Name: Allen Fernandes

GSEC Practical Assignment
Version 1.4b – Option 1
Date submitted: April 17, 2004

Title: The layered approach in achieving high availability

TABLE OF CONTENTS:

1.0 Abstract:

This paper is written with focus on showing how the layered approach plays an important role towards achieving high avai lability.
I will start with a short background, followed by explaining the broad view necessary to be considered, both from a user and a resource perspective. Will then start to explain the layered approach I think is necessary in achieving high availabili ty. I will also provide helpful quick tips; some specific to availability and others more general, but important ones to be considered. Finally I will end with a short conclusion.

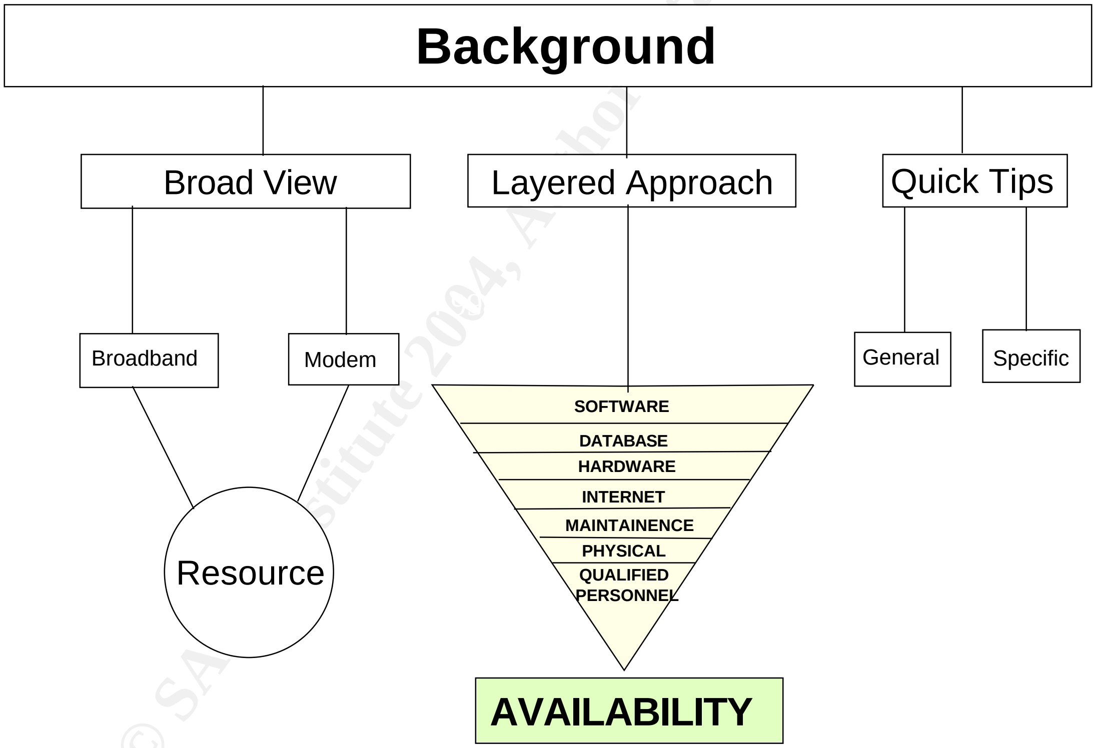


Figure 1: Abstract Diagram

2.0 Background:

In today's economy the availability aspect and its implications is a key factor that needs to be considered very carefully from the very start and not as an afterthought that can be added at the end of the process. Availability often considered by many as a technology solution that can be put in place once and then forgotten about, is not true as it is only a small piece of the puzzle in achieving high availability [1]. People and Processes are also common key factors, which play an important part in this puzzle. The global economic growth that now heavily depends on the communication medium for its very existence relies in turn on having its resources highly available. Certain vertical sectors, for example, cannot allow computers that are responsible for monetary transactions to stop functioning for even a few minutes as this could result in loss of revenues that could cost these companies from thousands to millions of dollars, depending on the total amount of time the business is brought down to a standstill causing the loss of availability.
In order to achieve the level of availability that is required by you, there is one important factor, you need to take into consideration, no matter what solution you choose or how big or small your company is.

- You need to have calculated the cost of downtime in minutes or hours relative to the loss in value. Knowing this will help you in determining your measures that will be implemented by seeing the cost benefit ratio.

3.0 Broad View:

Availability can be viewed from both the end user and the resource site perspective, both of which play an important role when talking about availability.
A resource, be it a web service made available to users on the Internet is dependent on 3 key factors for its success. First is the resource itself that needs to be accessible to be of any use. Second is the person or process that needs the use of that resource. Third is the Internet that is responsible for the communication channel between the person and the resource. All the 3 factors have their limitations and need to be considered and understood in order to achieve high availability. Having a resource highly available by itself is of little use if it cannot be utilized for the purpose it was created.
Example 1 Resource limitation: A resource may be designed to allow 300 users simultaneous access; this is a limit that the resource has. However if 1000 users try to access and use this resource, then the resource will not be able to provide its services the way it was designed to perform.
Example 2 Resource 100% available: A web site that sells computers takes all the necessary steps to make its resources 100% available for those that would like to purchase computers from them. However if people that want to purchase but cannot connect, due to uncontrolled limitations to the company's web site, does not help the company to make a sale. Thus only making a resource available does not guaranty its use.
The User and Resource limitation factors will be discussed in this chapter. The Internet limitation factor is discussed in chapter 4 the Internet layer.
The user limitations are addressed in the way they obtain Internet access, be it broadband or dial-up access.

[1] Michael Hotek SQL Server Magazine, October 2003 p – 20.

The resource limitations are addressed by understanding all the components responsible for making it available. The internal network will also be described briefly as it is often a part that is connected to the resource site.

