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 17 shows one way to configure disk storage devices on a server that will support a small workgroup.

Table 17. Disk storage groups for a small workgroup

Physical Disk	Partition	File System(s)	Purpose
Disk1	C:	\WIN32	Windows system files
Disk1	D:	\Program Files	OnDemand and other applications
Disk2	E:	\arsdb, \arsdbpri, \arsdbarc	OnDemand database and log files
Disk3	F:	\arscache1	Cache storage
Disk4	G:	\arsacif1, \arstmp	Data indexing and loading; 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 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 18 provides an overview of RAID implementations.

Table 18. RAID implementations

RAID Level	Description	Protection	Performance
RAID 0	Data striping on multiple disk drives.	Poor; single disk failure.	Best; read and write requests can be met by any disk.
RAID 1	Disk mirroring.	Good; any disk can fail and data is still accessible.	Good; read request can be met by any disk.
RAID 3	Disk striping with parity disk, using interleaved bytes.	Good; if any disk fails, data can be accessed by using information from other disks and parity disk.	Good for large data transfers.
RAID 4	Disk striping with parity disk, using interleaved sectors.	Good; if any disk fails, data can be accessed by using information from other disks and parity disk.	Good for large data transfers.
RAID 5	Disk striping with distributed parity data.	Good; if any disk fails, data can be accessed by using information from other disks and parity information.	Good for small block sizes.
RAID 5 Orthog –onal	Disk striping with distributed parity data, using dual controllers.	Best; if any disk fails, data can be accessed by using information from other disks and parity information, with additional protection from any single disk controller failure.	Good for small block sizes; improved performance because of use of dual controllers to read and write data.

Important: For most OnDemand systems, we recommend that you use the RAID 1 or RAID 5 implementations.

RAID for cache storage

A typical use of RAID storage on an 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 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. We strongly recommend RAID storage for customers that do not maintain copies of reports in archive 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 OnDemand database

You can store the 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.

Configuring RAID storage on UNIX servers

This section provides information about configuring an IBM 7135 RAIDiant Array, a fault tolerant disk solution for the IBM RS/6000 family of workstations. For information about configuring other types of RAID devices, see the product information.

The IBM 7135 RAIDiant Array Model 210 can be configured with six banks of five disks and up to 135 GB of storage capacity. With RAID 5, the drives are configured into six logical units of five drives. Each logical unit appears as a physical volume to the operating system. You can assign one or more of the physical volumes to a volume group. You can create logical volumes and file systems. For example, a fully populated 7135-210 can be configured into two cache storage file systems by assigning the disks to two volume groups, creating one logical volume for each volume group, and creating one file system for each logical volume.

IBM 7133 Serial Disk System

The IBM 7133 Serial Disk System Advanced Models D40 and T40 provide high performance and enhanced availability for I/O intensive applications on UNIX and Windows servers. The 7133 Advanced Models use Serial Storage Architecture (SSA), the highest performing IBM disk subsystem, supporting up to 3,000 I/O operations per second per adapter and a data transfer rate of up to 160 MB per second. SSA includes a loop design so that a single adapter failure will not cause loss of access to data. If there is a disk failure, the hot swappable disks can be removed without loss of communication between the adapter and the other disks on the loop.

Each 7133 Advanced Model can be configured with up to 16 disks and 291 GB of storage. The 7133 Advanced Models can be populated with 18.2 GB, 9.1 GB, and 4.5 GB IBM Ultrastar™ disk drives with auto-docking capabilities. Drive capacities can be intermixed, providing the flexibility to build storage environments ranging in capacity from gigabytes to terabytes.

The 7133 Advanced Models can be attached to IBM RS/6000, HP, and Sun workstations and selected Intel-based servers. The 7133 Advanced Models support multi-host attachment, with one or more servers sharing the disk subsystem, so that if one server fails, the others can continue to process data.

IBM 3527 SSA Storage Subsystem

The IBM 3527 SSA Storage Subsystem for PC Servers uses Serial Storage Architecture technology to deliver outstanding performance, highly scalable capacity, and enhanced availability. The 3527 SSA Storage Subsystem can sustain data rates as high as 35 MB per second in RAID arrays, to support the most demanding storage applications.

The 3527 subsystem mini-tower design features five optional hot-swap drive bays that can hold any combination of 4.5 GB or 9.1 GB SSA disk drives. This results in a total storage capacity of up to 45.5 GB per enclosure. As many as 48 drives can be attached to each of two loops on a single adapter. This provides a maximum of 98 drives with 873 GB of storage on PC Server Models 320, 325, 330, and 520. For larger storage applications, PC Server Models 704 and 720 support two adapters per server, for a maximum of 192 drives totaling 1.75 TB of storage.

The 3527 subsystem uses SSA's bi-directional loop architecture to ensure that data access is never interrupted by a broken cable or failure of a single disk drive. All drives are self-configuring with optional hot-swap capability, for each service or replacement without shutting down the server. The SSA RAID Adapter for PC Servers helps protect against loss of data access using RAID 1 or 5. The system is quickly returned to normal operation once the failed drive has been replaced or data path restored.

Profile of ESS architecture

This section describes some distinguishing features of ESS and how using it can benefit DB2 work. You can find ESS details in *IBM Enterprise Storage Server*, SC24–5465. You can also find more information about ESS from the IBM Storage Systems Web site at:

```
http://www.storage.ibm.com
```

The ESS lets you simplify and centralize the management of your storage resources across your data processing enterprise, while providing the highest levels of data protection and performance:

- ESS allows a storage administrator to allocate, share, and reassign storage across multiple operating systems, including various types of UNIX servers and Windows servers. ESS also provides security mechanisms to prevent one host system from accessing another host system's data. Storage administrators are provided with a Web-based interface to simplify and perform tasks without involving a customer engineer.
- ESS can become an integral part of your storage area network by connecting to application and database servers through a variety of protocols, such as ESCON, SCSI, and fiber channel fabrics. A single ESS can provide storage ranging from as small as 420 GB up to 11 TB. With the use of a fiber channel fabric, a network of database or application servers can share multiple ESS systems to expand on these already large storage capabilities.
- ESS is designed to handle the performance requirements of a wide variety of applications, such as file servers, online transaction processing, business intelligence, e-commerce, and more. This superior performance is the result of a variety of factors, which include large read and write cache, high speed microprocessors and memory buses, latest technology SSA disk drives and attachment, and intelligent cache management functions.
- ESS provides extensive data protection and redundancy throughout the system design: extensive error checking and correction throughout its components; RAID protection for disk failures; mirrored copies of data in cache, including use of battery-backed, non-volatile storage with a battery life of seven days; data scrubbing and error correction capabilities of the disk drives; and a high-availability design to protect against internal processor or cache failures. Routine maintenance functions are also made easier by call-home capabilities.
- ESS provides a set of copy services to allow customers to non-disruptively maintain additional replicas of their data, to help them protect against user or application errors, as well as major disasters. These services help customers maintain these data replicas while minimizing the down-time associated with normal backup windows.

Part 3. Planning information

This section of the book is a planning source for 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 of the book describes activities that we recommend OnDemand administrators perform to plan for the installation of OnDemand and prepare for the operation of OnDemand.

