Technology Review#2002-4

Business Continuity Planning

Padmavathy Ramesh

July 2002

Contents

List of Illustrations

1 Introduction

1.1 Continuity Planning

Business entities today exist in a highly competitive world. They are constantly innovating to meet their business objectives of providing essential and unique services to their customers. Technology advances have enabled them to achieve their varied strategies. And yet, the threats of disaster, on account of business interruption, are not extinct – in fact, they have also evolved along with the technology. Business interruption does happen – but what is of significance is, how much of the consequences of such interruptions can the business afford? Business Continuity Planning is the act of proactively working out a way to prevent, if possible, and manage the consequences of a disaster, limiting it to the extent that a business can afford.

1.2 Need

There are various threats and vulnerabilities to which business today is exposed. They could be:

- catastrophic events such as floods, earthquakes, or acts of terrorism
- accidents or sabotage
- outages due to an application error, hardware or network failures

Some of them come unwarned. Most of them never happen. The key is to be prepared and be able to respond to the event when it does happen, so that the organization survives; its losses are minimized; it remains viable and it can be "business as usual", even before the customers feel the effects of the downtime. An effective Business Continuity Plan serves to secure businesses against financial disasters. The bonus — customer satisfaction, enhanced corporate image and no dip in the market share.

2 Elements of Business Continuity Planning

2.1 Initiation

The first step is to obtain the commitment of the management and all the stakeholders towards the plan. They have to set down the objectives of the plan, its scope and the policies. An example of a decision on scope would be whether the target is the entire organization or just some divisions, or whether it is only the data processing, or all the organization's services. Management provides sponsorship in terms of finance and manpower. They need to weigh potential business losses versus the annual cost of creating and maintaining the Business Continuity Planning. For this, they will have to find answers to questions such as how much it would cost or how much would be considered adequate.

Broadly, the objective of the Business Continuity Planning (BCP) for a business can only be – to identify and reduce risk exposures and to proactively manage the contingency. The specific objectives that a BCP can set will be described in the subsequent sections.

The final outcome of the BCP exercise is:

- a set of measures to prevent disasters
- a BCP operational team, trained to handle the situation
- a plan that provides a roadmap when disaster strikes – a plan that is sufficient and complete, detailing what needs to be done with each element that falls within the plan's scope.

The discussions that follow are mainly in the context of IT services provided by an organization. They do not deal with the safety management of the firm's personnel, in case of a disaster.

2.2 Risk Assessment

Risk assessment is the exercise of identifying and analyzing the potential vulnerabilities and threats. The sources of risks could be:

- community-wide hazardous events
- accidents or sabotage causing extreme material disaster
- security threats, network and communication failures
- disastrous application errors

Each of these areas should be looked at in the light of the business and the exact possible source located. For each source identified:

- the magnitude of the risk and
- the probability of its occurrence

must be evaluated to judge the extent of risk exposure. Risk exposure is the easiest way to know how much attention needs to be paid to a source of risk.

Planning is done for both — prevention and control. Accidents and sabotage can be prevented using measures of physical security and personnel practices. Vulnerability assessment and reviews of existing security measures can throw up areas where access control, software and data security, or backups are required. Application errors can be prevented by effective reviews and testing during the software releases.

If needed, the expertise of external agencies can easily be called upon to analyze, devise and put in place some of the preventive measures.

The tougher part is to come up with activities for controlling the effects of disaster, and this necessitates a detailed business impact analysis.

The end result of the Risk Assessment should be a risk-benefit analysis statement giving the exact threats, and the estimated exposure together with the contingency and mitigation actions required, and also the benefits arising out of covering the risk. This statement should also delineate any assumptions or constraints that exist.

Often, this exercise will show that the complete physical disaster has a remote probability of occurring and application crashes, or security break-ins are very frequent. However, only having a procedure for handling catastrophic disasters without a plan for application failure or vice versa is not advisable. The solution is to prepare a BCP for the worst-case, i.e., complete destruction of the site providing the services. Any other outage can then be easily tackled using a sub-set of the main plan.

2.3 Business Impact Analysis

Business Impact Analysis (BIA) is essentially the process of identifying the critical business functions and the losses and effects if these functions are not available.

It involves talking to the key people operating the business functions in order to assess:

A. Impact

- how vital the function is to the overall business strategy
- how long the function could be inoperative without any impact or losses
- how the rest of the business would be affected by its outage – the operational impact
- what the revenue lost due to its outage would be – the financial impact
- whether its outage would result in violation of Service Level Agreements (SLAs), regulatory requirements, any contractual liabilities, and penalty, or whether it would create legal issues – the regulatory and legal impact
- whether it would affect relationships with customers – loss of customer confidence
- whether it would affect the market rates – decline in market rates
- whether it would affect the industry ranking – loss of competitive edge
- whether it could result in losing future sales – loss of opportunities
- what the maximum/acceptable/permissible outage would be

B. Requirements for recovery

- what the resources and records required would be to continue the function
- what the bare minimum resource requirements would be
- which of the resources would be from external sources
- what other business functions it would be dependent upon, and to what extent

- what other business functions would depend on it and to what extent
- upon which external business/suppliers/vendors it would be dependent, and to what extent
- which SLAs and measures for continuity these external businesses/suppliers/vendors would follow
- what the backup needs would be
- what the time and effort required to recreate up-to-date data from the backups would be
- what precautions or verifications would need to be taken or done for recovering without a test environment

Based on these discussions, it will be possible to **classify** the business functions as:

a) **Critical functions** – If these business functions are interrupted or unavailable for some time, it can completely jeopardize the business and cause heavy damages to the business.

b) **Essential functions** – Those functions, whose loss would seriously affect the organization's ability to function for long.

c) **Necessary functions** – The organization can continue functioning; however, absence of these functions would limit their effectiveness, to a great extent.

d) **Desirable functions** – These functions would be beneficial; however, their absence would not affect the capability of the organization.

Based on their recovery needs, organizations can come up with standard ***recovery time frames*** for the above classifications. For example, Critical functions: < 1 day, Essential functions: 2-4 days, Necessary functions: 5-7 days and Desirable functions: > 10 days.

This impact analysis helps to rank the business functions and come up with an order in which they should be brought up. In other words, it defines ***recovery priorities***.

BIA helps define the ***recovery objectives***. In the course of this study, it might be possible to discover that when resuming operations after a disaster, it is enough to recover to a limited capacity, i.e., recover to the extent of handling 40 percent of the usual workload within 24 hours.

It will also be possible to define in detail the ***resource requirements*** for making a business function operational after disaster or interruption. This will include infrastructure, manpower, documents, records, machines, phones, fax machines, whatever is needed – with complete specifications. Having adequate details is important, since in the event of disasters, there is bound to be some amount of panic and it may not be possible to come down to such details.

The team and managers actually involved in the day-to-day operations of the business functions would be the best people to talk to during the impact analysis, as they would certainly know the details of the functions. Moreover, they can perform a brainstorming exercise on how an outage of their function would affect the revenue objectives, market position and customer expectations, or how they could restore normal operations, or what resources they would require to operate in normal mode.

Interdependence between various functions (internal and external) is crucial information obtained as part of the analysis. While consolidating the information gathered from the questionnaires/discussions and ranking the functions to derive the recovery priority, one must not overlook functions, which by themselves are low priority, however, have some critical functions depending on them. By virtue of this dependence, they also become important.

Cost considerations are not to be ignored during this exercise. Things to be kept in mind are:

- Revenue losses and opportunity losses will be directly proportional to the time taken for recovery

- Cost of a recovery strategy will be inversely proportional to the time permitted for recovery

- Cost of the possible recovery strategy must be compared with the actual loss due to the outage before accepting the strategy. If the solution proposed costs much more than the projected losses, it will not be possible to justify the investment to the management.

When presenting the findings of the business impact analysis, the results must also be expressed in business terms. Quantifying the impact, possibly in terms of money, will catch the attention of the management. Stating the impact in terms of time will help in proposing concrete recovery goals. Stating the requirements in technical terms will help planning the recovery strategies. Ultimately, the business impact analysis must justify the continuity plan and aid selection of the best possible recovery strategy within the budget.