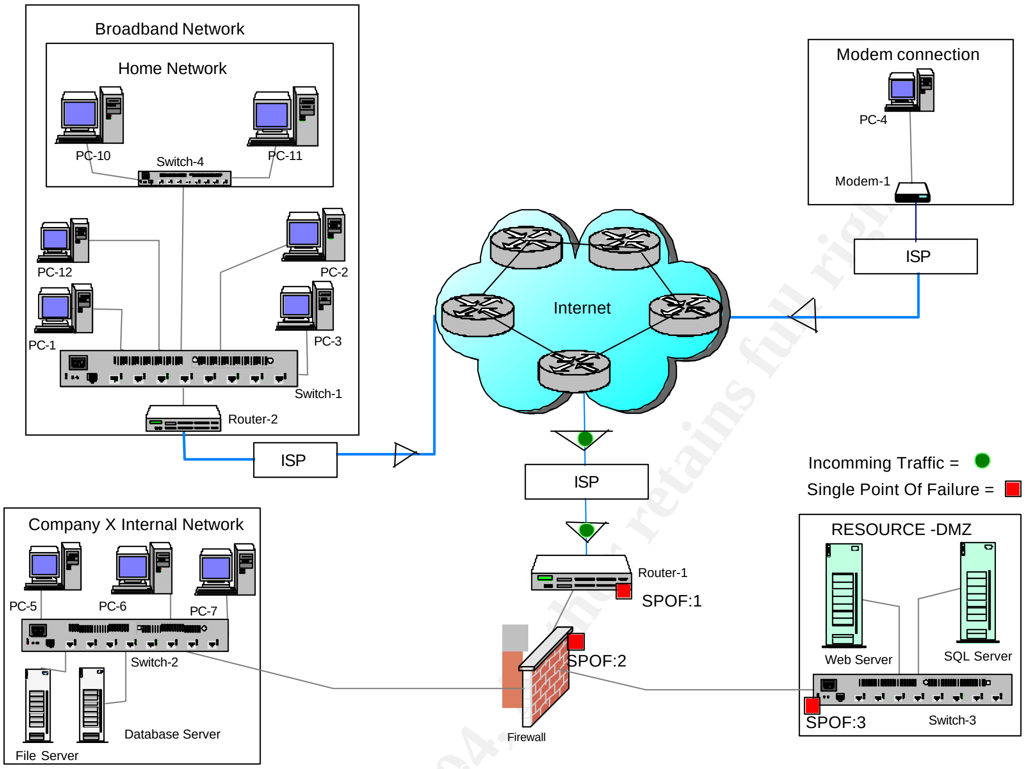


Figure 2: Broad view diagram

The following 4 aspects contribute towards the broad view.

3.1 Broadband Network Connection:

- A typical broadband network as shown in Figure 2 could be described as follows. All users are connected via a switch and thus make up a local broadband network. As all users each get a dynamic IP address allocated on start up, each user has the possibility to communicate with the other users on the local broadband network. However if users want to communicate with a web server in the Resource -DMZ they will first need to login to the ISP'S login server and on being authenticated, allowed access out to the Internet.
  A home user may like to share his connection with all other members of his family each having a computer. This would require the home user to buy a switch or hub and in this way share his one connection as shown in the home network. Each user in the home network also receives an IP address. Another solution for a home user would be to purchase a home firewall device with Network Address Translation (NAT), usually these devices are relatively cheap and also provide an inbuilt switch. This device does create a level of protection as the computers connected behind the

firewall now are allocated private addresses and thus separated from the outside network. However if this device is compromised, all your systems become vulnerable. Keep your firewall updated with vendor patches for security vulnerabilities. The recommendation for home users having a broadband connection is to have some type of software or hardware firewall at a minimum.
Local access bandwidth is in the range of 10 MB to 100 MB, howeve r Internet access is in the range of 1 to 10 MB that is often offered by many ISP'S.
What we need to understand is that the Internet access bandwidth is shared with all the users on the local network. The sharing of this bandwidth is something to consider. If 5% of the users on the network start to overload the network, it causes an availability problem affecting the balance 95% of the users. A good example is when a user that is connected 24 hours seven days a week and is downloading music files, as is ver y common today, soon starts to occupy and cause degradation of the bandwidth.
You also need to understand that even if you are not connected to the Internet your system still can be exploited by many attacks. The reason being, that once you start your comp uter and have its network cable connected to the local broadband network you are vulnerable to attacks from other computers on that network. As attacks are mostly automated today, it is no longer a question about if you will be attacked but rather, if you will be able to detect that your system has been compromised.
Employees working from home should consider, both the bandwidth as well as the security issues. Your company may be depending upon you being able to deliver some important document or on you u pdating your critical server, but if you are having problems due to bandwidth, you may not be able to get the job done. Broadband conne ctions are also vulnerable to DDOS attacks, resulting in total loss of a vailability.[2]
Having a broadband connection has i ts advantages but also it's disadvantages, so make sure you understand the consequences of a broadband connection.

3.2 Dial -Up Modem Connection:

- Dial up user connection: A dial up connection would require a user to have a modem to be able to dial into a m odem pool of an ISP. The user in this type of connection does not have the problems as described in the broadband connection. If the user is disconnected his computer is not vulnerable to outside attacks. However the user is limited to the modem speed, and does not have the high bandwidth provided by broadband. Dial up users get an IP a ddress automatically assigned on establishing a connection and lose that address on disconnection. A problem for dial up users to consider is that, if their systems are compr omised when connected to the Internet, an attacker can install software that continuously keeps the connection open, even without the user knowing. This could result in big bills for the user, as dial up connection fees are usually based upon the time, whi ch i t is utilized.

[2] David Berlind March 18 2004.

3.3 Resource Site Connection:

- Resource site connection: Both broadband and dial -up users connect to the resource site in order to achieve its requirements. A resource site is connected via its ISP to a router, which in turn is connected in most cases to a firewall. The firewall usually has 3 interfaces, the first is connected to the router, and the second is usually connected via switch/hub to a Demilitarized Zone (DMZ) that has the resources that are available for users to access via the Internet. The third interface is connected via a switch/hub to the company's local network.
- Component 1, Router [3]: As all traffic from the ISP comes through this point, this is the first Single Point Of Failure (SPOF: 1). All routers have some type of software installed that usually at some time or other, found to have some type of software bug. These bugs once discovered are then exploited and can cause the entire site to come to a standstill.
- Component 2, Firewall: Firewall is the second SPOF: 2. The main function of the firewall is to reduce the window of opportunity and only expose the services that need to be accessible. This device is expected to perform miracles or at least have the perception that it could solve all our security problems, however, this is not correct, it is only one layer in the defence -in-depth solution. Even this device is dependent on its own software that also results, at some time, to be containing software bugs causing it to be exploited once these bugs are discovered.
- Component 3, Switch for the DMZ: This is the third SPOF: 3 and is a layer 3 device that has the sole purpose to create a star network for all the other computers that are to reside in the DMZ. Switches are also vulnerable to software bugs discovered and exploited. Address Resolution Protocol (ARP) poisoning [4] is a well know technique to disrupt the correct functioning of the switch. A hub could be used instead of a switch but is considered less secure and should not be used in the DMZ
- DMZ [5]: Normally in this zone we place our Web, SQL servers. Each of these is in the zone, depending if the resource is to be made available to the public Internet. The web server is usually installed on a server with a operative system(Windows/Unix/Linux) and then the web server application(Internet Information Server/Apache) is installed. We then have the web server providing some type of web access to users, this could be html pages or asp pages depending on what solution the designers have designed this site to deliver its content to the users.
- The computer that is performing the server functions has 3 critical components. The first is the Central Processor Unit (CPU) responsible for performing all the processing; this component needs continuous monitoring to avoid saturation that causes processing to be put on hold. Second is Random Access Memory (RAM) also a key component, this component also needs continuous monitoring as it causes an undesired affect when the memory is not enough and it is then forced to use virtual memory, this results in what is known as paging and excessive paging causes degradation. Third, is the hard disk that is responsible for storing data, is sensitive to fragmentation also resulting in degradation.

