

Mission-Critical Network Planning

September 2009

A decade ago, end users were typically internal to the company. Now, heavy reliance on the Internet has created an environment in which the bulk of users are outside of the company. Employees can't do much more than complain or file a trouble ticket. Customers can take their business elsewhere. Matthew Liotine's book entitled Mission-Critical Network Planning¹ addresses this new complexity.

A better title for his book might have been Mission-Critical Continuity Planning. In his title, Liotine includes in the term "network" any component that might affect the continuity of a mission-critical application – communication networks, servers, storage, data centers, power, and HVAC.

Each of these topics deserves a full volume (or more) on its own. Consequently, Liotine cannot delve into great detail in his 400-page book on any of them. Rather, he gives an excellent overview of all of the pertinent technologies involved in mission-critical systems, pointing the way to making them redundant for high availability. Each topic is supported by extensive references to other publications that will fill in the detail if needed.

As a result, this book is not intended to be a design guide for high availability. Rather, it provides the background that one new to the field can use to become rapidly conversant in the various technologies involved in high-availability design. It is especially useful to managers who are faced for the first time with converting their legacy systems to high-availability systems capable of recovering from a fault in minutes rather than hours or days.

Principles of Continuity

The fundamental requirement to achieve continuity is redundancy. All single points of failure must be eliminated. This requires that all critical components be backed up by an equivalent component that, though not necessarily identical, can provide the same level of service as the primary component.

Simply providing a backup component is not enough. First, means must be provided to detect a fault in a primary component. There must then be an adequate failover process to bring the redundant element into service. Redundant elements should be geographically dispersed to avoid dual outages due to some common disaster; and they should not share any common resource, the failure of which could take down both the primary component and its backup.

¹ Matthew Liotine, *Mission-Critical Network Planning*, Artech House; 2003.

During the failure, the backup component represents a single point of failure. Therefore, the repair of failed components must be done as quickly as possible, and the procedures to return them to service with minimal impact on users must be in place.

The author describes several types of redundancy:

- kN , in which N resources are replicated as a group k times..
- $N+k$, in which k backups are provided for N resources.
- $N+k$ with kN redundancy, which are combinations of the above.

He also distinguishes between fault-tolerant systems (which provide automatic recovery with no data loss), fault-resilient systems (which provide automatic recovery with some data loss), and high-availability systems (which minimize disruptive impact).

He then proceeds to define the metrics of availability, including RTO (the time to restore service), RPO (the amount of data that can be lost), the mean time before failure or the mean time to failure (MTBF, MTTF), the mean time to repair (MTTR), and availability (the proportion of time that a system is up). With respect to availability, he presents the equations to calculate the availability of serial components, in which the failure of any one component causes the system to fail, and parallel components, in which the system is operational so long as any of the components are operational.

Other metrics include cost metrics (the cost of downtime), capacity metrics (utilization, bandwidth, and overhead), and performance metrics (latency, response time, data loss, data errors, and throughput).

The exposure metric measures the risk and the cost of threats, and is the product of the frequency of the threat, the probability that the threat will cause a problem, and the cost if a problem occurs.

Communication Networks

Network Topologies

Chapters 4, 5, and 7 deal with communication networks. Various network topologies are described, including point-to-point, bus, multipoint (a collection of point-to-point links that share common nodes), star, and string.

These elemental topologies can be combined into more sophisticated topologies such as meshes, rings, tiered, edge, and peer-to-peer networks. Meshes may either be fully connected, in which every node is connected by a point-to-point link to every other node, or may be only partially connected. A ring network is essentially a string network, that is, a series of point-to-point links, that closes upon itself.

A tiered network is hierarchical in that one or more master nodes connect to slave nodes, which act as masters for lower-level slave nodes, until the end nodes are reached. To provide alternate paths in the event of a node or link failure, a master may connect to multiple slave nodes, and nodes at the same level may connect with each other.

Edge networks are the end nodes of a network and are typically the users or another network. Edge networks can be the most vulnerable points in a network because they are often nonredundant.

Peer-to-peer networks allow peer services to interact. These networks can be any combination of the above networks.

Communication protocols are layered, with each layer feeding and being fed by its neighboring layers. Layer 1 is the physical medium such as copper or fiber. Layer 2, the data-link layer, is used to transmit data over the medium. Ethernet, ATM, and SONET are layer 2 protocols. Layer 3, the network layer, is the networking protocol that allows messages or packets to be delivered to specific addresses. IP is a layer 3 protocol. Layer 4, the transport layer, ensures proper ordering of messages and guarantees message delivery. TCP is a layer 4 protocol.