2.4 Strategies

Business Continuity Planning should include strategies on:

- Prevention

- Response

- Resumption

- Recovery

- Restoration

Prevention aims at lessening the chances of the disaster happening.

Response is the reaction when the event occurs. It must stem further damage, assess the extent of damage, salvage the business entity's reputation by providing appropriate communication to the external world and indicate a possible recovery timeframe.

Resumption involves resuming only the time-sensitive business processes, either immediately after the interruption or after the declared Mean Time Between Failures (MTBF). All operations are not fully recovered.

Recovery addresses the startup of less time-sensitive processes. The time duration of this naturally depends on the time taken for resumption of the time-sensitive functions. It could involve starting up these services at an alternate location.

Restoration is the process of repairing and restoring the primary site. At the end of this, the business operations are resumed in totality from the original site or a completely new site, in case of a catastrophic disaster.

2.4.1 Prevention

Strategies for prevention would include both deterrent and preventive controls.

- Deterrent controls reduce the likelihood of the threats.
- Preventive controls safeguard the vulnerable areas to ward off any threat that occurs and reduce its impact.

Having these measures in place is always more cost-effective than attempting recovery after the interruption. The aim should be to cover as many as possible of the risks identified, using deterrent and preventive controls, so that the recovery strategy has to work only on the residual risks.

A wide variety of such controls exist. Some of the common ones are described below.

(a) Security at the premises — It is a deterrent control and exists in the form of barriers to protect the location and prevent accidental or unauthorized entry. It could also involve manned or technology-driven surveillance at the location.

(b) Personnel procedures — Areas housing the critical resources could be restricted zones where only authorized people are allowed to enter after some means of identification are provided. The means of identification can be varied depending on the technology used for the identification process.

(c) Infrastructure-related — This includes having an appropriate sized UPS, backup power, air conditioning, smoke/fire detectors, fire extinguishers, waterproofing, fire resistant containers for vital records and backups and also monitoring weather forecasts.

(d) Software controls — The most common of these are authentication, access control, anti-virus, encryption, firewall, and intrusion detection systems.

(e) Storage and recovery related — Frequent backups. The various mechanisms will be discussed later in this paper. Offsite storage of vital records and backups later contribute to the resumption and recovery process.

The above list distinctly highlights one aspect: most of the safeguards are closely related to the security policy and practices in an organization.

Business firms will want to ensure the availability and safety of their assets (which includes information). Their security policy addresses these objectives and provides guidelines for usage and management of their assets. Armed with knowledge of the firm's assets, their layout and the risk assessment results, the firm can come up with the necessary controls needed to implement the security policy. These controls or security practices must be reviewed from time-to-time and also be tested to see whether they are penetrable by all categories of people, i.e., by people having valid access, by having complete knowledge of the systems or by a complete outsider. Any of them can misuse the access. The reviews will help enrich or strengthen the measures.

Having a security policy, putting preventive safeguards in place, monitoring the system for intrusions and ensuring action against those who violate it, is itself a deterrent control. Planning for prevention is an exercise that must be done carefully. It has to ensure that the mechanisms used are neither very restrictive, nor would they constitute a bottleneck, nor cause an availability problem, nor allow undesirable/easy access and usage.

2.4.2 Response

The first reaction to an interruption would be to inform all the relevant people about the interruption. If it is an impending interruption about which there is a prior warning, then this notification can be done in advance. Timely notification is important, since it may provide an opportunity to stem any further damage. In a situation where there is adequate time to perform a shutdown, a switchover or an evacuation, it may even completely prevent damage. This, however, requires the presence of diagnostic or detective controls. Such controls either continuously scan themselves for a symptom of interruption (network, servers) or collect such information from external sources (natural calamities).

The exact notification procedure must be laid down. It involves clearly documenting who is to be notified, how, by whom, and also the escalation mechanism.

A notification call tree within the BCP team is set up. Here, the initial notification is sent to a set of people, who in turn, inform the next set of people, and so on. People belonging to this call tree have different roles. The type of information and amount of detail provided as a part of the notification depends on the role of the person. The following groups would be involved:

- Management — would need to be informed of the status. It has the powers to authorize the emergency response and further actions. The management will also deal with the press, public, customers and shareholders.

- Damage Assessment Team — would assess the damage and rate the severity of the interruption.

- Technical Team — would serve as the key decision-makers for further activities of the BCP.
- Operations Team — would execute the actual operations of the BCP.

It is also important to state an alternative for each contact. In case the primary person is not available or traceable, the backup person is to be notified. Notification can be done using various tools: pager, SMS, phone, and email. The team is equipped appropriately.

The Damage Assessment Team is among the earliest (along with the management) to be notified of the event. They would be required at the site at the earliest to evaluate the extent of the damage inflicted. In case the site itself has been subject to damage, then they should start their work as soon as an entry is allowed. (Of course, if the calamity is as great as on September 11th 2001, then it is obvious that it is a disaster of the greatest severity.)

The assessment should be done against a plan that is closely related to the business continuity priorities. This means that they should be aware of the area in the site and processes that are crucial to the business. This would help them prioritize their examination and also focus adequately on the critical areas. This team needs to look at:

- the cause of disruption
- whether there is scope to stem additional damage
- infrastructure and equipment damage
- services affected
- vital records damaged
- what can be salvaged
- what needs repair, restoration and replacement
- requirements for insurance claims, if applicable

Armed with this input (provided by the Damage Assessment Team) on the severity of damage to facilities and the extent to which the business is inoperable, the Technical Team can work ahead. Some of the questions faced by them are:

- Is it a disaster? Of what degree?
- When will the impact be felt?
- What is the extent of time to repair/resume/restore?
- Where does one begin?

The BCP must have a set of predefined parameters based on the Business Impact Analysis and their continuity goals to evaluate the information available on the damage.

These parameters should differentiate between an interruption and a disaster, and also rate the severity of the event. What the Technical Team uses here is a decision support mechanism based on these parameters before they declare a disaster (of any appropriate scale).

While the Damage Assessment Team and Technical Team are working, the rest of the BCP team is placed on alert for a possible activation of the continuity plan. The type and extent of the disaster declared would indicate which portions of the BCP need to be implemented. Accordingly, the BCP team is notified and resumption activities are started.

An optional step in the emergency response (the first action, in fact) is to move to safety all personnel on the premises and alert the police, fire service and hospitals. This is a step required only if the interruption is of the nature of an accident, act of sabotage or natural calamity.

2.4.3 Resumption

The focus shifts to the command centre once the BCP has been activated. This is a location different from the normal business facility. It is from here that the resumption, and subsequently, the recovery activities, are coordinated. The centre will have adequate communication facilities, PCs, printers, fax machines and office equipment to support the activities of the team.

The first decision to be taken is – whether the critical operations can be resumed at the normal business site or at an alternate site. In situations when access to the primary site is denied or the site is damaged beyond use, the operations could move to an alternate site.

Alternate sites can be of the following kinds:

(a) **Cold Site** – A facility that is environmentally conditioned, but devoid of any equipment. It is ready for all the equipment to move in, i.e., it has telephone points, power supply, and UPS facility, among others. It takes a little time to make this site operational. Using a cold site implies that the business entity has contracts with the providers of all the necessary equipment. These contracts are specifically for a business resumption scenario and therefore will have clauses on the time within which the setup will be completed.

(b) **Hot Site** — It is an alternate facility having workspace for the personnel, fully equipped with all resources and stand-by computer facilities needed to recover and support critical business functions after a disaster. It is a fully equipped site where the BCP team moves in to start work without further delay.

(c) **Warm site** — It is a partially equipped hot site and the data is not too old.

(d) **Mobile site** — It is a portable site with a smaller configuration. It can be positioned near the primary site, thus saving travel for the key staff.

(e) **Mirrored Site** – It is identical in all aspects to the primary site, right down to the information availability. It is equivalent to having a redundant site in normal times and is naturally the most expensive option.

At the alternate site (or primary site, if still usable), the work environment is restored. Communication, networks, and workstations are set up. Contact with the external world can now be resumed. It is possible that an organization might choose to function in the manual mode until the critical IT services can resume. If the recovery alternative (described in a later section) permits, the critical functions can also be resumed in the automated mode very quickly.