[3] Advanced Information Assurance Handbook, page - 91

[4] Vulnerability Note VU#399355

[5] Advanced Information Assurance Handbook, page – 85

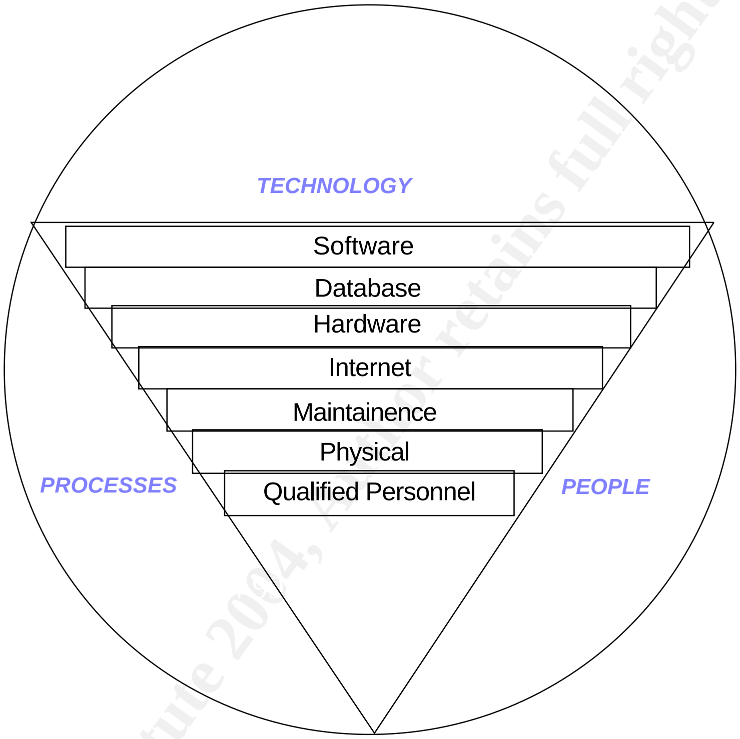
- The server operating system has a continuous demand of keeping itself updated with all the patches coming out on a regular basis. This, mildly stated is a headache but more importantly it requires a standard procedure to be implemented each time a patch is released in a test environment before actually implementing the patch on a production server. This action requires dedicated time each time a new patch is released, and it is time that is the most vital factor as should we not be able to update the patch in time before an exploit is released we are in a vulnerable position.
- The web server application has also the same problems described above namely patch management. However it has additional problems that are specific to the application. Request overload can cause the server application to freeze. Unexpected requests and malformed requests also can cause the application to go into an unstable state. Worst scenario is when the application is compromised and results in access to the underlying operating system.

3.4 Internal Company Network:

- Internal network of the domain providing the resource is where the company maintains all its crown jewels. It is this network that we often try to focus on, to keep extremely secure by implementing many different security solutions to avoid this network being accessed by the external Internet. However due to the fact that we need to give the internal employees access to the Internet does put us in a more vulnerable state as users are now subjected to directed attacks in many ways whether it is through email or through the web browser. Administrators suddenly have not only a single point to concentrate on in securing its network but also need to secure and control the various end users computers to monitor for other ways of attacks entering the network. Even after taking the measures to secure the end users pc, the internal networks could still be compromised if users have laptops that are used both at home and in office, securing these portable devices further widen the spectrum of making this network secure. Attacks directed towards end users in the internal network can have a serious impact on the availability of the resources in this network.

4.0 Layered approach: When your goal is to achieve high availability you need to think about all the layers that contribute towards making your goal possible. The combination of all the layers is a representation of Technology, Processes and People.

Achieving High Availability



4.1 Software Layer:

.

In today's environment software plays a major role. In whatever direction we turn we soon realize its importance, a few examples would be traffic lights, trains, medical equipment, cash automatic machines, atomic plants, energy plants, business transactions over the Internet, etc, etc the list just does not end. The result of software failure can be very extensive depending on the vertical market and dependency of the software to meet your requirements.

The following are a few of the factors that drive business expansion with focus towards global market strategies, resulting in increased dependency on software. [6]

- Faster Central Processor Unit (CPU)
- More complicated and advanced software programs
- Fast growing Internet expansion

[6] Gary McGraw April 2003.

With software now being responsible for e -business and business -to-business solutions the potential risk that now is exposed could have consequences that could actually put the companies very existence at a substantial risk.
A few of the things to consider that justify the cost of making your software more stable and reliable in the early stages of the Software Development Life Cycle (SDLC) would be as follows.

- Monetary loss: Due to a software problem the famous online site EBay, lost $4 million in revenue as a result of interruptions in its services, for about 22 hours.
- Company reputation: C D Universe's reputation was advers ely affected when a hacker exploited its software that resulted in obtaining 300 000 credit card numbers and making this information available to the public Internet.
- Liability costs: The pharmaceutical distributor FoxMeye r wanted to sue SAP and Anderson Consulting for $500 million each for the enterprise software failure that it believed was responsible for its bankruptcy.
- Productivity loss: The Standish Group estimated that unreliable software resulted in a loss of 45% of system downtime, that cost US companies $100 billion in loss productivity.

A step in the right direction for making software more reliable is to make the infrastructure for software testing more rigid from the very start of the SDLC. A part of the SDLC is having measures in place for detection, locating and rectifying bugs. However these measures do have a price tag associated with it and account for between 50 to 75 percent of the total development cost. Even with these measures implemented, bugs will st ill be found in the production life cycle of the software. Workarounds will be found to solve bug issues that occur when the software is running in the environment it was developed to run in.
Tools that meet a certain standard, accepted by both the develo per and user community, and the metrics for software testing would address some of the issues that torment the software market. In order to be able to capitalize on this measure we need to understand the following. [7]

- Software attributes

The International Organization for Standardization (ISO) has adopted the ISO 9126 that is accepted and widely in use, it defines six main attributes and 21 sub characteristics. The six main attributes are F unctionality, Reliability, Usability, Efficiency, Maintainability and Portability.

- Software metrics.

Having standards alone describing the attributes, is a part in the process; we also need to have metrics to make sure that the software actually measures to those standards.

- Software testing.

Running software on a computer a nd then comparing the results against predefined criteria, allows us to evaluate if the software is working correctly in accordance with the established criteria. Testing tools can aid us in checking the performance of the software in different configurati ons and thus reduce the cost encountered in rectifying bugs after the software is in production.

[7] The Economic Impacts of Inadequate Infrastructure for Software Testing May 2002.

Examples of software bugs that caused availability issues:

Example 1[8]: A software bug that caused a system overload situation in FirstEnergy Corp. on the 14th of august resulting in a blackout that caused serious electricity availability disturbance both in USA and Canada.
An alarm reporting system ceased to function correctly due to simultaneous demand for specific data by many systems. Instead of handling the requests in a sequence giving one system at a time access to the data. The emergency management system out of function and no knowledge of its failure kept the company's system operators handicapped, as they were unable to see the degradation of their electrical systems and take the appropriate steps to immediately rectify the problem.
Example 2: On the 8TH OF June the New York Stock Exchange market (NYSE) interrupted retail content due to a flaw in the routine software upgrade causing loss of availability for about 90 minutes. Of the 30 stocks that make-up the Dow Jones Industrial Average 10 of the stocks were unable to trade a single share.

