Telephone System Availability

Mitel System Engineering Group

May 2008
Table of Contents

1 Introduction ............................................. 1
2 Availability and Business Continuity ...................... 3
  2.1 The Seven-Layer Business Continuity Model .......... 4
3 MTBF, MTTF, MTTR, and Availability .................... 4
  3.1 System Availability, What is Required? .............. 6
  3.2 Improving System Availability ........................ 7
4 Mitel PBX Hardware Availability .......................... 9
  4.1 How is Product MTTF Determined? .................. 9
  4.2 What is a Critical Failure Versus a Non-Critical Failure? 9
  4.3 What are the Critical Hardware MTTF Values for Mitel Products? 10
  4.4 What are the Availability Values for Mitel Products? 10
  4.5 How Many Spare Parts Should Be Kept On-Site? ........ 11
  4.6 High Availability Servers ........................... 11
5 Mitel PBX Software Availability .......................... 12
  5.1 Inherent Stability ................................... 12
  5.2 Service Recovery Time ................................ 12
  5.3 Fault Detection ..................................... 12
  5.4 Mitel PBX Software Availability Figures ............ 13
  5.5 High Availability Servers ........................... 13
6 Data Network Availability .................................. 14
  6.1 LAN Reliability ...................................... 15
  6.2 WAN Reliability ...................................... 16
  6.3 Additional Information .............................. 16
7 Power .................................................... 17
  7.1 System Powering Considerations ...................... 17
8 Geography ................................................ 18
  8.1 The Advantages of Geographical Distribution .......... 18
  8.2 Multiple Locations and Disaster Recovery ............ 19
  8.3 Additional Information .............................. 19
9 Process .................................................. 19
10 People ................................................... 20
11 Key Conclusions ........................................ 20
Table of Contents (con’t.)

12 Availability Solution Examples ........................................... 21
  12.1 Base Availability – 3 ‘9s: Non-Resilient Controller ................. 21
  12.2 Medium Availability – 4 ‘9s: Resilient Controllers and Redundant LAN .... 23
  12.3 High Availability – 5 ‘9s: Resilient Controllers, Redundant LAN And WAN .......... 24
  12.4 Alternate High Availability – 5 ‘9s: Resilient Controllers, Redundant LAN, And Survivable Remote Gateway ................................................. 25
  12.5 Geographically Distributed Call Control ................................ 27

13 Glossary ................................................................................. 29

14 Appendix A – Availability Formulas ......................................... 31
  14.1 Series Operation Versus Parallel Operation .............................. 31
1 Introduction

The introduction of IP based telephony systems has brought about many changes to a world formerly based on TDM telephony. These changes have had many affects ranging from how end users interact with their phones to how system administrators manage their telephony systems.

However, one of the most important changes brought about by the introduction of IP based telephony is the fact that IP technology can offer increased phone system availability relative to TDM based technology. The improved availability offered by IP technology in turn can be leveraged to realize substantial improvements in business continuity.

Because IP networks are ‘connectionless’, IP based telephony systems and sub-components can be deployed architecturally in ways that are just not possible with TDM based systems. More specifically, IP based telephony systems can be deployed in a geographically distributed fashion that was just not feasible with TDM technology.

The diagrams below highlight three different deployment scenarios:

- Highly Centralized
- Hybrid of central control with some remote control
- Fully Distributed

Both TDM and IP systems support highly centralized redundant systems that are physically co-located.

Centralized redundant TDM and IP systems can also be extended over communication links to support remote offices in the hybrid configuration. However, TDM based systems must rely on dedicated PSTN links, or physical tied links, to connect to the remote offices, whereas as IP based systems can be connected via IP based networks which offer the inherent advantages of fault tolerance and location independence.

In TDM Enterprise based systems, connections between geographically dispersed components are made over either dedicated point to point links or PSTN links, these TDM connections have the following limitations:

- Dedicated point to point TDM links do not allow for call rerouting when a link fails, furthermore these links have physical distance limitations.
- PSTN links can be used to overcome distance limitations, however these links can be expensive and the ability to reroute calls during a link failure is not possible.

IP technology offers the opportunity to provide a fully distributed telephony system across the Enterprise.

Certain vendors continue to emphasize the centralized call control approach, whereas the Mitel IP solution can be equally deployed in either scenario.
The advantages of geographically distributed telephone systems are explored in more detail later in this white paper; also a more in depth discussion can be found in the Mitel White Paper entitled Geographically Distributed IP Phone Systems, DK117894.

The Mitel Product Information Note entitled IP Infrastructure Architecture, EM003647 discusses the traditional PBX architecture, the new distributed IP telephony architecture and the benefits that are offered by the distributed IP model versus that of the traditional model.

While IP based telephony makes use of the advantages offered by the underlying data networking technology to improve system availability, network designers and administrators must be cognizant of the fact that the data network is now a major part of the telephone system.

In the past, the reliability rating of traditional PBX hardware and software was the most significant factor influencing telephone system availability. Now with IP-based telephony systems, there are many more factors, in addition to PBX hardware reliability, that can have an influence on overall telephone system availability.