Chapter 6. Reports and other data

This section contains information that can help you plan for the reports that you will be storing into OnDemand. You can use the information to help determine the hardware configuration that you need to support your OnDemand system. We 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 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 OnDemand system, including the storage and network configuration, to support your applications and users:

- Will you operate a single OnDemand server or a network of 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 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 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

Note: An input file cannot exceed 4 GB in size.

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. OnDemand provides Adobe PDF viewing software so that users can view PDF documents stored in OnDemand.
- 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.

In addition to the types of data listed above, 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 OS/390 environment, OnDemand allows application programs that produce 1403 or 3211 data streams 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.[6] For example, users can retrieve and view customer statements that 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 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 OS/390 system, then the indexing program must gather the resources into a resource group so that the resource group can be transferred to the OnDemand server on which you plan to load the data. If data will be indexed on an OnDemand server, then the resources must be resident on the OnDemand server (or be accessible from the OnDemand server) on which you plan to index and load the data.

Indexing data

One of the main operations that you do with OnDemand is to index reports. When you index a report, 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, OnDemand displays a list of search fields, which represent the database fields. To perform a query, the user enters values in the search fields. OnDemand compares the values from the search values with the values in the database fields and retrieves the items that match the query.

6. Viewing of bar code objects is not currently supported. However, you can view bar codes that have been rendered using fonts.

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 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.

Index information can be added to reports at the same time that the application program generates the print data, or the output print data can be processed by an indexing program. The information and examples in the sections that follow assume that you will use ACIF to index your reports. ACIF is a post-processing program that tags key information found in the report data stream and builds the index values using parameters that you specify. You can index reports with ACIF on an OnDemand server or an OS/390 system. See "Indexing data with other programs" on page 72 for information about other methods that you can use to index reports.

Indexing data with ACIF

ACIF is a powerful tool for indexing the print data streams of OS/390 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 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 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. 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.

The reports that you process with ACIF generally fit into one of two categories:

- Document. For reports made up of logical items, such as statements, bills, policies, and invoices.

- Report. For reports that (typically) contain line data, with sorted values on each page, such as a transaction log or general ledger.

Reports that do not contain logical items or sorted line data can usually be indexed with the report indexing method.

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, OnDemand returns a list of the items that match the user's query and transfers the individual items to the OnDemand client program for viewing and printing. 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 15 shows an example of a report file and document indexing.

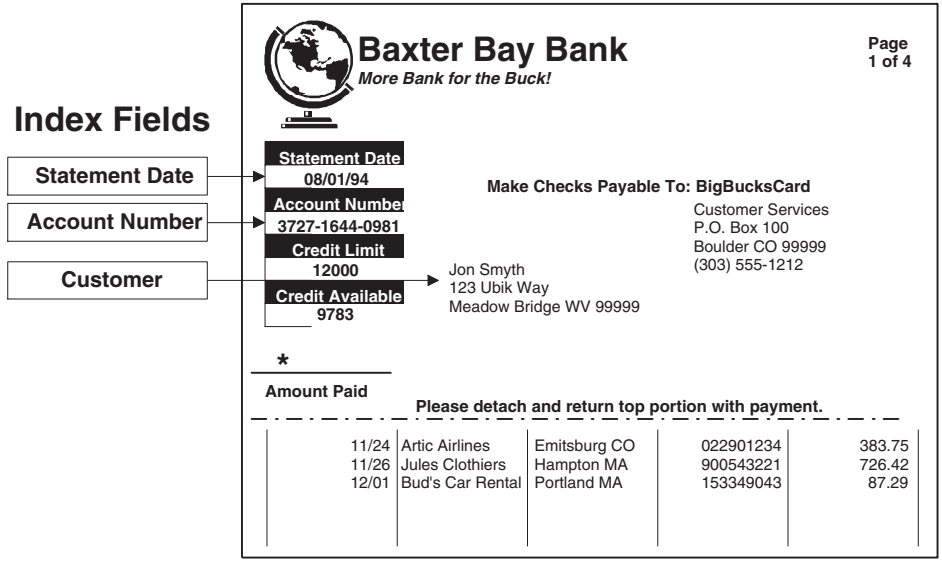

Figure 15. Document indexing method

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. 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, OnDemand returns a list of the items that match the query. When the

user selects an item for viewing, OnDemand performs a text search within the item for the value specified by the user. The OnDemand client program displays the first page that contains the value specified by the user. 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 16 shows an example of a report file and report indexing.

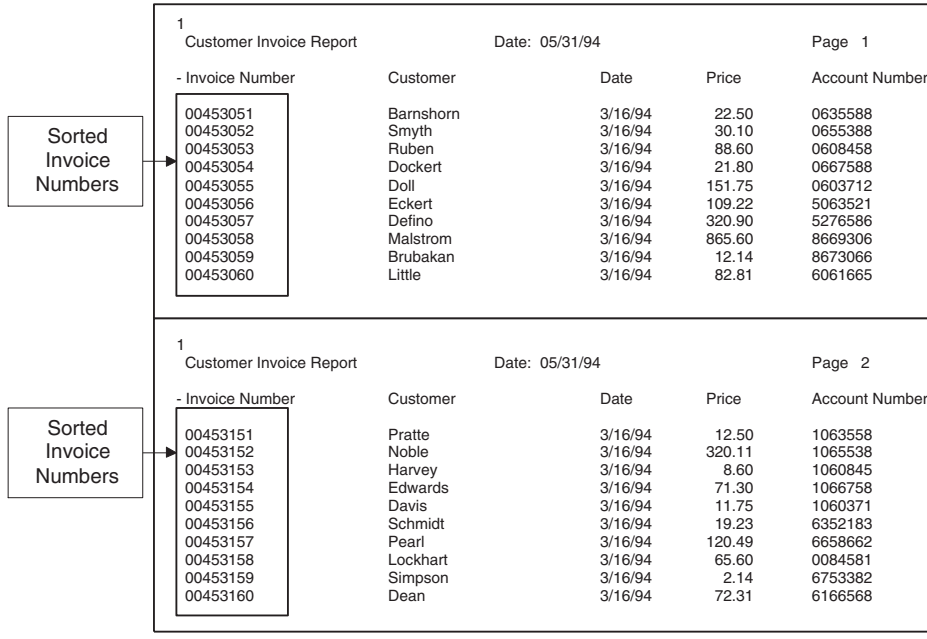

Figure 16. Report indexing method

Indexing data with other programs

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. OnDemand provides the PDF Indexer so that you can extract index data from and generate index information about Adobe Acrobat PDF input files. You can use the Meta2AFP Transform to process Xerox print data streams, converting them to AFP and generating index data. OnDemand also provides a Generic Indexer program so that you can index almost any other type of input data.

The following topics provide additional information::

- Generate the index data in the application program that generates the report

- Use AFP applications that automatically generate index data, for example, the IBM Document Composition Facility (DCF)
- Use the AFP API to generate indexed AFP data
- Insert AFP records into the output data stream that can be reliably extracted by ACIF
- Use the OnDemand PDF Indexer program to index Adobe PDF files
- Create index data for the OnDemand Generic Indexer program

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 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 Application Programming Interface

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 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.

Indexing PDF input files

The OnDemand PDF Indexer is a utility 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 OnDemand. The PDF Indexer processes PDF input files. A PDF file is a distilled version of a PostScript file, adding structure and efficiency. A PDF file can be created by Acrobat Distiller or a special printer driver program called a PDFWriter. The PDF Indexer supports PDF Version 1.2 input and output files. The *Indexing Reference* provides details about the PDF Indexer program and shows examples about how to use it to process PDF input files.

