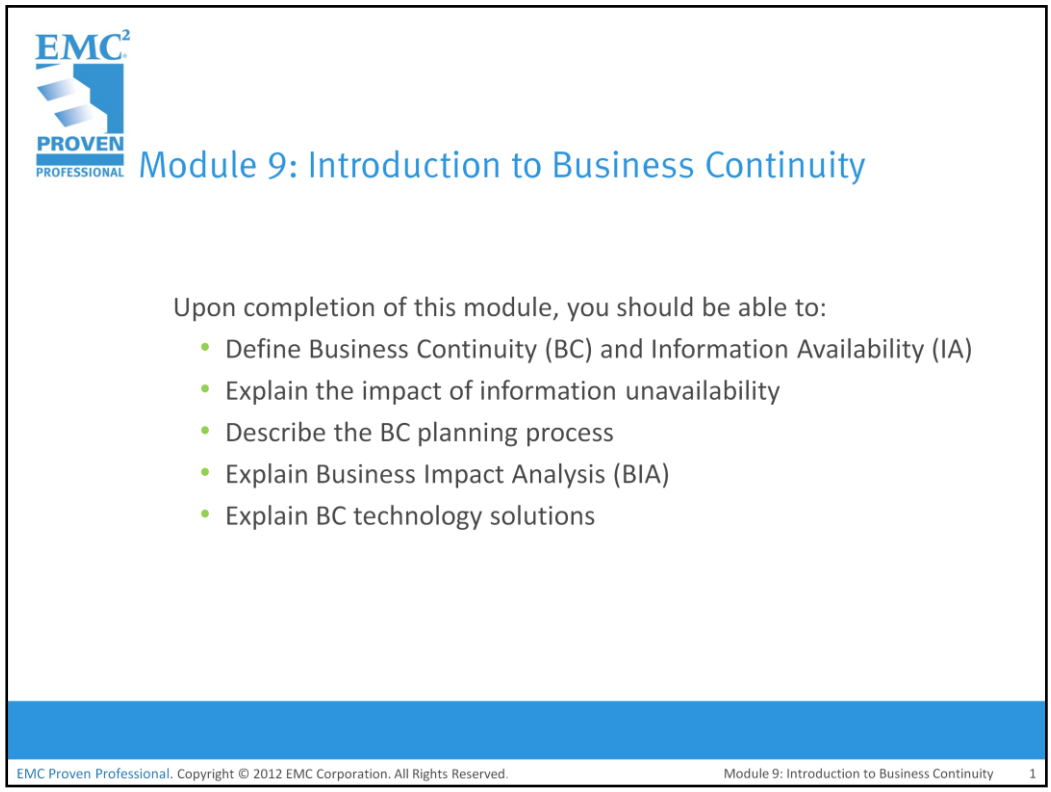
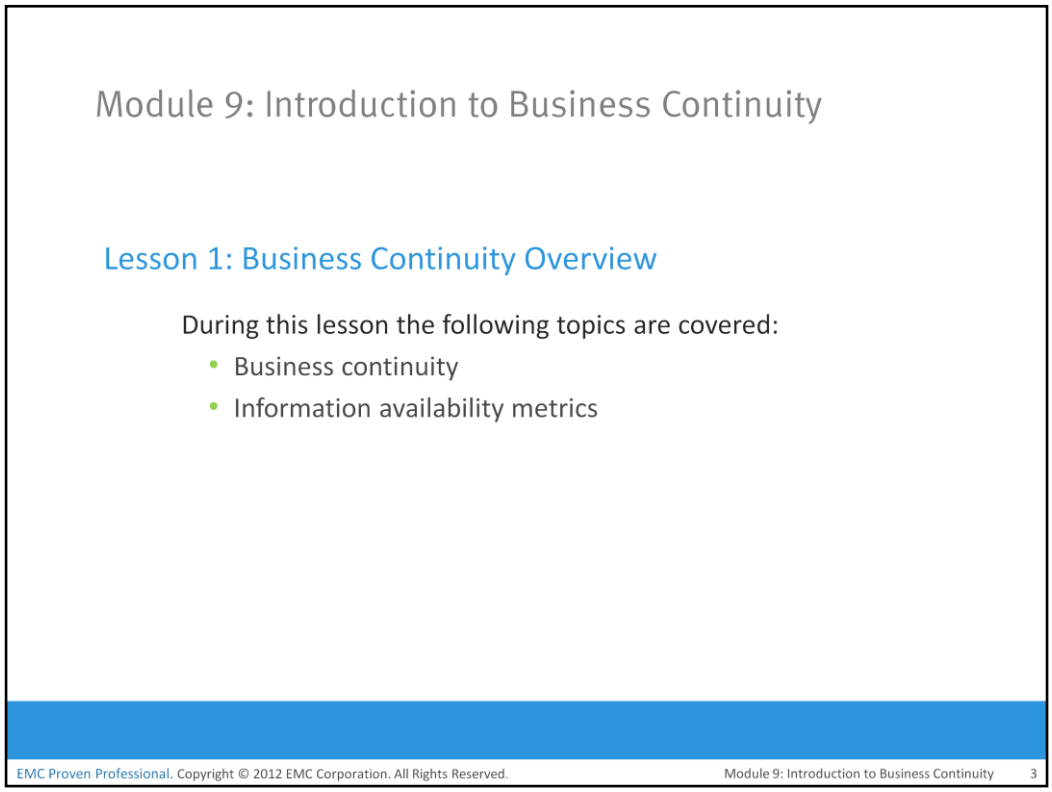
Module – 9