Therefore, successful deployment and operation of voice over IP (VoIP) telephone systems requires the usage of a network service model that is a blend of the traditional telephone network service model and the high availability data network service model. The customer’s requirements for availability and the cost will dictate what the most appropriate balance between these two network service models will be.

This white paper discusses the factors that should be taken into consideration, so that an appropriate network service model can be applied to meet the customer’s needs. In particular, sections four through 10 detail the various levels of the seven-layer business continuity model.

In some cases Mitel has published more detailed discussions that are specific to a particular level of the seven-layer business continuity model, these White Papers, Product Application Notes and Guidelines are available at Mitel OnLine. The Mitel Application Note entitled List of Telephone System Availability Documents DK117891 provides a list of applicable documents and illustrates how they are inter-related.
2 Availability And Business Continuity

Telephone system reliability, redundancy, and resiliency are all inter-related and directly influence system availability. A highly reliable, non-redundant system will provide users with a high level of availability. Redundant and resilient design techniques can provide users with a higher reliability system than non-redundant systems by providing continued availability, even when a system failure occurs.

The seven-layer business continuity model, which is introduced in the following section, can be used to assess and define the business availability requirements for different businesses. For example, a telephone system designed for a nationally based financial institution may need to provide high levels of availability at all levels of the business continuity model, while a small retail store may find a telephone system that provides lower levels of availability perfectly acceptable.

2.1 The Seven-Layer Business Continuity Model

When discussing telephone system availability, some vendors make the mistake of only discussing the reliability of their own hardware components. This narrow-minded view ignores the fact that a chain is only as strong as its weakest link. To truly understand a customer’s system availability requirements, Mitel uses a seven-layer business continuity model, which is shown below (Figure 1).

The seven-layer business continuity model shows that telephone system availability is dependent on many different factors and that overall system availability should not be based solely on the reliability of the PBX hardware. Instead, the seven-layer business continuity model provides an overall view that addresses all components that influence system availability. These components are:

- **PBX hardware** – The PBX hardware forms the foundation of the business continuity model.
- **PBX software** – This is software that runs on the PBX hardware (e.g., the operating system, call control software, application software). It can have a major impact on system availability.
- **Data Network** – The data network includes the data networking hardware and related protocols (e.g., Layer 2 switches, routers, networking protocols). Overall system availability is dependent on the data network availability.
- **Power Distribution** – This fundamental requirement needs to be taken into consideration so that if required, equipment can continue to be powered even under fault conditions (e.g., uninterruptible power systems, generators).
- **Geography** – System availability can be enhanced when the network design takes geographical distribution of equipment and personnel into account, particularly when the business is dispersed across multiple locations and/or cities.
- **Process** – Company processes related to maintenance and repair need to be considered, since these processes can have a direct effect on availability.
- **People** – The availability of maintenance and repair personnel and their impact on system availability needs to be considered, i.e., whether or not repair and maintenance personnel are located on site or off site.
Seven-Layer Business Continuity Model

| People          | Process          | Geography     | Power Distribution | Data Network | PBX Software | PBX Hardware |

Figure 1: The Seven-Layer Business Continuity Model

It is quite straightforward to quantify availability figures for the first four lower layers of the seven-layer business continuity model, but less easy for the higher levels, since these levels are less dependent on actual equipment and software.

What value of availability is needed and at what level of the model this availability is needed is dependent on the level of business continuity required. The business continuity requirements for small- and medium-sized businesses (SMBs) are typically met by addressing the first four lower layers of the seven-layer business continuity model. For much larger businesses and businesses with critical operations, however, the upper layers of the seven-layer business continuity model also need to be considered. In this case, insurance companies often have the necessary information to evaluate the availability of a service at the higher levels of the model, e.g., risk due to fires, hurricanes, floods, and earthquakes. It is very possible that a particular business may have different business continuity requirements within their organization. For example, a central office phone system may need to provide a very high level of availability at the core of the network, a medium level of availability for commercial customers, and a lower level of availability for residential customers.

3 MTBF, MTTF, MTTR, and Availability

There are a number of metrics that are used by manufacturers to convey the reliability and repairability of individual components. These metrics can also be extended to convey the reliability and repairability of a complex grouping of components, such as a system or a network.

Manufacturers may quote calculated or demonstrated values. Calculated values are derived from a theoretical analysis and demonstrated values are derived from actual field data and/or accelerated life testing.

These metrics are:

- **MTTF or Mean Time to Failure:** This is the statistical length of time that a component will operate (uptime) before a failure occurs.
The meaning of MTTR is dependent on the context:

- **Mean Time to Repair** refers to hardware failures and the length of time that will elapse from the time a component failure was detected to the component being restored to service (repair time).
- **Mean Time to Recovery** refers to software failures and the length of time that will elapse from the time that a system software failure is detected to the time that the system’s operational integrity has been restored or recovered.

**MTBF or Mean Time Between Failures**: This is the statistical length of time elapsed between sequential component failures. This metric also includes the time it takes for the component to be restored to service. In other words, MTBF is equal to the MTTF plus the MTTR (or uptime + repair time).