Each layer has its own protection and recovery mechanisms. Each higher layer protects against unrecoverable failures in the lower layers and notifies higher layers of network problems.

Network Continuity

Local Area Networks

Local area networks (LANs) are a layer 2 technology. They are predominant for collocated networks today because of their low cost, ease of use, and support by standards and a plethora of mature products. The most common LAN layer 2 protocol is Ethernet.

Ethernet is a bus topology. It cannot on its own provide redundant data paths nor is it designed to carry connection-oriented traffic. Data collisions are expected and may cause frequent retransmissions as the load on the LAN increases beyond 20% or so. LAN segments should be joined by switches and not hubs to minimize collisions, as switches will block traffic to LAN segments for which the traffic is not destined.

If redundant LANs are configured, they should be used in a load-sharing arrangement to improve performance and to facilitate failover should one LAN rail fail.

A virtual LAN, commonly known as a VLAN, is a group of hosts with a common set of requirements that communicate as if they were attached to the same LAN regardless of their physical location.² A VLAN has the same attributes as a physical LAN, but it allows for end stations to be grouped together even if they are not located on the same network. Reconfiguration can be done through software instead of physically relocating devices. Virtual LANs control traffic by segmenting it. Bursty or streaming traffic can be segregated so as not to affect the quality of other traffic.

Other LAN devices include transceivers which connect devices to cable or fiber. Many transceivers can connect a device to two links for backup. Media translators interconnect fiber or copper links. Network adapters, also called Network Interface Cards (NICs), allow hosts and users to connect to the network. Since NICs can be a single point of failure, each host or user should connect to the network via two independent NICs.

Wide Area Networks

Wide area networks (WANs) are used for connecting end-points that are hundreds or thousands of miles apart. Today's Layer 2 protocols for a wide area network include frame relay, ATM (Asynchronous Transfer Mode), and SONET (Synchronous Optical Network).

Frame relay provides permanent virtual channels (PVCs) for customers that behave as dedicated connections even though they use a public network. Though backup PVCs can be configured, their value may be questionable since the carrier already has significant redundancy in its network. If the network goes down, so do both redundant PVCs.

² Virtual LAN, *Wikipedia*.

ATM provides higher bandwidth and better tracking of circuit availabilities than Frame Relay. However, it is more complex and expensive.

SONET provides optical networking over fiber. It may be arranged in a linear topology, though it is usually implemented as a ring. In this way, if a node or a link fails, connectivity is still provided in the reverse direction. Thus, SONET rings are inherently redundant.

SONET uses optical fiber to connect nodes. The earlier fiber links using a single light frequency are being replaced with Wave Division Multiplexing (WDM), in which many signals are sent simultaneously over the same fiber connection by using different light wavelengths, or colors. If WDM is used, auto-repair is even more critical since traffic volume is so much higher.

Much of today's WAN traffic is carried via the layer 3 Internet Protocol (IP) over the Internet. The Internet is a global partially-connected mesh network in which traffic is directed via routers. The layer 3 routers are intelligent devices that communicate with their neighbors to continually monitor the topology of the worldwide Internet so that they can choose the best paths and can route around failures.

There are many ways used to navigate an IP network. For routing within a company's network, the RIP (Routing Information Protocol) selects paths with the fewest hops (routers in the path). The OSPF (Open Shortest Path First) protocol selects the most efficient path, including hops, link bandwidths, and other parameters.

The VRRP (Virtual Router Redundancy Protocol) and the equivalent Cisco proprietary protocol, HSRP (Hot Standby Routing Protocol), support automatic failover of redundant routers to which users directly connect.

For communicating between networks in the Internet backbone, the most common protocol is the Border Gateway Protocol (BGP).

Virtual private networks (VPNs) can be implemented over the public Internet by encapsulating IP packets in the layer 2 packets (ATM, Frame Relay) or in other IP packets. This technology is also called "tunneling." A VPN acts as a private link that can have its own security. A VPN is a viable alternative as a backup to a dedicated WAN channel.

When backing up a WAN, it is important that different carriers be used for each WAN, and that these carriers not use some other common carrier whose failure could take down both the primary and backup WAN circuits. Even better, an entirely different technology such as satellite or microwave can be used for the backup channel.

Metropolitan Area Networks

Metropolitan area networks (MANs) lie between WANs and LANS. They are used to connect end points in a campus or city. They typically comprise Gigabit Ethernet over multiple SONET rings.