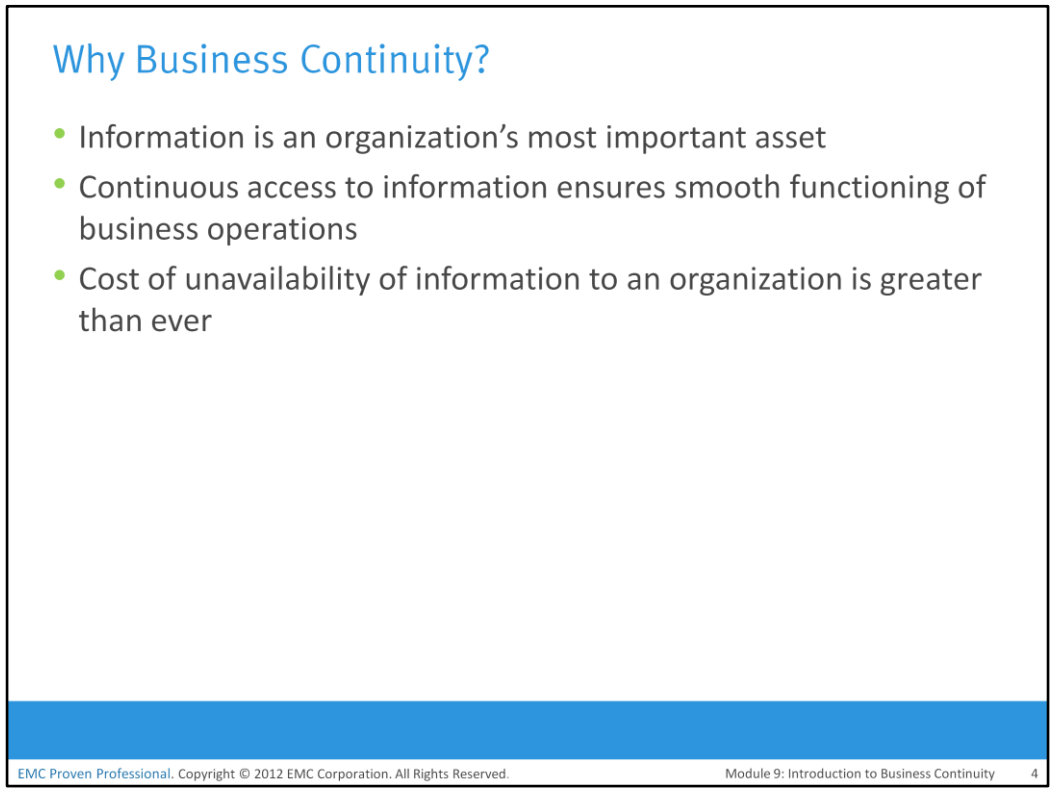
Introduction to Business Continuity



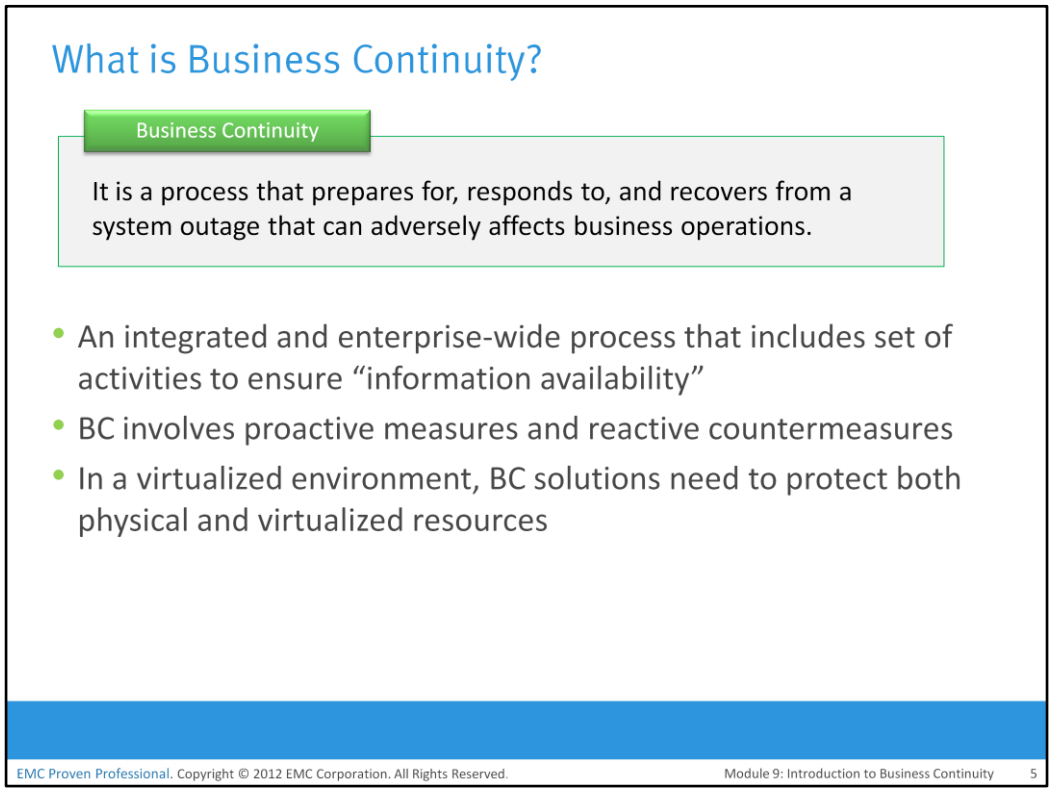
This module focuses on the importance of Business Continuity, the factors that can affect Information Availability, and the consequences of information unavailability. This module also details on the BC planning process and BC technology solutions, specifically on eliminating single points of failure.



This lesson covers the importance of business continuity to an organization, factors that can affect information availability and the consequences of information unavailability. Also this lesson focuses on information availability metrics namely mean time between failure (MTBF) and mean time to repair (MTTR).

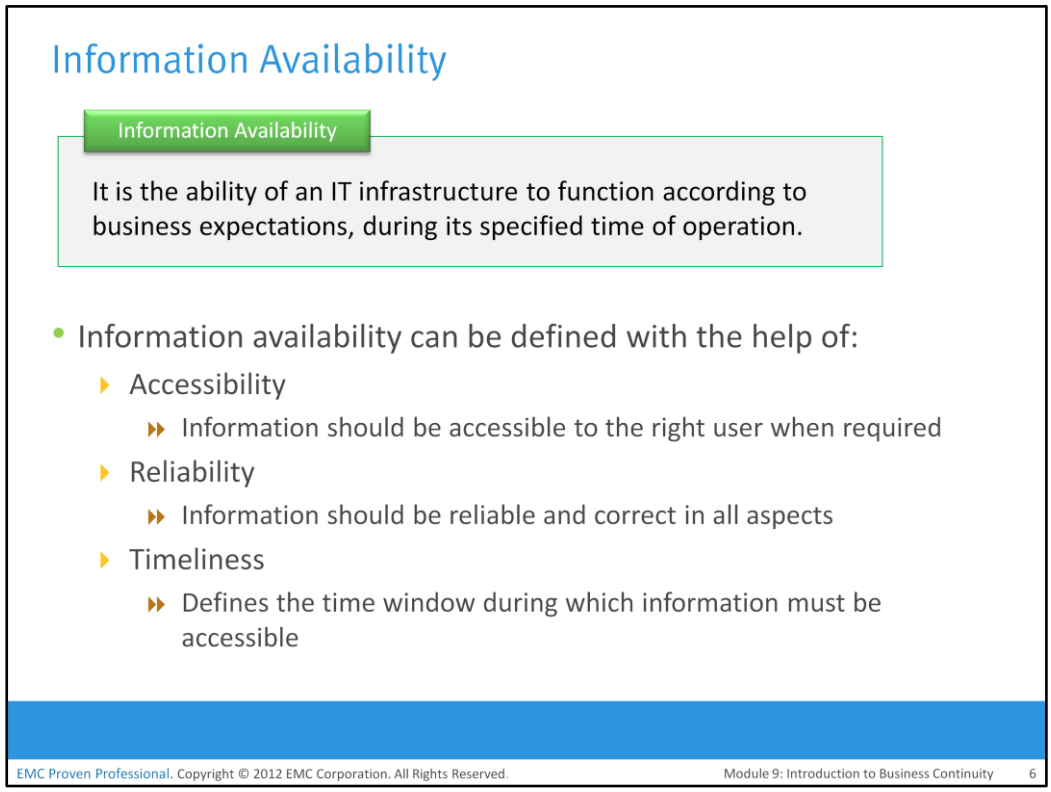


In today's world, continuous access to information is a must for the smooth functioning of business operations. The cost of unavailability of information is greater than ever, and outages in key industries cost millions of dollars per hour. There are many threats to information availability, such as natural disasters, unplanned occurrences, and planned occurrences, that could result in the inaccessibility of information. Therefore it is critical for businesses to define appropriate strategies that can help them to overcome these crises. Business continuity is an important process to define and implement these strategies.



Business continuity (BC) is an integrated and enterprise-wide process that includes all activities (internal and external to IT) that a business must perform to mitigate the impact of planned and unplanned downtime. BC entails preparing for, responding to, and recovering from a system outage that adversely affects business operations. It involves proactive measures, such as business impact analysis, risk assessments, BC technology solutions deployment (backup and replication), and reactive measures, such as disaster recovery and restart, to be invoked in the event of a failure. The goal of a BC solution is to ensure the "information availability" required to conduct vital business operations.

In a virtualized environment, BC technology solutions need to protect both physical and virtualized resources. Virtualization considerably simplifies the implementation of BC strategy and solutions.

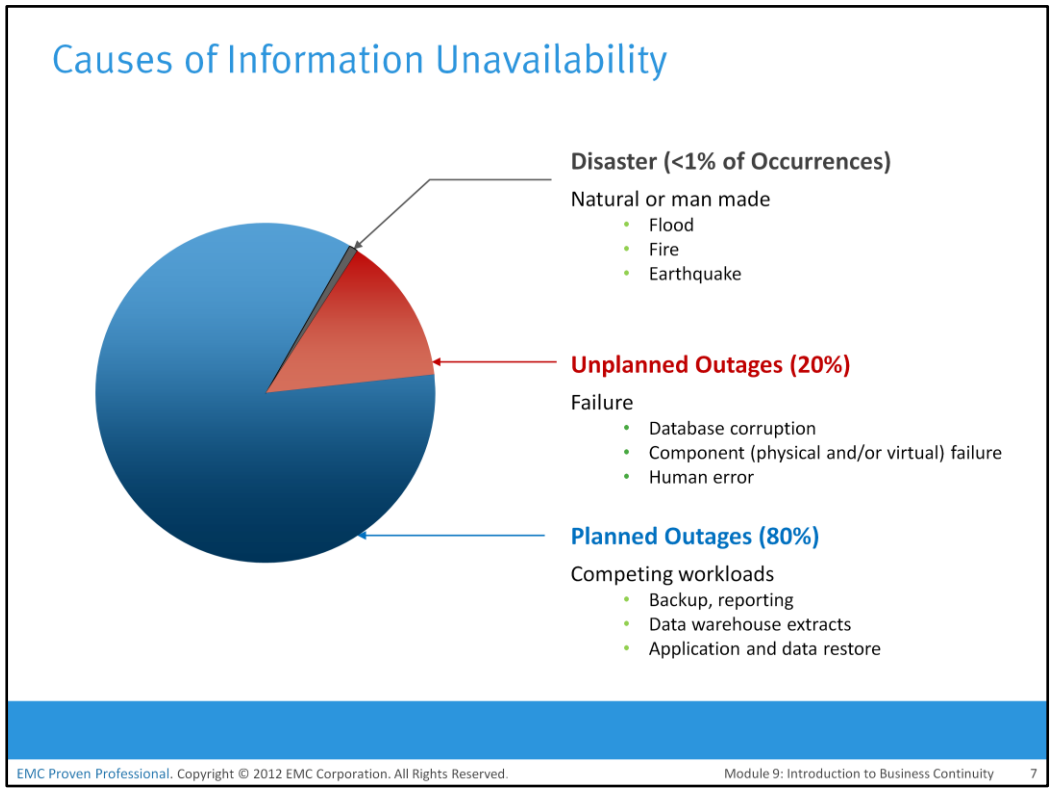


Information availability (IA) refers to the ability of an IT infrastructure to function according to business expectations during its specified time of operation. IA ensures that people (employees, customers, suppliers, and partners) can access information whenever they need it. IA can be defined in terms of the accessibility, reliability, and timeliness of the information.

