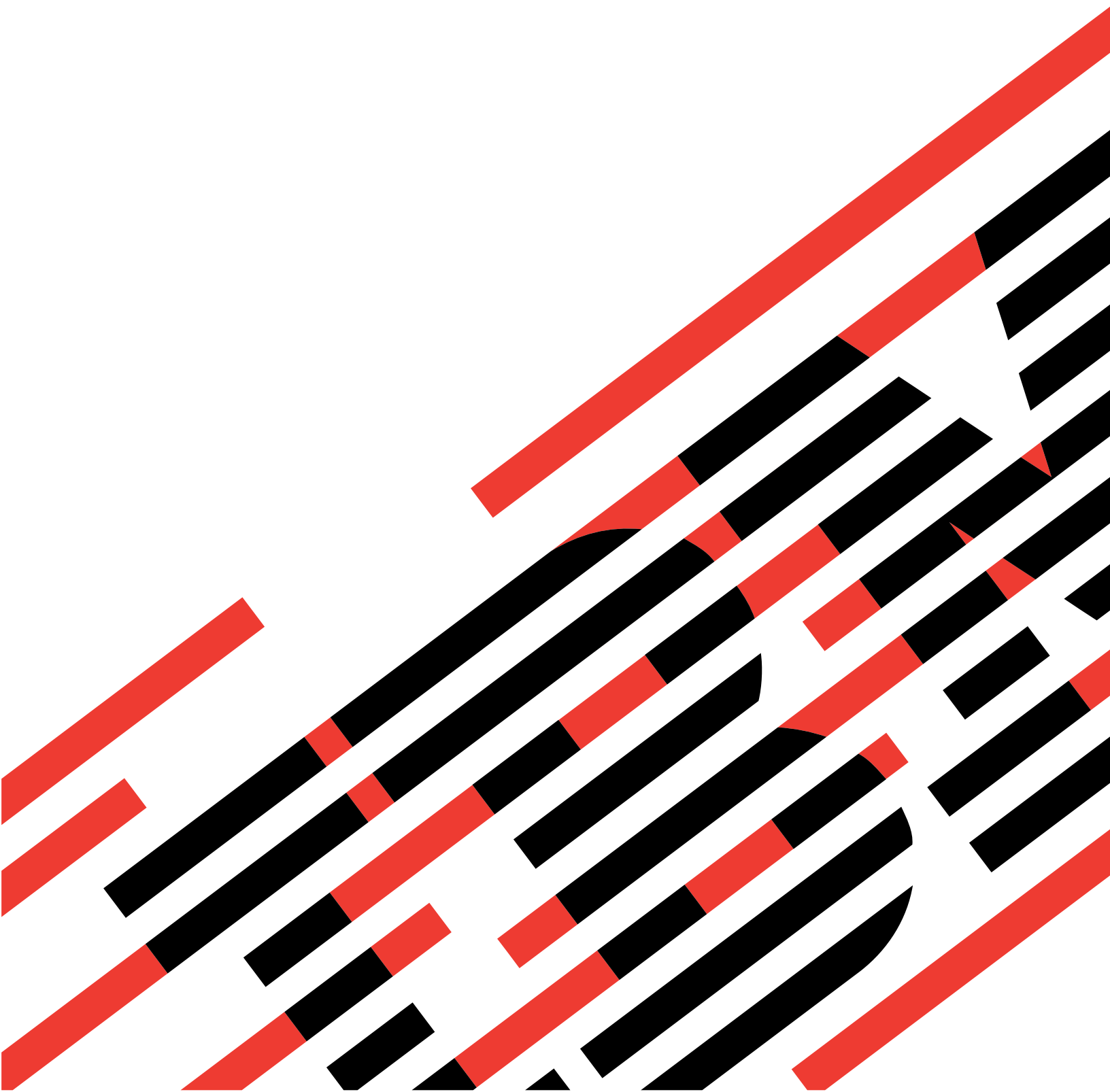




@server

iSeries

Availability roadmap for your iSeries server





@server

iSeries

Availability roadmap for your iSeries server

Contents

Availability roadmap for your iSeries server	1
Availability concepts	2
Estimate the value of availability	3
Decide what level of availability you need	5
Prevent unplanned outages	5
Prevent unplanned outages: Prepare for disk failures	5
Prevent unplanned outages: Plan for power loss	6
Prevent unplanned outages: Use effective systems management practices	7
Prevent unplanned outages: Prepare the space for your server	8
Shorten unplanned outages	8
Shorten unplanned outages: Reduce the time to restart your server	9
Shorten unplanned outages: Recover recent changes after an unplanned outage	9
Shorten unplanned outages: Recover lost data after an unplanned outage	10
Shorten planned outages.	11
Shorten planned outages: Shorten backup windows	11
Shorten planned outages: Shorten software maintenance and upgrade windows	14
Shorten planned outages: Shorten hardware maintenance and upgrade windows	14
Availability for multiple servers: Clusters	15
Highly available programs	15
Related information for the availability roadmap	16

Availability roadmap for your iSeries server

In today's fast-paced Internet environment, it is crucial that your data and applications be available to you when you need them. If your customers cannot access your Web site because your server is down, they may go to your competitors instead.

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 servers within the same company may have different availability needs. The purpose of this topic is to guide you through the world of iSeries availability and help you decide which availability tools are right for your business. 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, there are some things you need to understand. Read the following for more information:

Availability concepts

This topic contains definitions of the availability terms used throughout this topic.

Estimate the value of availability

This topic helps you evaluate how much an outage could cost your company.

Decide what level of availability you need

This topic helps you decide what level of availability is required for your company.

Once 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 your server. There are several different ways of approaching availability, based on your setup and the types of outages that you are anticipating, as follows:

Availability solutions for a single server

The iSeries has an incredible record for reliability. But, if you have very high availability needs for your server, there are some tools you can use to ensure that you meet your availability goals.

Prevent unplanned outages

This topic describes how to avoid the unplanned outages that you have some control over, and how to prepare for the ones you cannot control.

Shorten unplanned outages

This topic describes what you can do to ensure that unplanned outages, if they occur, are as short as possible.

Shorten planned outages

This topic describes how you can shorten the planned outages that you control.

Availability solutions using multiple servers