Generating index data using the Generic Indexer

OnDemand provides the Generic Indexer program so that you can create index data for files that cannot be indexed using other methods, such as ACIF. For example, you can create an index file for the Generic Indexer program that describes a set of input files that contain data in the TIFF image format. The index file contains the index fields and values for each file that you want to process. The index file also describes where the Generic Indexer program can find documents within a file. OnDemand will create a row in the database for each index record contained in the index file. Users can search the database using any combination of the index fields that were defined in the index file. The *Indexing Reference* describes the generic index file format.

Indexing reports using date fields

To store data in the system, each report must be indexed with a date field. When querying the database, OnDemand uses the date in a report to determine one report's data from another. 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 OnDemand does not contain a date, you can use the date that the report was loaded into the system.

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.

Note: OnDemand is Year 2000 ready. When OnDemand is used in accordance with its associated documentation, it is capable of correctly processing, providing, or receiving date data within and between the twentieth and twenty-first centuries, provided all other products (for example, software, hardware, and firmware) used with the product properly exchange accurate date data with it.

Running ACIF on OS/390 Systems

Indexing reports on OS/390 systems can provide certain benefits in a production archival process, including:

- The potential to balance processing and resources across the OS/390 system and OnDemand servers. If you index data as a step in the jobs that generate reports for OnDemand, you can transmit report and index data directly to one or more OnDemand servers. The OnDemand servers can load the data into the appropriate application groups. Additional storage and processing cycles to do indexing on the 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 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 OnDemand server, you must either transmit the individual resources to the 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 OnDemand servers

Indexing reports on an 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 *Installation and Configuration Guide* for your server describes how to configure the OnDemand data indexing and loading programs to run on a regular schedule, automatically processing input data.
- You can off-load part of the OS/390 processing to an OnDemand server, by running the indexing program on the OnDemand server. If your OnDemand system consists of several 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 7. OnDemand objects

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

Overview

When you install and configure the 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, OnDemand adds an application group table structure and other control information to the database.

OnDemand uses a set of objects to describe the database tables, fields, and data that make up the system. When you define an object to OnDemand, such as an application group, 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, 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 OnDemand system open a folder to query and access reports that are stored on the system. A folder is the primary OnDemand object that users deal with. A folder provides users the means to search for and retrieve data stored in 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.

An application group represents the index and report data that you load into OnDemand. The 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 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 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 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 OnDemand, such as application groups, applications, and folders.

OnDemand uses properties to describe the appearance, behavior, and internal structure of the objects that make up an OnDemand system. For example, Display Format is a property of a folder field that determines how 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. 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, 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 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.

OnDemand maintains information about the name of the folder and its structure in the 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 OnDemand through properties and values grouped in categories. A category is a set of related properties. 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.

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 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, 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, 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.

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. 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.

OnDemand requires a *segment field* for each application group that you define. 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. 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. 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. 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. 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. 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 OnDemand should store the index data for the reports that you load into the application group. OnDemand provides two methods that you can use to determine how index records are loaded into the database and how users can query the application group:

- Multiple Loads per Database Table

With this method, each time that you load a report into the application group, 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. 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. OnDemand automatically segments the application group data when it grows beyond a certain size. 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. We recommend that you 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.

- Single Load per Database Table

  With this method, each time that you load a report into the application group, OnDemand stores the index records into a new database table. You can define a *report field* for the application group so that users can easily search for and retrieve a specific version of a report. We recommend that you use this method to organize your database when the users that search for data stored in the application group need to query a specific version of a report.

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

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.) 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.

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, 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. 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 OnDemand application describes the physical characteristics of a report, processing instructions for the indexing and data loading programs, and information about how OnDemand displays and prints pages of a report. You can specify default settings for viewing and printing pages of a report at the 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, 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. OnDemand uses the information that you specify to properly interpret the data for viewing.

The OnDemand application is where you specify information to the indexing and data loading programs, such as the technique that OnDemand uses to compress the report file, the parameter used to index the data, and information that OnDemand uses to process index data before loading index records into the database. 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 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 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 OnDemand with a userid. OnDemand authenticates userids and determines the usage and administrative authority available to the user based on the log on userid. An OnDemand userid does not necessarily have to identify an individual user. However, for accounting purposes, you probably want to assign an OnDemand userid to each person that uses the system.

OnDemand automatically creates the ADMIN userid when you initialize the system. The ADMIN userid 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.

OnDemand groups are a means to organize users by function, authorization, or any other purpose you might require. When you define an 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 19 on page 84, a user assigned to both groups can access the Student Bills and Student Transcripts folders.

Table 19. Group permissions

Group	Folders
Accounting	Student Bills
Admissions	Student Transcripts

However, there are exceptions to this rule. See information about permissions in the *Administrator's Guide* for details.

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

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

Folder permissions

You can set folder permissions at the folder, group, and user levels. Setting permissions at the folder level provides all 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 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 OnDemand client programs and can modify the folder field information with the administrator interface. 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. 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 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 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 OnDemand displays report file pages. Users can define their own logical views with 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 OnDemand, you assign names to the various objects.

Note: If you install 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.

When naming a user, the name that you specify:

- Can contain from one to twenty 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, OnDemand converts lowercase letters in a user name to uppercase (for example, laguarde is stored as LAGUARDE)

Note: If you define a logon user exit, then you can determine the characteristics of userids on your system.

When creating a password, the value that you specify:

- Can contain from one to twenty characters (bytes)

  Note: When creating a password, the value that you specify can be a maximum of 20 characters. However, the password authentication that is built into OnDemand 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 using the logon user exit. Contact the IBM support center for more information about the logon user exit.
- By default, OnDemand converts lowercase letters in a password to uppercase (for example, laguarde is stored as LAGUARDE)

Note: If you define a logon user exit, then you can determine the characteristics of passwords on your system.

When naming a group, the name that you specify:

- Can contain from one to twenty 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, 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, 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, OnDemand doesn't create a unique name (for example, rDate is the same as rdate)
- Must be unique to the application group
- Cannot be any of the 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, 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 sixty characters (bytes)
- Can be mixed case; however, 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, 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:

- Must be a valid printer queue name on the library server

Data types and field types

When you define an application group, 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, 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 OnDemand what kind of data can be stored in the field.

When you define a folder to 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 20 lists the types of application group and folder fields supported by OnDemand.

Table 20. 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.
Decimal	Contains numbers between -10^{307} and 10^{308} with up to 15 significant digits. You can store currency values in a decimal field, and use the precision attribute to format the decimal places.
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. 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.

Table 20. Application group and folder field types (continued)

Field Type	Description
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 21 lists additional types of fields that are supported in folders.

Table 21. 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 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.

Table 21. Additional folder field types (continued)

Field Type	Description
Text Search	Used to find documents that contain a non-indexed word or phrase. A match occurs and an item is added to the document list when one or more lines in a document contain the word or phrase exactly as specified. The search string can contain letters, numbers, special symbols, such as the % and #, and any other printable character. A folder can have one text search field. **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.

Chapter 8. Storage requirements

Overview

Estimating storage requirements for an OnDemand system begins with understanding and documenting end-user requirements for storing and accessing data. “Chapter 6. Reports and other data” on page 67 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, 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. You would typically configure cache storage with the fastest disk devices, where they can provide the most benefit to your users. 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 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 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 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 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 OnDemand server.
- Temporary storage space for loading a report on the 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.