Accessibility: Information should be accessible to the right user when required.

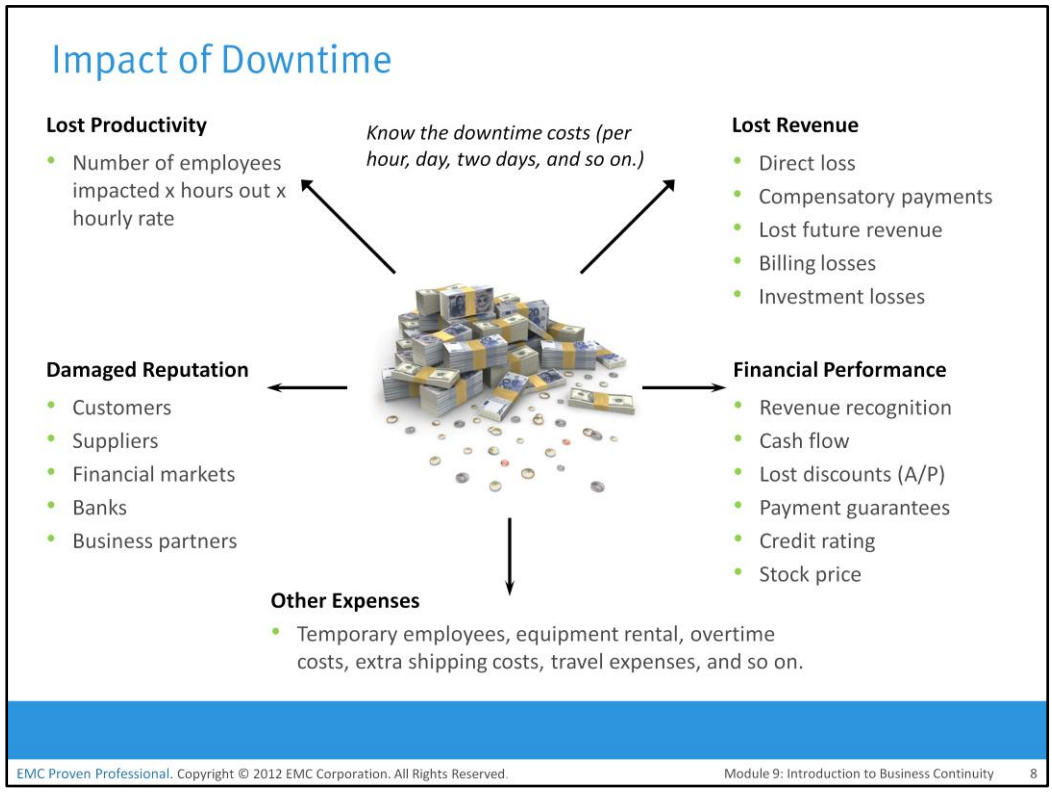
Reliability: Information should be reliable and correct in all aspects. It is "the same" as what was stored and there is no alternation or corruption to the information.

Timeliness: Defines the time window (a particular time of the day, week, month, and year as specified) during which information must be accessible. For example, if online access to an application is required between 8:00 am and 10:00 pm each day, any disruptions to data availability outside of this time slot are not considered to affect timeliness.



Various planned and unplanned incidents result in information unavailability. *Planned outages* include installation/integration/maintenance of new hardware, software upgrades or patches, taking backups, application and data restores, facility operations (renovation and construction), and refresh/migration of the testing to the production environment. *Unplanned outages* include failure caused by human errors, database corruption, and failure of physical and virtual components.

Another type of incident that may cause data unavailability is natural or man-made disasters, such as flood, fire, earthquake, and so on. As illustrated in figure on the slide, the majority of outages are planned. Planned outages are expected and scheduled but still cause data to be unavailable. Statistically, the cause of information unavailability due to unforeseen disasters is less than 1 percent.



Information unavailability or downtime results in loss of productivity, loss of revenue, poor financial performance, and damages to reputation. Loss of productivity include reduced output per unit of labor, equipment, and capital. Loss of revenue includes direct loss, compensatory payments, future revenue loss, billing loss, and investment loss. Poor financial performance affects revenue recognition, cash flow, discounts, payment guarantees, credit rating, and stock price. Damages to reputations may result in a loss of confidence or credibility with customers, suppliers, financial markets, banks, and business partners. Other possible consequences of downtime include the cost of additional equipment rental, overtime, and extra shipping.

The business impact of downtime is the sum of all losses sustained as a result of a given disruption. An important metric, *average cost of downtime per hour*, provides a key estimate in determining the appropriate BC solutions. It is calculated as follows:

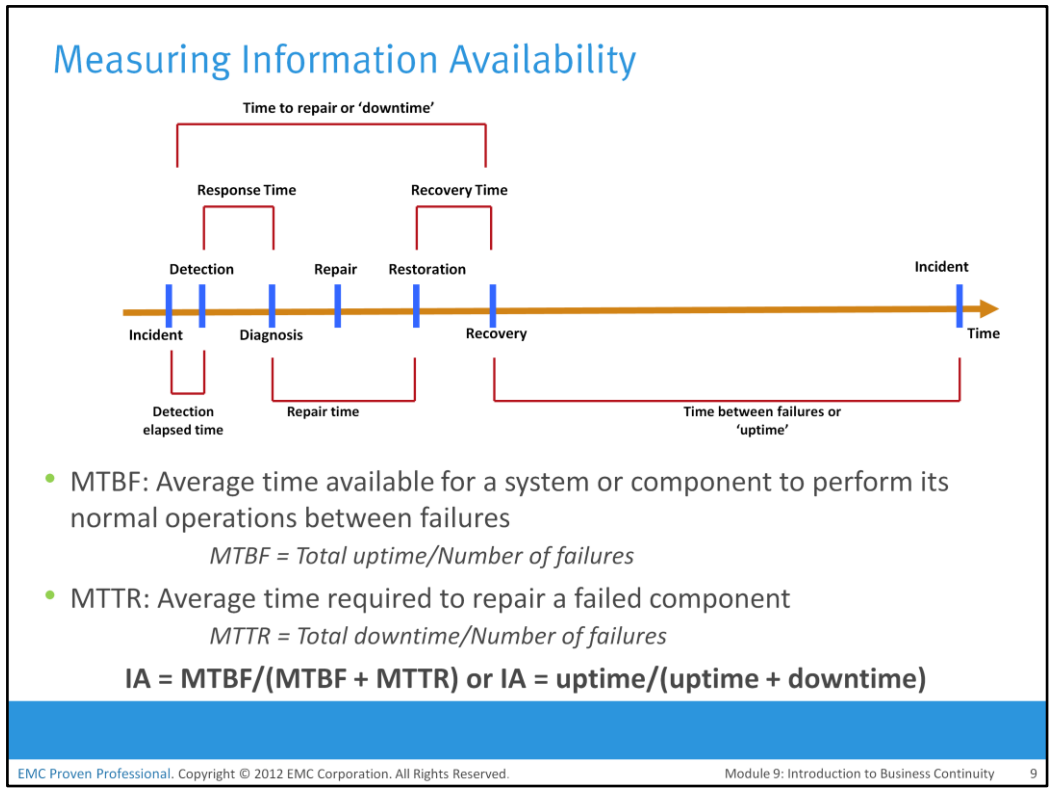
Average cost of downtime per hour = average productivity loss per hour + average revenue loss per hour

Where:

Productivity loss per hour = (total salaries and benefits of all employees per week) / (average number of working hours per week)

Average revenue loss per hour = (total revenue of an organization per week) / (average number of hours per week that an organization is open for business)

The average downtime cost per hour may also include estimates of projected revenue loss due to other consequences, such as damaged reputations, and the additional cost of repairing the system.



Information availability relies on the availability of both physical and virtual components of a data center. Failure of these components might disrupt information availability. A failure is the termination of a component's ability to perform a required function. The component's ability can be restored by performing an external corrective actions, such as a manual reboot, a repair, or replacement of the failed component(s). Proactive risk analysis, performed as part of the BC planning process, considers the component failure rate and average repair time, which are measured by MTBF and MTTR:

Mean Time Between Failure (MTBF): It is the average time available for a system or component to perform its normal operations between failures. It is the measure of system or component reliability and is usually expressed in hours.

Mean Time To Repair (MTTR): It is the average time required to repair a failed component. MTTR includes the total time required to do the following activities: detect the fault, mobilize the maintenance team, diagnose the fault, obtain the spare parts, repair, test, and restore the data. MTTR is calculated as: Total downtime/Number of failures

IA can be expressed in terms of system uptime and downtime and measured as the amount or percentage of system uptime:

IA = system uptime / (system uptime + system downtime)

Where *system uptime* is the period of time during which the system is in an accessible state; when it is not accessible, it is termed as *system downtime*.

