IBM i
7.5

Availability
Availability roadmap



Note

Before using this information and the product it supports, read the information in “Notices” on page 19.

This edition applies to IBM i (product number 5770-SS1) and to all subsequent releases and modifications until otherwise indicated in new editions. This version does not run on all reduced instruction set computer (RISC) models nor does it run on CISC models.

This document may contain references to Licensed Internal Code. Licensed Internal Code is Machine Code and is licensed to you under the terms of the IBM License Agreement for Machine Code.

© Copyright International Business Machines Corporation 1998, 2022.
US Government Users Restricted Rights – Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Contents

Availability roadmap....1
What's new for IBM i 7.5....1
PDF file for Availability roadmap....1
Availability concepts....2
Estimating the value of availability....3
Deciding what level of availability you need....3
Preventing unplanned outages....5
Preparing for disk failures....5
Preparing for power loss....7
Using effective systems management practices....8
Preparing the space for your system....8
Shortening unplanned outages....9
Reducing the time to restart your system....9
Recovering recent changes after an unplanned outage....10
Recovering lost data after an unplanned outage....11
Reducing the time to vary on independent disk pools....12
Shortening planned outages....12
Shortening backup windows....12
Performing online backups....13
Backup from a second copy....14
Backing up less data....14
Shortening software maintenance and upgrading windows....15
Shortening hardware maintenance and upgrade windows....16
High availability....16
Related information for Availability roadmap....17

Notices....19
Programming interface information....20
Trademarks....20
Terms and conditions....21

Availability roadmap

The topic collection guides you through IBM i availability and helps you decide which availability tools are right for your business.

Availability is the measure of how often your data and applications are ready for you to access when you need them. Different companies have different availability needs. Different systems or different applications within the same company might have different availability needs. It is important to note that availability requires detailed planning. These availability tools are only useful if you have implemented them before an outage occurs.

Before you can really start to plan for availability on your system, you should become familiar with the basic availability concepts, understand the costs and risk associated with outages, and determine your company's needs for availability. After you have a basic understanding of availability concepts and know what level of availability you need, you can start to plan for that level of availability on a single system or on multiple systems within a cluster environment.

What's new for IBM i 7.5

Read about new or significantly changed information for the Availability roadmap topic collection.

Updates to Availability roadmap information

Minor updates and clarifications were made to the Availability roadmap topic.

How to see what's new or changed

To help you see where technical changes have been made, the information center uses:

- The » image to mark where new or changed information begins.
- The « image to mark where new or changed information ends.

In PDF files, you might see revision bars (|) in the left margin of new and changed information.

To find other information about what's new or changed this release, see the Memo to users.

PDF file for Availability roadmap

You can view and print a PDF file of this information.

To view or download the PDF version of this document, select Availability roadmap .

Saving PDF files

To save a PDF on your workstation for viewing or printing:

1. Right-click the PDF link in your browser.
2. Click the option that saves the PDF locally.
3. Navigate to the directory in which you want to save the PDF.
4. Click **Save**.

Downloading Adobe Reader

You need Adobe Reader installed on your system to view or print these PDFs. You can download a free copy from the Adobe Web site (www.adobe.com/products/acrobat/readstep.html).

© Copyright IBM Corp. 1998, 2022

Related reference

Related information for Availability roadmap
Product manuals, IBM Redbooks publications, websites, experience reports, and other information center topic collections contain information that relates to the Availability roadmap topic collection. You can view or print any of the PDF files.

Availability concepts

Before you plan for the availability of your system, it is important for you to understand some of the concepts associated with availability.

Businesses and their IT operations that support them must determine which solutions and technologies address their business needs. In the case of business continuity requirements, detailed business continuity requirements must be developed and documented, the solution types must be identified, and the solution choices must be evaluated. This is a challenging task due in part to the complexity of the problem.

Business continuity is the capability of a business to withstand outages, which are times when the system is unavailable, and to operate important services normally and without interruption in accordance with predefined service-level agreements. To achieve a given level of business continuity, a collection of services, software, hardware, and procedures must be selected, described in a documented plan, implemented, and practiced regularly. The business continuity solution must address the data, the operational environment, the applications, the application hosting environment, and the user interface. All must be available to deliver a good, complete business continuity solution. Your business continuity plan includes disaster recovery and high availability (HA).

Disaster recovery provides a plan in the event of a complete outage at the production site of your business, such as during a natural disaster. Disaster recovery provides a set of resources, plans, services, and procedures used to recover important applications and to resume normal operations from a remote site. This disaster recovery plan includes a stated disaster recovery goal (for example, resume operations within eight hours) and addresses acceptable levels of degradation.

Another major aspect of business continuity goals for many customers is *high availability*, which is the ability to withstand all outages (planned, unplanned, and disasters) and to provide continuous processing for all important applications. The ultimate goal is for the outage time to be less than .001% of the total service time. The differences between high availability and disaster recovery typically include more demanding recovery time objectives (seconds to minutes) and more demanding recovery point objectives (zero user disruption).

Availability is measured in terms of *outages*, which are periods of time when the system is not available to users. During a planned outage (also called a scheduled outage), you deliberately make your system unavailable to users. You might use a scheduled outage to run batch work, back up your system, or apply fixes.

Your *backup window* is the amount of time that your system can be unavailable to users while you perform your backup operations. Your backup window is a scheduled outage that typically occurs in the night or on a weekend when your system has less traffic.