  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. 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 OnDemand system tables (control information and objects that you define to 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, OnDemand can divide the report into groups of a fixed number of pages and create one index row for each group of pages.[7] For reports that contain logical items, such as statements, and policies, 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

7. For sorted transaction data, the examples and calculations that follow assume that 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 *Indexing Reference* provides move information.

TSM to maintain DB2 archived log files, 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 media to the database.

The following components of the system require archive storage (optical and tape storage):

- Reports that you plan to store on archive media.
- Backup copies of reports stored on archive media. (For critical applications, some customers require that the system maintain two or more copies of a report on archive media.)
- Database archived log files, if you use TSM to maintain DB2 archived log files.
- Database backup image files, if you use TSM to maintain DB2 backup image files.

When you calculate archive storage requirements, you should also determine the number of storage volumes and libraries that you need to support your system. Most optical libraries can hold a large amount of data, and can hold even more, depending on the compression ratio achieved for your reports. For example, an IBM 3995-C68 optical library can hold up to 1.341 TB of (unformatted) of data. If OnDemand can achieve a 6:1 compression ratio on the reports to be stored in the library, then the library could hold over seven terabytes of report data.

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 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 OnDemand and most archive storage managers.

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

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. In general, OnDemand should only use the archive storage manager to maintain reports on optical or tape storage devices. When you load a report on the system, OnDemand can automatically store one copy of the report in cache storage and another copy of the report in archive storage. OnDemand also supports the option of storing reports in cache storage and then later migrating them to archive storage. However, we recommend that you always plan to copy reports to cache storage and archive media 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 media. Copying reports to cache storage and archive storage at the same time also eliminates the need for you to migrate reports to archive media.

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. 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, OnDemand reclaims the space taken by expired documents.

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 OnDemand and the archive storage manager. For example, the Life of Data and Indexes, which is used by OnDemand, and the Retention Period, which is used by TSM, should specify the same length of time.

Data compression

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 OnDemand can achieve has a significant impact on the amount of space required to store reports.

The compression ratios that 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, we recommend 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 *Administrator's Guide* for more information about the ARSADMIN program.

Calculating disk storage requirements

System software

Most OnDemand servers require approximately 2 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 OnDemand server software.

Download

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 OS/390 system to an IBM Content Manager OnDemand for Multiplatforms object server. Many customers download report data during the day when their application programs generate the reports, but do not load the data into 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.) We recommend 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 17. 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 OnDemand server is:

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

Temporary space for indexing

OnDemand requires temporary storage space on disk to index reports. The temporary space required by 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. For IBM Content Manager OnDemand for Multiplatforms servers, you also need to know where you will index the reports: on an OS/390 system or on the OnDemand server.

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

- For IBM Content Manager OnDemand for Multiplatforms servers, if you plan to index reports on an OS/390 system, then no temporary space is required.
- If you plan to index reports on the OnDemand server:

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

Figure 18. 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 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 will vary greatly based on requirements such as the number of reports that you store on the system, the compression ratio that OnDemand can achieve, and 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 media. 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 media.

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 OnDemand will 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 19. 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
```

OnDemand database storage

When you load a report into the system, 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, OnDemand can create one database row for every item found in the report. For reports that contain sorted transaction data, 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 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 OnDemand.)

There are four major factors that determine the amount of disk space required for the 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 OnDemand maintains the index data in the database

Table 22 lists the types of index fields supported by OnDemand and the number of bytes required to store a value in each type of index field.

Table 22. Index field types and sizes

Field Type	Field Size (DB2)	Field Size (Oracle)	Field Size (SQL Server)
Small Integer	2 bytes	22 bytes	2 bytes
Integer	4 bytes	22 bytes	4 bytes
Big Integer	8 bytes	Not Supported	8 bytes
Decimal	8 bytes	22 bytes	8 bytes
String (Fixed)	1 – 254; 1 byte per character declared, even if partially used	1 – 254; 1 byte per character declared, even if partially used	1 – 254; 1 byte per character declared, even if partially used
String (Variable)	1 – 254; 1 byte per character plus 4 bytes overhead; unused characters do not consume storage	1 – 254; 1 byte per character plus 4 bytes overhead; unused characters do not consume storage	1 – 254; 1 byte per character plus 4 bytes overhead; unused characters do not consume storage
Date	4 bytes	22 bytes	4 bytes
Time	4 bytes	22 bytes	4 bytes
Date/Time	4 bytes	22 bytes	4 bytes
Date/Time (TZ)	4 bytes	22 bytes	4 bytes

Calculating the size of the database

You can use the following calculations to determine the space required in the OnDemand database to hold the index data for a report. In general, we recommend that you add 10 to 20 percent to the calculated space requirements. The calculations can be used for reports that contain logical items and reports that contain a sorted transaction value.

Note: The formulas that follow were derived from information provided with the database manager products that work with OnDemand. See the product information for details.

```
TableSize = ( Sum of column lengths )

IndexSize = ( 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 20. Calculating database storage space for DB2

- Table 22 on page 101 lists the sizes of the various data types supported by OnDemand.
- `Index n length` is the size of a database field for which you want 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.
- OnDemand 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 21 shows the calculation that you can use to detemine database space requirements when the database manager is Oracle.

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

IndexSize = ( 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 21. Calculating database storage space for Oracle

Figure 22 on page 103 shows the calculations that you can use to detemine database space requirements when the database manager is SQL Server.

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

IndexSize = ( 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 22. 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 21 on page 102; if SQL Server is the database manager, use the calculations from Figure 22.

1. The following example illustrates how to calculate the database storage space required for a report that contains logical items, such as statements. In the example, we plan to index one million items per month and keep the index data in the database for 24 months. Table 23 lists information about the database fields.

Table 23. 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 = ( 4 + 8 ) + ( 12 + 8 ) ) = 32

DatabaseSize = ( ( 48 + 40 ) * 1.5 ) + ( 32 * 2 ) ) = 196
                * 1,000,000 = 196000000
                  * 24  =  4704000000
```

OnDemand requires 4.7 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. 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 OnDemand maintains the index data in the database.

In the example, we plan to index one million pages per month, in groups of 100 pages, and keep the index in the database for 24 months. Table 24 lists information about the database fields.

Table 24. 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
Page Number	Integer	4 bytes	Filter

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

IndexSize = ( ( 4 + 8 ) + ( 10 + 8 ) + ( 10 + 8 ) ) = 48

DatabaseSize = ( ( 28 + 40 ) * 1.5 ) + ( 48 * 2 ) ) = 198
                 * ( 1,000,000/100 ) = 1980000
                    * 24  =  47520000
```

OnDemand requires 47.5 MB of magnetic disk to store 24 months of report index data in the database.

Database log file storage

The 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 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, 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, 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 OnDemand, you set parameters that determine the number and size of the log files and where 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 23. Calculating primary log storage space

- You can find the calculations for `TableSize` and `IndexSize` in "OnDemand database storage" on page 100.
- 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 32, and the Number of indexed items added to the database is 50,000.
OnDemand requires approximately 39 MB of primary log space for the

```
PrimaryLogSpace =  ( ( ( 48 + 40 ) * 1.5 ) + ( 32 * 2 ) ) = 196
                       * 50000 = 9800000
                       * 4   =   39200000
```

sample report.