4.2 Database Layer:

Tackling this layer can be one that demands your attention very early in the stage of the database design. A well-designed database with focus on actual resource specific availability is key at this point. Typical issues arise when SQL queries competing for access put locks on the underlying tables and rows in a table and cause serious record access problems. Performance problems not taken care off usually result in saturation of CPU resources. Queries not optimised and background jobs taking much of the CPU time slice, instead of giving the queries the time needed to provide access, are often forgotten or not thought about, leave the database prone to availability problems. Also a baseline created and stress testing performed before putting the database in production provides a great help in tracking performance issues.

4.3 Hardware Layer:

This is one of the issues considered by many that solve all availability problems, get more hardware, get better performance hardware and all the new technology, be it firewalls, Intrusion detection or Intrusion Prevention can offer and you will be able to provide high availability. However this is certainly not true, sure it is part in making your resources available but it is not the silver bullet many believe it to be.
The following are some of the hardware solutions that could certainly provide better availability.

- RAID[9]: Redundant array of independent disks provides fault tolerance by implementing data redundancy. As data is written to more than one disk, data redundancy protects the data in situations when a single hard disk fails. We usually have two options, software or hardware RAID. Windows server operating system provides software RAID implementations, however this solution does not provide fault tolerance during the time period that a failure

[8] Kevin Poulsen April 7 2004
[9] Microsoft Windows 2000 Server Documentation, RAID disk arrays

occurs and it's rectification. Hardwa re RAID a preferred solution offers high availability as it provides better performance with it's faster disk I/O, hot swapping of disks that fail to function, hot sparing that automatically replaces a failed disk with an online spare in the system.

- UPS: Uninterruptible Powe r Supply/ Separate power feed: A UPS provides a specific amount of power for a specific amount of time, this is usually used to enable normal shut down of systems in cases when power gets disrupted. This solution is limited If availability is a concern you may need to consider a generator for longer power outages. However if the requirement demands highest availability possible then a secondary feed coming in from an alternative power station and preferably from another supplier would definitely be somethin g to consider.
- Clustering[10]: A solution that does provide high availability. Basically, clustering consists of two or more servers with a common shared external storage unit. Should one server fail the second server takes control and is then responsible for providing the services that were being delivered by the first server. The user connected to the server that failed does not even notice the change over to the second server. Clustering operates at the sever level and thus ensures that the entire SQL serve r is available.
- Log shipping[11]: Basically this solution copies the main server's backup files to a secondary server, and then performs a restoration on the secondary server when manually initiated. The advantage this solution has over clustering is that it does not share any hardware components and thus eliminates the single point of failure that the clustering solution entails when utilizing the shared drive array. The downside of this solution is that it does not have any automatic failure detection in ord er to initiate a failover to the second ary server. Another point to consider is the lag time duration caused by the restoration procedure before the secondary server can be made available. This solution also requires the client to connect to the secondary server by a different name and thus lacks transparency. Another feature of log shipping is the use of the WITH STANDBY option for restore that allows the secondary server to be able to provide read only services parallel with the main server still functioning. This provides a level of scalability as the secondary server can be used for reporting or other read operations performed by the main server and thus reduces the load on the main server. Log shipping works at a more granular level and ensures that onl y one database in the server is available.
- Replication[12]: A replicated technology consists of a Publisher the primary server, a Subscriber the secondary server also know as the standby server and a Distributor, the server responsible for delivering transactions to the secondary server. Transactions are delivered from the publisher to the distributor and then the distributor delivers the transactions to the subscriber. This process makes it possible for the application to fail over to the secondary server ver y fast, in spite of the fact, that not all transactions might be committed on the secondary server. It is because the distributor guaranties the delivery of all transactions. Replication can give you instantaneous failover under ideal conditions. The secon dary server is available 100% and can

[10] Bri an Knight SQL Server Magazine, October 2003 p - 27

[11] Michael Hotek SQL Server Magazine, October 2003 p - 24

[12] Michael Hotek SQL Server Magazine, October 2003 p - 24.

provide read operations even during the time the distributor is delivering transactions. As it could be used to offload the primary server by providing read access, it does give us scalability. Replication binds multiple machines (Publisher > Distributor > Subscriber) and a failure in one could result in failure of another. This is caused by transactions not getting committed to the next in the chain of servers. This is a factor to consider when implementing replication.

4.4 Internet layer:

Internet availability is one of the availability factors that most people can relate to, as we are continuously reminded through the mass media hype, every time the Internet is burdened by different attacks causing the Internet infrastructure to reach a saturation point. This layer is difficult to control as it has many dependent parties that contribute in providing this function. From users located in any part of the world connecting through a series of ISP's and finally entering the company's router, is a difficult problem to control as we cross different boundaries each having control over its own domain.
The Internet when designed in the early 70's was designed to meet much less demanding requirements than what we today are using it for, both the TCP/IP protocol and the infrastructure have been stretched to the limit and we will soon reach a point of maximum overload if we do not take actions in the near future to redesign and make changes that provide for these new demands we have today, and for what we will be needing in the foreseeable future. Yes the Internet community is aware of this issue and is working hard to solve these problems. However the complexity and the infrastructure that spans different countries each having its own legislation also need to be considered.

- **Internet Service Provider (ISP):** This is a major player that has the ability to control the flow of traffic into your network. It is one that you do not have any significant control over and is entirely in the hands and at the mercy of your ISP and other ISP'S connected in the chain of flow. Even with your limited control to solve this issue, you still need to see that the following is in place as this is what you can and should strive to achieve before you are affected by an attack.

1. Have the contact name, telephone number, fax number and email address to your ISP in case you have an incident. Also a reserve backup contact name is often very useful in cases when the main contact is unavailable
2. You must try to get the ISP to put it in writing in your contract, to what level of support you can expect to receive from them on reporting an incident. It is often when signing a new contract or renewing your present contract that you could get your ISP to take your concerns seriously and make them agree to putting the measures to be taken into your contract. Sales personal often have a tendency to promise a lot when trying to get you to sign the contract and will literally say anything only to get the contract signed. However, when the time of need comes to make good on those promises it often results in disappointment and could cost a company huge losses. So, make sure everything is written down in your contract. The level of support you need

Depends on what value y ou put on availability. In your business impact analysis you should have calculated how much "downtime" would cost your company, in minutes, hours or days and it is this value that decides the level of support needed to meet your requirements. The higher d emands you put on your ISP the more it will cost you but if your needs justify the cost it will help in getting the right support level to satisfy your concerns.

3. You and your contact person at the ISP need to agree on what will be needed from one another i n such cases before an incident occurs. This will help a great deal the day you have an incident and it will speed up the process to being helped if the relevant information is provided. Procedures should be documented to avoid panic and to make recovery a s smooth as possible.

- **Attacks against your resources:**

Denial Of Service (DOS) and Distributed Denial Of Service (DDOS) are two types of attacks that focus on, bandwidth consumption, resource exhaustion or a combination of both.