**NOTE**: While it is often seen that an MTBF figure is being quoted, in reality what is likely being requested is MTTF. Since MTTR is usually much smaller than MTTF, it is common to confuse MTTF and MTBF figures. In practice, the differences are negligible; however, when the MTTR is a significant component, system availability will be impacted.

The figure below (Figure 2) illustrates the relationships between MTTF, MTTR, and MTBF.

![Figure 2: MTTF, MTTR, and MTBF](image)

While metrics such as MTTF, MTTR, and MTBF are necessary for analyzing component and system reliability, some vendors mislead their customers by focusing exclusively on the MTBF ratings for their product.

Mitel understands that what really matters to the end users of the product is system availability.

System availability expressed as a percentage of time is defined by:

- $P\% \text{ Available} = \frac{\text{MTTF}}{\text{MTBF}} = \frac{\text{uptime}}{\text{uptime} + \text{repair time}}$

System unavailability expressed as a percentage is defined by:

- $P\% \text{ Unavailable} = \frac{\text{MTTR}}{\text{MTBF}} = \frac{\text{repair time}}{\text{uptime} + \text{repair time}}$

It is a common practice within the industry to refer to the reliability level of a product, or system, as 5 '9s or 5 x 9. What is actually being referred to, however, is the availability expressed as a percentage for a particular product or system. When a product has an availability of 5 '9s it is in fact available 99.999% of the time.
For Example

A product has an MTTF or uptime of 2.5 years and an MTTR or repair time of 10 hours. What is the availability of this product?

- If uptime = 2.5 years and repair time = 10 hours
- There are 8,760 hours in one year so repair time can be expressed as 10 / 8,760 hours
- (P% Available) = 2.5 / (2.5 + (10 / 8760)) = 0.9995 x 100%
- (P% Unavailable) = 1 – P% Available = 0.0005 x 100%

This system is available 99.95% of the time, or in other words, the system is providing 3 '9s availability. Alternately, this system is unavailable 0.05% of the time.

It should be noted that this availability figure only considers the "product", it does not consider the effect of other equipment, such as power and LAN equipment on overall availability.

3.1 System Availability: What is Required?

Different businesses will have different system availability requirements. To determine the system availability requirements for a particular business, the following business requirements need to be established:

- Which services have an availability requirement?

A business will have certain availability requirements for basic telephone operation and the same business may have different availability requirements for the voice mail system or advanced telephony features.

Availability requirements for all services need to be defined so that the each of the layers in the seven-layer business continuity model can be designed to meet these requirements.

- What level of service outage is acceptable?

The business continuity requirements need to be understood so that an acceptable level of service outage can be defined. For example, a business that needs phones operational from nine to five, five days a week will accept a different level of system outage than a business that needs phones operational twenty-four hours a day, seven days a week.

Quite often, hardware products such as a PBX can achieve better than 5 '9s of availability, but it must be remembered that the overall system availability is defined by the weakest link. The weakest link in a system could be located at any layer of the business continuity model. This is why it is important to look beyond the MTBF rating of a single hardware component.

Older standards that specify the requirements for PBX or central office (CO) availability, in general, only considered the hardware components and only addressed the core of the network.

Newer guidelines suggest that the core should provide at least 5 '9s availability, however, the needs of individual users may be satisfied with less than 5 '9s availability.
For example, telephone companies will typically provide their residential customers with guarantees that when a fault occurs, service will be restored by midnight of the next working day (weekends don’t count). This translates to a 99 percent level of availability. Telephone companies usually offer their business customers with multiple lines a 99.9 percent level of availability.

3.2 Improving System Availability

One method that can be used to improve overall system availability is to deploy certain components in parallel. To illustrate this point, we will take two components that have the same level of availability and then we will compare the level of availability attained when these components are placed in series with each other and in parallel with each other.

NOTE: This section presents a high level view of the benefits realized with parallel operation over serial operation, for detailed calculations refer to Appendix A.

Serial Operation

When two components are placed in series with each other, both components must work for the system to be available. As shown below (Figure 3) the overall availability of this system deteriorates compared to the availability of the individual components.

Parallel Operation

When two components are placed in parallel with each other, there are two possible states that will allow the system to be available. These states are:

• Both components are available
• One component is available

As shown below, the overall availability of this system improves, compared to the availability of the individual components.

It should be noted that this improved level of availability can only be maintained if the failed component is repaired or replaced, and the following sections highlight this.
Parallel Operation: With and without service recovery
The following examples illustrate why service recovery must be performed to continue to realize the benefits of operating components in parallel:

• Parallel Operation without Service Recovery
  • If one component fails and isn’t recovered, the system will be operating without redundancy
  • With no service recovery, the level of system availability will be 99.98 percent
  • This overall level of availability is only marginally better than the availability obtained with just a single component

![Figure 5: Parallel Operation without Service Recovery](image)

• Parallel Operation with Service Recovery
  • If one component fails, but is recovered, system redundancy will be maintained
  • With a service recovery system, the level of system availability will be 99.999975 percent
  • This level of availability is significantly better than the availability obtained with a single component

![Figure 6: Parallel Operation with Service Recovery](image)

In conclusion, it should be noted that high availability is only maintained if the faulty component is repaired or a service is recovered. If nothing is done upon failure, then availability only marginally improves with parallel components.