Primary log storage space for the system

Important: It is critical that you allocate enough disk space for the primary logs. 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 TSM to maintain the DB2 archived log files, we recommend 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" on page 105.
- Estimate the maximum number of reports that 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 105. 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 24. Calculating primary log storage space

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

By default, 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 TSM 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 112).

We strongly encourage you to 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, 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, we recommend 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 25. 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, 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 26. 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 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 media, 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

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. We strongly encourage you to define a dedicated file system where OnDemand can store the temporary print files. We recommend that at least 500 MB of free space be available in this file system at all times. If your storage configuration permits, we recommend 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 media, then you do not need to allocate temporary storage space for importing the migrated index data.

OnDemand requires temporary work space to import migrated index data from archive media into the database. You must allocate enough disk space to support the maximum number of concurrent import requests that the server must manage. We strongly encourage you to define a dedicated file system where 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. We recommend at least 500 MB of free space be 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 we made for the sample reports in "OnDemand database storage" on page 100 and making some assumptions about how the data is stored in OnDemand, we need approximately 500 MB of space to import one application group table. If we need to import two application group tables to satisfy a query, then the import program requires at least 1 GB of temporary disk space.

Calculating 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 media 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 27. 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 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 OnDemand system can maintain a backup (second) copy of reports that you store on archive media. 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 OnDemand uses to maintain the backup copy depends on the archive storage manager that you use. For example, with TSM, you can configure a copy storage pool. With this method, TSM maintains a backup copy of data stored in a primary storage pool independently and transparently to OnDemand. TSM automatically retrieves the backup copy if the primary copy becomes damaged, lost, or unusable.

If you need OnDemand to maintain a backup copy of your reports, double the archive storage space that you calculated in "Report storage space" on page 109.

Storage for database backup images

Note: If you do not plan to use TSM to maintain DB2 backup image files, you do not need to allocate space for the backup image files on archive media.

The storage pool where TSM 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 TSM 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 TSM must maintain a backup image for each table that you migrate.
- TSM can maintain multiple copies of each backup image. For example, for added protection, you may want TSM to maintain two copies of each backup image.

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 TSM 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, we recommend that you configure TSM 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 TSM maintains the database backup images. If TSM 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 we listed above, you may need to maintain hundreds of backup image files in TSM-managed storage. We strongly encourage you to contact TSM 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 TSM can achieve on the backup image files.

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

For example, if the maximum size of the database is 5.4 GB and you need TSM 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 =  (5.4 GB  *  .33)
                         *  2
                           *  2   =  7.2 GB
```

OnDemand requires approximately 7.2 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 we want TSM 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
```

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 TSM to maintain DB2 archived log files, you do not need to allocate space for the archived log files on archive media. Allocate disk storage space for the archived log files instead (see "Archived log storage" on page 107).

The storage pool where TSM 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 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, we recommend 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 28. 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
```

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, then we recommend 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 we listed above, you may need to maintain hundreds of archived log files in TSM-managed storage. We strongly encourage you to contact TSM 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 TSM maintains the archived log files. If TSM 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

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 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 29. Calculating migrated index storage space

For example, if the database size is 202 MB 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  =  ( 202 MB  * .33 )
                     *  ( 84  -  24 )  =  4 GB
```

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 109), approximately 244 GB of archive storage space is required. Assuming that the formatted capacity of a 5.25 inch optical storage volume is about 5.2 GB , then approximately 48 storage volumes would be required to hold the data. An IBM 3995 Model C68 optical library can hold up to 258 storage volumes at a time. Therefore, a single IBM 3995-C68 optical library can hold the entire seven years of data (without the need to replace full storage volumes with empty ones).

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 TSM to maintain DB2 backup image files and archived log files, we recommend that you dedicate a library for that purpose.
- If you need to maintain a backup copy of the reports that you store on archive media, and the archive storage manager that you are using supports it, we recommend 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 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 25. Report that contains logical items, part 1 of 4

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

Table 26. 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 25	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 OS/390 or OnDemand server	OnDemand server

Table 27. Report that contains logical items, part 3 of 4

Disk Storage Requirements in Bytes	
Storage Component	Storage Requirement
Base system software	2,000,000,000
Data download	480,000,000
Indexing	600,000,000
Cache	8,712,000,000
OnDemand database	4,704,000,000
Database logs	516,000,000
Archive storage manager database and logs	47,040,000
Server print	500,000,000
Importing migrated index data	500,000,000
Total Disk Storage Required (Bytes)	18,059,040,000

Table 28. Report that contains logical items, part 4 of 4

Archive Storage Requirements in Bytes	
Report data	244,000,000,000
Migrated index data	4,000,000,000
DB2 log files	0
DB2 backup image files	0
Total Archive Storage Required	248,000,000,000 (bytes)

Note: 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 29. Report that contains transaction data, part 1 of 4

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

Table 30. Report that contains logical items, 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 (bytes)	200,000,000
Largest Cycle - number of pages	50,000
Largest single report file size (bytes)	200,000,000
Group of indexed pages	100
Number DB columns from Table 29	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 OS/390 or OnDemand server	OnDemand server

Table 31. Report that contains logical items, part 3 of 4

Disk Storage Requirements in Bytes	
Storage Component	Storage Requirement
Base System Software	2,000,000,000
Data download	240,000,000
Indexing	300,000,000
Cache	0
OnDemand database	47,520,000
Database logs	172,000,000
Archive storage manager database and logs	8,400,000
Server print	500,000,000
Importing migrated indexes	0
Total Disk Storage Required (Bytes)	3,267,920,000

Table 32. Report that contains logical items, part 4 of 4

Archive Storage Requirements in Bytes	
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 OnDemand. Complete the worksheets to calculate the storage requirements for the report. See "Calculating disk storage requirements" on page 98 and "Calculating archive storage requirements" on page 109 for the formulas that you can use to calculate the storage requirements.

Report that contains logical items

Table 33. Report that contains logical items, part 1 of 4

Database Columns			
Column Number	Name	Index or Filter	Bytes
1			
2			
3			
4			
5			
6			

Table 34. 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 33	
Life of data in days	
Number of days to cache data	
Number of days to keep indexes in database	
Compression ratio	
Index on OS/390 or OnDemand server	

Table 35. Report that contains logical items, part 3 of 4

Disk Storage Requirements in Bytes	
Storage Component	Storage Requirement
Base system software	2,000,000,000
Note: The base system software requirement of 2 GB is per server	
Data download	
Indexing	
Cache	
OnDemand database	
Database logs	
Archive storage manager database and logs	
Server print	
Importing migrated indexes	
Total Disk Storage Required (Bytes)	

Table 36. Report that contains logical items, part 4 of 4

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

Report that contains transaction data

Table 37. Report that contains logical items, part 1 of 4

Database Columns			
Column Number	Name	Index or Filter	Bytes
1			
2			
3			
4			
5			
6			

Table 38. Report that contains logical items, 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	
Number DB columns from Table 37	
Life of data in days	
Number of days to cache data	
Number of days to keep indexes in database	
Compression ratio	
Index on OS/390 or OnDemand server	

Table 39. Report that contains logical items, part 3 of 4

Disk Storage Requirements in Bytes	
Storage Component	Storage Requirement
Base system software	2,000,000,000
Note: The base system software requirement of 2 GB is per server	
Data download	
Indexing	
Cache	
OnDemand database	
Database logs	
Archive storage manager database and logs	
Server print	
Importing migrated indexes	
Total Disk Storage Required (Bytes)	

Table 40. Report that contains logical items, part 4 of 4

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

Chapter 9. Backup and recovery

Overview

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

- OnDemand software
- OnDemand server information, created or modified during installation, configuration, and ongoing operation of OnDemand
- The OnDemand database
- The TSM 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 TSM to maintain the DB2 backup image files and archived log files. This capability means that you do not have to manage these files on disk.