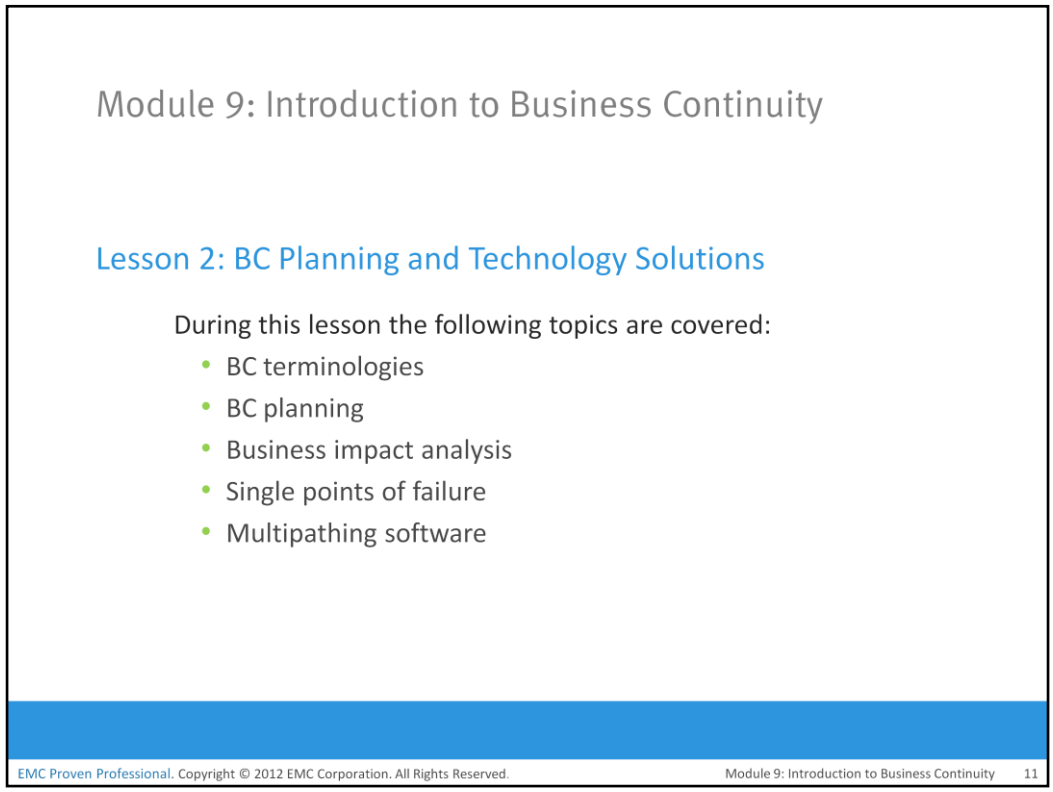
In terms of MTBF and MTTR, IA could also be expressed as: IA = MTBF / (MTBF + MTTR)

Availability Measurement – Levels of '9s' Availability

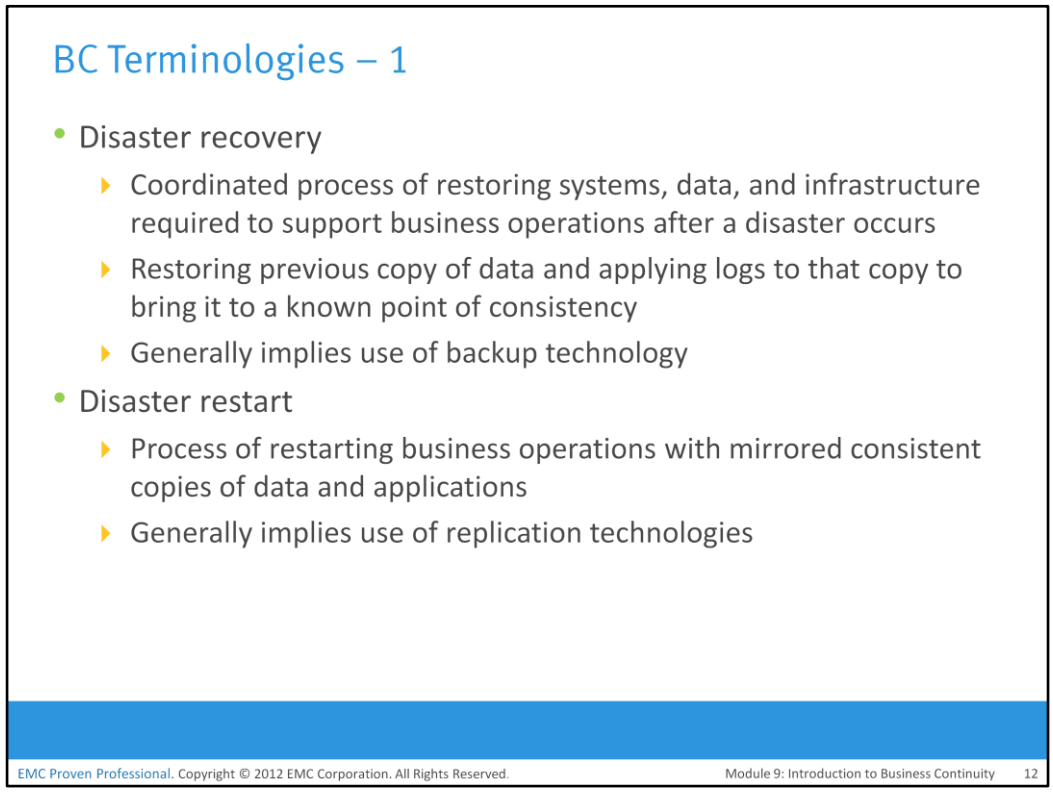
Uptime (%)	Downtime (%)	Downtime per Year	Downtime per Week
98	2	7.3 days	3hrs, 22 minutes
99	1	3.65 days	1 hr, 41 minutes
99.8	0.2	17 hrs, 31 minutes	20 minutes, 10 secs
99.9	0.1	8 hrs, 45 minutes	10 minutes, 5 secs
99.99	0.01	52.5 minutes	1 minute
99.999	0.001	5.25 minutes	6 secs
99.9999	0.0001	31.5 secs	0.6 secs

Uptime per year is based on the exact timeliness requirements of the service. This calculation leads to the number of "9s" representation for availability metrics. Table on the slide lists the approximate amount of downtime allowed for a service to achieve certain levels of 9s availability.

For example, a service that is said to be "five 9s available" is available for 99.999 percent of the scheduled time in a year (24 × 365).

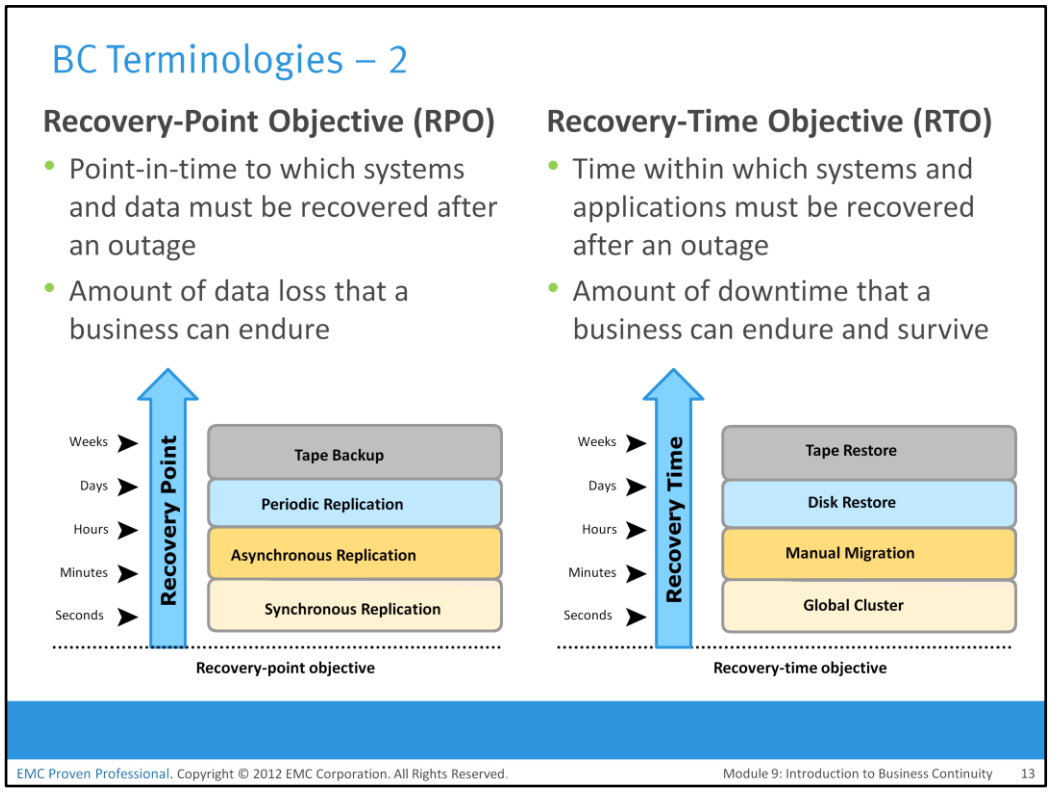


This lesson covers various BC terminologies and BC planning. This lesson also focuses on eliminating single points of failure and multipathing software.