The preceding examples used generic components to illustrate the availability improvements offered by parallel operation versus serial operation. In a VoIP system, these components could be Layer 2 ethernet switches, routers, communication links, or any other area of the network where a single point of failure has been identified.

Mitel’s resiliency solution allows two or more Mitel 3300 Controllers to be deployed in parallel.
4 Mitel PBX Hardware Availability

Seven-Layer Business Continuity Model

| People | Process | Geography | Power Distribution | Data Network | PBX Software | PBX Hardware |

PBX hardware is the first level of the seven-layer business continuity model. This section discusses Mitel product hardware availability.

4.1 How is Product MTTF Determined?
To determine the hardware MTTF for a product, Mitel uses the following reference data and analytical methods:

- MIL-STD217 (U.S. Department of Defense)
- HRDS (British Telecom)
- Telcordia SR-332 (Bellcore TR-332)
- Telcordia GR-512 (Bellcore GR-512)
- Measurements of actual product returns from the field
- Component manufacturer’s data

This approach includes semiconductor complexity, electro-mechanical complexity, environmental stresses, and manufacturing quality control.

4.2 What is a Critical Failure versus a Non-Critical Failure?
In general, a critical failure is a failure that prevents the product from performing its primary function. A non-critical failure does not. For example, if an indicator LED burns out, it will not prevent the user from placing phone calls and service can be scheduled at a time that will not impact the customer’s business. On the other hand, a microprocessor failure most likely will prevent a user from placing phone calls.

Mitel bases product MTTF on critical failures rather than non-critical failures; and defines a critical failure as a failure that prevents more than 10 percent of users from placing phone calls.
4.3 What are the Critical Hardware MTTF Values for Mitel Products?
Mitel is constantly striving to improve the reliability of the product. As a result of this the critical MTTF values for Mitel product keeps changing. The availability examples provided in this white paper are based on MTTF values that were valid at the time of publication.

For the most recent MTTF values for Mitel products please refer to the Mitel Application Note entitled 3300 ICP System Reliability and Availability Summary, DK116756 which can be found at Mitel OnLine.

4.4 What are the Availability Values for Mitel Products?
The following section shows availability numbers for a non-redundant MXe Controller, an MXe Controller with redundant power supplies and disk drives, and two controllers operating in resilient mode. The availability numbers for the various configurations show how the use of resilient controllers can greatly increase the availability figure and allow for less stringent MTTR requirements.

The values in Table 3 are based on a four-hour MTTR and the values in Table 4 are based on a one-hour MTTR.

Table 3: Mitel Product Availability, Four-Hour MTTR

<table>
<thead>
<tr>
<th>Device</th>
<th>Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>MXe Controller, without redundancy</td>
<td>99.996%</td>
</tr>
<tr>
<td>MXe Controller, with redundancy</td>
<td>99.997%</td>
</tr>
<tr>
<td>Mitel Single Line IP phone</td>
<td>99.999%</td>
</tr>
<tr>
<td>Mitel Multi-line IP phone</td>
<td>99.998%</td>
</tr>
<tr>
<td>Resilient MXe Controller</td>
<td>99.999999%</td>
</tr>
</tbody>
</table>

Table 4: Mitel Product Availability, One-Hour MTTR

<table>
<thead>
<tr>
<th>Device</th>
<th>Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>MXe Controller, without redundancy</td>
<td>99.999%</td>
</tr>
<tr>
<td>MXe Controller, with redundancy</td>
<td>99.999%</td>
</tr>
<tr>
<td>Mitel Single line IP phone</td>
<td>99.999%</td>
</tr>
<tr>
<td>Mitel Multi-line IP phone</td>
<td>99.999%</td>
</tr>
<tr>
<td>Resilient MXe Controller</td>
<td>99.999999%</td>
</tr>
</tbody>
</table>

From the preceding information, it can be seen that the repair time or MTTR has an influence on availability figures. A one-hour MTTR, although desirable, may not be affordable or even necessary. It must be kept in mind that MTTR is the total time elapsed from detection of the failure through to service recovery.

In situations where off-site service personnel are required to perform the repair, the service person’s travel time will be included in the MTTR. In this case, a one-hour MTTR usually cannot be achieved.
To achieve a one-hour MTTR, the following requirements must be met:

- There must be dedicated personnel on site 24/7 that are qualified to undertake the repair
- Spare parts and assemblies must be located on-site

While achieving a one-hour MTTR is possible, it can be prohibitively expensive. Most business continuity models can adequately be met with a four-hour MTTR.

The preceding data also shows that 8.9s of availability can be attained with a four-hour MTTR when using two controllers in a resilient configuration and 9.9s of availability can be attained with a one-hour MTTR when using two controllers in a resilient configuration. In other words, the use of resilient controllers allows high levels of availability to be attained with an industry standard MTTR of four hours.