2.4.4 Recovery

At the site of recovery (either primary or alternative), the operating system is restored on the stand-by system. Necessary applications are restored in the order of their criticality. When the applications to serve the critical functions are restored, data restoration from backup tapes or media obtained from the offsite storage can be initiated.

Data must also be synchronized i.e., to rebuild data accurately to a predetermined point of time before the interruption. The point to which the restoration is done depends on the requirements of the critical services. Business data comes from different sources, each of which must be reconstructed to reach the desired state of data integrity. The synchronized data must be reviewed and validated. This is mandatory because under such disastrous circumstances, it is possible that there is no test environment available and that applications will resume directly in the production environment. It is therefore necessary to have a clear method, strategy or checklist to perform this validation exercise.

Once the data has reached a reliable state, transactions that have been accumulating since the disaster can be processed and all the critical functions can then resume. Gradually, other services of the business can also begin functioning.

Some of the steps described above are not required for certain recovery strategies. The mechanism of the recovery strategy itself is the reason for it. A description of the technical alternatives is covered along with the recovery goals in subsequent sections.

2.4.5 Restoration

Even while the recovery team is supporting operations from the alternate site, restoration of the primary site for full functionality is initiated. In case the original building/work area or primary facility is beyond repair, then a new site is restored. It is possible that the team members of the recovery and restoration team are common.

It must be ensured that the site has the necessary infrastructure, equipment, hardware, software and communication facilities. It is necessary to test whether the site is capable of handling full operations. The operational data must then be uploaded at this site and the emergency site gradually dismantled.

Planning for all activities described above will include defining a time span within which they must be executed. This time duration is defined keeping in mind the recovery goals of the organization. The BCP team must remember that if at any point of time, they exceed this planned time, then the contingency must be escalated to the command centre at once, and immediate solutions must be worked out, or else they might miss their recovery targets.

2.5 Goals Definition

At the end of the phase of Risk Assessment and Business Impact Analysis, what stand out are the essentials to keep the business moving. Classification of the business services is available in terms of services that are:

- critical
- essential
- necessary
- desirable

This makes the Continuity Priorities clear. Goals can now be quantified in terms of:

- **Recovery Time Objective** (RTO) – maximum permissible outage time
- **Recovery Point Objective** (RPO) – the furthest point to which data loss is permitted
- Performance degradation on account of any measures introduced as a part of BCP
- Risks involved in the case of any measures introduced as a part of BCP
- Cost of implementing the BCP

These will drive the operational details of the BCP.

3 Technical Requirements

3.1 Available Options

3.1.1 Storage and Server Solutions

a) ***Conventional Backup*** – is the method of backing up various servers and shipping the tapes to a safe alternate location.

b) ***RAID*** – is an effective solution for redundancy. Based on the needs, RAID of an appropriate level can be chosen.

c) ***Remote Journaling*** – is the process of collecting the writes to the logs and journals and transmitting them to a remote site. It can be done in real time, i.e., by simultaneously transmitting the writes, or can be achieved by extracting the writes and periodically transmitting them. It *does not update* the database but only sends the logs so that recovery can be achieved to the point of last transmission. This can mean nearly no loss of data. *Remote Journaling* is not a stand-alone method and needs a starting point on which the logs are applied.

d) ***Electronic Vaulting*** – transmits data electronically and automatically creates the backup offsite. Automation implies that no transportation of backups needs to be done, as data reaches offsite faster, and backups can be performed more frequently without the need for manual supervision.

Vaulting at the hot site gives a very good RPO. When used with *Remote Journaling* – also sent to the hot site, it drastically reduces the outage time, thus achieving good RTO too. Electronic vaulting requires adequate bandwidth. If the transmission costs are a hurdle, then the vault can be at a more local secure site. It would help greatly to have the alternate site fairly near the hot site, as it reduces the time taken to transport the backup to the hot site at the time of disaster. If the vault is not at the hot site, then the organization can arrange to periodically ship the backups to the hot site. Not only will this improve the availability, it will also safeguard against a situation when a calamity affects transport to the hot site.

e) ***Disk Replication (Mirroring, Shadowing)*** – is the process of writing the data onto both the primary server and the replicated server. It protects from disk failures and provides an up-to-date copy of the data, thus providing an excellent RPO. If the replicated server is placed at the hot site, only logs of uncommitted work need to be run for complete recovery, i.e., the RTO is reduced to practically nothing.

How do they work? There is a "heartbeat" exchanged between the primary and the replicated server at periodic intervals. When the replicated server fails to receive one, it signals a possible problem in the primary server. This alert can cause either a manual planned switch to the replicated server or can be set up for an auto switch.

There are two options here: mirroring and shadowing.

- ***Mirroring*** is a synchronous process, i.e., the changes are applied on the replicated server synchronous to the primary server. This is the best method when the requirement is for "zero data loss". A drawback of this mechanism is that it can cause performance degradation and requires adequate bandwidth. Ideally, it is not used over long distances, when transaction volumes are high, when the bandwidth available is too little or the network latencies are high.

- ***Shadowing*** is an asynchronous process. The changes are collected in the form of logs. These logs *are periodically applied* on the replicated server. The RPO of this method comprises the last transmission and application of the changes, and the RTO consists of the time taken to apply the remaining logs. This method provides an RPO and RTO that is worse than *Mirroring*; but that does not affect the performance of the primary server and can function even with smaller bandwidth or with networks having latency.

Data replication solutions may be used at the database, file system, operating system or application level, based on the needs. When choosing replication as a solution, one needs to remember that replication has to be planned for each server that matters and that any corruption or loss at the primary server is also replicated at the replication server.

f) ***Clustering*** – is considered a solution for high availability. Its purpose is to use a secondary server to provide access to applications and data when the primary server fails. If it is a local cluster, then the two servers exchange a "heartbeat" periodically to indicate their status. Clusters can be *symmetric* or *asymmetric*.

In ***asymmetric clusters***, the secondary server remains inactive under normal circumstances. Both, the primary and secondary servers have access data and applications using shared, or mirrored storage. In the event of disaster, the secondary server becomes active, its access to the storage and network gets activated and it takes over the functions of the primary server.

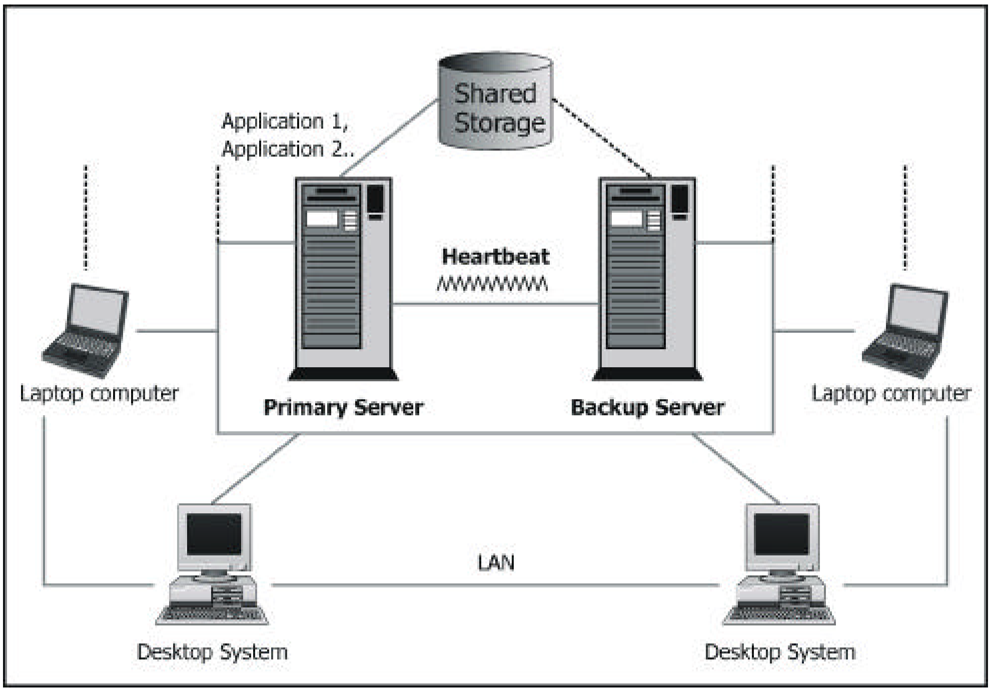

Figure 1: Asymmetric Cluster

In ***symmetric clusters***, the two servers are always active, each one running its set of applications. In the event of the failure of one server, the other takes over the functions of the failed server as well. It could also be that both servers run the same applications. In normal times, they serve as a cluster for load balancing, and in the event of disaster, they take the entire load.