An *unplanned outage* (also called an unscheduled outage) is typically caused by a failure. You can recover from some unplanned outages (such as disk failure, system failure, power failure, program failure, or human error) if you have an adequate backup strategy. However, an unplanned outage that causes a complete system loss, such as a tornado or fire, requires you to have a detailed disaster recovery plan in place in order to recover.

High availability solutions provide fully automated failover to a backup system to ensure continuous operation for users and applications. These HA solutions must provide an immediate recovery point and ensure that the time of recovery is faster than a non-HA solution.

Unlike with disaster recovery, where entire systems experience an outage, high availability solutions can be customized to individual critical resources within a system; for example, a specific application instance. High availability solutions are based on cluster technology. You can use clusters to avoid the impacts of both planned and unplanned outages. Even though you still have an outage, the business

function is not impacted by the outage. A *cluster* is a collection of interconnected complete systems used as a single, unified resource. The cluster provides a coordinated, distributed process across the systems to deliver the solution. This results in higher levels of availability, some horizontal growth, and simpler administration across the enterprise. For a complete solution, you must address the operational environment, the application hosting environment, application resilience, and the user interfaces in addition to providing the data resilience mechanisms. Clusters focus on all aspects of the complete solution. The integrated cluster resource services enable you to define a cluster of systems and the set of resources that should be protected against outages. Cluster resource services detect outage conditions and coordinate automatic movement of critical resources to a backup system.

Estimating the value of availability

You need to build a business case for improving availability when you are asked to justify the cost of additional hardware to support availability.

To estimate the value of your information services, follow these steps:

1. **Develop a list of the major services and solutions that your system provides.** Your system exists so that users and solutions can accomplish tasks that are critical to the operation of your business. The systems provide solutions to a business function. If the system is unavailable, the business function cannot be completed or is significantly degraded to the point of causing the business lost revenue or increased expenses.
2. **Assess how much it costs you when these services are unavailable.** Each application or service has a direct effect on business functions. You need to determine how these business functions would be affected and what would be the overall cost to your business if these services were unavailable.
3. **Look at direct costs versus indirect costs.** Direct costs are losses that can be traced directly to a system being unavailable. Indirect costs are those that are incurred by another department or function as a result of an outage.
4. **Consider tangible costs versus intangible costs.** Tangible costs can be measured in currency. However, there are other costs that are not measured with money, such as market share, lost opportunity, and good will.
5. **Analyze fixed costs versus variable costs.** Fixed costs are those that result from a failure and are the same, regardless of the length of the outage. Variable costs are those that vary, based on the length of the outage.

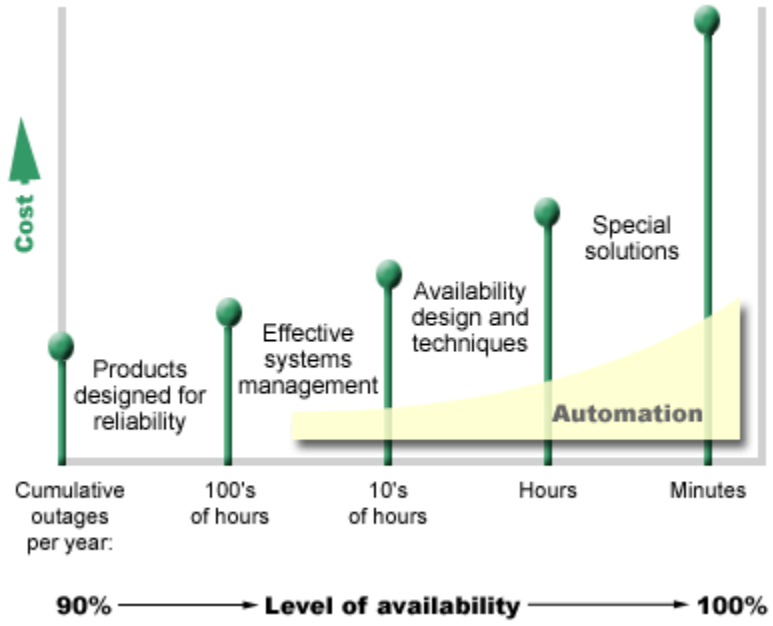
Deciding what level of availability you need

After understanding availability at a basic level, it is important to assess your individual availability needs. Higher availability is more costly than a lower-level availability. You must balance your needs and services with the overall cost of implementing and maintaining these availability solutions.

You want to be sure that you have analyzed your business needs thoroughly in order to decide what level of availability you can afford to maintain. To decide what level of availability you need, consider the following questions:

Do you have any applications that require 100% availability?

In most cases, you can achieve a high level of availability by implementing sound processes and systems management practices. The closer you need to be to continuous availability, the more of an investment you must make. Before you make that kind of investment, you should be sure that you require that level of availability. The following figure shows how different techniques can improve availability, but can increase the price you must pay for it.



If your requirements for levels of availability increase, you can consider multiple system availability solutions, such as clusters.

How much downtime is acceptable to you?

It is helpful to know the amount of downtime that each level of availability represents. The following table shows the amount of downtime expected for different levels of availability.

Level of availability	Downtime per year
90%	36.5 days
95%	18.25 days
99%	3.65 days
99.9%	8.76 hours
99.99%	50 minutes
99.999%	5 minutes

Along with knowing how much downtime is acceptable to you, you need to consider how that downtime might occur. For example, you might think that 99% availability is acceptable if the downtime is a series of shorter outages that are distributed over the course of a year, but you might think differently about 99% availability if the downtime is actually a single outage that lasts three days.