Disaster recovery: This is the coordinated process of restoring systems, data, and the infrastructure required to support ongoing business operations after a disaster occurs. It is the process of restoring a previous copy of the data and applying logs or other necessary processes to that copy to bring it to a known point of consistency. After all recovery efforts are completed, the data is validated to ensure that it is correct.

Disaster restart: This is the process of restarting business operations with mirrored consistent copies of data and applications.



Recovery-Point Objective (RPO): This is the point-in-time to which systems and data must be recovered after an outage. It defines the amount of data loss that a business can endure. Based on the RPO, organizations plan for the frequency with which a backup or replica must be made. An organization can plan for an appropriate BC technology solution on the basis of the RPO it sets. For example:

RPO of 24 hours: Backups are created at an offsite tape library every midnight. The corresponding recovery strategy is to restore data from the set of last backup tapes.

RPO of 1 hour: Shipping database logs to the remote site every hour. The corresponding recovery strategy is to recover the database to the point of the last log shipment.

RPO in the order of minutes: Mirroring data asynchronously to a remote site.

RPO of zero: Mirroring data synchronously to a remote site.

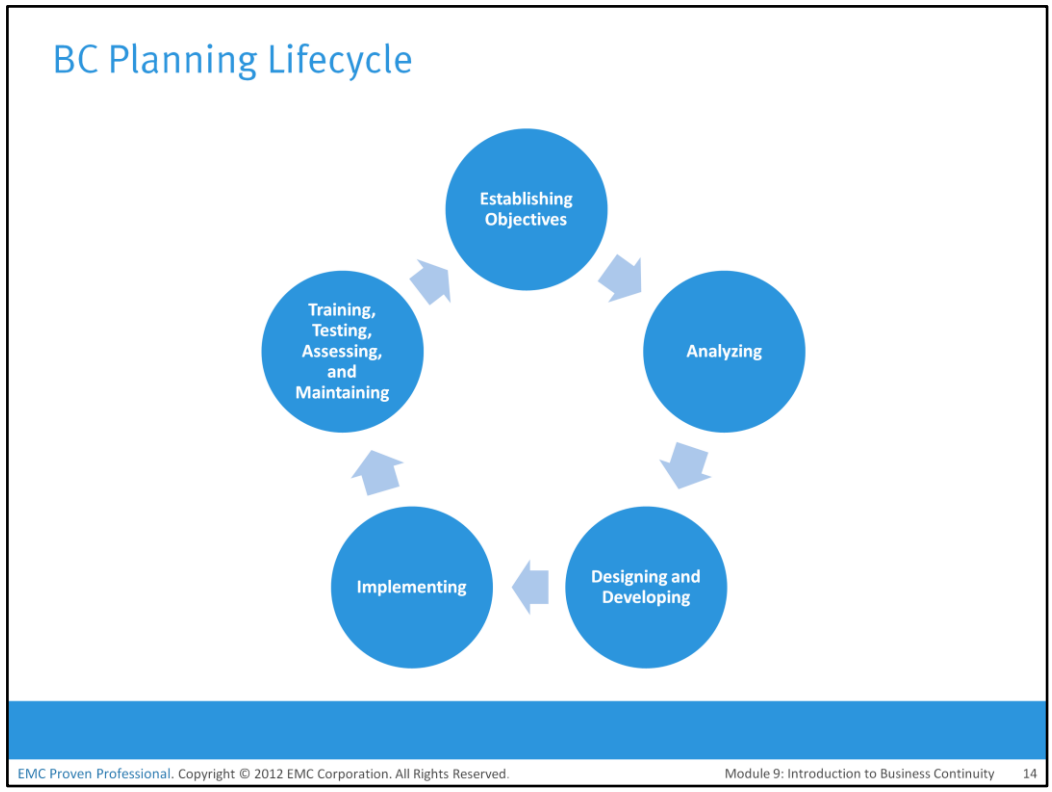
Recovery-Time Objective (RTO): The time within which systems and applications must be recovered after an outage. It defines the amount of downtime that a business can endure and survive. Some examples of RTOs and the recovery strategies to ensure data availability are listed below:

RTO of 72 hours: Restore from tapes available at a cold site.

RTO of 12 hours: Restore from tapes available at a hot site.

RTO of few hours: Use disk-based backup technology, which gives faster restore than a tape backup.

RTO of a few seconds: Cluster production servers with bidirectional mirroring, enabling the applications to run at both sites simultaneously.



BC planning must follow a disciplined approach like any other planning process. Organizations today dedicate specialized resources to develop and maintain BC plans. From the conceptualization to the realization of the BC plan, a lifecycle of activities can be defined for the BC process. The BC planning lifecycle includes five stages:

1. Establishing objectives

- Determine BC requirements.
- Estimate the scope and budget to achieve requirements.
- Select a BC team that includes subject matter experts from all areas of the business, whether internal or external.
- Create BC policies.

2. Analyzing

- Collect information on data profiles, business processes, infrastructure support, dependencies, and frequency of using business infrastructure.
- Conduct a business impact analysis.
- Identify critical business processes and assign recovery priorities.
- Perform risk analysis for critical functions and create mitigation strategies.
- Perform cost benefit analysis for available solutions based on the mitigation strategy.
- Evaluate options.

Cont..

3. Designing and developing

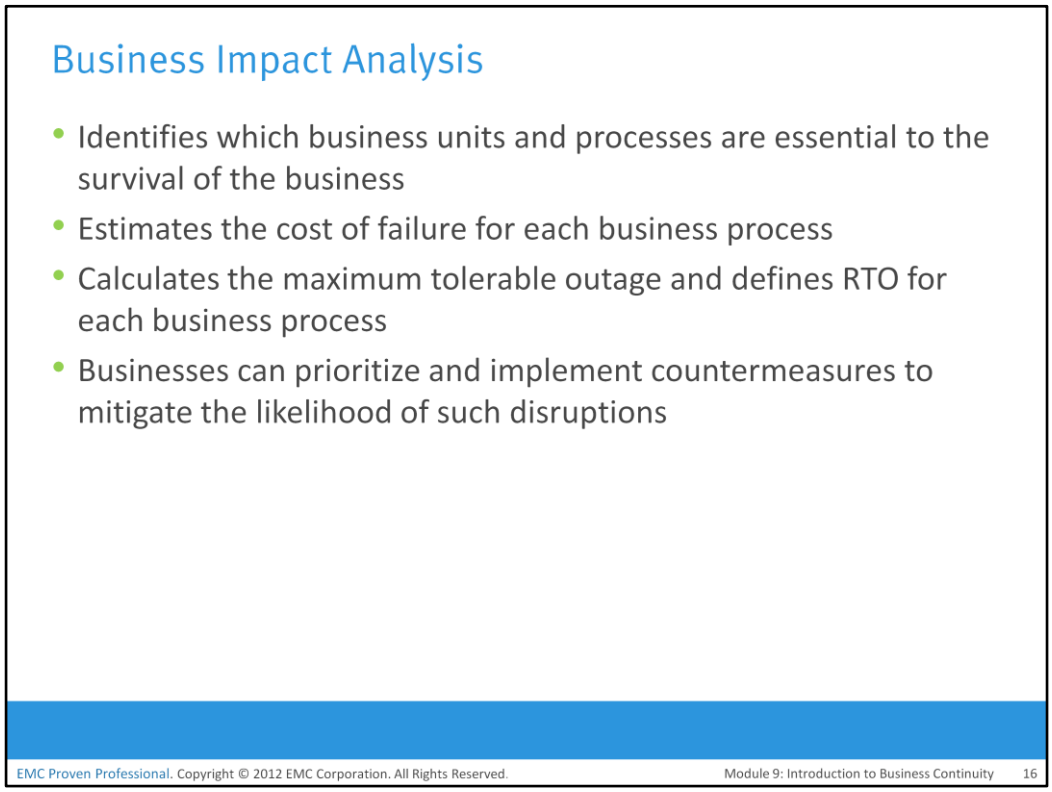
- Define the team structure and assign individual roles and responsibilities. For example, different teams are formed for activities, such as emergency response, and infrastructure and application recovery.
- Design data protection strategies and develop infrastructure.
- Develop contingency solution and emergency response procedures.
- Detail recovery and restart procedures.

4. Implementing

- Implement risk management and mitigation procedures that include backup, replication, and management of resources.
- Prepare the DR sites that can be utilized if a disaster affects the primary data center.
- Implement redundancy for every resource in a data center to avoid single points of failure.