Network Access

Network access for voice services is typically via an in-house PBX (private branch exchange). Today's PBXs are often digital, and some now even use VoIP (voice over IP) to communicate voice. They connect to analog or digital links to the central office and provide redundancy and failover.

PBXs often comprise part of the network for data services as well. Especially if data is being carried, it is important that the technology used to access the carrier's backbone network is redundant with fast failover. Ideally, redundant links should be aggregated and used as a single

logical link. If a link fails, it is simply removed from the aggregated group until it has been repaired.

Wireless access is being used more and more for data connections. This technology ranges from in-house wireless LANs to cellular networks and microwave (in rural areas).

Satellite links are a good choice for backup. Geostationary satellites provide the best quality of service, but introduce significant latency because they are so high (20,000 miles). Medium and low earth orbit satellites introduce less latency, but channel quality may be less due to the need to switch from one satellite to another as they move relative to the earth.

Servers

Chapters 6, 8, and 9 deal with the servers that provide the processing functions in a mission-critical network.

Server Topologies

A common topology for redundant servers is the cluster. In this regard, the author defines a "cluster" in the very broad sense. A cluster is a group of interrelated computers that work together to perform a variety of tasks. If one fails, another can take over. Clusters can include multiprocessing clusters in which multiple CPUs in a single system share the load, fault-tolerant systems in which two or more computers back each other up with instant failover, or server clusters using two or more less expensive servers to achieve high performance and availability.

In whatever form, all processors in a cluster require access to the same data. This must be provided redundantly to ensure availability. Redundant storage topologies include mirrored storage in which the same data is recorded on two different storage media, RAID storage (redundant arrays of independent disks) in which data is striped across several disks in such a way that it can be recreated should any one disk fail, and network storage.

Failover within a cluster can range from hot failover to warm failover to cold failover. This gives a range of failover times and system costs, with the faster failover technologies costing more.

Wide-area clusters add disaster tolerance to the reliability equation. An example given by the author is the IBM Parallel Sysplex system.

Given a cluster, it is important to be able to distribute the workload across the processors in the cluster. This can be done with network load balancers that distribute traffic among the several hosts via the network. Alternatively, it can be a load balancer that distributes traffic to local hosts.

Traffic can be distributed randomly or according to predefined rules such as server load or transaction type.

A major class of servers is the web farm. Proper web design is important to maximize web access performance. Encryption, graphics, site design, the extent of distributed applications, and the ISP are all factors in the performance of a web site. Site navigation should be intuitive from the user's viewpoint rather than a reflection of the organization of the company.

When choosing an ISP, it is important to know the level of availability provided by the ISP. For high availability, an ISP should connect to redundant backbone providers. Common services such as DNS servers (Domain Name Services), which translate URLs to IP addresses, should be redundant so as not to become a single point of failure.

Caching servers can be used to cache frequently-used web pages or other data to improve access time. Forward proxy caching puts cache servers near the end-users. Reverse-proxy servers put caching in front of the servers. Forward proxies can save bandwidth, and reverse proxies can act as firewalls. Caching servers reduce response times. They also increase redundancy because they can provide responses to data requests even if the primary servers are down.

Mission-Critical Platforms

The availability of server clusters is significantly enhanced by the use of fault-tolerant servers. Fault-tolerant servers provide immediate failover with four to five 9s of availability (50 minutes to 5 minutes of downtime per year). Fault-tolerant systems are designed to survive any single fault without losing any data.

Fault-resilient servers provide four 9s availability, but may lose some data. Rather than instant failover, they rely upon a warm standby that can take many minutes to be put into service. The warm standby has been kept up-to-date with the most recent data, but must have its applications started and connected to the database.

High-availability systems provide three 9s of availability, or about eight hours of downtime per year. They typically rely on a cold standby that needs to have the database loaded or brought up to date.

Even more important to system availability is the software architecture. It should provide fault management in the form of event monitoring, system state checkpointing, fault detection, alarming, dynamic reconfiguration, and logging.

Network components should be highly reliable, redundant, and serviceable. The network and server platforms should be included in a network management system that typically uses SNMP (Simple Network Management Protocol) to monitor all elements of the network.

All system power should be redundant, from component power supplies to independent primary power feeds. Power lines should be conditioned and secured so that there can be no inadvertent shutdown of power. Power components should be hot-swappable so that power does not have to be shut down to replace a faulty component.

Software

The author categorized applications as mission-critical, mission-necessary, mission-useful, and mission-irrelevant. Important software must be well-behaved under load. It must monitor its components and recover from internal or external faults.