It is also possible that the primary and secondary servers do not share the storage. This is done when the two servers are running independent applications. The backup server is provided access to a server's storage as soon as it goes down. Another option is to mirror each server's storage onto the other server's storage. In the event of disaster, the backup server can access the mirror.

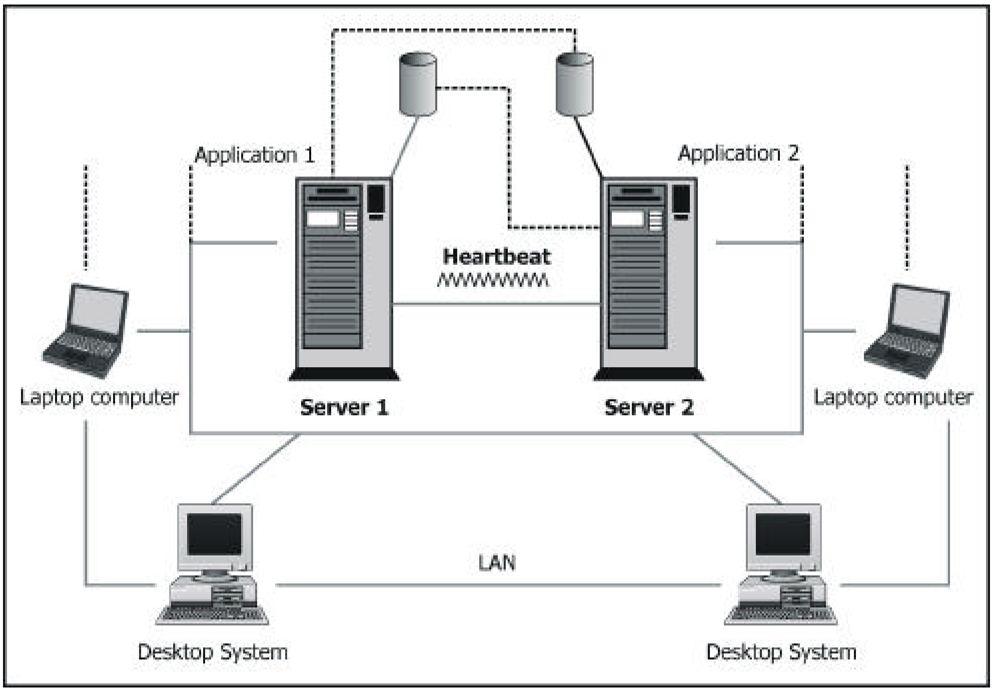

Figure 2: Symmetric Cluster

It is evident that a clustering solution has the best RPO and RTO. In addition to being a high availability solution, it also provides the possibility of load balancing. But it is an expensive solution and cost benefit analysis is suggested before taking it up.

g) ***Standby Operating System*** — is the process of having the operating system loaded and ready in a disk that can be attached to the machine at the alternate site. This method, when used with other techniques, can save the time and effort required getting the operating system ready in the backup server, after a disaster.

h) ***Storage Virtualization*** — provides a single logical view of all the storage devices irrespective of the actual configuration or physical location. It minimizes the differences between the underlying systems and simplifies the administration of storage. In the conventional scenario, it is not always possible to match the storage needs with the application type; e.g., file servers consume space faster than mail servers. Also, any major increase in storage usage requires an entire hardware upgrade. In storage visualization, since the logical and physical views of the storage are separated, it is possible to easily upgrade, replace or add more storage without affecting the server availability. It causes true sharing of data as it allows heterogeneous servers to connect to the storage. It also offers the benefits of *serverless backup and recovery*. Virtualization complements the Network Attached Storage (NAS) and Storage Area Network (SAN).

Network Attached Storage (NAS) — a storage device with a built-in network interface that can be plugged into the network to provide access to data. It supports all the file service protocols to share files across systems, and is easy to install and maintain. Enterprise systems' storage of this form are proprietary, therefore every time an upgrade is needed, it is necessary to go back to the manufacturer and also be tied to a specific vendor. Data transfer and backups need to use the network.

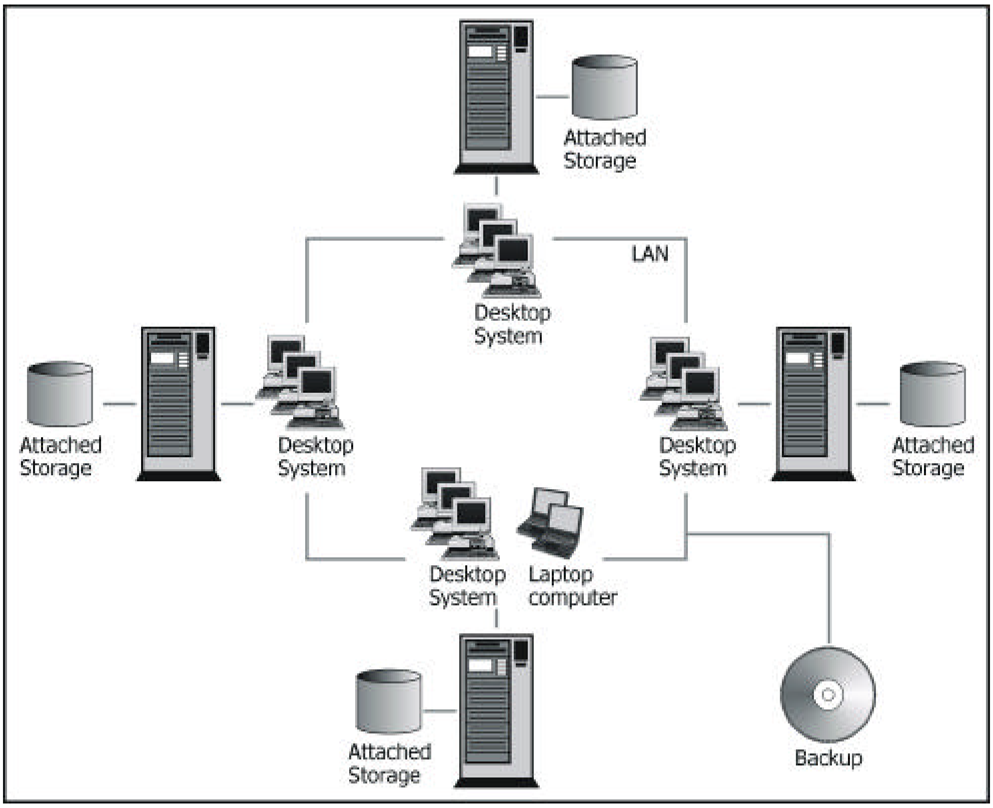

Figure 3: Network Attached Storage (NAS)

Storage Area Network (SAN) — is a dedicated, high-speed network based on the fibre channel, switches and hubs that connect many heterogeneous servers to storage devices. In effect, the storage devices are removed from their servers and are available to many servers across a network. Besides, it increases the manageability of storage and allows easy addition of storage or servers. It further removes the vendor dependence. The largest chunk of the LAN traffic has been observed to be pertaining to backup, mirroring, "heartbeat" and disaster recovery-related activities. With SAN, all these housekeeping activities are done off the LAN and happen over the SAN fibre. This frees the server power and the network for applications.