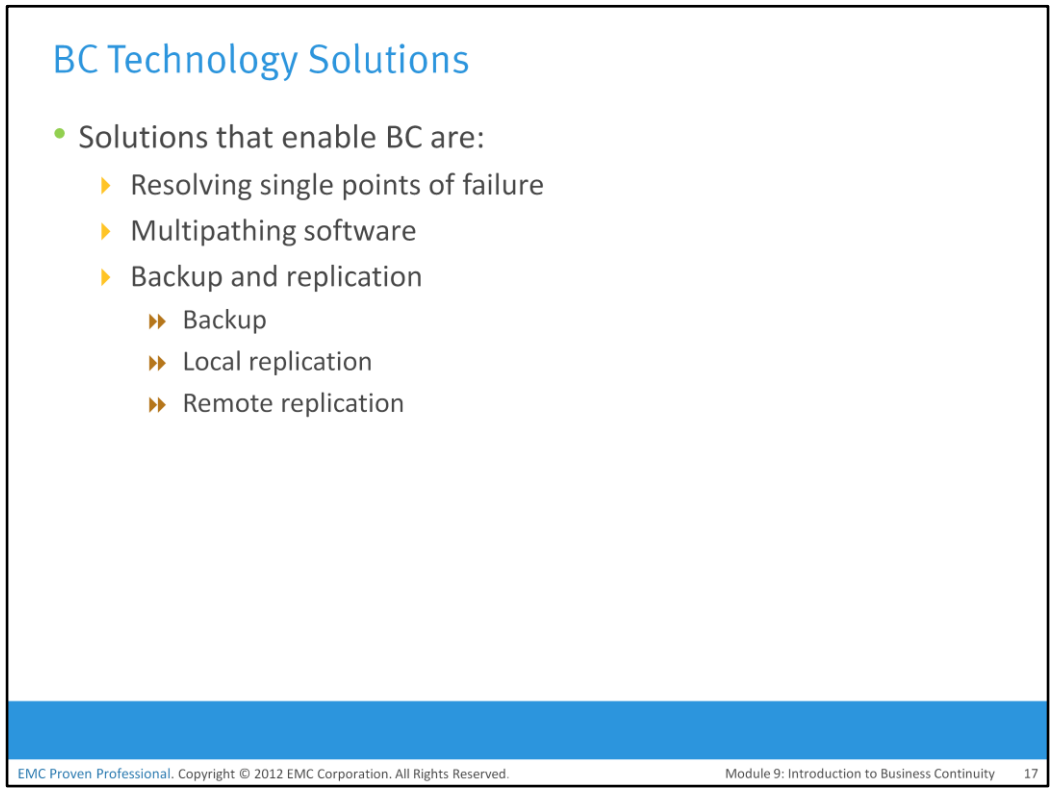
5. Training, testing, assessing, and maintaining

- Train the employees who are responsible for backup and replication of business-critical data on a regular basis or whenever there is a modification in the BC plan.
- Train employees on emergency response procedures when disasters are declared.
- Train the recovery team on recovery procedures based on contingency scenarios.
- Perform damage-assessment processes and review recovery plans.
- Test the BC plan regularly to evaluate its performance and identify its limitations.
- Assess the performance reports and identify limitations.
- Update the BC plans and recovery/restart procedures to reflect regular changes within the data center.



A *business impact analysis* (BIA) identifies which business units, operations, and processes are essential to the survival of the business. It evaluates the financial, operational, and service impacts of a disruption to essential business processes. Selected functional areas are evaluated to determine resilience of the infrastructure to support information availability. The BIA process leads to a report detailing the incidents and their impact over business functions. The impact may be specified in terms of money or in terms of time. Based on the potential impacts associated with downtime, businesses can prioritize and implement countermeasures to mitigate the likelihood of such disruptions. These are detailed in the BC plan. A BIA includes the following set of tasks:

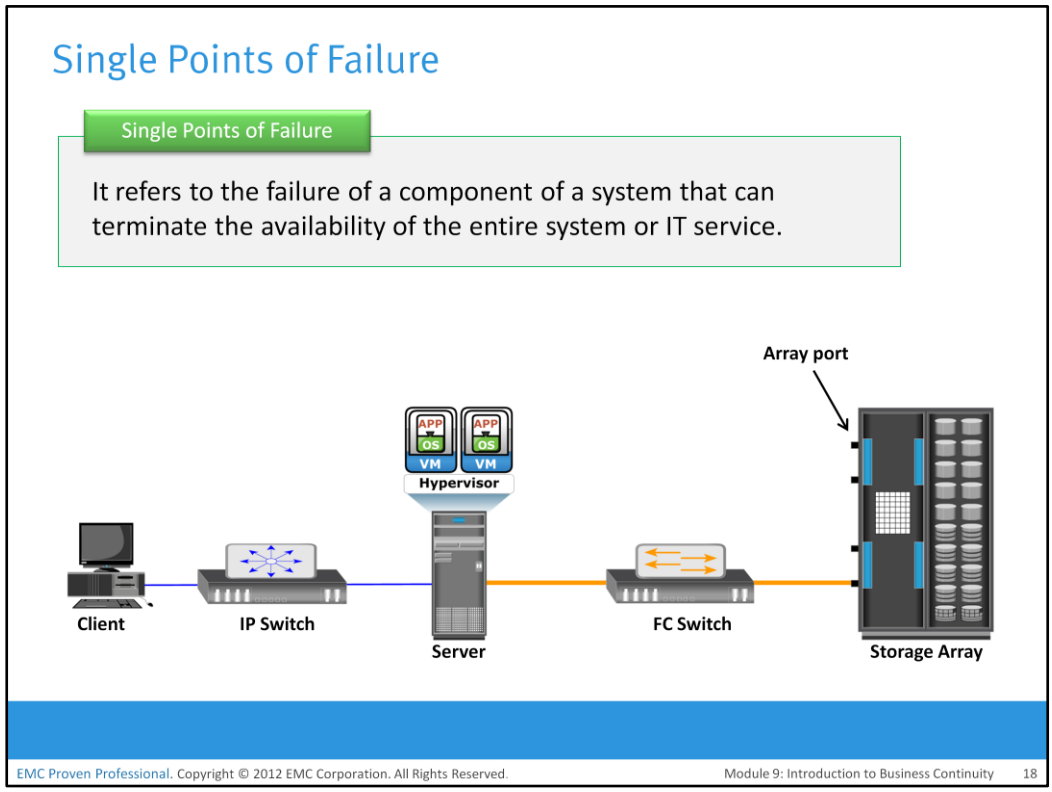
- Determine the business areas.
- For each business area, identify the key business processes critical to its operation.
- Determine the attributes of the business process in terms of applications, databases, and hardware and software requirements.
- Estimate the costs of failure for each business process.
- Calculate the maximum tolerable outage and define RTO for each business process.
- Establish the minimum resources required for the operation of business processes.
- Determine recovery strategies and the cost for implementing them.
- Optimize the backup and business recovery strategy based on business priorities.
- Analyze the current state of BC readiness and optimize future BC planning.



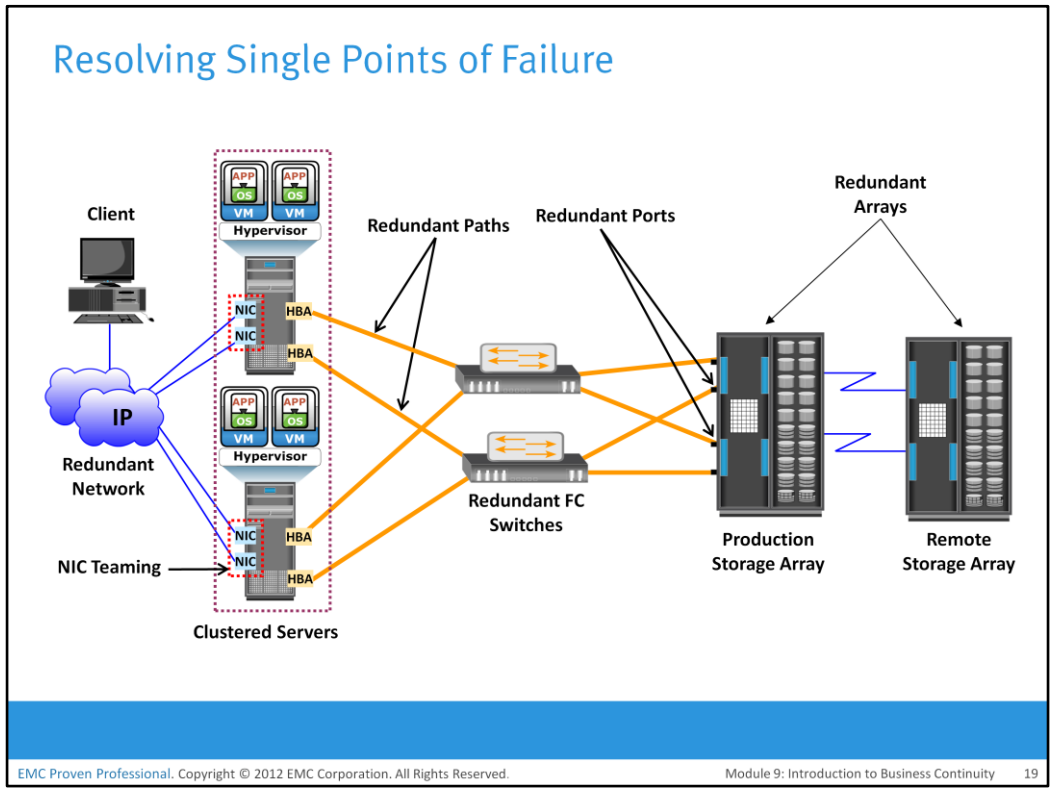
After analyzing the business impact of an outage, designing the appropriate solutions to recover from a failure is the next important activity. Following are the solutions and supporting technologies that enable business continuity and uninterrupted data availability:

- Resolving single points of failure
- Multipathing software
- Backup and replication

Note: Backup and Replication will be discussed in forthcoming modules.