You also need to consider when a downtime is acceptable and when it is not. For example, your average annual downtime goal per year might be nine hours. If that downtime were to occur during critical business hours, it might have an adverse effect on the bottom line revenue for your company.

What level of access do your customers need to your business?

It used to be that customers and business partners accessed your business from 9 a.m. to 5 p.m., so it was realistic to expect that your system only had to be available during those hours. However, the Internet and a diverse global marketplace have changed that expectation; customers and business associates might expect to have access to your company's data at any time of the day or night. Your working hours might be hours or even days different from your global business partner or customer.

You must determine what your customer expectations are, and what is realistic with regard to those expectations, as you determine what level of availability you will maintain.

Preventing unplanned outages

One way to approach availability is to try to prevent unplanned outages. You can use these different methods to ensure that your system experiences as little unplanned downtime as possible.

Note: If you have a single system and you cannot achieve the level of availability you need using these strategies, you can consider clusters.

Preparing for disk failures

Because your data is spread across your disks, it is important that you consider how to protect your data if one of those disks fails. Disk protection helps ensure the availability of data stored on the disks.

Disk storage is the storage that is either internal to your system or is attached to it. This disk space, together with your system's main memory, is regarded by your system as one large storage area. When you save a file, you do not assign it to a storage location; instead, the system places the file in the location that ensures the best performance. It might spread the data in the file across multiple disk units. When you add more records to the file, the system assigns additional space on one or more disk units. This way of addressing storage is known as *single-level storage*.

In addition to internal disk storage, you can also use IBM System Storage products to attach a large volume of external disk units. These storage products provide enhanced disk protection, the ability to copy data quickly and efficiently to other storage servers, and the capability of assigning multiple paths to the same data to eliminate connection failures. For additional information about IBM System Storage products and to determine whether this solution is right for you, see disk storage.

Another way of preparing for disk failures is by using DS8000® Full System HyperSwap®, which is part of IBM PowerHA for i Express Edition. DS8000 Full System HyperSwap is a single system solution that uses two IBM System Storage Servers and does not require a cluster. IBM i provides the ability for the system to switch between the DS8000 servers for planned and unplanned storage side outages without losing access to the data during the switch. For more information about DS8000 Full System HyperSwap, see PowerHA® data replication technologies.

Device parity protection

Device parity protection allows your system to continue to operate when a disk fails or is damaged. When you use device parity protection, the disk input/output adapter (IOA) calculates and saves a parity value for each bit of data. The IOA computes the parity value from the data at the same location on each of the other disk units in the device parity set. When a disk failure occurs, the data can be reconstructed by using the parity value and the values of the bits in the same locations on the other disks. Your system continues to run while the data is being reconstructed.

The IBM i supports two types of device parity protection:

RAID 5

With RAID 5, the system can continue to operate if one disk fails in a parity set. If more than one disk fails, data will be lost and you must restore the data for the entire system (or only the affected disk pool) from the backup media. Logically, the capacity of one disk is dedicated to storing parity data in a parity set consisting of 3 to 18 disk units.

RAID 6

With RAID 6, the system can continue to operate if one or two disks fail in a parity set. If more than two disk units fail, you must restore the data for the entire system (or only the affected disk pool) from the backup media. Logically, the capacity of two disk units is dedicated to storing parity data in a parity set consisting of 4 to 18 disk units.

RAID 10

With RAID 10, the system can continue to operate if one of the disk units in the pair should fail, the other unit in the set would be able to sustain the functions of the failed unit. If both units fail, you must restore the data for the entire system (or only the affected disk pool) from the backup media.

RAID 10 is expected to provide significant performance advantages over the RAID 5 and RAID 6 protection. RAID 10 can only be started on internal storage IOAs, which support the function.

Write cache and auxiliary write cache IOA

When the system sends a write operation, the data is first written to the write cache on the disk IOA and then later written to the disk. If the IOA experiences a failure, the data in the cache might be lost and cause an extended outage to recover the system.

The auxiliary write cache is an additional IOA that has a one-to-one relationship with a disk IOA. The auxiliary write cache protects against extended outages due to the failure of a disk IOA or its cache by providing a copy of the write cache which can be recovered following the repair of the disk IOA. This avoids a potential system reload and gets the system back online as soon as the disk IOA is replaced and the recovery procedure completes. However, the auxiliary write cache is not a failover device and cannot keep the system operational if the disk IOA (or its cache) fails.

Hot-spare disks

A disk designated as a *hot-spare disk* is used when another disk that is part of a parity set on the same IOA fails. It joins the parity set and rebuilding the data for this disk is started by the IOA without user intervention. Because the rebuild operation occurs without having to wait for a new disk to be installed, the time that the parity set is exposed is greatly reduced. See Hot spare protection for additional information

Mirrored protection

Disk mirroring is recommended to provide the best system availability and the maximum protection against disk-related component failures. Data is protected because the system keeps two copies of the data on two separate disk units. When a disk-related component fails, the system can continue to operate without interruption by using the mirrored copy of the data until the failed component is repaired.

Different levels of mirrored protection are possible, depending on what hardware is duplicated. The level of mirrored protection determines whether the system keeps running when different levels of hardware fail. To understand these different levels of protection, see Determining the level of mirrored protection that you want.