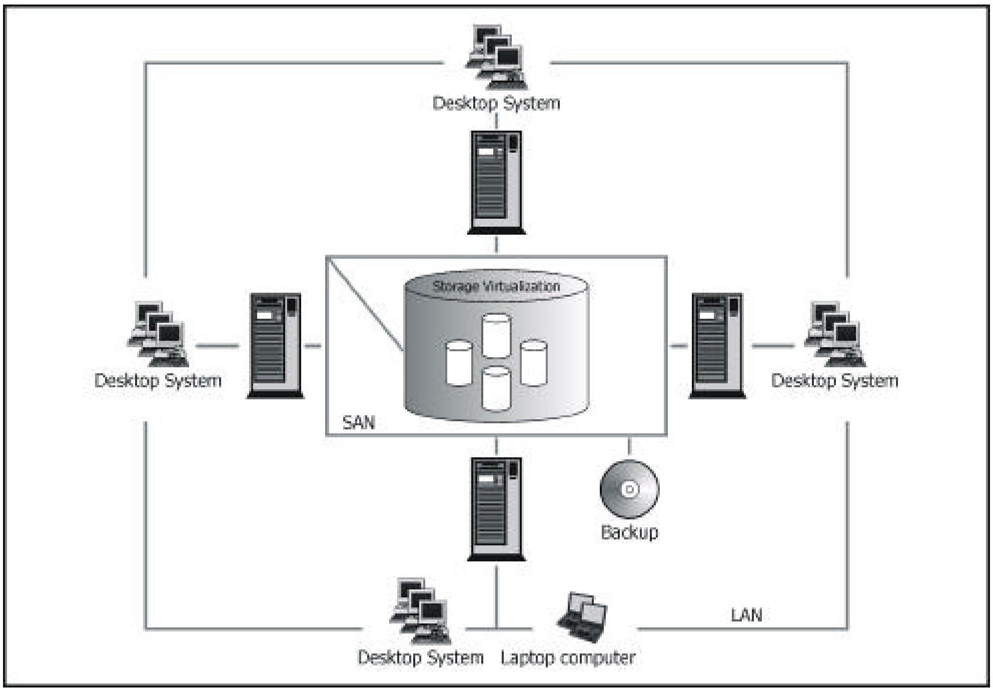

Figure 4: Storage Area Network (SAN)

The many-to-many server-to-storage connectivity makes it easier to plan for continuity. Applications of a server that go down can easily be taken over by its backup server, since connecting to the storage is not a problem. As a result, a very good RTO can be provided. Since backups and replication/mirroring activities happen reliably and rapidly over the SAN, the desired disaster recovery copies are available and easy to access, its RPO too, is excellent.

3.1.2 Network Solutions

a) ***Hot Network Nodes*** – implies having an operational network at the alternate facility. This can be regularly monitored for function. It saves the time to set up and test the networks after disaster. This method is used along with other techniques.

b) ***Virtual Private Networks (VPN)*** – are used for WAN recovery. VPN runs over public networks and allows access to a corporate network. Once an entry into the corporate network is accomplished, it is like using their Intranet via Internet. The tunnel request to the destination server from the Internet first authenticates the user. Besides this, encryption and the Internet tunneling protocol serve as security measures. In the event of disaster, VPN allows the recovery team to work on the recovery server even before they reach the alternate site.

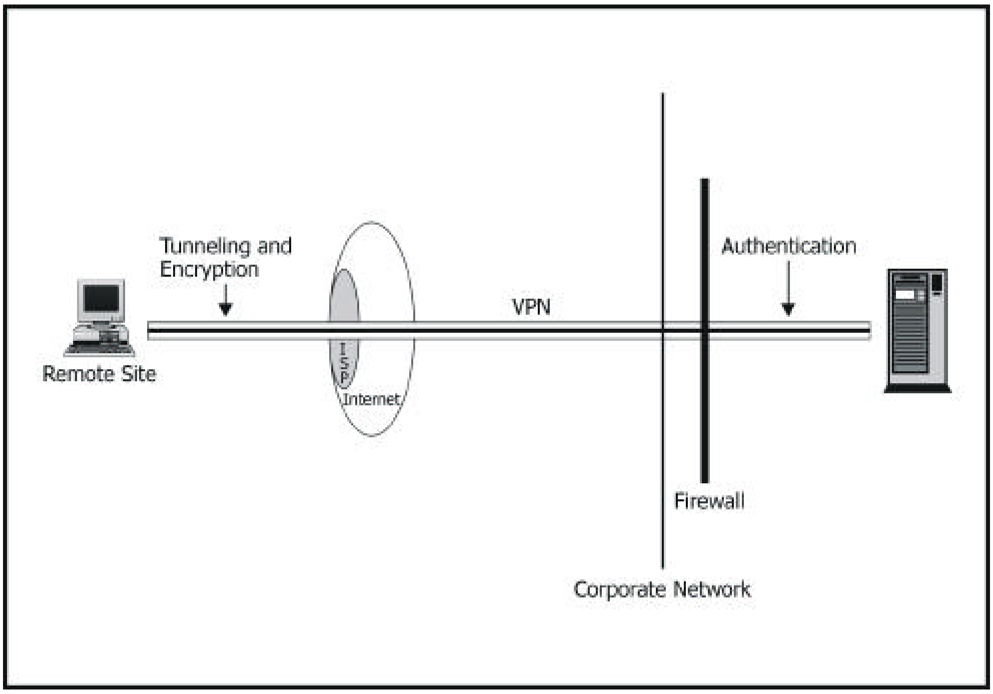

Figure 5: Virtual Private Network (VPN)

3.2 A Comparison

A comparison of the availability provided by the various continuity options can be done.

1. Conventional Backup — requires shipping tapes to site, restoring the operating system, staging and restoration of data, and recovering transactions.

2. Standby Operating System — requires shipping tapes to site, staging and restoration of data, recovering transactions.

3. Electronic Vaulting — requires restoring the operating system, staging and restoration of data, recovering transactions.

4. Remote Journaling — requires shipping tapes to site, restoring the operating system, staging and restoration of data and recovering transactions. The time for all these activities depends on the base on which the logs and journal will be applied. It is generally done in combination with vaulting.

5. Shadowing and Mirroring – requires switchover, recovering transactions.

6. Clustering and Virtualization – requires switchover

The rating, as far as availability goes, is fairly obvious. Cost is also a factor that influences selection. This list gives the solutions in the order of increasing cost.

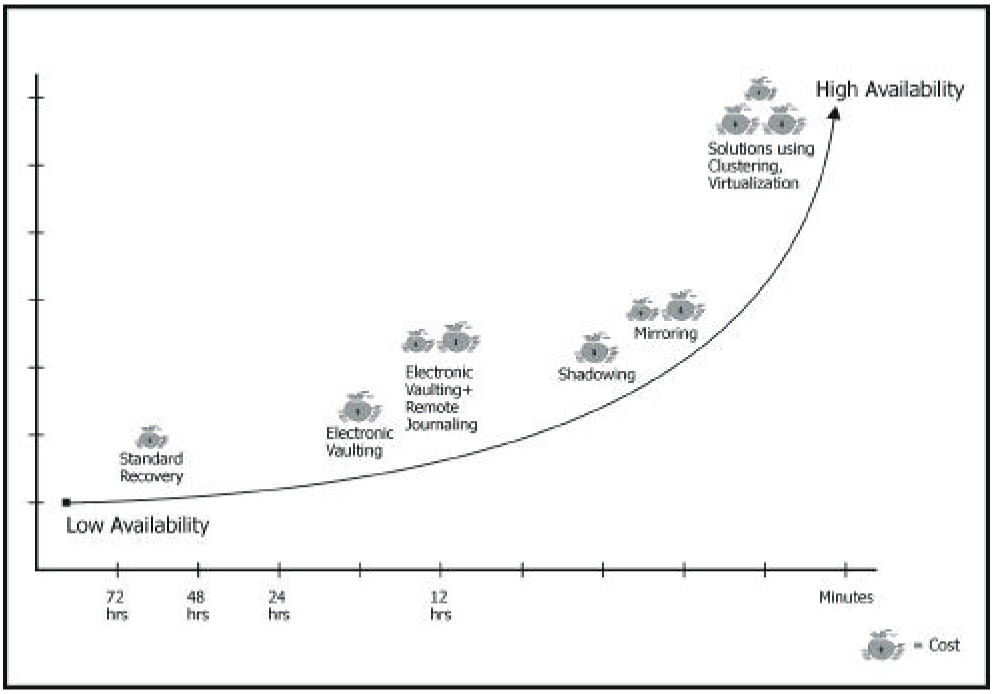

Figure 6: Availability Comparison

The BCP goals (Refer to section 2.5) would be given a ranking based on the continuity priorities. This ranking, in turn, drives the choice of the technology used to achieve them. At times, there can be conflicting goals. For instance, the primary goal is to have a good RPO, but without a fair amount of performance degradation and cost. For example, disk mirroring is the answer to RPO, but it may not measure up to the goals of performance or cost. In such a situation, the priorities must be weighed to arrive at a best-fit option.