A *single point of failure* refers to the failure of a component that can terminate the availability of the entire system or IT service. The figure depicts a system setup in which an application, running on a VM, provides an interface to the client and performs I/O operations. The client is connected to the server through an IP network, and the server is connected to the storage array through an FC connection. In this setup, each component must function as required to ensure data availability, the failure of a single physical or virtual component causes the unavailability of an application. This failure results in disruption of business operations. For example, failure of a hypervisor can affect all the running VMs and virtual network, which are hosted on it. In the figure on the slide, several single points of failure can be identified. A VM, a hypervisor, or an HBA/NIC on the server, the physical server itself, the IP network, the FC switch, the storage array port, or even the storage array could be a potential single point of failure.

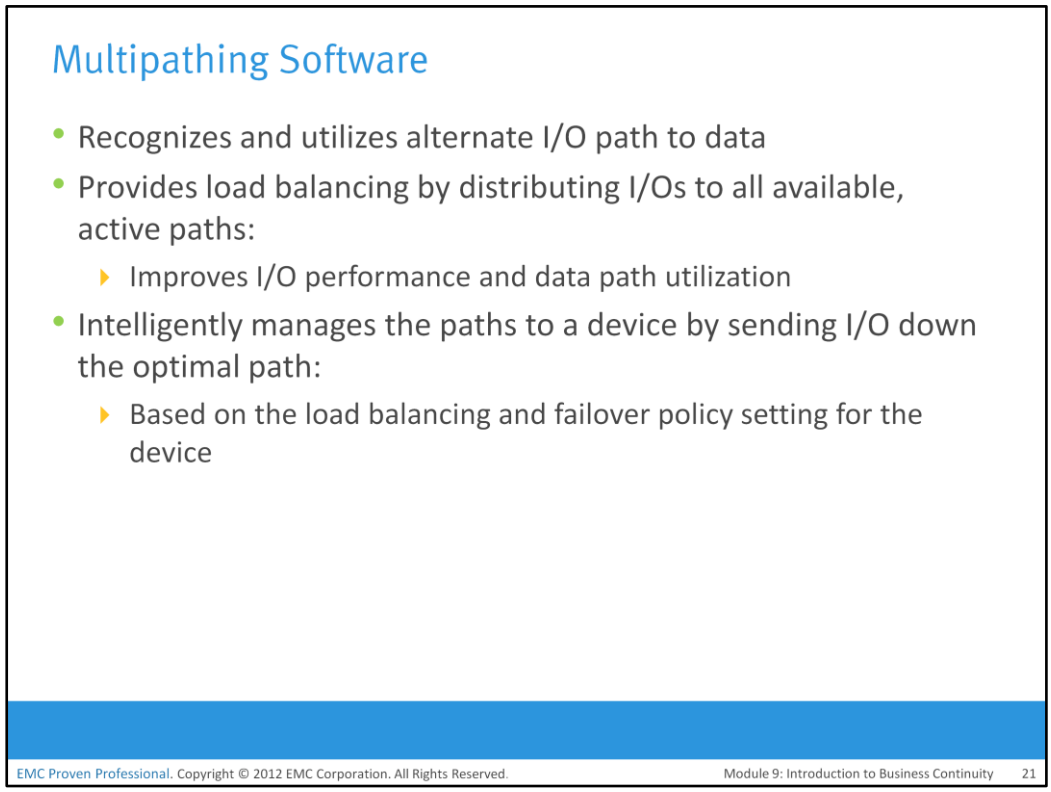


To mitigate single points of failure, systems are designed with redundancy, such that the system fails only if all the components in the redundancy group fail. This ensures that the failure of a single component does not affect data availability. Data centers follow stringent guidelines to implement fault tolerance for uninterrupted information availability. Careful analysis is performed to eliminate every single point of failure. The example shown in figure on the slide represents all enhancements in the infrastructure to mitigate single points of failure:

- Configuration of redundant HBAs at a server to mitigate single HBA failure.
- Configuration of NIC teaming at a server allows protection against single physical NIC failure. It allows grouping of two or more physical NICs and treating them as a single logical device. With NIC teaming, if one of the underlying physical NICs fails or its cable is unplugged, the traffic is redirected to another physical NIC in the team. Thus, NIC teaming eliminates the single point of failure associated with a single physical NIC.
- Configuration of redundant switches to account for a switch failure.
- Configuration of multiple storage array ports to mitigate a port failure.
- RAID and hot spare configuration to ensure continuous operation in the event of disk failure.

Cont..

- Implementation of a redundant storage array at a remote site to mitigate local site failure.
- Implementing server (or compute) clustering, a fault-tolerance mechanism whereby two or more servers in a cluster access the same set of data volumes. Clustered servers exchange *heartbeat* to inform each other about their health. If one of the servers or hypervisors fails, the other server or hypervisor can take up the workload.
- Implementing a VM Fault Tolerance mechanism ensures BC in the event of a server failure. This technique creates duplicate copies of each VM on another server so that when a VM failure is detected, the duplicate VM can be used for failover. The two VMs are kept in synchronization with each other in order to perform successful failover.

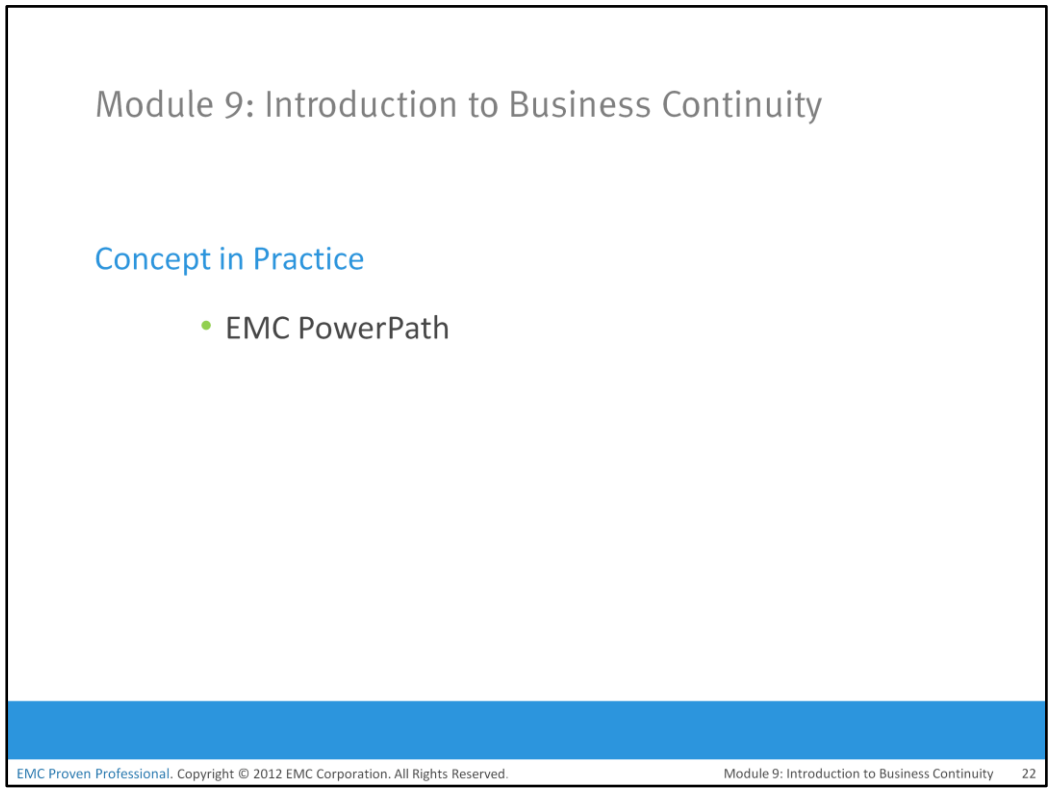


Configuration of multiple paths increases the data availability through path failover. If servers are configured with one I/O path to the data, there will be no access to the data if that path fails. Redundant paths to the data eliminate the possibility of the path becoming a single point of failure. Multiple paths to data also improve I/O performance through load balancing among the paths and maximize server, storage, and data path utilization.