1. Bandwidth consumption a ttacks will utilize all available bandwidth of a particular network and this can be performed on a local or re mote network, however the later is one that is most commonly used.

2. Resource exhaustion attacks focus on utilizing the system resources such as Central Processor Unit (CPU) utilization, memory or other specific system resource to a point that causes the system to crash or the process to get hung and thus making the resource unavailable to legitimate users.

A factor that is common to both of the ab ove attacks is that they try to manipulate the protocol and thus provides the possibilities of simultaneously affecting a large number of systems.

Attacker: One or more individuals that for some reason have a motive and a goal, takes the opportunities avai lable and creates a means for exploiting a target.

Motives: These could vary, but below are few of the most common motives attackers have.

1. Personal: A person or persons, that possess personal vendetta against a victim.
2. Prestige. A demonstration of one's sk ill in order to draw publicity.
3. Material. Used to make financial gain from using illegal and unethical methods
4. Political. Used for political reasons
5. By products. These are the side affects of malicious code, such as worms and viruses

Goal: The ultimate go al would be to cause damage by exploiting the target's resources or infrastructure and resulting in loss to the victim.

Opportunities: Easily available tools on the Internet provide the attacker with the opportunities to accomplish hi s goals. A few of the tools commonly used are, Trinoo, Tribe Flood Network (TFN) [13], Stacheldraht [14], TFN2K, WinTrinoo, and Mstream [15]

Means: Utilizing the opportunities the attacker creates the means by compromising a large amount of poorly secured computers on the Internet, that t hen make up the attack network to be used in achieving his end goal. The a ttacker installs special program on the compromised machines, and thus enables the attacker to control the computers remotely.

[13] David Dittrich Oct ober 21 1999

[14] Dav id Dittrich December 31 1999

[15] David D ittrich May1 2000.

Target: Resource made available for it's legitimate use rs.

Attack Scenario's:
DOS: Denial Of Service is a type of resource attack that an attacker first starts to take control of a large amount of poorly protected hosts on the Internet and when possessing control over these hosts the attacker can choose to u se the comprised hosts to do his dirty job of exhausting a resource.

- As DOS attacks are considered a one to one attack, a solution to flooding attacks could be solved by making the target have extra resources available. So if the target has more resources than the attacker, the attack will have a less impact.

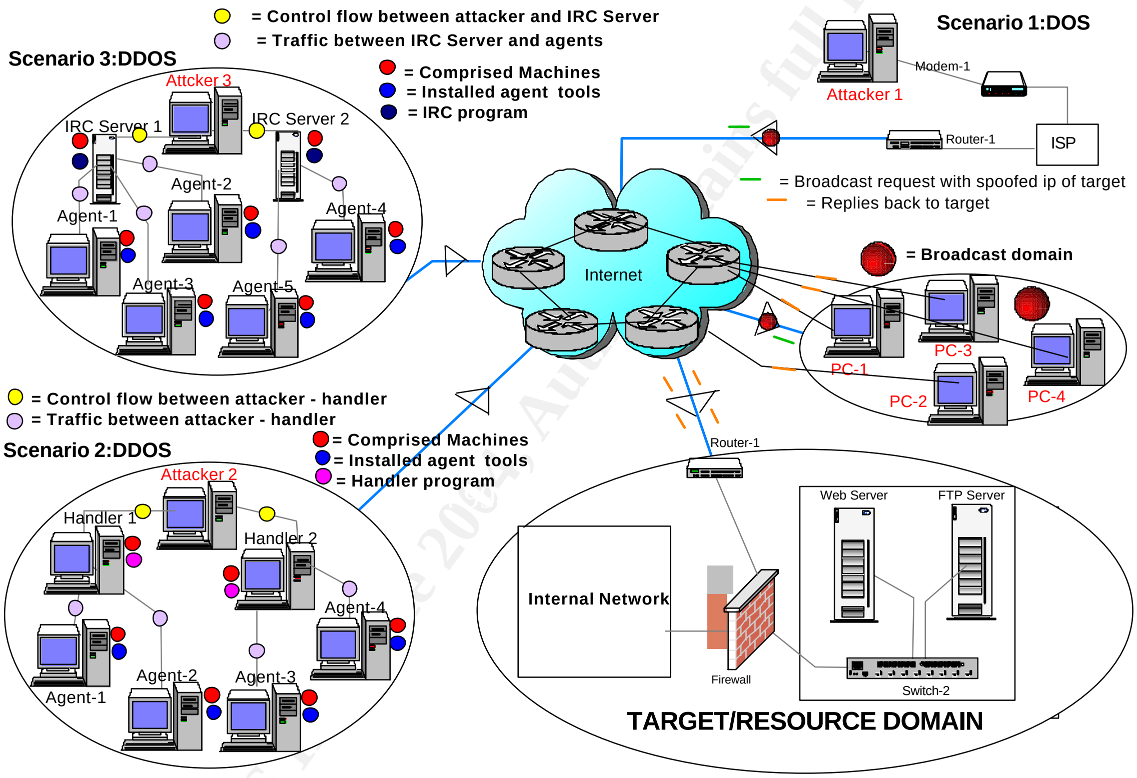


Figure 4: Scenario's diagram

Scenario 1: Figure 4
Smurf was classified as one of the more serious DOS attacks mainly due to the way it used amplification to make its affect so devastating. A Smurf attack has 3 parties involved, first the attacker (Attacker 1) that is the only one, actively involved, second some network (Broadcast domain) that can be used as the amplification network and third the target. The attacker starts by sending spoofed ICMP E CHO request packets with the target's IP address as its source address to a broadcast address on any network on the Internet that allows the directed broadcast functionality. Once the systems on the amplification network receive the request they will all s end replies to the source IP address it believes sent the request. The result is easy to imagine. Consider this example, say the amplification network has 1000 computers on its

network and all 1000 send replies to the victim due to the first initial request sent by the attacker thus causing the victim to get 1000 requests, this would generate a 1/1000 ratio.

- A counter measure for this type of attack could be accomplished by adding a rule in your border router to disable directed broadcasts functionality. This measure however will only protect your site from being used as an amplifying network.

Scenario 2: Figure 4
DDOS: This is a many to one attack scenario and consists of 3 parties. The first is the attacker (attacker 2) that is the active part and needs a person or persons sitting behind that client to perform the attack. Second are the hosts that the attacker succeeds in compromising, and on gaining access is able to install a special program that performs the distribution part of the attack on behalf of the attacker, these compromised hosts are called handlers or masters. Third are the hosts, also compromised by the attacker, but the special program installed on these compromised hosts, performs another function, and that is, it gets its instruction from the handler to perform some type of attack against the target, these compromised hosts are called the agents.
Communication between the attacker -handler/masters and handler -agents is usually known as the control traffic of the network that could be TCP, UDP, ICMP or a combination of the three if so determined by the attacker. However the communications between the agents and the target is referred to as flood traffic and this traffic flow could use any of the major point to point techniques namely, TCP SYN Flooding, UDP flooding, ICMP flooding and Smurf attacks or even a combination variation.