3.3 Implementation

The type of the target IT system or Business Process also decides the technical means of achieving the BCP goals. It is important to formulate the actual method based on the results of the business impact analysis.

(a) Desktop computers

They are not often subject to contingency planning. However, if necessary, desktops can also be planned for.

Users of the desktops may be directed to take backups of their data. If this is not a very acceptable option, then a networked disk can be used. Instructions are circulated pertaining to the folders in the networked disk in which the required data should be saved. The networked disk is backed up at a predefined frequency. The backup of the networked disk is taken to an offsite storage site.

It would help the restoration process greatly if the system and application configurations of the desktops are known. For this purpose, the documentation on each desktop should be maintained. If possible, all machines in a business unit can be of similar configuration and setup, or the minimum setup for all machines of the same type can be the same. This way, while restoring the desktops, the basic setup can be quickly done and followed by recovering from the individual's backup on the networked disk.

Desktops containing highly sensitive data can have power-on passwords or even use encryption.

(b) Software and their licenses

These are procured initially at a cost and so must be backed up and stored at an offsite storage location.

(c) LANS

Restoration requires that the configuration of the network and all the devices used be documented. Preventive measures in the form of network security controls and identification of single points-of-failure need to be done. A possible contingency strategy can be remote access using virtual private networks (VPN). VPNs provide a secure, centrally-managed remote access over a broad range of browser-based and browser-less applications. Hot network nodes can also provide real-time switchover. For these solutions, data bandwidth will need to be scaled appropriately.

(d) Servers

The loss of servers can be devastating, as they host all the applications and support several users. This is the most important area to focus upon while performing contingency planning. Adequate documentation of the system and applications hosted must be kept to enable ease of recovery. Adequate security measures must also be implemented to prevent malicious code attacks.

Backup of the data and application must be stored offsite. A policy to label the media, and retain the backup and test retrievals must exist. A backup schedule, stating when full backup and when incremental or differential backup should be taken, also needs to be formalized.

Besides the conventional backup (scores lowest on availability), strategies for implementing redundancy, and replicating data or other storage solutions can also be planned.

Redundancy can be implemented using RAID. It increases reliability by spreading data storage across multiple disks and the drive swapping can be done without bringing down the system.

Remote journaling, electronic vaulting and disk replication (either mirroring or shadowing) provide various solutions for replicating data.

Alternate storage solutions can be in terms of storage virtualization, NAS and SAN.

(e) Websites

Websites are often the entry point for hackers, so necessary security controls must be implemented. Documentation of the website, and its configuration, along with the application code, would speed recovery. During recovery, it is possible that the website is recovered from an alternate site with an IP address different from the usual. So care should be taken to anticipate this during the initial design and development of the site. Clustering across two servers is also a popular option that offers dual benefits: load balancing at normal times, and failover during disaster.

(f) Business Process

A thought can be given to build redundancy into the business process itself. Having this will help to counter a scenario when a failure occurs in the normal process flow. Planning for redundancy in a business process is best done right at the time of designing the process. However, it is possible that the need for it comes to light only while assessing it for continuity. The mechanism could just be switching to manual operations until the technology component is up and running, but the means to switchover smoothly must be in place.

Consider a firm providing financial services. Their services can be availed of by customers only over the phone. They have to safeguard themselves against shutting down customer interaction completely, during telecom failures. Therefore, they would provide the customer with more than one phone number, use more than one call centre at fairly distant locations and use more than one service provider. This firm may have to consider introducing another way of interacting with customers. They could also consider postal services and SMS as alternative means of communication.

Consider some examples of how a bank can build in redundancy in order to retain business or not lose too much of it when failures occur.

- Selling a demand draft or a fixed deposit to an account holder – Banks can offer both phone banking and net banking. If their net services are not accessible, the customer can always use the phone and vice versa. This can hold true for many of their services. In addition, there can be emergency help phone numbers that the customers can reach.

- Withdrawal from an ATM – Generally, in case of failure of ATM machines, the bank will inform its customers about the nearest branch or ATM that they can use; however, if it is a switch failure, then a stand-by switch (if it is financially feasible) is needed to overcome the situation. Another possibility is that the bank provides a cash delivery service that can be ordered using the Internet or phone.

Banks can have arrangements with other banks to provide some of their services. This can be planned for locations where the nearest bank/ATM is itself distant or the phone numbers are not reachable. Banks can choose to charge for this service.

Building in redundancy always involves costs. As a result, business firms mainly implement these measures for the most frequently used services or the most profitable processes. To come up with a list of processes where they can invest in continuity measures, they have to study the usage patterns and analyze their product/fee structure. Feedback from the customers is of use here. It will give them an insight into what the customer perceives as the most needed process and the painful or problematical areas. They will also need to regularly communicate with their customers to make them aware of the options they have.

A completely generic solution for building in redundancy in business processes may not be possible. The mechanism for redundancy will vary from business to business and also depend on the factors that emerge during the risk assessment and business impact analysis.

4 Software Development Life Cycle Approach to BCP Activities

A common practice is that business continuity planning is thought about after the system is either designed, or more commonly, after it is operational. Retrofitting the BCP measures can make it a costly exercise. If possible, BCP should be integrated into the IT project's life cycle. This will ensure a complete coverage by the measures designed as part of the BCP and naturally will be more cost-effective.

Requirement Analysis is the time when the expectations from the system are gathered. Along with the functionality desired, continuity expectations may also be observed. For example:

- A system working on very sensitive data *cannot afford any data loss*
- A system accessible across the globe must have *high availability and minimum possible outage*
- An unmanned system or one operating in remote conditions must be able to *diagnose a disaster, send notification and attend to itself until recovery is initiated by its base station*
- A system for which migration or conversion is being studied *already uses some BCP measures*

Such clues help in arriving at the BCP objectives and their success criteria.

Analysis and Design is the phase when each requirement is studied in depth and translated to an operation in the system. Failover, load-balancing, replication, or defensive mechanisms for fault tolerance can all be designed during this phase. All these are common measures for continuity but they are not very easy to introduce in a system that is already operational.

While design for each business scenario is done, continuity strategies are also to be designed and evaluated. The technical means for achieving continuity are designed as a part of the technical and application architecture. Costs and benefits of the measures are analyzed at this point. If more than one alternative exists, an evaluation will indicate the better option. If only one strategy is visible, then ways to reduce its negative impacts are studied.

Implementation consists of construction and testing of the IT solution. The software and/or hardware pieces for business continuity are also built/plugged into the system during this phase. While testing the system being developed, it is most important to plan different scenarios to test the continuity measures. The test results are the proof that the planned measures are sufficient and complete. If gaps are observed, then they must be closed and the testing exercise repeated.

Now that the visualized continuity measures have been verified, the operational details of the BCP can be written. The plan is enriched with equipment details, vendor details (if any), contact lists, etc., at this stage.

Maintenance is when the system is operational. It is important to ensure that the BCP is still relevant, the measures put in place by the BCP are also operational and the BCP team still remains capable of handling the interruptions. The BCP must also reflect the effects of any changes done to the system as part of normal maintenance. A mock test run of the BCP can be scheduled at regular intervals to make sure that the BCP is still sufficient and complete.

Retirement is the exercise of phasing out a system. The BCP measures in place are also dismantled with the system. A system could be retired either because it is being replaced by another system or because it has served its purpose. If it is the former, then care must be taken to ensure that there are continuity measures in the new system also, and no gaps have been introduced with the system change. And if it is the latter, then the experience serves to provide best practices and lessons learnt that could be utilized in other projects and systems of the business.

It is recommended that IT projects integrate BCP activities into their SDLC. At the least, each phase of the project can identify the continuity requirements (as much as is visible at that point of time) and document them as a part of the Non-functional Requirements. This can serve as the starting point for the BCP exercise.