If you require a level of availability that is higher than what you can achieve with a single server, you should consider clusters. See the following for more information on clusters:

Availability for multiple servers: Clusters

This topic describes how clusters can help provide nearly 100 percent availability for your critical applications and data.

For detailed information on availability and the availability tools mentioned in this topic, see Related information.

Availability concepts

Before you can plan for the availability of your iSeries server, it is important for you to understand some of the concepts associated with this topic.

Availability is measured in terms of **outages**, which are periods of time when the server 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 server, or apply fixes.

An **unplanned outage** (also called an unscheduled outage) is usually 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.

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

There are several levels of availability. These levels differ in the type and duration of outages that they tolerate. These levels are as follows:

- **Highly available.** The server delivers an acceptable or agreed-upon level of service during its scheduled period of operation. The goal is to have the server available when the customer needs it.
- **High availability.** The server delivers an acceptable or agreed-upon level of service during its scheduled period of operation. The goal is to have no unplanned outages; there may be some planned outages.
- **Continuous operations.** The server delivers an acceptable or agreed-upon level of service 24 hours a day, 365 days a year. The goal is for the server to operate without any planned outages; there may be some unplanned outages.
- **Continuous availability.** The server delivers an acceptable or agreed-upon level of service 24 hours a day, 365 days a year. The goal is to have no planned or unplanned outages.

The following diagram shows how these different levels of availability relate to each other and what kinds of businesses are suited to each level.



*

For more information on these concepts and how they fit together, see the topics in Related information.

Estimate the value of availability

No one would argue the importance of availability. However, when asked to justify the cost of additional hardware to support improved availability, many people do not know how to build a case. The following steps describe how to estimate the value of your information services:

1. **Develop a list of the major services that your server provides.** Your server exists so that end users can accomplish tasks. To correctly assess the value of your server to the organization, you must identify what the server helps people do.
2. **Assess how much it costs you when these services are unavailable.** Each application or service has an affect on those who use your server. You need to determine which users are affected and how they are affected.
3. **Look at direct costs versus indirect costs.** Direct costs are losses that can be traced directly to a server 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 are those that 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.