### 4.5 How Many Spare Parts Should be Kept On-Site?

The annualized failure rate (AFR) can be used to determine how many spare parts or units should be kept on site for repair purposes. The AFR for an item is calculated by finding the reciprocal of the unit’s MTTF expressed in years. The AFR indicates the probability of failure in any given year.

As an example, a manufacturer indicates that a power supply has an MTTF rating of 500,000 hours. How many failures can be expected over one year (8,760 hours in one year)?

\[
\text{MTTF (in hours)} = \frac{\text{MTTF (in years)}}{\text{Hours in one year}} \\
500,000 = \frac{\text{MTTF (in years)}}{8,760} \\
\text{MTTF in years} = \frac{1}{57} \times 100 = 1.75\%
\]

Based on the above calculations, it can be expected that 1.75 percent of the power supply units will fail in the average year. This will therefore provide guidance regarding the number of spare units that will need to be kept in inventory.

### 4.6 High Availability Servers

When there are requirements for an application to provide high levels of availability, the availability of the server itself needs to be addressed. When selecting a server for a high availability application, the following points should be considered:

- Using a RAID-based disk subsystem that supports hot swapping will increase server availability and offer data preservation in the event that a catastrophic disk failure occurs.
- Cooling fans typically have a low MTTF value, and as a result, can greatly reduce the server’s availability. To increase the server’s availability, select a server that supports fan redundancy and allows the fans to be hot swapped.
The server should support hot swappable redundant power supplies, preferably each power supply should have its own AC power input.

Redundant network connections to the server need to be considered to eliminate a single point of failure.

5 Mitel PBX Software Availability

Seven-Layer Business Continuity Model

- People
- Process
- Geography
- Power Distribution
- Data Network
- PBX Software
- PBX Hardware

This section discusses Mitel’s PBX software and what effect it has on availability.

PBX software can have a significant effect on PBX availability, both in terms of inherent stability and service recovery time in the event of a PBX hardware failure or a network failure. To determine availability the areas mentioned in this section need to be considered.

5.1 Inherent Stability

This refers to a product’s ability to deal with issues that may arise during normal operation, such as memory fragmentation and software logic errors.

5.2 Service Recovery Time

This includes the time it takes to detect a fault and the time it takes to recover from the fault. The service recovery time can have a significant effect on availability. When calculating availability, Mitel evaluates the following parameters:

- PBX failure detection time
- PBX fail-over performance
- PBX boot time of operating system
- Boot time of IP phones

5.3 Fault Detection

When calculating availability of systems using resilient controllers, Mitel uses 45 seconds as an average case value for the time it takes an IP phone to detect that there is a fault with the primary PBX or that communicating with the PBX is not possible due to a network fault.
5.4 Mitel PBX Software Availability Figures

Service recovery time for a system is dependant on many variables and this variability will in
turn have an impact on the level of system availability. The following two examples illustrate
the effect of this variability.

Example 1

The following table shows the percent of availability that will be achieved for different
quantities of phones when resilient controllers running Release 8.0 software are used. The
values are based on a primary PBX failure occurring once per year and an average phone
fault detection time of 45 seconds.

<table>
<thead>
<tr>
<th>Number of Phones</th>
<th>Service Restoration Time</th>
<th>Fault Detection Time</th>
<th>Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,400</td>
<td>140 seconds</td>
<td>45 seconds</td>
<td>99.9994%</td>
</tr>
<tr>
<td>1,000</td>
<td>95 seconds</td>
<td>45 seconds</td>
<td>99.9995%</td>
</tr>
<tr>
<td>300</td>
<td>30 seconds</td>
<td>45 seconds</td>
<td>99.9997%</td>
</tr>
</tbody>
</table>

Table 5: Availability with Release 8.0 Resilient Controllers

Example 2

In a non-resilient system, if a system reboot is required once a year, the system availability will
be 99.99 percent based on a 10 to 30 minute service restoration time.

5.5 High Availability Servers

High availability servers can be used as PBX platforms and they can also be used to run
telephony applications. There will be some subsystems within the server that are not
redundant, such as the CPU and memory systems. To cover such cases, it may be necessary to
deploy a backup or redundant server to meet the availability requirements. Server redundancy
can be implemented in the following ways:

- A warm standby backup server: With this model, the backup server is not actively running
  the application, but the server is kept current of database changes and software updates
  that occur on the primary server. This server could typically be on line within 30 minutes
to take over in the event that the primary server has failed. With this model, application
  availability of 99.998 percent can be achieved.

- An active standby or load-sharing backup server: This model will provide an almost
  immediate application switchover (within seconds) from the primary server to the backup
  server in the event that the primary server has failed. With this model, software application
  availability of 99.9995 percent or better can be achieved.

Although the two cases above provide similar hardware availability capabilities, the
configuration and fault / recovery software also plays a role in achieving the best result.
Application availability is also dependent on the server’s fault detection and recovery
mechanisms and the particular application’s behavior during a switch over scenario.
6 Data Network Availability

Seven-Layer Business Continuity Model

<table>
<thead>
<tr>
<th>People</th>
<th>Process</th>
<th>Geography</th>
<th>Power Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Network</td>
<td>PBX Software</td>
<td>PBX Hardware</td>
<td></td>
</tr>
</tbody>
</table>

This section discusses data network availability and considers the impact that data network availability can have on system availability.