Server software

If a media failure or some other unforeseen event occurs, you may be required to restore the 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. We recommend that you register 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 OnDemand, you specified information that customized 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. We recommend that you backup the control files or Registry immediately after you have verified the installation of OnDemand. In addition, if you periodically make changes to the OnDemand server information, we recommend 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 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 NT, you can use the NTBACKUP, 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 TSM, to backup files, including the Registry, maintain the backups, and assist with recovery. TSM 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.

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 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.

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, we recommend that you backup the OnDemand database after each report file load. If your schedule does not permit you to run the backup command after each load, we recommend 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, we recommend 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 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 *Installation and Configuration Guide* for your server provides details about how to configure the system to support table spaces.

Database backup

OnDemand provides the ARSDB program so that you can create backup images of the 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, we recommend that you stop the 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, 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 TSM to manage DB2 log files, then the policy domain determines when archived log files are eligible to be removed.

If your production schedule allows, we strongly encourage you to 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. We recommend that you write offline

backup images to removable media or TSM-managed storage. 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 *Installation and Configuration Guide for UNIX Servers* shows how to use the cron facility to create online backups of the database with the ARSDB program automatically on a regular schedule. The *Installation and Configuration Guide for Windows Servers* shows how to use the Windows server configuration program to schedule online backups of the database.

The *Administrator's Guide* provides details about the ARSDB program and its parameters and options.

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

Using TSM to maintain backup image files

If you use DB2, you can configure TSM 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 command 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 TSM-managed storage.

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

See the *Installation and Configuration Guide* for details about how to configure the system to use TSM to maintain the backup image files. See the *Administrator's Guide* for details about how to use the ARSDB program to backup table spaces to TSM-managed storage.

Database logging

The database manager uses transaction logging to record information about changes to the 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 OnDemand database can be recovered to any point in time.

Using TSM to maintain DB2 archived log files

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

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

See the *Installation and Configuration Guide* for details about how to configure the system to use TSM 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 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 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 OnDemand database before the catastrophic failure.

The 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 OnDemand database.

TSM database

TSM 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, TSM updates its database and stores a copy of the report on a storage volume. When you define archive storage devices and register client nodes, TSM updates its database. When TSM 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 TSM in storing objects on and retrieving objects from the optical and tape storage volumes that it manages.

We strongly encourage you to *mirror* the TSM database. When you mirror the database, TSM replicates the database onto different physical storage. TSM automatically keeps track of and refreshes both copies of the database. When you configure physical storage so that TSM 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, TSM 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 TSM 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 TSM with OnDemand. In addition, we recommend 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 TSM with OnDemand. In addition, there may be restrictions on the number of incremental backups that can be taken between each full backup. See the TSM documentation for details.

In general, we recommend that you plan to backup the database after you load reports into the system and after TSM 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.

TSM 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 TSM information for details about scheduling administrative commands and automating database backups.

If you are using TSM on a UNIX server, OnDemand provides the ARS_ADSM program to create a full backup of the database. If you invoke the ARSLOAD program from a shell script, then you can add a step to the script that runs the ARS_ADSM program after loading reports into OnDemand. The *Installation and Configuration Guide for UNIX Servers* shows how to run the ARS_ADSM program to create a full backup image of the database automatically on a regular schedule. The *Administrator's Guide* provides details about the ARS_ADSM program, parameters, and options.

Recovery log

The recovery log is critical to the operation of TSM. If the recovery log is unusable, then TSM 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 TSM 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 TSM 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.

TSM 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 TSM 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, TSM updates the database and makes an entry in a device configuration history file. To restore the database, TSM 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 TSM 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 TSM. After you define the volume to TSM, the server synchronizes the volume with the database.

Recovering using backup copies of the database

TSM 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, TSM 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 TSM information for details about recovering data.

Reports

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 TSM.

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 we recommend 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 RAID storage subsystems. RAID storage can provide excellent availability, allowing users to access reports even if a disk or controller fails. “Chapter 5. Disk storage” on page 57 provides information about using RAID storage with OnDemand. However, RAID storage is not fail-safe. There may be situations when, because of multiple disk or controller failures, users cannot access reports. We encourage you to always maintain a backup copy of reports in archive storage.

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 TSM, you would define a *copy storage pool*. With a copy storage pool, TSM manages a backup copy of files that are stored in a primary storage pool independently and transparently to 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 TSM database. A copy storage pool must reside on the object server where the primary storage pool resides. TSM 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.

Part 4. Appendixes

Appendix. 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 the IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe on 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 Directory 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:

IBM World Trade Asia Corporation
Licensing
2-31 Roppongi 3-chrome, Minato-ku
Tokyo 106, 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
Department MYGA Building 001L
6300 Diagonal Highway
Boulder, CO 80301–9191
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 information 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.

Trademarks and service marks

AFP, AIX, DB2, DB2 Universal Database, Enterprise Storage Server, ESCON, IBM, Infoprint, OS/390, POWERparallel, Print Services Facility, RS/6000, SP, S/390, Ultrastar, and WebSphere are trademarks of International Business Machines Corporation in the United States, other countries, or both.

Lotus is a registered trademark of Lotus Development Corporation.

Tivoli is a registered trademark of Tivoli Systems Inc. in the United States, other countries, or both.

Adobe, the Adobe logo, Acrobat and the Acrobat logo are trademarks of Adobe Systems Incorporated, which may be registered in certain jurisdictions.

Intel and Pentium are trademarks of Intel Corporation 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.

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

Glossary

This glossary includes definitions from the following sources:

- Definitions reprinted from the *American National Dictionary for Information Processing Systems* , copyright 1982 by the Computer Business Equipment Manufacturers Association (CBEMA), are identified by the symbol (A) following the definition. Copies can be purchased from the American Standards Institute, 1430 Broadway, New York, New York 10018.
- Definitions reprinted from a published section of the International Organization for Standardization's (ISO) *Vocabulary—Information Processing* or from a published section of the ISO *Vocabulary—Office Machines* are identified by the symbol (I) following the definition. Because many ISO definitions are also reproduced in the *American National Dictionary for Information Processing Systems* ISO definitions may also be identified by the symbol (A).
- Definitions reprinted from working documents, draft proposals, or draft international standards of ISO Technical Committee 97, Subcommittee 1 (Vocabulary) are identified by the symbol (T) following the definition, indicating that final agreement has not yet been reached among its participating members.
- Definitions are also reprinted from the *CCITT Eighth Plenary Assembly Red Book, Terms and Definitions* and working documents published by the International Telegraph and Telephone Consultative Committee of the International Telecommunication Union, Geneva, 1985.
- Definitions that are specific to IBM products are so labeled, for example, "In OnDemand," or "In MVS™."

The following cross references are used in this glossary:

Contrast with. This refers to a term that has an opposed or substantively different meaning.