For help in calculating how much downtime costs you, you can use the IT Cost of Downtime Calculator



Another helpful reference is *So you want to estimate the value of availability?* (GG22-9318). You can order this publication through the IBM Publications Center .

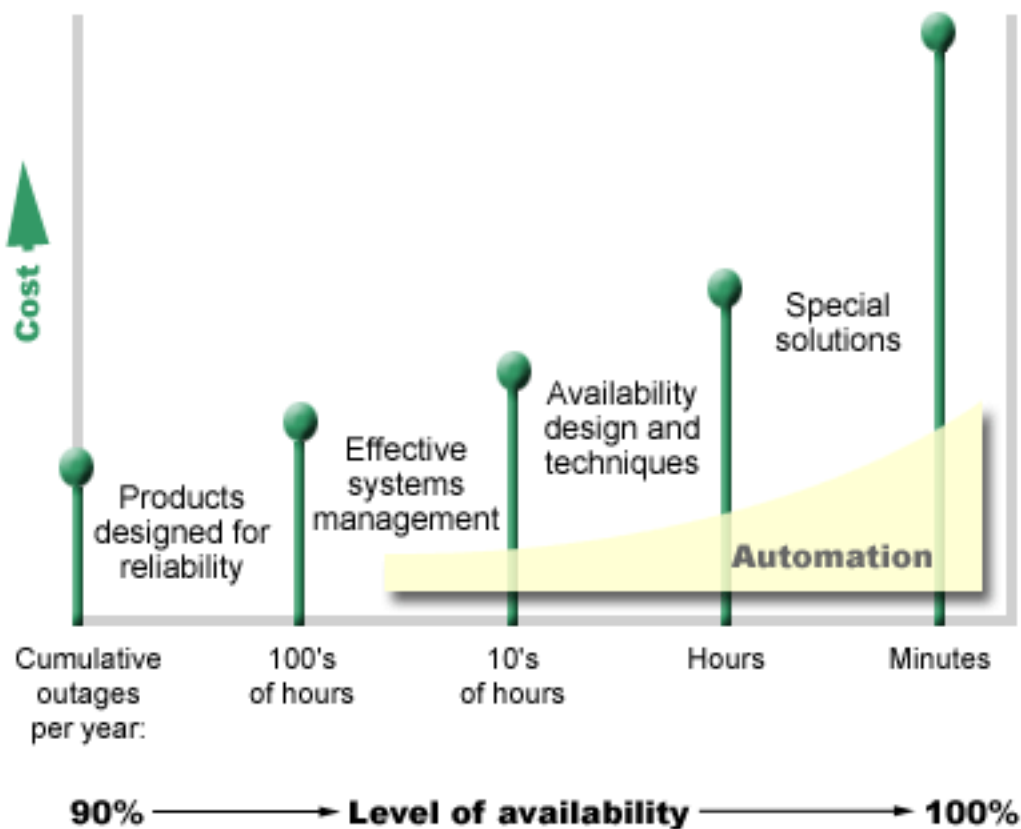
Decide what level of availability you need

Now that you understand some of the concepts behind availability and have figured out how much downtime costs you, you should also be aware that availability comes at a price. The higher the level of availability you need, the higher that price will be. So, 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 have to 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 have to pay for it.



How much downtime is acceptable to you?

It may help you to know what amount of downtime each level of availability represents. The following table shows the amount of downtime you should expect 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

Along with knowing how much downtime is acceptable to you, you need to consider how that downtime may occur. For example, you may 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 may think differently about 99% availability if the downtime were actually a single outage that lasts 3.65 days.

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

It used to be that customers accessed your business from 9 a.m. to 5 p.m., so it was realistic to expect that your server only had to be available during those hours. However, the Internet has changed that expectation; customers may expect to have access to your company's Web site at any time of the day or night. You have to 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.