Scenario 3: Figure 4
DDOS: More sophisticated DDOS today are using the Internet Relay Chat (IRC) network and protocols for its control traffic. Its agents establish outbound connections to a standard service port, normally used by a legitimate service, and thus avoid detection by network scanners searching for listening ports. Attackers (attacker 3) connect to the IRC servers using normal communication channels, and are thus able to control all its agents
The IRC servers keeps track of all agents IP addresses, and is responsible for the communication between the attacker and the agents.

Flooding attacks: are achieved by sending the target a huge volume of malicious packets, and it is this huge volume, that makes traffic analysis a tedious and time consuming task.

Measures against DDOS:

- Identify the type of DDOS attack (ICMP, UDP OR TCP) that is bombarding your network, is the first step to take. You will need a Sniffer to capture your traffic, save to a file, and then you can use some tool to analyze your captured saved file. Tools provide a good view of your traffic by showing you which protocols are consuming your bandwidth.

- D-WARD[16] (DDOS Network Attack Recognition and Defence) is a Distributed DDOS defence system that is deployed near the source-end network. It has two goals.

  1. Control outgoing traffic to a target and detect if the traffic is of DDOS nature. If it finds DDOS traffic it stops the traffic.

  2. Provide better service to legitimate traffic between the deploying network and the target, even during the time an attack is occurring.

  Advantages of Source-end DDOS defences are as follows.

  1. Reducing traffic congestion: As the traffic is restricted close to the source it stops the malicious traffic from overloading the Internet infrastructure. Thus resulting in better use of the resources available.

  2. Minimum damage. Normal defence techniques, like rate limiting affect all traffic to the target, resulting in legitimate users also not being able to utilize the resource. However with source end, the malicious traffic is restricted at the source and thus does not affect the legitimate traffic.

  3. Quicker trace back. The source-end defence systems can trigger alerts to administrators for machines that are compromised and can thus check all the machines in its protected network.

  4. Advanced detection capabilities. The use of advanced detection strategies enables it to accomplish its goals.
- Black hole filtering[17]: An option that can be used by the ISP by forwarding bad traffic to an imaginary interface also know as Null0, as this is an invalid interface all traffic routed to this interface is dropped.
- Ingress Filtering: Controls the flow of traffic into the network and thus reduces the possibility of spoofed and invalid IP addresses into the network.
- Egress Filtering [18]: Controls the flow of traffic out of the network and thus reduces the possibility of the domain being used as part of a zombie network.
- Baseline of your traffic consumption and the type of traffic should be captured and ready to compare with the results caused by an attack to see and help the process of identification.
- Having a redundant network connection with another IP address range is also a good defence if the main network is under attack.
- Rate limit some network traffic: Also called traffic shaping:

  Allow you to control the bandwidth being utilized by a specific type of Traffic. However you need to consider a side affect in implementing this, it could also be limiting legal traffic, if the attacking traffic is using legitimate traffic for its attack.

  Cisco uses Committed Access Rate (CAR) [19] a feature built into its IOS software.

  This feature can be used both as a proactive method and also as reactive method.
- Host Auditing tools: Detects the presence of known DDOS client tools and server binaries in your system. OBS: Sensitive to updating the signatures otherwise useless.
- Network Auditing tools: Detects the presence of DDOS agents running on hosts in your network.

[16] Jelena Mirkovic 2003 D-WARD
[17] Matthew Tanase January 7 2003
[18] Egress Filtering v 0.2 Feb 29 2000
[19] Cisco CAR RATE LIMITING.

- Emergency Data Centers: Offsite centers to go into function when the main site is under attack and thus reduces the availability to its core resource.

Vulnerability attacks : Can be exploited by a small amount of specific packets th us reducing the volume in the attack. They can be focused on software or service vulnerabilities

- Solution to software vulnerability attacks is relatively simple, as you can patch the vulnerability if of course there is one available. Note: Patches can sometimes cause other problems, so you need to test the patch in a test environment before you apply it on your production servers.
- Identifying the specific packets and then blocking the packets can solve Service vulnerability attacks. However if you need that type of traffic in your network then it might not be an option for you.

:

4.5 Maintenance Layer:

The maintenance layer is composed of three major parts. Key success to this layer is having all your plans, procedures and routines updated on a regular ba sis, Having someone responsible for all changes made that affect the plans, is an important step to take, if you want your plans to be useful in time of need. Outdated plans reduce your time period of recovery.

1. Business Continuity Plan (BCP)

A BCP deals with the restoration of the business processes or the continued operation of business processes. Organizational processes could operate without computers, example, orders can be taken by telephone and written by hand on a order form. With a BCP, the company could reduce the impact a disaster could have on the normal business operation, and thus keep the revenue streams coming in.

2. Disaster Recovery Plan (DRP)

A DRP covers the restoration of the critical information systems that support the business processes. It is usually more technical in nature and helps in carrying out emergency procedures, by helping people.

3. Normal maintenance procedures.

The daily routines keep track on systems by monitoring for performance and comparing with a baseline to look for s ignificant changes. Backup routines performing the regular daily and weekly backup are also part of maintenance. Verifying your backups can be restored correctly is important and regular controls should be performed to ensure this.

Key factors in the diff erent phases of creating a good BCP and DRP: [20]
Phase 1: Project Initiation.

- Business case needs to be made, as management usually does not understand all the technical aspects that are required to be in place. Following are steps to include in a business c ase.

[20] Shon Harris ´ CISSP Certification Package. Business Continu ity & Disaster Recovery Planning.

    a. Showing benefits, of all the suggested measures and most importantly what each measure cost, would enable management to make a business decision if the suggested measures result in profitability. A risk analysis needs to be performed.
    b. Regulations and laws need to be complied with, as some laws dictate that companies MUST have a BCP.
    c. Liabilities need to be considered.
    d. Business partnerships need to be evaluated, for having a BCP/DRP in place. Example, A car manufacture depends on its partner if it receives the motor from them, thus making sure your partner also has the required plans in place is very important
    e. Statistics, example less than 5% of businesses are prepared for a disaster, however 65% would go out of business if they were closed down for more than a week. Sept 11 was a wake up call for many companies, as we soon realized the consequences of not having a BCP/DRP plan in place.

- Management's approval is vital.
  a. Resources and a budget are required to make your plans feasible.
  b. Team member's selection should be management's task
  c. Testing is considered management's responsibility.

Phase 2: Business Impact Analysis (BIA)
A BIA is a type of risk assessment but with a focus on disasters and disruptions. Following are the steps in a BIA.

  a. Identify the key functions of a company.
  b. Identify the resources that the key functions depend upon.
  c. Categorize the resources, and make sure all new resources are also categorized when implemented later. You could categorize your resources example, Normal = 7 days, Important = 72 hours, Urgent = 24 hours, and critical = x min or x hours
  d. Calculate the Maximum Tolerable Downtime (MTD) for every function and resource.
  e. Collection of knowledge about how the company works. All department managers should be interviewed and information gathered, relative to each department requirements should be documented. Understanding how the different departments interact is essential to making recovery possible after a disaster.
  f. Identify Vulnerabilities.
  g. Identify Threats, could be man-made (such as a strike), technical (a system crashes) or natural disasters (a flood).
  h. Solutions need to be made to address the threats and vulnerabilities
  i. Develop a loss criteria list to calculate the probable level of damage
  j. Scope needs to be defined by both the BIA team and management. It is not possible to focus on all types of threats.
  k. A report should now be presented to management that enables them to now make the final decision and give approval for the creation of the plans.