An IP telephony system is built upon LAN and WAN networking infrastructure. If the IP telephony system is required to provide a high level of availability, then the underlying LAN and WAN infrastructure needs to be able to provide an equal or better level of availability. An accepted practice for designing LANs is to use a hierarchical approach and distribute the LAN functions into three layers or tiers. These layers are called the Core, the Distribution Layer and the Access Layer. The following diagram depicts a hierarchical LAN.

![Hierarchical Network Design](image)

Figure 7: Hierarchical Network Design
In a hierarchical network, end devices such as IP phones and desktop computers connect to the access layer. Access layer devices will typically be Layer 2 ethernet switches. The distribution layer is used to connect access devices together and to connect access devices to the core. The distribution devices are typically high performance Layer 2 ethernet switches. The core connects distribution devices together and provides connections to shared network resources such as servers and internet gateways. The core is usually a mixture of Layer 2 and Layer 3 devices.

When data networks are designed using a three-tier hierarchy, the following levels of availability can be attained:

- A three-tier network with no redundancy can achieve an availability of 99.96 percent
- A three-tier network with a redundant core can achieve an availability of 99.98 percent
- A three-tier network with a redundant core and distribution level can achieve an availability of 99.995 percent
- A three-tier network with complete redundancy can achieve an availability of 99.99999 percent

6.1 LAN Reliability

LAN design can be as varied as the customers that are using the LANs. As a result, the reliability of LANs can range across a wide spectrum. Simple inexpensive LAN designs may have numerous single points of failure, while other LANs might be designed with full redundancy. For the following reasons, enterprise LANs are usually very reliable:

- Redundant switching equipment is used at the core and the distribution levels of the network
- Multiple connection paths are used to provide Layer 2 and Layer 3 redundancy
- Servers are deployed with redundant load-sharing connections
- Data is stored on a separate storage area network that typically uses RAID-configured devices
- LAN redundancy is supported through the use of networking protocols such as RSTP, VRRP (virtual router redundancy protocol), and LACP (link aggregation control protocol)
- LAN hardware is capable of providing availability of 5 ‘9s or better

LAN Hardware Availability

Typical LAN hardware availability figures for individual units are:

- 99.991 percent - for a Layer 2 switch (consumer grade)
- 99.995 percent - for a Layer 2 switch (enterprise access grade)
- 99.986 percent - for a Layer 2 switch (enterprise distribution grade)
- 99.982 percent - for a Layer 2 / Layer 3 switch (enterprise core grade) (see note)
NOTE: Despite the fact that Layer 2 core and distribution switches have a lower individual availability, when these units are used in a redundant configuration, they provide much higher system availability. They also generally provide more capability and a larger number of connections than access switches, and are therefore more complex devices.

6.2 WAN Reliability

WANs are usually considered less reliable than LANs for the following reasons:

- The WAN can potentially be a single point of failure
- The WAN link is subject to a number of outside factors, e.g., clouds on a radio link, aircraft reflections, road maintenance, water ingress, and frost heave
- WAN availability is dependent on a third-party provider. The availability of the WAN link is defined in the service level agreement (SLA). In some cases when an outage occurs, it may be more economical for the service provider to pay the cash penalties rather than to restore the link to service.

The following techniques can be used to eliminate single points of failure in the WAN:

- Multiple WAN links and routers can improve the link availability. Routing protocols such as VRRP and HSRP (hot standby routing protocol) ensure that one or both links are available
- When building redundancy into a WAN, the use of more than one service provider is recommended to avoid a situation where the service provider becomes a single point of failure
- The use of different physical media for implementing WAN links can provide another level of resiliency, e.g., fibre, copper, and wireless
- The WAN links should use different building entry points
- The WAN links should follow different geographical paths

WAN Link Availability

- Depending on the SLA, a WAN link that is implemented with a single physical link may typically provide 99.95 percent availability.
  A WAN link that is implemented with redundant physical links can provide 99.99998 percent availability.

6.3 Additional Information

Additional reading on network design and how it impacts telephone system availability can be found in the Mitel White Paper entitled Network Design for Availability, DK117893, this document is available at Mitel OnLine.
7 Power

Seven-Layer Business Continuity Model

<table>
<thead>
<tr>
<th>People</th>
<th>Process</th>
<th>Geography</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Power Distribution

<table>
<thead>
<tr>
<th>Data Network</th>
<th>PBX Software</th>
<th>PBX Hardware</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If the telephone system needs to remain operational during a power outage, then the powering of all necessary network devices must be carefully considered. There are no single power sources that provide 5 '9s of availability. By some estimates, the North American power grid can be expected to provide an average availability of 3 '9s, however, grid power availability is dependant on many factors that vary from one geographical location to the next, such as the frequency and severity of electrical storms. If a telephone system availability figure of 5 '9s is required, then backup power systems need to be part of the network and system design.