The distribution of software across multiple systems adds a significant complexity to software design and management. Often, an application is distributed because users are separated from the application servers via the Internet.

Application performance should be monitored and managed. This can be done on a transaction basis for transaction processing systems, on a file transfer basis for transport applications, and on a streaming basis for video or other bulk-transfer applications.

Performance management can be accomplished via test transactions, by checkpoints inserted into the program to measure response times, by application agents which monitor performance, or by traffic inference by gathering measurements from the network environment.

An application should be designed with fault recovery in mind, whether this recovery is manual or automatic.

Storage

Chapter 10 of Liotine's book deals with storage systems. There are a variety of storage technologies that provide compromises between availability, performance, and cost. The first step is to establish the mission-critical requirements for each class of storage. This includes the consistency requirements and the RPO – how much data loss is acceptable following an outage.

Data Replication

Data replication technology can provide the highest data availability and integrity. With data replication, any change to a source database is immediately replicated in real time to a target database that can be hundreds or thousands of miles away. Replication may be hardware-based or software-based. Software data-replication techniques can maintain the integrity of the target database so that it can be used for query purposes. Hardware technology does not guarantee target database consistency until replication has been terminated.

Replication can be asynchronous or synchronous. Asynchronous replication happens after the fact and therefore has no impact on the application. However, following a source-node failure, any data in the replication pipeline may be lost.

Synchronous replication guarantees no data loss. However, applications are delayed as database updates are completed across the network.

A less expensive method of replication is log journaling. With this method, only the transaction journal containing the database changes is replicated. The journal can then be replayed to update the target database.

Backup Strategies

The classic method for database durability is tape or virtual disk backup. Full backups are taken periodically (such as weekly), and incremental backups are taken more frequently (perhaps daily). The incremental backup only contains changes since the last full or incremental backup.

However, database recovery typically takes hours or perhaps even days for very large databases. Even worse, the backup may fail if one or more tapes are unreadable.

Storage Systems

RAID

RAID is a commonly used storage system for high availability. A RAID system includes several disks (typically five) in which one disk is redundant. If any one disk is lost, the data can be reconstructed from the surviving disks. There are several levels of RAID which are described by the author.

Tape

Tape is the common method for long-term storage. Liotine describes several tape technologies such as DLT (digital linear tape) and helical scan tape. Tapes should be stored off-site. There are many vaulting services for doing this.

Storage Area Network (SAN)

A storage area network (SAN) is a particularly powerful storage system. It comprises multiple storage devices on their own network typically controlled by redundant servers acting as highly intelligent controllers. The data on a SAN can be accessed simultaneously by several applications over fibre channel connections. The SAN can be geographically remote from some or all of the application servers within the distance limitations of the fibre channels.

SANs can be backed up by remote SANs using data replication to achieve disaster tolerance.

Network-Attached Storage (NAS)

Network-attached storage (NAS) is similar to direct-attached storage in which storage devices are attached directly to the application servers, except that with NAS, the storage devices are attached to the production network (typically a LAN) via simple controllers. A NAS device can contain multiple disks, including RAID arrays. Because NAS storage is not managed by a sophisticated server, simultaneous user access to particular data objects and backup via replication are not so common.

Storage Management

Storage management facilities can be extensive. They should provide disk and tape backups, data reorganization, disk defragmentation, replica management, capacity planning, mobile device management, data integrity via comparison of primary and backup databases, and storage capacity optimization via hierarchical storage.

In addition to data backup, the storage management facilities must support data restoration and recovery. This is a much more difficult task than data backup. The database must be restored to a point-in-time via full and incremental backups and then rolled forward via the journal logs if possible to recover lost data.

Facilities

Chapters 11 and 13 discuss continuity issues with the data center facilities.

A single facility is a single point of failure. Therefore, for true high availability, an enterprise must operate at least two data centers that back each other up. This means that there must be redundant networks for user access to all data centers and distribution networks to interconnect the data centers.

Cable Plant

The internal cable plant is often the most difficult to troubleshoot and to repair. Therefore, particular attention must be paid to its design and implementation.

Redundant cables should be routed via separate cable runs. Power cables should be routed separately from data cables. Entry points for redundant power and data cables must be made through separate points in the building.

Copper is the most common cabling for LANs. However, it has poorer performance than fibre. It provides less bandwidth and is susceptible to electrical noise and crosstalk. Fiber is more sensitive to improper installation. Kinks in the fiber cable and poor optical terminations can cause fiber faults.