You can duplicate the following disk-related hardware:

- Disk unit
- Disk controllers
- I/O bus unit
- I/O adapter
- I/O processors
- A bus
- Expansion towers

Hot-spare disks

A disk designated as a *hot-spare disk* is used when another disk that is mirror-protected fails. A hot spare disk unit is stored on the system as a non-configured disk. When a disk failure occurs, the system exchanges the hot spare disk unit with the failed disk unit. The exchange of a mirrored subunit with the hot spare disk unit does not occur until mirror-protection has been suspended for 5 minutes and the replacement disk has been formatted. After the exchange occurs, the system synchronizes the data on the new disk unit. See Hot spare protection for additional information.

Independent disk pools

With independent disk pools (also called independent auxiliary storage pools), you can prevent certain unplanned outages because the data on them is isolated from the rest of your system. If an independent disk pool fails, your system can continue to operate on data in other disk pools. Combined with different levels of disk protection, independent disk pools provide more control in isolating the effect of a disk-related failure as well as better prevention and recovery techniques.

PowerHA replication technologies

IBM PowerHA SystemMirror for i product offers different replication technologies, which protect against disk outages. Some examples of these technologies are Geographic mirroring, Metro Mirror, and Global Mirror. Some technologies are a combination of IBM System Storage copy services technology and IBM i clustering technology, which protect against disk outages.

For more information on PowerHA technologies, see PowerHA data replication technologies.

Multipath disk units

You can define up to eight connections from each logical unit number (LUN) created on the IBM System Storage products to the input/output processors (IOPs) on the system. Assigning multiple paths to the same data allows the data to be accessed even though some failures might occur in other connections to the data. Each connection for a multipath disk unit functions independently. Several connections provide availability by allowing disk storage to be used even if a single path fails.

Related information

Disk management

Preparing for power loss

To ensure that your system is available when you need it, you need to make sure that it has an adequate supply of power and that it is protected in the event that power is lost.

Power requirements

Part of the planning process for your system is to ensure that you have an adequate power supply. You need to understand your system's requirements and then enlist the aid of a qualified electrician to help install the appropriate wiring, power cords, plugs, and power panels. For details on how to ensure that your system has adequate power, see Power Systems->*your-specific-Power® system*->*your-specific-model*->Planning for the System->Site and Hardware Planning->Planning for power.

Redundant power supplies

Some systems are designed with a fully redundant power system. A redundant power supply is a feature that prevents an unplanned outage by providing power if one power supply fails.

Some systems support a dual power cable feature. This allows you to provide power to the system from two different external power sources. This can avoid an unplanned outage if one power source fails.

Uninterruptible power supplies

Even when you have an adequate power supply, there are still times when you might lose power, such as during a storm. To prevent unplanned outages that result from losing power, you might need to invest in hardware specifically designed to keep your system running when power is lost. One such piece of hardware is an *uninterruptible power supply*. You can use an uninterruptible power supply to provide auxiliary power to your processor, disks, system console, and any other devices. Uninterruptible power supplies provide the following advantages:

- Let you continue operations during brief power outages (brown outs).
- Protect the system from voltage peaks (white outs).

- Provide a normal end of operations in case of an extended power outage, which can reduce your recovery time when you restart your system. You can write a program that helps you control your system's shutdown in these conditions.

Make sure that your uninterruptible power supplies are compatible with your systems.

For details on how to ensure that your system has uninterruptible power supplies, see Power Systems->*your-specific-Power system*->*your-specific-model*->Planning for the System->Site and Hardware Planning->Planning for power->Uninterruptible power supply.

Generator power

To prevent a long outage because of an extended power failure, you might consider purchasing a generator. A generator goes a step further than an uninterruptible power supply in that it enables you to continue normal operations during longer power failures.

Using effective systems management practices

One of the simplest ways to prevent unplanned outages is to ensure that you are doing everything you can to keep your system running smoothly. This includes performing basic preventive maintenance and systems management tasks that help your system perform at its peak.

Many of these systems management tasks can be automated, which helps you prevent failures that might occur because of human error or an oversight.

One way you can help ensure that the availability of your system is to monitor its performance and react promptly to any problems that you encounter. You can use the collection services and monitors functions in Management Central to actively monitor and track the performance of your system. You can be notified of any problems that jeopardize the availability of your system in time to react and prevent an unplanned outage.

Fixes are also an important systems management component that can help you keep your system available. When problems are discovered in IBM i or other IBM software, IBM issues a fix (also known as a PTF, or program temporary fix) to solve the problem. You need to be aware of fixes and install them on your system to ensure that your system is operating at its optimal level. You should create a fix management strategy and make checking for and applying fixes part of the routine maintenance for your system. For help in determining a strategy for preventive maintenance based on your IBM i environment and applications, see Guide to fixes .

In addition, you should also develop a strategy of when and how new applications are added to your systems. Before adding a new application or updating an existing application, ensure that any software, hardware, or other dependencies are understood and in place. It is also important to plan and test these new or changed applications before introducing them into a production environment to help avoid potential outages or other unexpected impacts.

Related information

Performance

Maintain and manage IBM i and related software

Guide to fixes

Preparing the space for your system

One way to prevent unplanned outages is to ensure that the space where you put your system promotes availability. Many physical and environmental factors contribute to how your system performs.

The first thing you need to do is to familiarize yourself with your system. Different system models have different requirements with regard to the conditions they are in, so you should be aware of what your system needs. For more information about the physical characteristics of each model, see Power Systems->*your-specific-Power system*->*your-specific-model*->Planning for the System->Site and Hardware Planning->Hardware specification sheets->Server specifications. For information about the

physical characteristics of the hardware that you can attach to your system, see Hardware specification sheets.