5 Developing a Plan

5.1 Typical Contents

This section provides a view of the contents of a typical business continuity plan. There could be separate plans for each of these: prevention, response, resumption, recovery, and restoration or each of these could constitute different chapters in the same plan. If there are going to be separate plans, then the essential items, detailed in section 5.1.1, form a part of each of them.

5.1.1 Essential Items

- ***Objective***

The purpose for writing the plan is stated. It also states what phases the firm intends to have and their interpretation of the objective of each phase.

- ***Scope***

The divisions or business operations that fall within the scope of the continuity planning are stated here. If the continuity planning is only for certain calamities and not for total disaster, then a mention of the special scenarios being handled must be made.

- ***Pre-requisites/Assumptions and Limitations***

The assumptions based on which the plan was formulated are stated here. In some ways, these also form the pre-requisites for the continuity plan to be a success. For example, that the server backups available will not be older than 'n' hours or that a team, which is trained to handle the recovery operations will be available on call, or, that the alternate site, with valid vendor contracts, will be ready for use within 'x' hours, are some assumptions.

If the plan has some limitations, then they should also be documented.

- ***Team***

The organization/hierarchy of the BCP team, their sub-teams, roles and responsibilities are stated here. It also mentions which level comes in at which phase of the process, i.e., response, recovery, etc. For example, if there are separate plans for each phase, the teams responsible for the relevant phase must be shaded/colour-coded to indicate the same.

- ***Goals***

As a policy, the firm will have RPO and RTO goals, which are stated to customers and shareholders. These are stated here, along with the performance goals.

5.1.2 Preventive Safeguards

The preventive measures to be implemented as a part of the continuity planning are detailed here. They can be organized in terms of:

- Surveillance
- Access Control
- Authentication
- Anti-virus
- Filters
- Intrusion Detection systems
- Backup plans

(Refer to section 2.4.1 for more details.)

5.1.3 Emergency Response

- ***Needs for the initial response***

The resource requirements for the response phase are listed here along with the detailed configuration and quantity required. If hard copies of documents and forms are needed, they must also be mentioned here.

- ***Notification Call Tree***

It is a good practice to mention the position titles rather than the names of people in the notification list. This minimizes maintenance of the plan itself. An appendix can contain the contact list. The list will clearly state the position, whether primary/alternate, timeout duration for response, address, phone/cell numbers, email and any other possible modes of contact.

- ***Damage Assessment***

The procedure used to assess the damage is detailed: e.g., all the things that must be inspected, the kinds of evaluation to be done—whether salvage/replace, the reporting to be done, and the time required, among other activities.

- ***When to declare a disaster***

Conditions that must be met for the event to be declared a disaster are mentioned here. The disaster scenarios should also be mentioned. This will make it easy to initiate the resumption and recovery processes.

- ***Plan Activation Criteria***

The circumstances under which the business continuity plan is activated are outlined here: e.g., the network or server will not be up before 24 hours, or the primary site cannot be restored before 48 hours, etc.

(Refer to section 2.4.2 for more details.)

5.1.4 Resumption

The procedure for transition from the emergency response to business resumption is given here. The process of making decisions regarding operations, concerning where and how they would be deployed, and the activities to be performed and to what extent, are described. Activities are assigned to the different sub-units in the BCP team and each group performs its assigned tasks. (Refer to section 2.4.3 for more details.) This part of the plan is also called the Business Resumption Plan (BRP).

5.1.5 Recovery

The procedures to perform recovery are stated here. This part of the plan is called the Disaster Recovery Plan (DRP).

There could be many ways to organize this part of the document. One possible way is to list all the recovery goals (in terms of RPO, RTO, target server/network). For each of these, the target division of the organization involved is mentioned, followed by the identification of the team/role and their tasks. It is also possible to organize this section on the basis of the department/division. Either way, it must ensure that all the BCP objectives are covered. (Refer to section 2.4.4 for more details.)

This section of the plan must be more like an operations manual. It should be a simple sequence of instructions that can be followed to perform recovery. Any dependencies among the activities must also be clearly stated. It must be fairly detailed to avoid mistakes that can result in time loss.

5.1.6 Restoration

The steps for restoring the original site for business are described here. Responsibilities are marked against each team/role. The process of performing the parallel run with the alternate recovery site along with the procedures for comparing results from the alternate site with the restoration-in-progress site are described. The criteria for switching to the original site and dismantling the alternate site are also stated. (Refer to section 2.4.5 for more details.)

Appendices can be used to mention the vendor contracts, and the business continuity measures of these vendors, information that is needed because the success of the BCP depends on the availability of the vendors. Furthermore, a description of the alternate site facilities, addresses and phone numbers of the control centre, and the contact list for notification could be presented in the appendices.

5.2 Testing

The business continuity plan needs to be tested for adequacy. It is an exercise that must be carried out periodically. Ignoring this exercise would mean that the plan gets tested only when disaster actually strikes. This is certainly not a risk that any firm can take.

It is important to test the business continuity plan every time the plan is revised, a system is added to the production, a system in production is changed, a scheduling change occurs in the business activities, a process falls in the scope of the plan changes, a requirement arises for reporting on the adequacy of the plan, or when reporting on the preparedness of the business continuity team.

Testing the BCP is the only way to see if the goals set have been met. These tests will throw up inconsistencies, incorrect information (if any) and points where the actual and expected results differ. The team can brainstorm on the gaps found and revise the plan accordingly. This would lead to yet another test cycle.

This exercise of testing the plan also serves as training for the team. It is possible that at the time of disaster, the team may not be able to refer to the plan. Moreover, there would be panic and human response could be affected. At such moments, the earlier "drills" would serve to remind them of their activities and provide some level of confidence.

Planning a BCP test will involve defining the following:

- Test Scenario — defining the disaster that is to happen as a part of the testing.
- Test Plan — defining the audit schedule, the set of test scenarios, the type of exercise, or the participants, i.e., the primary team, or a mix of primary and alternate teams.

The test exercise can be a *checklist exercise* or a *tactical exercise*.

Checklist exercise involves a structured walkthrough. The team comes together with a prior knowledge of the test scenario. Each member plays a designated role and walks through the activities assigned to him/her in the continuity plan.

Tactical exercise involves an actual simulation. There will be a coordinator for each test, who will announce the intermediate events for the scenario — as if they were happening. The team goes through the entire plan, performing all the activities, from notification to restoration. If this exercise is a planned or notified exercise, then it will be generally designed in such a way that the activities of the entire team are covered. Surprise simulations can also be performed. It is usually the last type of exercise in the testing of the continuity plan and gives a picture of the actual preparedness of the team.

In short, while testing the BCP, the following activities are performed.

- Prepare a test plan, choose the test scenario(s) and state the expected results
- Execute the plan
- Document the test results
- Review the actual results and report gaps and/or slippages
- Circulate the results and the report among the team
- Identify the changes that are to be made to the BCP to cover gaps and overcome observed slippages
- Train the team (this is an activity performed whenever the BCP undergoes an update)

The **Test Plan** would normally contain the following information.

- Test plan identification
- Test scenario(s) selected
- Type of exercise, e.g., walkthrough, surprise simulation, etc.
- List of participants
- Sections of the BCP operations that are to be executed
- Expected results and expected timeframes for achieving them
- Actual results and actual elapsed time
- Gaps and slippages observed
- Recommendations

5.3 Maintenance

Maintenance of the BCP is an often-ignored activity. Preparing and testing the plan does give an assurance that there is a means of getting back to normal operations when disaster strikes; however, unless the plan keeps pace with the changes in the business processes and the technology implementing the business, it will not prove reliable.

The BCP must be reviewed periodically. It is mandatory to review it when a system is added to the production, a system in production is changed, a process that falls in the scope of the plan changes, or there is a change in the schedule of the business activities. Besides these events, a change in the contact person list can also trigger an update.

Review of the BCP may also happen with the intention of improvement, e.g., in the course of documenting the lessons learnt during the testing exercises, the organization revising its continuity goals and deciding to move up on the “availability” spectrum, or when an alternative method of doing things has been evaluated to give better results. So, maintenance of the BCP is done with the intention of both change and improvement.

Every revision to the plan must be followed by a distribution to the BCP team, a training update and a testing exercise.

Note: The resources involved in the business continuity activities — people and equipment — are also subject to maintenance; people by means of training and test runs, and the equipment, by means of regular maintenance procedures. Only if they are in good shape will they prove reliable and dependable when the moment to perform arrives.