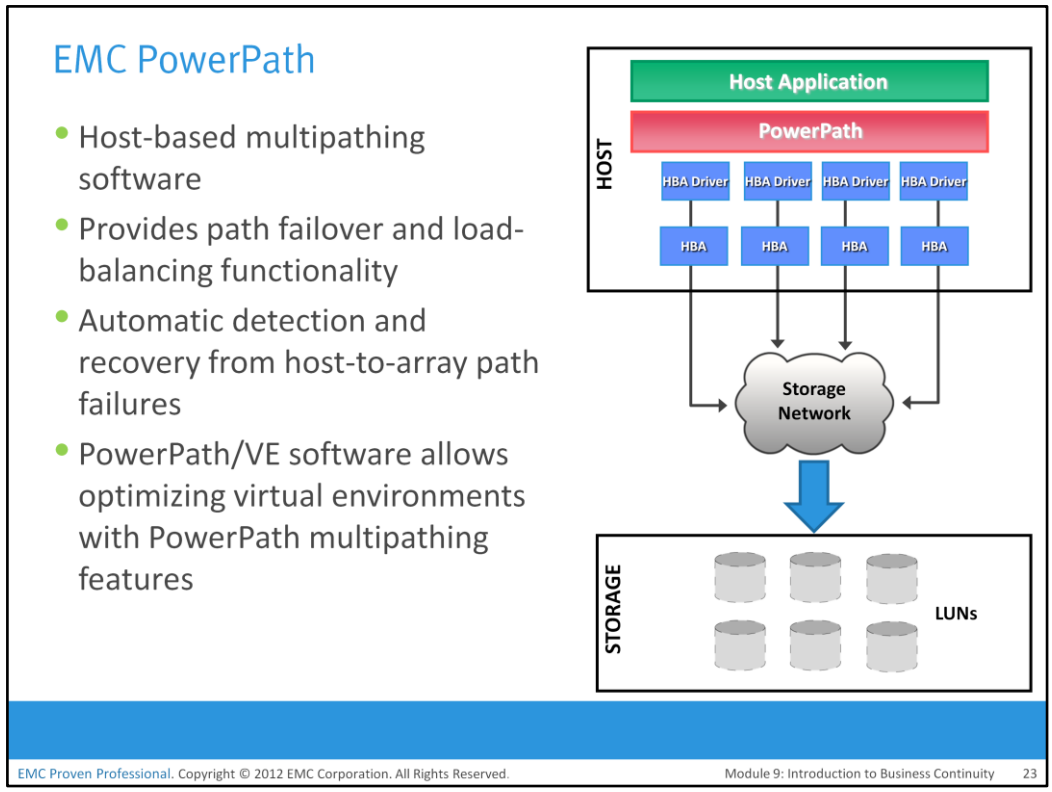
In practice, merely configuring multiple paths does not serve the purpose. Even with multiple paths, if one path fails, I/O does not reroute unless the system recognizes that it has an alternative path. Multipathing software provides the functionality to recognize and utilize alternative I/O paths to data. Multipathing software also manages the load balancing by distributing I/Os to all available, active paths.

Multipathing software intelligently manages the paths to a device by sending I/O down the optimal path based on the load balancing and failover policy setting for the device. It also takes into account path usage and availability before deciding the path through which to send the I/O. If a path to the device fails, it automatically reroutes the I/O to an alternative path.

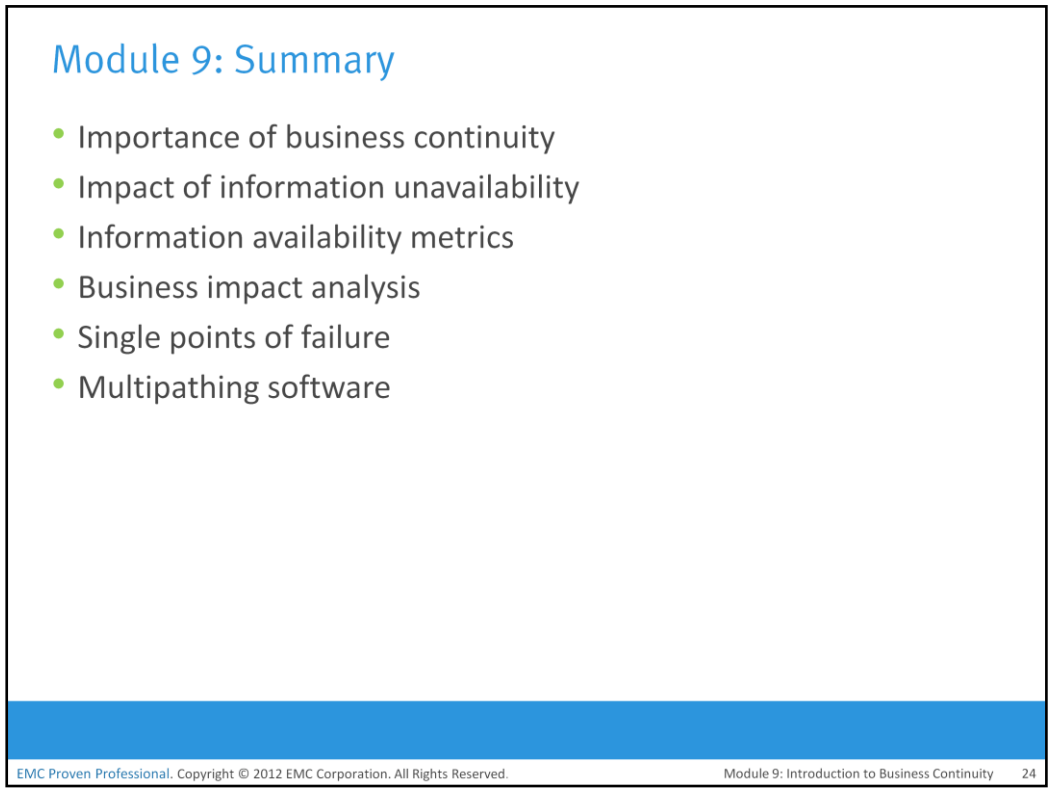
In a virtual environment, multipathing is enabled either by using the hypervisor's built-in capability or by running a third-party software module, added to the hypervisor.



The concept in practice section covers EMC PowerPath.



EMC PowerPath is host-based multipathing software. Every I/O from the host to the array must pass through the PowerPath software, which allows PowerPath to provide intelligent I/O path management. PowerPath provides path failover and dynamic load balancing. PowerPath/VE software allows optimizing virtual environments with PowerPath multipathing features.



This module covered the importance of business continuity, impact of information unavailability, and information availability metrics. This module also focused on business continuity planning and business impact analysis. Further, this module detailed on single points of failure and multipathing software.

BC entails preparing for, responding to, and recovering from a system outage that adversely affects business operations. Information unavailability or downtime results in loss of productivity, loss of revenue, poor financial performance, and damages to reputation. Information availability metrics are MTBF and MTTR. MTBF defines average time available for a system or component to perform its normal operations between failures. MTTR defines the average time required to repair a failed component. A business impact analysis identifies which business units, operations, and processes are essential to the survival of the business. A single point of failure refers to the failure of a component that can terminate the availability of the entire system or IT service. Multipathing software provides the functionality to recognize and utilize alternate I/O paths to data. Multipathing software also manages the load balancing by distributing I/Os to all available, active paths.

Check Your Knowledge – 1

- Which statement is true if the recovery-point objective (RPO) of an application is 2 hours?
  - A. Time to resume application operations must be less than 2 hours
  - B. Time to resume application operations must equal to 2 hours
  - C. No more than 2 hours of production data can be lost
  - D. Mean time between application failure is 2 hours

- Which allows grouping of two or more physical NICs and treating them as a single logical device?
  - A. NIC streaming
  - B. NIC porting
  - C. NIC teaming
  - D. NIC zoning

Check Your Knowledge – 2

- Which best describes recovery-time objective (RTO)?
  - A. Point-in-time to which data must be recovered after an outage
  - B. Time available for a system or component to perform its normal operations between failures
  - C. Time within which systems and applications must be recovered after an outage
  - D. Amount of data loss that a business can endure

- Which expression represents availability of a system in terms of MTBF and MTTR?
  - A. MTTR/(MTBF x MTTR)
  - B. MTBF/(MTBF x MTTR)
  - C. MTTR/(MTBF + MTTR)
  - D. MTBF/(MTTR + MTBF)

Check Your Knowledge – 3

- A department requires access to the database application from Monday to Friday, 9 AM to 5 PM. Last Thursday at 1 PM the application crashed and it took six hours to fix the problem. What was the availability of the application during last week?
  - A. 85%
  - B. 90%
  - C. 95%
  - D. 100%

Exercise 1: MTBF and MTTR

- A system has three components and requires all three to be operational for 24 hours from Monday to Friday. Failure of component 1 occurs as follows:
  - Monday = No failure
  - Tuesday = 5 am to 7 am
  - Wednesday = No failure
  - Thursday = 4 pm to 8 pm
  - Friday = 8 am to 11 am

Calculate the MTBF and MTTR of component 1.

Exercise 2: Information Availability

- A system has three components and requires all three to be operational from 8 am to 5 pm business hours, Monday to Friday. Failure of component 2 occurs as follows:
  - Monday = 8 am to 11 am
  - Tuesday = No failure
  - Wednesday = 4 pm to 7 pm
  - Thursday = 5 pm to 8 pm
  - Friday = 1 pm to 2 pm

Calculate the availability of component 2.