7.1 System Powering Considerations

- In cases where equipment has redundant power inputs, each of the power inputs should be powered from its own dedicated AC branch circuit.
- In cases where redundant equipment is used (i.e., two routers), the equipment should be powered from separate dedicated branch circuits.
- Backup power systems need to be in place to maintain power to networking equipment when there is an AC power outage. This can be achieved with an appropriately sized uninterruptible power supply (UPS).
- IP phones should be powered from a central location via power over ethernet (PoE). The powered ethernet switch itself will require back up power from either a UPS or another backup power source.
- In the event of a grid power outage, loads must be switched back onto grid power in a controlled fashion once grid power has been restored, so that the power system is not overloaded with sudden demand.
- In some cases, AC power can be provided from different suppliers. When this is done, each AC feed should enter the building using different entry points.
- In addition to meeting the customer’s business continuity requirements, the backup power system must meet local requirements for emergency (i.e., 911, 999, 112) telephone service.
Some equipment may require a controlled environment for proper operation. In such situations, heating and cooling systems will also require backup power.

General design recommendations and detailed product powering information can be found in the Mitel 3300 IP Communication Platform Product documentation.

8 Geography

7-Layer Business Continuity Model

<table>
<thead>
<tr>
<th>People</th>
<th>Process</th>
<th>Geography</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Power Distribution</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Data Network</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PBX Software</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PBX Hardware</td>
</tr>
</tbody>
</table>

8.1 The Advantages of Geographical Distribution

For a large business or a business providing a critical resource, it may be necessary to distribute the operation between multiple sites. In the event of a single failure, areas that remain operational can continue to provide service. An example might be a large multinational company or a financial institution that may have local data backup, but also a remote backup facility in another city or country.

The distribution of personnel, network and system functionality across multiple sites can offer resiliency and protection against many site specific perils such as:

- Power outages
- WAN link outages
- Fire, floods, and severe weather
- Earthquakes and natural disasters
- Civil unrest and acts of war
- Labor disputes
- Environmental disasters
- Disease and pandemics

In cases where a business requires the utmost in continuity, the business will maintain a backup location as part of their disaster recovery plan. The network and data storage facilities at the backup location will, to a certain degree, mirror the infrastructure at the primary location.
8.2 Multiple Locations and Disaster Recovery
The Mitel PBX can be configured as both a central and distributed system. A distributed phone system can provide backup from multiple geographic locations, such as a dedicated disaster recovery site, or a designated secondary controller at an alternate location. IP networks are essentially self-healing, so loss within one location can be quickly offset by setting up at a different location. Such a possibility also allows a disaster recovery controller to potentially be mobile and not tied to one location.

8.3 Additional Information
Geographically distributed systems are discussed in more detail in the Mitel White Paper entitled Geographically Distributed IP Phone Systems, DK117894 and the Mitel Product Information Note entitled IP Infrastructure Architecture, EM003647.

9 Process

Seven-Layer Business Continuity Model

People
Process
Geography
Power Distribution
Data Network
PBX Software
PBX Hardware

Companies should evaluate their internal processes to ensure that their business continuity goals will be met. There should be processes and procedures in place to address the following:

- Problem reporting and tracking
- Preventative maintenance
- A method for returning failed units to service and a process to maintain the spare units inventory
- Database backup schedules and a method of safely storing and recovering backups
- Mechanisms that ensure that new security updates and software updates are verified and installed
- A process that ensures that data stored or mirrored at multiple sites remains synchronized.

The processes and procedures need to be robust enough to ensure that there is not a procedural breakdown for reasons such as employee absence, major weather occurrences, or disasters. A company needs to have disaster recovery programs in place that address both short-term and long-term recovery plans. The processes need to be tested on a regular basis to ensure that they are providing the expected service results and if not, updated; for it is too late to test after an event has occurred.
People are included in the seven-layer business continuity model because people are the business. People are also required to operate, troubleshoot, and maintain networks and business systems. Ensuring that there are adequate personnel to handle the business demands also needs to be considered. Business continuity plans need to address personnel issues such as:

- Who are the individuals that are required to operate, troubleshoot, and maintain the systems?
- What are the availability requirements for personnel, and are different shifts required to ensure coverage?
- Must personnel be on-site or can they be reached via pagers or cell phones? Must some personnel work off-site?
- The plan needs to consider employee vacations, illness, and training courses.
- Do personnel need to support multiple sites, and if so are travel times acceptable or should personnel be based at each geographic location?
- Do support personnel have access to the necessary tools, spare parts, and resources that they require to carry out their functions at all sites?