Prevent unplanned outages

One way to approach availability is to try to prevent unplanned outages. This topic describes different methods that you can use to ensure that your server experiences as little unplanned downtime as possible.

To prevent unplanned outages, you should do the following:

Prepare for disk failures

Disk failure is rare, but it is something that you can prepare for. This topic describes how you can do that.

Plan for power loss

Power loss is something that is unplanned and unavoidable, but you can prepare for when it happens. This topic describes how to do that.

Use effective systems management practices

This topic describes how monitoring system performance and managing system operations can contribute to better overall availability.

Prepare the space for your server

The physical space where your server resides must be prepared carefully to ensure that the conditions promote optimum availability. This topic describes what the physical conditions should be like for your server.

If you have a single server and you cannot achieve the level of availability you need using these strategies, you may want to consider clusters. For more information, see [Availability for multiple servers: Clusters](#).

Prevent unplanned outages: Prepare for disk failures

Disk storage is the storage that is either internal to your iSeries server or is attached to it. This disk space, together with your server's main memory, is regarded by your server as one large storage area. When you save a file, you do not assign it to a storage location; instead, the server places the file in the location that ensures the best performance. It may spread the data in the file across multiple disk units, if that is the

best option. 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**.

Because your data is spread across your disks, it is important that you consider how to protect your data in the event that one of those disks fails. The purpose of this topic is to describe the methods you can use to protect your disks. For additional information on the methods described here, see [Disk protection](#).

Device parity protection

Device parity protection allows your server 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 server continues to run while the data is being reconstructed.

For details on device parity protection, see [Device parity protection](#).

Mirrored protection

Mirrored protection is one way to protect your data in the event of a disk failure. 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 may 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. You can duplicate:

- Disk units
- Disk controllers
- I/O bus unit
- Disk I/O processors
- A bus

For details on mirrored protection, including how it works and how to plan for it, see [Mirrored protection](#).

Independent disk pools

Independent disk pools (also called independent auxiliary storage pools) enable you to prevent unplanned outages because the data on them is isolated from the rest of your server. If an independent disk pool fails, your server can continue to operate. For detailed information on how to use independent disk pools, see [Independent disk pools](#).

Prevent unplanned outages: Plan for power loss

To ensure that your server 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 server is to ensure that you have an adequate power supply. You need to understand your server's requirements and then enlist the aid of a qualified electrician to help install the proper wiring. For details on how to ensure that your server has adequate power, see [Determining your power requirements](#).

Battery backups

Some iSeries servers come with battery backups. Your battery backup unit provides 30 seconds of runtime. If power is not restored within 30 seconds, the system immediately goes into a controlled shutdown.

Redundant power supplies

Some iSeries servers are available with redundant power supplies. A redundant power supply is a feature that prevents an unplanned outage by providing power if one power supply fails.

Uninterruptible power supplies


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

- Let you continue operations during brief power outages (brown outs).
- Protect the server from voltage peaks (white outs).
- Provide a normal end of operations, which can reduce your recovery time when you restart your server. For information on how to write a program that will help you control your server's shutdown in these conditions, see [Control server shutdown using a power handling program](#).

For information on which uninterruptible power supplies are compatible with your iSeries server, see [Uninterruptible power supply systems](#).

Generator power

If you think you could experience an extended power failure, you might want to consider purchasing a generator. A generator goes a step further than a UPS in that it enables you to continue normal operations during longer power failures.


If you need help planning for the power needs of your iSeries server, see [Power Protection Services](#) .

Prevent unplanned outages: Use 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 server running smoothly. This includes performing basic preventive maintenance and systems management tasks that help your server perform at its peak. Many of these systems management tasks can be automated, which helps you prevent failures that may occur because of human error or an oversight.

One way you can help ensure the availability of your server 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 server. You can be notified of any problems that jeopardize the availability of your server in time to react and prevent an unplanned outage. For more information on how to plan for and manage the performance of your server, see [Performance](#).