After you are familiar with the physical characteristics of your system, be sure to consider the following information about the space where the system resides:

- **Location.** The physical location of your system can affect on how available it is. For example, if the room is not secure, your system might be vulnerable to vandalism or even someone accidentally dislodging the power cord. For more details about what precautions you need to take with regard to your system's location, see Power Systems->*your-specific-Power system*->*your-specific-model*->Planning for the System->Site preparation and physical planning.
- **Cables.** Cables are often overlooked, but without them your system will be unusable. You need to ensure that your cables are in good condition and are being used correctly. For more information about how to ensure that your cables do the job that you want them to do, see Power Systems->*your-specific-Power system*->*your-specific-model*->Planning for the System->Site and Hardware Planning->Planning for cables.
- **Environment.** The environment you provide for your system is also crucial to the availability. The environment includes such things as temperature, humidity, and other factors that can inhibit the performance of your system. For more detailed information about the environment you need to provide for your system, see Power Systems->*your-specific-Power system*->*your-specific-model*->Planning for the System->Site preparation and physical planning.

Shortening unplanned outages

Unplanned outages do occur, and a key to availability is to ensure that you can recover from the outages as quickly as possible.

The purpose of all recovery strategies is to get back to where you were before the outage occurred.

If you have a single system and you cannot achieve the level of availability you need using these strategies, you can consider clusters.

Reducing the time to restart your system

Unfortunately, systems periodically experience an unplanned outage. Use the availability tools described here to restart your system as quickly as possible after an unplanned outage.

Before your system turns off, it performs certain activities to ensure that your data is protected and that jobs are ended in a controlled manner. When you experience an unplanned outage, your system cannot perform these activities. Each of these tools should be used together to quicken the startup times for your system.

For more details on what happens when your system ends abnormally, see Starting and stopping the system.

Use the Reducing iSeries IPL Time experience report to learn how to control the time that it takes to start your system.

System-managed access-path protection (SMAPP)

An *access path* is the route that an application takes through a database file to get to the records it needs. A file can have multiple access paths, if different programs need to see the records in different sequences. When your system ends abnormally, such as during an unplanned outage, the system must rebuild the access paths the next time it starts, which might take a long time. When you use system-managed access-path protection, the system protects the access paths so that they do not need to be rebuilt when your system starts after an unplanned outage. This saves you time when you restart your system, which enables you to get back to your normal business processes as quickly as possible.

Journaling access paths

Like SMAPP, journaling access paths can help you to ensure that critical files and access paths are available as soon as possible after you restart your system. However, when you use SMAPP, the system decides which access paths to protect. Therefore, if the system does not protect an access path that you consider critical, you might be delayed in getting your business running again. When you journal the access paths, you decide which paths to journal.

SMAPP and journaling access paths can be used separately. However, if you use these tools together, you can maximize their effectiveness for reducing startup time by ensuring that all access paths that are critical to your business operations are protected.

Protecting your access paths is also important if you plan to use a high availability product, such as IBM PowerHA SystemMirror for i, to avoid rebuilding access paths when you failover to a backup system.

Independent disk pools

When a system is started or restarted, you can start each independent disk pools individually. By starting each independent disk pool separately, the system can be made available more quickly. You can prioritize the workload so that the critical data becomes available first. You can then vary on independent disk pools in a specific order based on this priority.

Related information

Starting and stopping the system
System-managed access-path protection
Reducing iSeries IPL Time
Example: Make independent disk pool available at startup

Recovering recent changes after an unplanned outage

After an unplanned outage, your goal is to get your system up and running again as quickly as possible. You want to get back to where you were before the outage occurred without having to manually reenter the transactions.

This might involve rebuilding some of your data. There are a few availability tools that you can use that help you get back to where you were before the outage more quickly.

Journaling

Journal management prevents transactions from being lost if your system ends abnormally. When you journal an object, the system keeps a record of the changes you make to that object.

Commitment control

Commitment control helps to provide data integrity on your system. With commitment control, you can define and process a group of changes to resources, such as database files or tables, as a single transaction. It ensures that either the entire group of individual changes occurs or that none of the changes occurs. For example, you lose power just as a series of updates are being made to your database. Without commitment control, you run the risk of having incomplete or corrupted data. With commitment control, the incomplete updates are backed out of your database when you restart your system.

You can use commitment control to design an application, so that the system can restart the application if a job, an activation group within a job, or the system ends abnormally. With commitment control, you can have assurance that when the application starts again, no partial updates are in the database because of incomplete transactions from a prior failure.

Related information

Journal management
Commitment control

Recovering lost data after an unplanned outage

You might lose data as a result of an unplanned outage, such as a disk failure. The most extreme example of data loss is losing your entire site, such as what might happen as a result of a natural disaster.

There are a few ways that you can prevent your data from being lost in these situations or at least limit the amount of data that is lost.

Backup and recovery

It is imperative that you have a proven strategy for backing up your system. The time and money you spend on creating this strategy is more than recovered should you need to restore lost data or perform a recovery. After you have created a strategy, you must ensure that it works by testing it, which involves performing a backup and recovery, and then validating that your data was backed up and restored correctly. If you change anything on your system, you need to assess whether your backup and recovery strategy needs to be changed.

Every system and business environment is different, but you should try to do a full backup of your system at least once a week. If you have a very dynamic environment, you also should back up the changes to objects on your system since the last backup. If you have an unexpected outage and need to recover those objects, you can recover the latest version of them.