Phase 3: Restoration alternative strategies
The recovery alternative strategy is when an organization researches and evaluates all the available alternatives for business resumption after a disaster. Below are some of the alternatives.
Facilities alternatives:

a. Hot site. This is fully equipped with all the necessary software and hardware and is ready for use in a very short period of time. This is a expensive solution, but does provide the fastest recovery time of all the other solutions available. A point to be aware of when choosing this solution is that it does not have propriety software or legacy software/hardware in place.
b. Warm site. This is partially equipped and will require you to bring your own systems with you. It could be ready for use in days and is a solution to choose when your business recovery can wait for the time it takes to get business running.
c. Cold site. This is basically an empty facility with only the wiring and electricity. This solution could take weeks to get business back up and running.
d. Mobile hot site . A rolling hot site is usually a vehicle such as a trailer that is fully equipped with electricity, communication links, network and computers and all other necessities essential to recovery.
e. Redundant site is the most expensive solution as the company owns the redundant site, and is only utilized when a disaster occurs. The company is responsible for keeping it updated.
f. Reciprocal agreement provides an inexpensive alternate backup site option. This solution should not be used as a primary alternative, but could be used as a backup alternative to your primary. However you should be aware that reciprocal agreements are not legally binding, however it could be made enforceable if special agreements drafted and put in place to ensure its legality.

Offsite automated backup alternatives:

a. Disk shadowing is a mirroring technology that normally writes data from one disk to another but in the case of an offsite backup alternative data is being written on another disk that is offsite.
a. Electronic vaulting is a technology that does batch backup automatically. It collects all the changes made during a period of time and then sends the collection of updates at a later time.
b. Remote logging (journaling) is a real time update of only the changes over a telecommunication link.

Phase 4: Recovery development plan
This phase is the actual creation of the development plans. Procedures should be documented by the responsible individuals, which are performing the same procedures as part of their normal job function. Service level Agreements (SLA), Offsite facility agreements, and other necessary agreements that need to be committed are part of this phase.

Phase 5: Testing and Maintenance

Plans are worthless unless tested, and verified that they perform their function. A scope for the testing should be made and testing should be based on a defined scenario. As the network environment is dynamic, our plans soon become outdated and need to reflect all our changes that could have an adverse effect on the plans. Roles should be assigned to individuals Test should be carried out on a regular basis, with focus to train people to take the correct steps and get an understanding of the procedures that they will need to perform, when such plans are activated. Testing also provides a means of deter mining if there are any missing procedures, which could affect the entire plan. Mistakes will occur and we would like to learn from the mistakes occurring during testing. Following are suggestions on the different steps to take in sequence in order to make test recovery as smooth as possible.

a. Check lists should be created. Each department manager should have it's own checklist to verify that it covers all the necessary procedures to be put in place that allow his department to become functional. After each department has verified and made changes if necessary, then the planning team synchronizes all the department lists into one complete checklist plan.

b. Structured walk through is a type of meeting where all the department managers meet to v erify t hat checklist is accurate. Th e group will discuss the goal, scope and assumptions of the plan. They will also walk through various scenarios of the plan from start to finish and verify that nothing has been left out.

c. Simulation test is actually goi ng to run a test based, but is limited to the primary site. This test requires very good planning as it is in this step that the different roles responsible for the operational and support functions will put into practice the execution of the disaster recovery plan based on a defined scenario.

d. Parallel test is similar to a simulation test but it is a step further, in that it allows testing to some degree on the offsite location.

e. A complete interruption test. This is the last step in a testing drill, that shuts down the primary site and resumes working on the alternative site. This is a very risky test and should be performed after all the above tests are performed and verified.

4.6 Physical Layer: [21]

The physical layer is one of the layers not considered b y most to be related to availability, however, even this layer does contribute to improving availability. Normally what we are talking about here is protecting your physical location, and in particular your critical data center. Access should be monitored to the data center and office location in order to allow only authorised users to perform their job function. Neglecting to protect this layer, could easily allow unauthorised users to do things that could compromise both your security and your availabilit y. If one of your servers is stolen or if an unauthorised user spills coffee on one of your critical servers both actions cause an availability issue. Sure, you may have a procedure in place to take care of such problems, but consider the time it takes to get a new machine,

[21] Advanced Information Assura nce Handbook page – 9.

configure it, test it, and secure it the list continues. So, the bottom line is securing your physical location by allowing only authorised people to get their job done.

4.7 Qualified Personnel Layer: [22]

All the above -mentioned availability issues combine together to provide a high availability solution but still depend on the most critical aspect; qualified personnel. Incidents and security breaches are happening on a daily basis. Attacks are much more complicated today and need qualified personnel to be able to track and eliminate them. The entire IT environment with all the technologies in place cannot run by itself. All the complexities involved, demand that people are trained and kept updated to reflect the environment they are responsible for maintaining.
It is said that the Insider threats are the most dangerous to a company, compared to external threats. This most serious threat is not often addressed. The case of Timothy Lloyd a disgruntled employee of Omega Engineering is a real frightening but true story. Lloyd's destructive actions cost Omega a loss of $12 million at a minimum. What Lloyd did was plant a logic bomb to detonate, a few days after his dismissal. This caused so much damage that it nearly could have put Omega out of business. Could this have been avoided? Yes and No, if the employees had taken his threats seriously and understood the potential of the destruction he could cause, then yes it could have been avoided. However had he not made his threats known to others and done everything secretly, he still could have succeeded in carrying out his evil actions.
Awareness training to all employees allows them to obtain knowledge and gain confidence to respond in defence, with focus on the company's best interest. Employees need to understand social engineering attacks, and make them equipped to deal in the correct way when encountered. Training is a cost, just like hardware and software but in order to make our environment secure and available we need to calculate for this cost, as this is an important part in the success of maintaining our environment.

5.0 Tips:

5.1 Specific Tips:

1. Development team with focused goals, each with a priority and with management's awareness to related costs in producing a reliable and secure code.
2. Software testing routines for stability, security and availability performed by both internal and external consultants in order to get a non biased opinion of the results of testing
3. Evaluate and prioritise your availability risks. A software risk analysis should be performed.
4. Think of availability in layers of depth.
5. Monitoring of key resources with two different software -monitoring tools to avoid one monitoring tool having a software bug causing the entire system to come to a standstill.

[22] Humanware – the Killer App for Information Security.