Fixes are also an important systems management component that can help you keep your server available. When problems are discovered in iSeries programs, IBM issues a **fix** (also known as a PTF, or program temporary fix) to correct the problem. You need to be aware of fixes and install them on your server to ensure that your server 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 server. For

more information on how to obtain and apply fixes, refer to Use software fixes. For help in determining a strategy for preventive maintenance based on your iSeries environment and applications, try the Fix Maintenance Advisor  .

Prevent unplanned outages: Prepare the space for your server

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

The first thing you need to do is familiarize yourself with your server. Different server models have different requirements with regard to the conditions they are in, so you should be aware of what your server needs. For more information on the physical characteristics of each model, see Server specifications. For information on the physical characteristics of the hardware that you can attach to your server, see Hardware specification sheets.

Once you are familiar with the physical characteristics of your server, be sure to consider the following about the space where it resides:

- **Location.** The physical location of your server can have an impact on how available it is. For example, if the room is not secure, your server could be vulnerable to vandalism or even someone accidentally dislodging the power cord. For more details on what precautions you should take with regard to your server's location, see Location considerations.
- **Cables.** Cables are often overlooked, but without them your server would be unusable. You need to ensure that your cables are in good condition and are being used correctly. For more information on how to ensure that your cables do the job you need them to do, see General cabling considerations.
- **Environment.** The environment you provide for your server is also crucial to availability. The environment includes such things as temperature, humidity, and other factors that can inhibit the performance of your server. For more detailed information on the environment you need to provide for your server, see Environment reference.

Shorten unplanned outages

Unplanned outages do occur, and a key to availability is to ensure that when they do occur you can recover from them as quickly as possible. The purpose of all recovery strategies is to get back to where you were before the outage occurred.

The following topics describe what availability tools help you restart your server and recover data as quickly as possible after an outage occurs:

Restart the server

This topic describes what availability tools you can use to ensure that your server restarts as quickly as possible after an outage.

Recover recent changes

This topic describes what availability tools you can use to recover from an outage when you lose only those changes made shortly before the outage.

Recover lost data

This topic describes what availability tools you can use to recover from an outage that causes you to lose data. The loss may be a few files or can be the entire site, such as during a natural disaster.

If you have a single server and you cannot achieve the level of availability you need using these strategies, you may want to consider clusters. For more information, see Availability for multiple servers: Clusters.

Shorten unplanned outages: Reduce the time to restart your server

Before your server powers down, 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 server cannot perform these activities. For more details on what happens when your server ends abnormally, see [Start and stop the iSeries](#).

The purpose of this topic is to describe some availability tools that will help your server restart as quickly as possible.

System-managed access-path protection (SMAPP)

An **access path** is the route 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 server ends abnormally, such as during an unplanned outage, the server must rebuild the access paths the next time it starts, which can take a long time. When you use system-managed access-path protection, the server protects the access paths so they do not have to be rebuilt when your server starts after an unplanned outage. This will save you time when you restart your server, which will enable you to get back to your normal business processes as quickly as possible. For detailed information on SMAPP, see [System-managed access-path protection](#).

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 server. However, when you use SMAPP, the *server decides* which access paths to protect. Therefore, if the server does not protect an access path that you consider critical, you may be delayed in getting your business running again. When you journal access paths, *you decide* which paths to journal. For more details on journaling access paths, see [SMAPP and access path journaling](#).

Shorten unplanned outages: Recover recent changes after an unplanned outage

After an unplanned outage, your goal is to get your server 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 re-enter transactions. This may involve rebuilding some of your data. There are a few availability tools that you can use that will help you more quickly get back to where you were before the outage occurred.

Journaling

Journal management prevents transactions from being lost if your server ends abnormally. When you journal an object, the server keeps a record of the changes you make to that object. For detailed information on how to plan for and use journaling, see [Journal management](#).

Commitment control

Commitment control helps to provide data integrity on your server. It allows you to define and process a group of changes to resources, such as database files or tables, as a single transaction. Then, it ensures that either the entire group of individual changes occur or that none of the changes occur. 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 corrupt data. With commitment control, the incomplete updates would be backed out of your database when you restart your server.

You can use commitment control to design an application so 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 due to incomplete transactions from a prior failure.

For detailed information on how to plan for and use commitment control, see Commitment control.