6 Conclusion

Even though Business Continuity Planning appears to primarily deal with technology, it is equally associated with business. It is true that the operational aspect involves technology, but knowledge of technology alone is not sufficient for this exercise. It includes activities in risk management, crisis management, identification of business processes, impact analysis, cost benefit analysis, storage management, network management, continuity planning, recovery planning, training, communication and coordination. The team involved in business continuity planning should ideally be a cross-functional team with adequate domain knowledge, expertise in system and recovery management and skills in planning.

The following diagram summarizes business continuity planning.

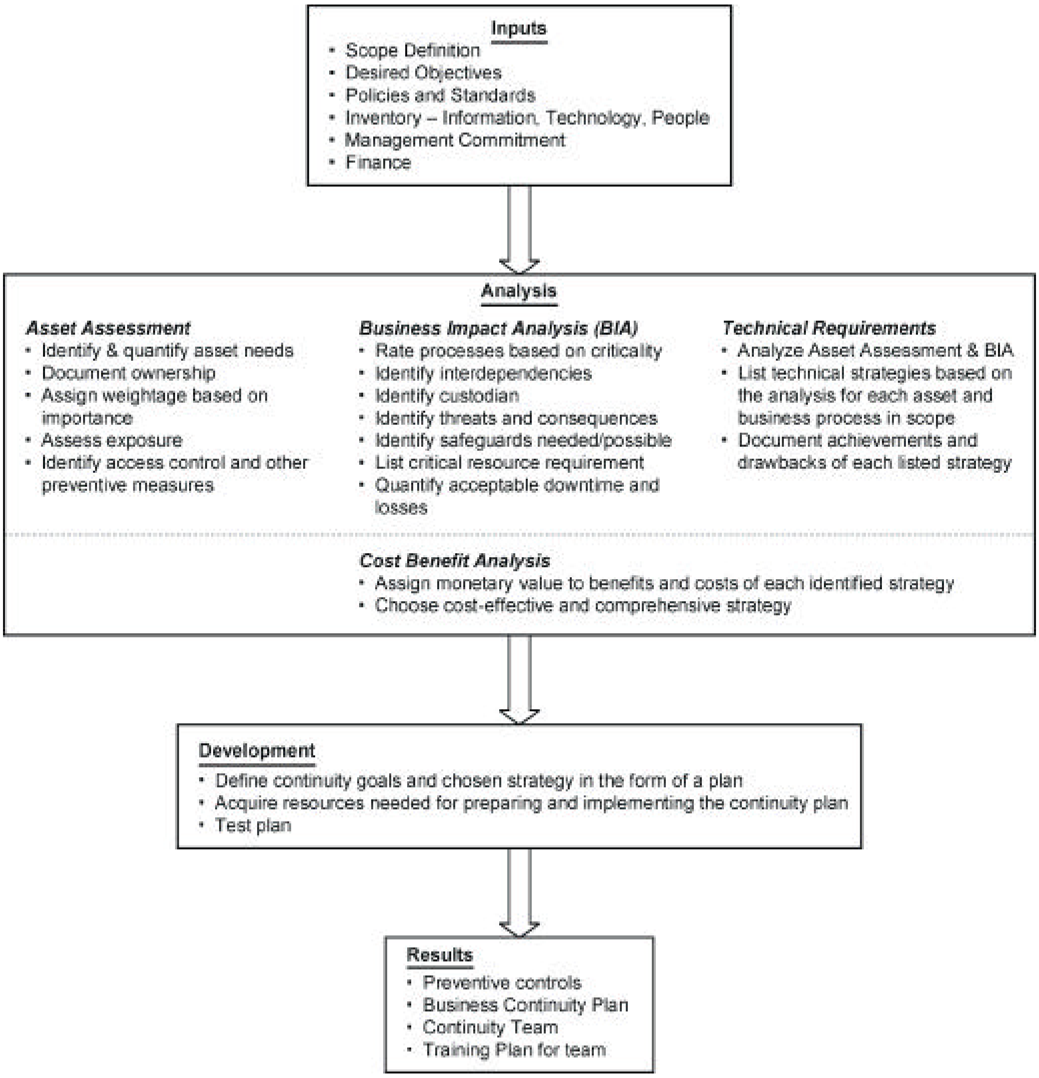
Inputs
• Scope Definition
• Desired Objectives
• Policies and Standards
• Inventory – Information, Technology, People
• Management Commitment
• Finance
Analysis
Asset Assessment
• Identify & quantify asset needs
• Document ownership
• Assign weightage based on importance
• Assess exposure
• Identify access control and other preventive measures
Business Impact Analysis (BIA)
• Rate processes based on criticality
• Identify interdependencies
• Identify custodian
• Identify threats and consequences
• Identify safeguards needed/possible
• List critical resource requirement
• Quantify acceptable downtime and losses
Technical Requirements
• Analyze Asset Assessment & BIA
• List technical strategies based on the analysis for each asset and business process in scope
• Document achievements and drawbacks of each listed strategy
Cost Benefit Analysis
• Assign monetary value to benefits and costs of each identified strategy
• Choose cost-effective and comprehensive strategy
Development
• Define continuity goals and chosen strategy in the form of a plan
• Acquire resources needed for preparing and implementing the continuity plan
• Test plan
Results
• Preventive controls
• Business Continuity Plan
• Continuity Team
• Training Plan for team

Figure 7: Business Continuity Planning - A Summary

Business Continuity Planning is not just relevant for business organizations that offer services using IT and that handle a lot of data. It is also significant for IT service providers. Their development centres and support units have a wealth of knowledge and all their past experiences are in the form of sources and documents on the servers and tape libraries. They have a dual responsibility — to plan for their own continuity, as well as that of their customers.

Planning for continuity is about being safe — safe from the consequences of events that one hopes will never happen; and the truth is — it is always better to be safe than sorry.

Appendix A: Sample Business Impact Analysis (BIA) Questionnaire

Process Name	
Identification	#BP-003
Custodian	
Description	
Acceptable Outage	
Importance to Business	Critical / Essential / Necessary / Desirable
Depends on	Process id. Int./Ext. (Low/Medium/High)
Dependent Processes	Process id. (Low/Medium/High)

Critical Resources

People

S. No.	*Role*	*Responsibility*	*Weightage (Scale 1 – 5)*
1.			
2.			
3.			
4.			
5.			
6.			
7.			

Hardware

S. No.	*Item*	*Configuration*	*Quantity*		*Acceptable Outage*	*Outage Impact*
			Normal Operations	*Critical Situations*		*Low/Medium/High*
1.						
2.						
3.						
4.						

Software

S. No.	*Item*	*Version*	*License*		*Acceptable Outage*	*Outage Impact*
			Type	*Quantity*		*Low/Medium/High*
1.						
2.						
3.						
4.						

Documents

S. No.	*Name*	*Type* *Paper/Electronic*	*Copies*	*Acceptable Outage*	*Outage Impact* *Low/Medium/High*
1.					
2.					
3.					
4.					

Business Impact

S. No.	*Area*	*Impact* *None/Low/Medium/High*
1.	SLA Violation	
2.	Industry Ranking	

Acknowledgements

I thank Dr. Gautam Shroff, S. Santhanakrishnan, C.S.R. Krishnan, Sanjeev Sachdeva, Hema Shriram, Shyam Chellaramani and Shashank Chowdhary for their comments and contributions to this work. I also thankfully acknowledge Payserv AG and Tata Consultancy Services for giving me my first exposure to BCP.

References

Disaster Recovery Journal: http://www.drj.com/drj2/drj2.htm

Gartner Research. “Business Continuity and Disaster Recovery Planning and Management: Perspective” October 8, 2001

Gartner Research. “Integrating BCP into IT Project Life Cycle” June 15, 2001

NIST Special Publication 800-34. “Contingency Planning Guide for Information Technology Systems”

Moore, Pat. “How to Plan for Enterprise-Wide Business and Service Continuity”: http://www.disaster-resource.com/content_page/articles.shtml

Sharp, John. “The Ten Certification Standards for Business Continuity Practitioners”: http://www.business-continuity.com/faqs/faq2.html