Synonym for. This indicates that the term has the same meaning as a preferred term, which is defined in its proper place in the dictionary.

Synonymous with. This is a backward reference from a defined term to all other terms that have the same meaning.

See. This refers the reader to multiple-word terms that have the same last word.

See also. This refers the reader to terms that have a related, but not synonymous, meaning.

A

access. To obtain data from or to put data in storage.

ACIF. Advanced Function Presentation Conversion and Indexing Facility

Acrobat. The Adobe viewer for PDF files. Acrobat is similar to the IBM AFP Workbench, that is, a stand-alone viewer. Acrobat also supports a robust set of APIs. It is through these APIs that Acrobat is integrated with the OnDemand client program.

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 rollbacks and recovery.

active policy set. In TSM, the policy set within the policy domain that contains the most recently

activated policy currently in use by all client nodes that have been assigned to that 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. See also Pel.

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

ADSM. ADSTAR® Distributed Storage Manager

ADSTAR Distributed Storage Manager. A program that provides storage management for archived files. See Tivoli Storage Manager.

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. AFP also includes creating, formatting, archiving, viewing, retrieving, and distributing information.

Advanced Function Presentation Application Programming Interface (AFP API). An AFP program shipped with PSF/MVS 2.1.1 and PSF/VM 2.1.1 that creates the AFP data stream from the COBOL and PL/1 high-level programming languages.

Advanced Function Presentation Conversion and Indexing Facility. A program shipped with 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 (AFP data stream). 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. (1) Advanced Interactive Executive (2) IBM's version of the UNIX operating system.

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

all-points-addressable (APA). The capability to address, reference, and position data elements at any addressable position in a presentation space or on a physical medium. An example of all points addressability is the positioning of text, graphics, and images at any addressable point on the physical medium. See also Picture Element.

all-points-addressable mode. Synonym for Page Mode.

alphabetic character. A letter or other symbol, excluding digits, 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, numbers, 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 for the purpose of establishing voluntary industry standards.

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.

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 containing the carriage control. ANSI carriage control characters are encoded in ASCII or EBCDIC.

APA. All points addressable

API. Application Program 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 OnDemand.

application group. A collection of one or more 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 could be placed in the same application group.

Application Program Interface (API). A formally defined programming language interface that is between a program and the user of a program.

archive copy group. In TSM, 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 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 TSM.

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. (A)

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

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

BCOCA. Bar Code Object Content Architecture

backend. In the AIX operating system, the program that sends output to a particular device. Synonymous with Backend Program.

backend program. Synonym for Backend.

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

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

BMP. Bitmap

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

C

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

CD-ROM. Compact disc read-only memory

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. (1) Pertaining to devices attached to a controlling unit by cables, rather than by telecommunication lines. (2) Synonymous with Local.

character. A letter, digit, or other symbol representing, organizing, or controlling data.

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.

click. To press the left mouse button while pointing to an object such as a command button or a toolbar button.

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 personal computer connected to a network running OnDemand software that can log on and query the library server, retrieve documents from OnDemand, and view and print documents.

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

client node. An application group that has been registered to the TSM server.

COBOL. Common business-oriented language. A high-level programming language, based on English, that is used primarily for business 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.

code point. A character within a code page.

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 OnDemand and TSM to exchange information.

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

compact disc read-only memory (CD-ROM). High capacity read-only memory in the form of an optically read compact disk.

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.

configuration. The process of describing to a system the devices, optional features, and program products that have been installed so that these features can be used. Contrast with Customization.

configuration file. A file that specifies the characteristics of a system or subsystem; for example, the operating system queueing system.

configure. To describe to a system the devices, optional features, and licensed programs installed on a system.

console. The main operating system display station.

constant. A data item with a value that does not change during the running of a program. Contrast with Variable.

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.

Content Manager. A comprehensive set of Web-enabled, integrated software solutions from IBM for managing information and making it available to anyone, anywhere.

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.

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.

copies. See Copy Group.

copy group. In TSM, 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

customization. The process of describing optional changes to defaults of a software program that is already installed on the system and configured so that it can be used. Contrast with Configuration.

customize. To describe the system, the devices, programs, users, and user defaults for a particular data processing system or network. Contrast with Configure.

D

daemon. In UNIX, a process begun by the root user or by the root shell that can be stopped only by the root user. Daemon processes generally provide services that must be available at all times, such as sending data to the printer. A daemon runs continuously, looking for work to do, performing that work, and waiting for more work. A daemon does not have a controlling terminal associated with it. The OnDemand data download program (ARSJESD) is an example of a daemon.

database. (1) The collection of information about all objects managed by OnDemand, including reports, groups, users, printers, application groups, storage sets, applications, and folders. (2) The collection of information about all objects managed by TSM, including policy management objects, administrators, and client nodes.

Database Managed Space (DMS). A type of DB2 table space. A DSM table space is managed by the database manager.

data set. Synonym for 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.

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 the printed output from a display station assigned to it.

default value. A value stored in the system that is used when no other value is specified. See also Default.

desktop printer. In this publication, an IBM LaserPrinter 4019 or 4029, or compatible printer.

device class. A named group of TSM 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 type. A type of TSM storage device. Each device class must be categorized with one of the following devices types: disk, tape, or optical disk.

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

dialog box. An application window on the display that requests information from the user.

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.

disk operating system (DOS). An operating system for computer systems that use disks and diskettes for auxiliary storage of programs and data.

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

DMS. Database Managed Space

document. (1) In 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. An IBM licensed program used to prepare printed documents.

domain. See Policy Domain or Client Domain.

DOS. Disk operating system

double-click. To rapidly press the left mouse button twice while pointing to an object.

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.

drag. To hold down the left mouse button while moving the mouse.

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. This is the default type of data encoding in an MVS environment. Contrast with ASCII.

EIP. Enterprise Information Portal

enqueue. To place items in a queue.

enter. (1) An instruction to type specific information using the keyboard. (2) A keyboard key that, when pressed, confirms or initiates the selected command.

Enterprise Information Portal. An IBM software product that provides a coordinated, Web-enabled entry point to what would otherwise be disconnected, incompatible data scattered across an enterprise.

Enterprise Storage Server. An IBM disk storage system that provides industry-leading availability, performance, manageability, and scalability. Virtually all types of servers can concurrently attach to the Enterprise Storage Server, including S/390, UNIX servers, and Windows servers. As a result, the Enterprise Storage Server is ideal for organizations with growing e-business operations that are being handled by multiple heterogeneous servers.

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 computer 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 failure data 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.

ESS. Enterprise Storage Server

Ethernet. A 10-megabit baseband local area network using CSMA/CD (carrier sense multiple access with collision detection). The network 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 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). 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. Synonym for Resource Object.

external object. Synonym for Resource Object.

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. In OnDemand, the end-user view of data stored in the system. Folders provide users a convenient way to find related information, regardless of the source of the information or where the data is stored.

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. Often synonymous with Character Set. See also Coded Font.

form definition (FORMDEF). A form definition is a resource used by OnDemand. A form definition 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). Synonymous with FORMDEF.

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. An MVS or OS/390 interface that allows communication between JES and a FSS and FSS applications. Download uses an FSI to communicate with the operating system and JES to process spool data sets created by application programs.

FSS. Functional SubSystem. An MVS or OS/390 subsystem comprised of programs residing in the same address space that provide JES-related functions. For example, a print programs that extend the scope of JES processing could be defined as a FSS.