Shorten unplanned outages: Recover lost data after an unplanned outage

You may 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 flood or tornado. 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 server; the time and money you spend creating this strategy is more than recovered should you need to restore lost data or perform a recovery. Once 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 server, you need to assess whether your backup and recovery strategy needs to change.


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

For guidance on how to create a backup and recovery strategy, see Plan a backup and recovery strategy. For instructions on how to perform backups on your server, see Back up your server. For information on how to restore your server, see Recover your server.

If you would like a solution to help you 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 retrieve 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 server fully in the event of 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 have to keep track of which backup items are on which volumes, and worry that you will accidentally write over active data. You can also track the movement of your media to and from offsite locations.

For detailed information on 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, contact IBM Business Continuity and Recovery Services .

Limit 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 have to recover the data that was stored in the disk pool that the failed disk unit was a part of.

For detailed information on disk pools, disk pool types, and examples of how to use disk pools for different purposes, see [Disk pools](#). For information on how to configure disk units and disk pools, see [Manage 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. 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. For detailed information, see [Independent disk pools](#).

The iSeries server also offers you the ability to divide one server into several independent servers. This technology is called **logical partitions**. 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 server, 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. Then, if another partition fails, that program is protected.

While logical partitions have many advantages, there are also some drawbacks that you have to consider. Logical partitions can make a full system recovery more complicated, which lengthens the amount of time that your server is unavailable.

For detailed information on logical partitions and how to use them, see [Logical partitions](#).

Shorten 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. The following topics describe the ways you can reduce outages that result from different types of maintenance:

Shorten backup windows

This topic includes information on how you can reduce the amount of time that your server is unavailable while you do backups.

Shorten software maintenance and upgrade windows

Time that you spend doing software maintenance and upgrades is time that your server is unavailable to your users. Learn how to shorten these times.

Shorten hardware maintenance and upgrade windows

Time that you spend doing hardware maintenance and upgrades is time that your server is unavailable to your users. Learn how to shorten these times.

If you have a single server and you cannot achieve the level of availability you need using these strategies, you may want to consider clusters. For more information, see [Availability for multiple servers: Clusters](#).

Shorten planned outages: Shorten backup windows

One way to reduce the amount of time that your server is unavailable is to reduce the amount of time needed for your backups. There are several things you can do to reduce the amount of time your backups take, including the following:

Improve tape performance

If you are using tape for your backups, it is logical that you can reduce your backup window by increasing the speed at which you can do your backups. This topic describes your options for doing that.

Perform online backups

There are certain objects and types of data that you can back up while they are in use, with little or no disruption to server operations.

Back up less data

You can also reduce your backup window by having less data to save when that window of time comes. You can back up some objects at different times during the week, which reduces the number of objects you have to save during your window.

Shorten backup windows: Improve tape performance

This article will describe how customers can reduce their backup windows by increasing the speed and capacity of the media and hardware they use.


High-performance hardware

One way to reduce your backup window is to use tape hardware that provides very high performance. It is logical that the faster you can back up data, the shorter your backup window is. One example of this high-performance hardware is the IBM TotalStorage^(TM) Enterprise Tape System 3590. For more

information on these high-speed tape drives, see [IBM TotalStorage Enterprise Tape System 3590](#) .

Another tape solution that can help you reduce your backup window is Linear Tape-Open (LTO). This technology offers performance that is comparable to the 3590 at a price that is likely to be more affordable for smaller enterprises. For more details on LTO and how you can use it with your iSeries, see [Linear](#)

[Tape-Open](#) .

For details on the rates of backup devices that are supported on your iSeries server and tips for improving performance, see the [Performance Capabilities Reference](#) .

Concurrent and parallel backups

Another way you can use tape hardware to improve the speed of your backups is to have multiple tape devices doing concurrent and parallel backups. These types of backups can considerably reduce your backup window and streamline your backup operations.

When you do a **concurrent backup**, you send multiple save jobs to multiple tape devices to process at the same time. For example, using a concurrent backup strategy, you might send one library or group of libraries to one tape device, and another set of libraries to a different tape device. Anything that you can save concurrently, you can restore concurrently, thereby reducing your downtime after a system failure by recovering multiple libraries or objects at the same time.