Power

Power problems are the most frequent source of service interruptions. An average system experiences one-hundred power disturbances per month. Thanks to proper power conditioning, most of these disturbances have no impact on operations. However, more serious power disturbances such as blackouts, brownouts, transients, and surges can impact operations.

Primary power availability tends to run between three 9s and four 9s. Therefore, to achieve high availability in the data center, it is wise to obtain primary power from two independent sources.

Backup power must be available. Backup power starts with a battery-operated UPS (uninterruptible power system) that will provide enough power to keep critical data center equipment operational until more permanent backup power can be provided. This is generally in the form of redundant generators powered by diesel, gasoline, propane, or natural gas. At least two fuel tanks should be provided to protect against fuel contamination, and fuel supplier contracts should be in place to guarantee a continued fuel supply during an extended outage.

A comprehensive power management system is required for power system shutdown and reboot, for the control of power system components, for usage reporting, and for quality monitoring and notification. It should be integrated with the network management system to provide a single point of control for all data center facilities.

Environment

In addition to power, cooling and humidity control must be carefully planned. The author describes and evaluates several air-flow configurations for cooling.

Fire protection planning is another important environmental consideration.

Redundant Site

There are several levels of availability that can be achieved with a backup data center:

- *Hot backup* – The backup site is ready and failover is rapid. It might even be participating by sharing the work load or by performing related work. Recovery can be performed in minutes.
- *Warm backup* – The data center contains all of the necessary equipment and maintains a reasonably current copy of the database. However, the database must be made current and applications started before the data center is operational. Recovery can be performed in hours.
- *Cold backup* – The data center site is available but is empty. Equipment must be brought in and placed into service, Recovery could take days or longer.
- *Mobile backup* – It may be possible to wheel in a mobile data center.³ Recovery can be done in days.

As an alternative to building and maintaining a redundant site, there are hosting services which provide facilities and/or servers on a shared basis. Also, many companies make collocation agreements with other companies or divisions within the same company.

³ Data Center in a Box, *Availability Digest*, July 2009.

In any event, a well-defined, documented, and tested recovery plan is a necessity for mission-critical systems.

Management

Chapter 12 deals with the management of the mission-critical system. Distributed system management and operations typically consume half the costs of running a mission-critical system.

Management is complicated by the fact that a large number of heterogeneous multiservice components are involved. Management includes communication management, configuration management, security management, system management, and software management. It is important to include all of these management components within a common enterprise model.

Because mission-critical networks are so complex, it is important that the management facility can self-discover the network topology of communication, processing, and environmental components and present these in an intuitive graphical context. It should monitor events, identify real or potential faults, isolate problems, initiate recovery actions, and aid in root-cause analysis.

It should be capable of managing congestion and rerouting workloads accordingly. It should provide capacity planning to avoid such congestion. The goal is to neither under-provision nor over-provision the network.

The specification and monitoring of Quality of Service (QoS) is extremely important. QoS policies must be established, and policing mechanisms instituted to ensure enforcement. SLAs (service level agreements) should be used where possible to guarantee performance of different portions of the network.

Changes cause outages. Therefore, change management is very important. Not only should changes be planned carefully, but they should be deployed incrementally with a plan for falling back to the original system if problems are encountered.

Testing

Testing of the mission-critical network is covered in Chapter 14.

The cost of correcting errors increases exponentially through the implementation phases. Errors are corrected most easily during the definition and design phases of a project. The cost of errors increases rapidly as the project progresses through the development, testing, acceptance and deployment phases.

Therefore, good test planning is essential. There are several levels of testing. Unit testing is done as each component becomes ready. Integration testing proceeds as components are integrated into a growing system. The system test is a full system test of a completed system before it is deployed. It should include backup and recovery testing. The acceptance test is run for the benefit of the end users to let them sign off on the system. It should include performance and load testing. During all of this, regression tests should be used to ensure that nothing has broken as the system is assembled.

When it comes time to cutover, there should be a means to return to the old system if severe problems should occur.

Testing is expensive. But not testing properly can lead to even greater expense as errors are corrected in a system nearing completion.

Summary

The author summarizes his book succinctly:

“Network continuity is a discipline that blends IT with reliability engineering, network planning, performance management, facility design, and recovery planning. It concentrates on how to achieve continuity by using preventive approaches, instead of relying solely on disaster recovery procedures.”

All that we would add is that the book is necessarily an overview of the myriad topics instrumental to an understanding of mission-critical system continuity. The author provides substantial references for those who want to delve deeper into any of these topics.