To help manage your backup and recovery strategy and your backup media, you can use Backup, Recovery, and Media Services (BRMS). *BRMS* is a program that helps you implement a disciplined approach to managing your backups, and provides you with an orderly way to recover lost or damaged data. Using BRMS, you can manage your most critical and complex backups, including online backups of Lotus servers, simply and easily. You can also recover your system fully after a disaster or failure.

In addition to these backup and recovery features, BRMS enables you to track all of your backup media from creation to expiration. You no longer need to keep track of which backup items are on which volumes, and worry that you might accidentally write over active data. You can also track the movement of your media to and from offsite locations.

For detailed information about the tasks that BRMS can help you perform, see Backup, Recovery, and Media Services.

For help in planning and managing your backup and recovery strategy, see Selecting the appropriate recovery strategy or contact Business continuity .

Limiting the amount of data that is lost

You can group your disk drives into logical subsets called *disk pools* (also known as auxiliary storage pools or ASPs). The data in one disk pool is isolated from the data in the other disk pools. If a disk unit fails, you only need to recover the data in the disk pool that contains the failed unit.

Independent disk pools are disk pools that can be brought online or taken offline without any dependencies on the rest of the storage on a system. This is possible because all of the necessary system information associated with the independent disk pool is contained within the independent disk pool. Independent disk pools offer a number of availability and performance advantages in both single and multiple system environments.

Logical partitions provide the ability to divide one system into several independent systems. The use of logical partitions is another way that you can isolate data, applications, and other resources. You can use logical partitions to improve the performance of your system, such as by running batch and interactive processes on different partitions. You can also protect your data by installing a critical application on a partition apart from other applications. If another partition fails, that program is protected.

Related information

Planning a backup and recovery strategy

Backing up your system

Recovering your system

Backup, Recovery, and Media Services (BRMS)
Disk pools
Disk management
Independent disk pool examples
Logical partitions
Restoring changed objects and apply journaled changes

Reducing the time to vary on independent disk pools

When unplanned outages occur, the data stored within independent disk pools is unavailable until they can be restarted. To ensure that the restart occurs quickly and efficiently, you can use the strategies described in this topic.

Synchronizing user profile name, UID, and GID

In a high-availability environment, a user profile is considered to be the same across systems if the profile names are the same. The name is the unique identifier in the cluster. However, a user profile also contains a user identification number (UID) and group identification number (GID). To reduce the amount of internal processing that occurs during a switchover, where the independent disk pool is made unavailable on one system and then made available on a different system, the UID and GID values should be synchronized across the recovery domain for the device CRG. Administrative Domain can be used to synchronize user profiles, including the UID and GID values, across the cluster.

Using recommended structure for independent disk pools

The system disk pool and basic user disk pools (SYSBAS) should contain primarily operating system objects, licensed program libraries, and few user libraries. This structure yields the best possible protection and performance. Application data is isolated from unrelated faults and can also be processed independently of other system activity. Vary on and switchover times are optimized with this structure.

This recommended structure does not exclude other configurations. For example, you might start by migrating only a small portion of your data to a disk pool group and keeping the bulk of your data in SYSBAS. This is certainly supported. However, you should expect longer vary-on and switchover times with this configuration since additional processing is required to merge database cross-reference information into the disk pool group.

Specifying a recovery time for the independent disk pool

To improve vary on performance after an abnormal vary off, consider specifying a private customized access path recovery time specifically for that independent disk pool by using the Change Recovery for Access Paths (CHGRCYAP) command rather than relying upon the overall system-wide access path recovery time. This will limit the amount of time spent rebuilding access paths during the vary on.

Related information

Recommended structure for independent disk pools

Shortening planned outages

Planned outages are necessary and are expected; however, because they are planned does not mean they are nondisruptive. Planned outages are often related to system maintenance.

Clusters can effectively eliminate planned outages by providing application and data availability on a second system or partition during the planned outage.

Shortening backup windows

A key consideration of any backup strategy is determining your backup window, which is the amount of time that your system can be unavailable to users while you perform your backup operations. By reducing

the amount of time it takes to complete your backups, you can also reduce the amount of time your system is unavailable.

The challenge is to back everything up in the window of time that you have. To lessen the impact the backup window has on availability, you can reduce the amount of time a backup takes using one or more of the following techniques.

Improved tape technologies

Faster and denser tape technologies can reduce the total backup time. See Storage solutions for more information.

Parallel saves

Using multiple tape devices concurrently can reduce backup time by effectively multiplying the performance of a single device. See Saving to multiple devices to reduce your save window for more details on reducing your backup window.

Saving to nonremovable media

Saving to nonremovable media is faster than removable media. For example, directly saving to a disk unit can reduce the backup window. Data can be migrated to removable media at a later time. Saving from a detached mirror copy or flash copy, can also reduce the backup window. See Virtual tape media for more information.

Performing online backups

You can reduce your backup window by saving objects while they are still in use on the system or by conducting online backups.

Save-while-active

The save-while-active function is an option available through Backup, Recovery, and Media Services (BRMS) and on several save commands. Save-while-active can significantly reduce your backup window or eliminate it altogether. It allows you to save data on your system while applications are still in use without the need to place your system in a restricted state. Save-while-active creates a checkpoint of the data at the time the save operation is issued. It saves that version of the data while allowing other operations to continue.