When you do a **parallel backup**, you back up items to two or more devices at one time. There are two methods of parallel backups: parallel save/restore support and multiple-library parallel support. Parallel save/restore support spreads each object across multiple resources. Multiple-library support spreads libraries across multiple resources, such that each library is backed up to a single resource. You can use both of these methods to optimize the performance of your devices while keeping your resources balanced.

For more information on concurrent and parallel backups, see [Save to multiple devices to reduce your save window](#). You can also use Backup, Recovery and Media Services (BRMS) to perform these kinds of backups.

Automation

Not surprisingly, the backup window for most companies occurs during the night, when personnel are not normally on hand to perform backup operations. There are a few ways that you can automate your processes to both reduce the length of your backup window and to make human intervention unnecessary. This automation cuts costs and also helps prevent the possibility of human error or delay.

Backup, Recovery and Media Services (BRMS) enables you to schedule operations using a **job scheduler**, such as the Management Central job scheduler. For example, you could schedule an unattended full backup every Saturday night and then schedule maintenance tasks immediately following that backup. You can also schedule the movement of your backup media to offsite locations and keep track of its whereabouts during the move.

You can also use **media libraries** to improve backup processing. These devices hold a large number of volumes and make automated backups much easier because they change the tapes in the drive for you. There is no unnecessary time lost in removing and inserting volumes as the backup proceeds. For more information on how you can use tape libraries to improve your backup processing, see Manage tape libraries.

Shorten backup windows: Perform online backups

You can reduce the length of your planned outages by saving some objects while they are active.

This method of backing up objects is called **save-while-active** support. Save-while-active enables you to perform save operations with limited disruption to jobs and subsystems. When you use save-while-active, objects are unavailable for a short time at the beginning of the backup process until they reach a certain point, called a checkpoint. Once the backup reaches that checkpoint, users can work with those objects again. Save-while-active is a useful operation if you want to reduce your planned outage time. However, you should plan for it carefully and be aware of any restrictions that might affect your backup processing. For detailed information on save-while-active, see Save your server while it is active.

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. Backup, Recovery, and Media Services (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. For detailed information on Lotus online backups, see Backup, Recovery and Media Services.

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.

Shorten backup windows: Back up less data

This topic describes how you can reduce planned outages by backing up smaller amounts of data at different times. There are a few different ways that you can isolate the data on your server for this type of operation. You can use Backup, Recovery and Media Services 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. **Cumulative** backups save the changes to an object since the last full backup of that object. This is useful for objects that do not change very often, or do not change greatly between full backups. **Changes-only** backups save the changes to an object since the last time that object was backed up, regardless of whether that backup was a full backup or an incremental 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 so 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 will 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 server 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 server down. For details on how to back up these types of data, see [Manually save parts of your server](#).

Omit 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 may 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 server give you the option to omit items from the backup. For details on these commands, see [Commands to save parts of your server](#) and [Commands to save specific object types](#). You can also use Backup, Recovery and Media Services to omit items from a backup.

Shorten planned outages: Shorten software maintenance and upgrade windows

One way to ensure that your server stays available is to keep your software current. This process takes time and requires planning. The purpose of this topic is to describe how you can keep your server available by managing fixes and installing new releases.

Manage fixes

To reduce the amount of time your server 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. For recommendations on how to create a fix management strategy, see [Plan your fix management strategy](#).

Individual fixes can be **delayed** or **immediate**. Delayed fixes can be loaded and applied in two separate steps. They can be loaded while your server runs and then applied the next time you restart your server. Immediate fixes do not require you to restart your server in order for them to take effect, which eliminates the need for downtime. Immediate fixes may have additional activation steps that are described in full in the cover letter that accompanies the fix. For more information on how to apply fixes, see [Install fixes](#).

Install new releases

You are required to power down your server when you install a new version of OS/400. To minimize the amount of time a software upgrade takes, it is important that you plan your installation carefully. For information on the planning process, see [Plan to install the OS/400 release](#). For a checklist of the planning steps, see [Server planning: Software](#).