FTP. File Transfer Protocol

G

GB. Gigabyte

GIF. Graphic 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)

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

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

Graphical User Interface. A type of user interface that takes advantage of a high-resolution monitor, including some combination of graphics, the use of pointing devices, menu bars, overlapping windows, and icons.

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

Graphic Interchange Format (GIF). A bit-mapped color graphics file format for IBM and IBM-compatible computers. GIF employs an efficient compression technique for high resolution graphics.

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

hardware. The physical equipment of computing and computer-directed activities. The physical components of a computer system. Contrast with Software.

help. One or more files of information that describe how to use application software or how to perform a system function.

hex. Hexadecimal

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).

host. (1) The primary or controlling computer in the communications network. (2) See Host System.

host-based computer. (1) In a computer network a computer that provides end users with services such as computation and data bases 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.

indexing. (1) A process of segmenting a print file into uniquely identifiable groups of pages (a named collection of sequential pages) for later retrieval. (2) In ACIF, 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.

indexing with data values. Adding indexing tags to a MO:DCA-P document using data that is already in the document and that is consistently located in the same place in each group of pages.

indexing with literal values. Adding indexing tags to a MO:DCA-P document by assigning literal values as indexing tags, because the document is not organized such that common data is located consistently throughout the document.

Infoprint Manager. A sophisticated IBM print subsystem that drives AFP printers, PostScript printers, and PCL printers. Infoprint Manager is supported under AIX, OS/390, Windows NT, and Windows 2000. 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.

install. (1) To add a program, program option, or software program to the system in a manner such that it may be executed and will interact properly with all affected programs in the system. (2) To connect a piece of hardware to the processor.

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

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.

K-byte. Kilobyte

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

kilobyte (K-byte). 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. (T)

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 may 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 (also known as *resource name*).

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

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) Contrast with Page Printer.

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, 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 may 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.

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

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

LPD. Line Printer Daemon. In TCP/IP, the command responsible for sending data from the spooling directory to a printer.

LPR. Line Printer Requestor. In TCP/IP, a client command that allows the local host to submit a file to be printed on a remote print server.

LZW. See Lempel Ziv Welsh

M

M byte. Megabyte

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 TSM. 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.

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 may 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 TSM, 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). (1) A strategic, architected, device-independent data stream for interchanging documents. (2) A printing data stream that is a subset of the Advanced Function Presentation data stream.

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

mount. To make a file system accessible.

mouse. A hand-held locator that a user operates by moving it on a flat surface. It allows the user to select objects and scroll the display screen by pressing buttons.

N

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 OnDemand library server or object server and is connected to a TCP/IP network.

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

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

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 may contain one or more other structured fields whose content consists of one or more data elements of a particular data type. An object may be assigned a name, which may 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. Synonymous with Type 1 fonts.

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. (1) A collection of data that can be printed on one side of a sheet of paper or a form. (2) The boundary for determining the limits of printing. See also Logical Page and Physical Page. (3) Part of an AFP document bracketed by a pair of Begin Page and End Page structured fields.

page definition. A resource used by 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) Contrast with Line Printer.

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.

PAGEDEF. Page definition

parallel device. A device that can perform two or more concurrent activities. Contrast with Serial Device.

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.

PC. Personal Computer

PCL. Printer control language

PCX. Picture Exchange Format

PDF. Portable Document Format

permissions. Codes that determine the users that 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.

personal computer. A microcomputer primarily intended for stand-alone use by an individual. (T)

Picture Exchange Format (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. (1) To move the mouse pointer to a specific object. (2) 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 TSM, 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 TSM, 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.

press. To touch a specific key on the keyboard.

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 TSM stores archive copies of files.

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 Services Facility (PSF). A sophisticated IBM print subsystem that drives IPDS page printers. PSF is supported under MVS, VSE, VM, OS/2®, AIX, and is a standard part of the operating system under OS/400®. PSF 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, PSF 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.

Print Services Facility/2 (PSF/2). PSF/2 is an OS/2-based print server that drives IPDS page printers, as well as IBM PPDS and HP-PCL compatible printers. PSF/2 manages printer resources and provides error recovery for print jobs. PSF/2 supports distributed printing of AFP print jobs from PSF for AIX, PSF/MVS, PSF/VSE, PSF/VM, and OS/400. PSF/2 also supports printing from a wide range of workstation applications, including Microsoft Windows and OS/2 Presentation Manager®, as well as the ASCII, PostScript, and AFP data streams.

Print Services Facility for AIX (PSF for AIX). An IBM licensed program that produces printer commands from the data sent to it and it runs on the AIX/6000® operating system.

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

Printer Control Language (PCL). The data stream used by Hewlett-Packard LaserJet II and III and other compatible printers.

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. Print Services Facility

PSF/2. Print Services Facility/2

PSF for AIX. Print Services Facility for AIX

PTF. Program temporary fix

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

radio button. Round option buttons grouped in dialog boxes; one is preselected. Like a radio in an automobile, select only one button (“station”) at a time.

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. Contrast with Outline 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.

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 TSM, 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 appears on the display screen. 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 TSM.

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

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 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. A place in which 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 TSM 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. Synonymous with Character Rotation. (2) The number of degrees a character is turned relative to the page coordinates. (3) The orientation of the characters of a font with respect to the baseline.

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

S

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. OnDemand divides report files into segments to provide enhanced performance and maintenance.

segment table. A high-level index to index 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.

select. To pick a menu command or other object with a single click of the mouse.

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

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 OnDemand programs that store, retrieve, and maintain report files. OnDemand supports two types of servers: a library server an object server.

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

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 Space

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

spool file. (1) A disk file containing 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 synonym for the queuing system that pertains to its use for queuing print jobs.

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. (1) The current condition or state of a program or device. For example, the status of a printer. (2) The condition of the hardware or software, usually represented in a status code.

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 TSM domain on an 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 TSM, a named collection of storage volumes that is the destination for archived files.

storage pool volume. In TSM, 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 OnDemand server.

string. A series or set of alphabetic or numeric characters. A string can be composed of letters, numbers, and special 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.

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. Synonymous with Console.

system customization. Specifying the devices, programs, and users for a particular data processing system. See also Configuration.

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. Synonymous with Main Storage, but used in hardware to refer to semiconductor memory (modules).

system prompt. Synonym for command line. 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.

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

Tivoli Storage Manager. An IBM software program that provides archive storage management of data stored in an OnDemand system.

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 OnDemand client programs that support a graphical user interface.

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

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.

TSM. Tivoli Storage Manager

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

U

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.

UNIX operating system. An operating system developed by Bell Laboratories that features multiprogramming in a multi-user environment. The UNIX operating system was originally developed for use on minicomputers but has been adapted for mainframes and microcomputers.

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

user. A person authorized to logon to an OnDemand server.

user exit. (1) A point in an IBM-supplied program at which a user-defined program may be given control. (2) A programming service provided by an IBM software product that may 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 take 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.

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

wildcard. Search characters that represent other letters, numbers, or special characters. In 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

Numerics

A

B

C

D

E

F

G

H

I

J

L

M

N

O

T

U

W

Y

Printed in the United States of America
on recycled paper containing 10%
recovered post-consumer fiber.

GC27-0839-00

Spine information:

Content Manager OnDemand
for Multiplatforms

Introduction and Planning Guide

Version 7.1