Online backups

Another method of backing up objects while they are in use is known as an *online backup*. Online backups are similar to save-while-active backups, except that there are no checkpoints. This means that users can use the objects the whole time they are being backed up. BRMS supports the online backup of Lotus servers, such as Domino and QuickPlace. You can direct these online backups to a tape device, media library, save files, or a Tivoli® Storage Manager (TSM) server.

Note: It is important that you continue to back up system information in addition to any save-while-active or online backups you do. There is important system information that cannot be backed up using save-while-active or online backups.

Related information

Saving your system while it is active
Backup, Recovery, and Media Services (BRMS)

Backup from a second copy

You can reduce the backup window by conducting backups from a second copy of the data.

Note: If you are saving from a second copy, ensure that the contents of the copy are consistent. You may need to quiesce the application.

These techniques are as follows:

Saving from a second system

Some technologies, such as replication, make available a second copy of the data that can be used to shift the save window from the primary copy to the secondary copy. This technique can eliminate the backup window on the primary system since the backup processing is done on a second system.

System Storage FlashCopy

System Storage provides enhanced storage capabilities. This technology uses the FlashCopy® function on an independent disk pool. A single system storage server creates a point-in-time version of the independent disk pool where the host is not aware of the copy operation. The backup system can bring the copy online to perform a save operation or other processing.

The FlashCopy technology can be managed directly using System Storage commands, through the IBM PowerHA SystemMirror for i product, or through the IBM Lab Services Full System FlashCopy toolkit.

Related information

IBM System Storage and TotalStorage
Copy Services for IBM System i
High availability overview
High availability technologies
Implementing high availability
Simple Configuration Example for Storwize V7000 FlashCopy and PowerHA SystemMirror for i

Backing up less data

You can reduce your backup window by backing up smaller amounts of data at different times. There are several ways that you can isolate the data on your system for this type of operation.

Note: You can use Backup, Recovery, and Media Services (BRMS) to perform all of these kinds of backups.

Incremental backups

Incremental backups enable you to save changes to objects since the last time they were backed up. There are two types of incremental backups: cumulative and changes only. A *cumulative backup* specifies a backup that includes all changed objects and new objects since the last full backup. This is useful for objects that do not change often, or do not change greatly between the full backups. A *changes-only backup* includes all changed objects and new objects since the last incremental or full backup.

Incremental backups are especially useful for data that changes frequently. For example, you do a full backup every Saturday night. You have some libraries that are used extensively and you need to back them up more frequently than once a week. You can use incremental backups on the other nights of the week instead of doing a full backup to capture them. This can shorten your backup window while also ensuring that you have a backup of the latest version of those libraries.

Security data and configuration data

You can also reduce planned outages by isolating certain kinds of data and then backing it up separately. Security data includes user profiles, authorization lists, and authority holders. Configuration data includes information about how your system is set up, such as line descriptions, device descriptions, and configuration lists. These types of data are saved as part of a full system backup, but you can also save

them separately without having to shut your system down. For details on how to back up these types of data, see Manually saving parts of your system.

Omitting certain items

You can also reduce your backup window by reducing the number of objects you save or preventing objects from being saved more than once. You can do this by choosing to omit certain objects from a backup. For example, you might want to save all user libraries except for temporary libraries. You can choose to omit all temporary libraries from your backup, which will shorten the length of backup processing. Many of the commands that you use to save your system give you the option to omit items from the backup. For details on these commands, see Commands for saving parts of your system and Commands for saving specific object types. You can also use Backup, Recovery, and Media Services (BRMS) to omit items from a backup.

Data archiving

Any data that is not needed for normal production can be archived and taken offline. It is brought online only when needed, perhaps for month-end or quarter-end processing. The daily backup window is reduced since the archived data is not included.

Related information

Manually saving parts of your system
Commands for saving parts of your system
Commands for saving specific object types

Shortening software maintenance and upgrading windows

You can keep your system available by effectively planning and managing fixes and installing new releases.

Managing fixes

To reduce the amount of time your system is unavailable, you should ensure that you have a fix management strategy in place. If you stay current on what fixes are available and install them on a routine basis, you will have fewer problems. Be sure that you apply fixes as frequently as you have decided is appropriate for your business needs.

Individual fixes can be delayed or immediate. Delayed fixes can be loaded and applied in two separate steps. They can be loaded while your system runs and then applied the next time you restart your system. Immediate fixes do not require you to restart your system for them to take effect, which eliminates the need for downtime. Immediate fixes might have additional activation steps that are described in full in the cover letter that accompanies the fix.

Upgrading to a new release

You are required to turn off your system when you upgrade to a new version of IBM i. To minimize the amount of time a software upgrade takes, it is important that you plan your installation carefully. For example, in a cross site mirroring environment, the mirror copy can be upgraded first, then synchronize up and role swap. Then update the original production copy with the new release.

Related information

Planning your fix management strategy
Installing fixes
Upgrading or replacing IBM i and related software

Shortening hardware maintenance and upgrade windows

By effectively planning hardware maintenance and upgrades, you can greatly reduce and even eliminate the impact of these activities to the availability of your system.

There are times when you need to perform routine maintenance on your hardware or increase the capacity of your hardware. These operations can be disruptive to your business.

If you are performing a system upgrade, be sure that you plan carefully before you begin. The more carefully you plan for your new system, the faster the upgrade will go.

Concurrent maintenance

Many hardware components on the system can be replaced, added, or removed concurrently while the system is operating. For example, the ability to hot plug (to install a hardware component without turning off the system) is supported for Peripheral Component Interconnect (PCI) card slots, disk slots, redundant fans, and power supplies. Concurrent maintenance improves the availability of the system and allows you to perform certain upgrades, maintenance, or repairs without affecting the users of the system.