Shorten planned outages: Shorten hardware maintenance and upgrade windows

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. But, with some planning, you can greatly reduce or even eliminate some of these outages.

If you are performing a system upgrade, be sure that you do careful planning before you begin. The more carefully you plan for your new server, the faster the upgrade will go. For more details on the planning process, see [Upgrades](#).


Concurrent maintenance

Concurrent maintenance is a feature of the iSeries server that enables a service representative to repair or replace certain hardware components while your server is being used for normal operations. If you need to replace or upgrade an eligible component, you would be able to do so without disrupting your business.

Capacity upgrade on demand

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

For more information on capacity upgrade on demand, see [Capacity Upgrade on Demand](#) .


Availability for multiple servers: Clusters

The primary availability strategy for a multiple system environment is clusters. A **cluster** is a collection or group of multiple iSeries servers that work together as a single server. If your business needs high or continuous availability, clusters is the solution that you should consider.

Servers in a cluster work cooperatively to provide a single computing solution. You can have up to 128 servers in a cluster. This allows you to efficiently group your iSeries servers together to set up an environment that provides availability that approaches 100 percent for your critical applications and your critical data. This helps ensure that your critical servers and applications are always available. Clusters also provide simplified systems management and increased scalability to seamlessly add new components as your business grows.

While the benefits of clusters are numerous, the cost is significant. You have to weigh the cost of this solution against the cost of downtime on your server to decide whether to implement clusters in your business. For information on how to determine the cost of downtime in your business, see [Estimate the value of availability](#).


If you do choose to use clusters in your environment, it is important that you consider the types of applications that you use. There are some applications that are designed to withstand some of the effects of a failure. For more information on these applications, see [“Highly available programs”](#).

To learn more about clusters, including how it works and how to implement clusters in your business, see [Clusters](#). For more information on clusters in the context of other high availability solutions, see [High availability and clusters](#) .

Highly available programs

Your applications and your data are critical to your business. If you are using clusters, there are programs that you can use that are resilient in a system outage. You can design these applications yourself, but you can also purchase applications that meet the necessary criteria. If you want to design the programs

yourself, you need to understand what a resilient program is and what the levels of application availability are. For more information on these topics, see Cluster applications.

If you purchase applications, you can also be sure that they are highly available. ClusterProven^(R) is an IBM brand that identifies these kinds of applications. An application that is ClusterProven for iSeries continues to be available in the event of an outage (planned or unplanned). For more information on the criteria these programs must meet and a list of the programs that are ClusterProven, see High Availability and Clusters  .



Related information for the availability roadmap

Listed below are the iSeries manuals and IBM Redbooks^(TM) (in PDF format), and Web sites that relate to the Availability roadmap topic. You can view or print any of the PDFs.







iSeries Information Center topics

- Backup and Recovery
- Clusters
- Commitment control
- Disk management
- Journal management
- Logical partitions
- Storage solutions





Manuals




- Backup and Recovery 
- Backup, Recovery and Media Services for iSeries 

Redbooks

- Roadmap to Availability on the iSeries 400 
- High Availability on the AS/400 System: A System Manager's Guide 
- The System Administrator's Companion to AS/400 Availability and Recovery 
- Clustering and IASPs for Higher Availability 
- Striving for Optimal Journal Performance on DB2 Universal Database for iSeries 
- AS/400 Remote Journal Function for High Availability and Data Replication 


Web sites

- High availability and clusters 
- Backup, Recovery and Media Services 
- IBM Business Continuity and Recovery Services 
- IT Cost of Downtime Calculator 

- Logical Partitioning 
- Performance Capabilities Reference 
- Tape and Optical Storage 

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

1. Open the PDF in your browser (click the link above).
2. In the menu of your browser, click **File**.
3. Click **Save As...**
4. Navigate to the directory in which you would like to save the PDF.
5. Click **Save**.

If you need Adobe Acrobat Reader to view or print these PDFs, you can download a copy from the Adobe Web site (www.adobe.com/prodindex/acrobat/readstep.html) .



Printed in U.S.A.