6. Design solutions with fault tolerance and load balancing to meet your needs.
7. Create a Disaster Recovery Plan (DRP) and test on a regular basis to verify that your plan is in function and that it has been updated for changes in your environment that could affect your DRP.
8. Create a Business Continuity Plan (BCP) and test on a regular basis to verify your plan is in function and that it has been updated for changes in your environment that could affect your BCP.
9. Establish a good working relationship with your ISP.
10. Keep your IT staff well educated and budget for this cost.
11. If you believe that your company or network is under a DDOS attack or if it is being used as a "zombie" in an attack follow the CERT/CC analysis. [23]

5.2 General Tips:

1. Know what your crown jewels are. Each and every company needs to determine what are its most valuable assets and with a combined teamwork from the different departments ascertain the value of the assets. Use either or both qualitative and quantitative methods for evaluating your assets. Prioritise the assets and then you can determine the percentage of availability required and the levels of procedures that need to be in place to achieve your requirements.
2. Practise the principle of least privilege. The huge impact caused by macro viruses, could have a much less impact if the principle of least privilege is implemented. As 70% of users do not use the macro function it would be a good procedure to have macros disabled, when installation is done. Only users that require the macro function should have it enabled. This will reduce the impact to only 30% of your users using this function.
3. Prevention is essential but detection is a must. Even with all the layers in place to protect you, unforeseen events will sometimes affect you and if you have a good detection mechanism in place, this would enable you to react quickly and thus reduce the loss of availability. A good example of this, is about the famous security professional Timomura whose network was hacked by Kevin Mitnic. Even though Kevin was successful in gaining what he set out to accomplish, Timomura was able to track Kevin with the help of his detection system in place and thus resulting in helping the FBI in capturing Kevin and putting him in prison for his deeds.
4. Know your network: You need to know and understand all your network components and have a network diagram that reflects how your network is configured today and not how it was configured many years back. Understanding each component in detail allows you to locate the root of the problem. A good example of this is your system, be it a personal computer, a web server or whatever, you need to know all the services that are running on your system and require the knowledge of what programs are running these services. It is easy to create a screen dump of all your services just before it is put in the production environment, then you could always use this as a reference point to compare on a regular basis to see if there are additional services running, and thus take appropriate steps to investigate why these extra services are on your system that should not be there in the first place.

[23] Cert, Protect your Web server against common attacks.

Note: It is very important that you update your screen dump after all changes to your system no matter what; otherwise you will not have a reliable reference screen dump to compare with.
Use a tool to create a snapshot of your registry settings of all critical systems, before being put in a production environment and compare them later to see for unexpected changes. The original snapshot needs to be updated every time a patch or any other change is made otherwise the result of comparison will be misleading. You should also use tools that check for file modifications made to critical files on your systems. Note: Make sure your reference files are saved in a secure location and not on the working server itself.

6.0 Conclusion:

Yes availability is a simple word and easy to grasp, but it has many aspects to be considered, in order to make Information Technology more well accepted both from a technical and e nd user perspective. Sure this new technology is here to stay but the success depends mainly on how we achieve a balance between functionality and demands of both technology and human interactivity.

References:

High Availability Solutions
Michael Hotek
SQL Server Magazine
October 2003, page 20

In search of a cure for DDOS attacks
By David Berlind,
ZDNet Tech Update
March 18, 2004
http://techupdate.zdnet.com/techupdate/stories/main/In_search_of_a_cure_for_DDo_S_attacks__print.html

Advanced Information Assurance Handbook
March 2004, page 91
Chris May, Marie Baker, Derek Gabbard, Travis Good, Galen Grimes,
Mark Holmgren, Richard Nolan, Robert Nowak, Sean Pennline
http://www.cert.org/archive/pdf/aia-handbook.pdf

Vulnerability Note VU#399355
Title: Cisco IOS anc CatOS fail to properly validate ARP packets thereby overwriting device's MAC address in ARP table
http://www.kb.cert.org/vuls/id/399355

Advanced Information Assurance Handbook
March 2004, page 85
Chris May, Marie Baker, Derek Gabbard, Travis Good, Galen Grimes,
Mark Holmgren, Richard Nolan, Robert Nowak, Sean Pennline
http://www.cert.org/archive/pdf/aia-handbook.pdf

Making Essential Software Work: Why Software Quality Management Makes Good Business Sense
Gary McGraw
Software quality management magazine, Vol. 3, No. 2 – April 2003
http://www.sqmmagazine.com/issues/2003-02/essential.html

The Economic Impacts of Inadequate Infrastructure for Software Testing
Planning Report 02-3
Prepared by: Research Triangle Park, NC 27709
Prepared for: National Institute of Standards and Technology
http://www.nist.gov/director/prog-ofc/report02-3.pdf

Security Focus News: Tracking the blackout bug
By Kevin Poulsen
April 7 2004
http://www.securityfocus.com/news/8412

Windows 2000 Server Documentation
RAID disk arrays
http://www.microsoft.com/windows2000/en/server/help/default.asp?url=/WINDOWS2000/en/server/help/sag_FAULTunder_A.htm

Clustering SQL Server
Brian Knight
SQL Server Magazine,
October 2003, pages 27 -32

High Availability Solutions
Michael Hotek
SQL Server Magazine
October 2003, page 24

The “Tribe Flood Network” distributed denial of service attack tool
David Dittrich
October 21 1999
http://staff.washington.edu/dittrich/misc/tfn.analysis

The “stracheldraht” Distributed Denial of Service Attack Tool
David Dittrich
December 31, 1999
Global Incident Analysis Center
SANS Institute
http://www.sans.org/y2k/stacheldra ht.htm

The “mstream” distributed denial of service attack tool
David Dittrich
May 1 2000
http://staff.washington.edu/dittrich/misc/stacheldraht.analysis

D-WARD: Source-End Defence Against Distributed Denial-of-Service Attacks
By Jelena Mirkovic 2003
http://www.lasr.cs.ucla.edu/ddos/dward-thesis.pdf

Closing the Floodgates: DDOS Mitigation Techniques
Matthew Tanase
Security Focus
January 7 2003
http://www.securityfocus.com/infocus/1655

Egress Filtering v 0.2
Feb 29 2000
Global Incident Analysis Center
SANS Institute
http://www.sans.org/y2k/egress .htm

CAR RATE LIMITING
Cisco
http://www.cisco.com/en/US/tech/tk543/tk545/tk764/tech_protocol_home.html

Shon Harris´ CISSP Certification Package
Video Seminar
Domain: Business Continuity & Disaster Reco very Planning
Note: This DVD -ROM's package can be purchased at the following link.
http://www.logicalsecurity.com/products/training/cbtc.html

Advanced Information Assurance Handbook
March 2004, page 9
Chris May, Marie Baker, Derek Gabbard, Travis Good, Galen Grimes,
Mark Holmgreen, Richard Nolan, Robert Nowak, Sean Pennline
http://www.cert.org/archive/pdf/aia -handbook.pdf

Humanware – the Killer App for Information Security
A white paper for Information Security Management
Presented by the National Security Institute
Note: In order to get a copy of the article you need to go to the following link and fill a simple form.
http://nsi.org/SSWebSite/Human_wareWP.html

Protect your Web server against common attacks
Cert Coordination Center
http://www.cert.org/securit y-improvement/practices/p082.html