Capacity on Demand

With Capacity on Demand, you can activate additional processors and pay only for the new processing power as your needs grow. You can increase your processing capacity without disrupting any of your current operations.

Capacity on Demand is a feature that offers the capability to nondisruptively activate one or more central processors of your system. Capacity on Demand adds capacity in increments of one processor, up to the maximum number of standby processors built into your model. Capacity on Demand has significant value for installations where you want to upgrade without disruption.

Related information

Concurrent maintenance

Capacity on Demand

High availability

Whether you need continuous availability for your business applications or are looking to reduce the amount of time it takes to perform daily backups, IBM i high availability technologies provide the infrastructure and tools to help achieve your goals.

Most IBM i high availability solutions are built on IBM i cluster resource services, or more simply clusters. A *cluster* is a collection or group of multiple systems that work together as a single system. Clusters provides the underlying infrastructure that allows resilient resources, such as data, devices and applications, to be automatically or manually switched between systems. It provides failure detection and response, so that in the event of an outage, cluster resource service responds accordingly, keeping your data safe and your business operational.

The other key technology in IBM i high availability is independent disk pools. *Independent disk* pools are disk pools that can be brought online or taken offline without any dependencies on the rest of the storage on a system. When independent disk pools are a part of a cluster, the data that is stored within them can be switched to other systems or logical partitions. There are several different technologies, which can use independent disk pools, including switched disks, Geographic Mirror, Metro Mirror, and Global Mirror.

IBM PowerHA SystemMirror for i

IBM PowerHA SystemMirror for i is a licensed program that provides the following interfaces:

- PowerHA graphical interface in the IBM Navigator for i, which allows you to work with clusters and implement high availability.
- PowerHA commands for working with clusters, replication technologies, and administrative domains.

With the IBM PowerHA SystemMirror for i product, you can easily select, set up, and manage your high availability (HA) solution.

IBM PowerHA SystemMirror for i licensed program has been packaged into an express, standard, and enterprise editions. Below we have identified what is supported in each edition.

- IBM PowerHA for i Express Edition supports DS8000 Full System HyperSwap
- IBM PowerHA for i Standard Edition supports synchronous geographic mirroring, switched disk, LUN level switching, FlashCopy, and all express edition function.
- IBM PowerHA for i Enterprise Edition supports asynchronous geographic mirroring, Metro Mirror, Global Mirror, DS8000 HyperSwap with IASPs and all express and standard edition function.

Related information

High availability with IBM PowerHA
High availability overview
High availability technologies
Implementing high availability
IBM PowerHA SystemMirror for i interfaces

Related information for Availability roadmap

Product manuals, IBM Redbooks publications, websites, experience reports, and other information center topic collections contain information that relates to the Availability roadmap topic collection. You can view or print any of the PDF files.

Manuals

- Backup Recovery and Media Services for iSeries
- Recovering your system

IBM Redbooks

- i5/OS V5R4 Virtual Tape: A Guide to Planning and Implementation
- IBM i and IBM Storwize® Family: A Practical Guide to Usage Scenarios
- IBM i and IBM System Storage: A Guide to Implementing External Disks on IBM i
- Implementing PowerHA for IBM i
- Introduction to Storage Area Networks
- iSeries in Storage Area Networks: A Guide to Implementing FC Disk and Tape with iSeries
- Planning for IBM eServer i5 Data Protection with Auxiliary Write Cache Solutions
- PowerHA SystemMirror® for IBM i Cookbook
- Simple Configuration Example for Storwize V7000 FlashCopy and PowerHA SystemMirror for i

Websites

- Business continuity
- Guide to fixes
- IBM Backup Recovery and Media Services for i wiki
- IBM Capacity Backup for Power Systems

- IBM i
- IBM PowerHA SystemMirror for i
- IBM PowerHA SystemMirror for i wiki
- IBM Storage
- IBM Systems Lab Services and Training
- IBM Techdocs Library

  This site gives you access to the most current installation, planning and technical support information available from IBM pre-sales support, and is constantly updated. Look for the latest and greatest technical documents on high availability, iASPs, SAP, JD Edwards, etc.
- Performance Management on IBM i

Experience reports

- Reducing iSeries IPL Time

Other information

- Backup and recovery
- Backup, Recovery, and Media Services (BRMS)
- Capacity on Demand
- Commitment control
- High availability
- Disk management
- Journal management
- Logical partitions
- Storage solutions

Related reference

PDF file for Availability roadmap
You can view and print a PDF file of this information.

Notices

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

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

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

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

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

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

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

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

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

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

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

IBM Corporation
Software Interoperability Coordinator, Department YBWA
2800 37th Street NW
Rochester, MN 55901-4441
U.S.A.

© Copyright IBM Corp. 1998, 2022

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

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

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some 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 is for planning purposes only. The information herein is subject to change before the products described become available.

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

COPYRIGHT LICENSE:

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

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

© (your company name) (year). Portions of this code are derived from IBM Corp. Sample Programs.

© Copyright IBM Corp. _enter the year or years_.

Programming interface information

This Availability roadmap publication documents intended Programming Interfaces that allow the customer to write programs to obtain the services of IBM i.

Trademarks

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

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

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

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

Terms and conditions

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

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

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

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

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

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

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



Product Number: 5770-SS1