11 Key Conclusions

- System availability rather than MTBF is what matters to the end users of a system.
- Availability is a key factor that influences how customers perceive a business.
- IP telephony solutions availability depends on much more than just the PBX hardware; it relies on all the elements of availability. As a result, there are many areas that need to be considered.
- Mitel products can be successfully configured for varying levels of solutions availability, including 5 ‘9s availability.
Redundancy means multiple units have to fail for a service to be lost.
Remember that maintaining high availability implies ongoing recovery mechanisms, which may include hardware maintenance. Ensure that the maintenance contract takes ongoing and timely support into consideration.
If in doubt, ask "Is this a single point of failure?", "What if this X isn’t working, what happens to the service, is it lost?"
People often talk of MTBF, but really mean MTTF.
MTBF can be used for a number of calculations. What is it being used for? Consider service critical components when calculating availability.
There are many organizations dedicated to furthering the science of business continuity planning and many of these organizations offer reference libraries for their members that discuss business continuity solutions in detail. The system administrator should consider membership in one of these organizations.

12 Availability Solution Examples
The following examples show how different levels of system availability can be achieved.

12.1 Base Availability – 3 '9s: Non-Resilient Controller
Consider the following example. A company has 1,000 phones in total. There are 900 phones located at the headquarters and 100 phones located at the branch office. The company has one (non-resilient) MXe Controller without subsystem redundancy located at the headquarters. The headquarters LAN is non-redundant and the headquarters and the branch office are interconnected with a non-redundant WAN link.

Solutions Availability
This non-resilient controller configuration provides 3 '9s availability.

Solutions Availability = 99.91%

Configuration
• Non-resilient controller
• A single MXe Controller without subsystem redundancy
• Non-redundant LAN configuration
• Non-redundant WAN

Assumptions
• Three restarts per year with controller outage of 10 minutes (See Note 1)
• Six minutes to restart 90 percent of phones (See Notes 1 & 2)
• WAN SLA of 99.95 percent availability
Availability

- 0.99996 hardware for base controller
- 0.9999 total software outage of 48 minutes, three outages per year and reconfiguration times
- 0.9992 non-redundant LAN
- 0.9995 non-redundant WAN, assume three outages per year

NOTE 1: Restarts can occur for a number of reasons, for example service, maintenance, upgrades, software recovery, network reconfiguration, and network outages.

NOTE 2: Ninety percent of the phones will restart within two minutes. If there are three restarts per year, the total time per year for 90 percent of the phones to restart will be six minutes.

Figure 8: Base Availability
12.2 Medium Availability – 4 ‘9s: Resilient Controllers and Redundant LAN

In this example, a company has 1,000 phones in total. There are 900 phones located at the headquarters and 100 phones located at the branch office. The company has two MXe Controllers with subsystem redundancy located at the headquarters and these controllers are configured for resilient operation. The headquarters LAN is redundant and headquarters and the branch office are interconnected with a non-redundant WAN link.

Solutions Availability

This resilient controller configuration provides 4 ‘9s of availability.

Solutions Availability = 0.99999%

In this example there are a number of users that are reliant on the WAN connection. As a result the overall solutions availability figure is negatively impacted by the WAN link availability figure. In a LAN-only installation, availability of 5 ‘9s would be possible because all users would be connected to the controller via the LAN.

Configuration

- Resilient controllers used in a load-sharing fashion so that 500 phones are registered on each controller
- Two MXe Controllers with subsystem redundancy
- Redundant LAN configuration based on STP
- Non-redundant WAN
- Non-redundant LAN at branch office

Assumptions

- Three restarts per year with controller outage of 10 minutes (See Note 1)
- Forty seconds to fail-over 90 percent of phones (See Notes 1 & 2)
- WAN SLA of 99.95 percent availability

Availability

- 0.99999999 hardware for controller pair
- 0.99992 registration of 90 percent of phones
- 0.99999 redundant LAN
- 0.9995 WAN

NOTE 1: Restarts can occur for a number of reasons, for example, service, maintenance, upgrades, software recovery, network reconfiguration and network outages.

NOTE 2: As of 3100 ICP Release 8.0, 90 percent of the phones will fail-over within 40 seconds. If there are three restarts per year, the total time per year for 90 percent of the phones to fail-over will be 120 seconds or two minutes.
12.3 High Availability – 5 ‘9s: Resilient Controllers, Redundant LAN and WAN

In this example, a company has 1,000 phones in total. There are 900 phones located at the headquarters and 100 phones located at the branch office. The company has two MXe Controllers with subsystem redundancy located at the headquarters and these controllers are configured for resilient operation. The headquarters LAN is redundant and headquarters and the branch office are interconnected with redundant WAN links.

Solutions Availability

This resilient configuration provides 5 ‘9s of availability.

| Solutions Availability = 99.9992% |

Configuration

- Resilient controllers used in a load-sharing fashion so that 500 phones are registered on each controller
- Two MXe Controllers with subsystem redundancy
- Redundant LAN configuration based on RSTP
- Redundant WAN Links running VRRP
- Redundant LAN at branch office

Assumptions

- Three restarts per year with controller outage of 10 minutes
- Forty seconds to fail-over 90 percent of phones (See Notes 1 & 2)
- WAN SLA of 99.95 percent availability for each WAN link
TELEPHONE SYSTEM AVAILABILITY

Availability

- 0.99999999 hardware for controller pair
- 0.999992 registration time for 90 percent of phones
- 0.999999 redundant LAN
- 0.999996 dual WAN links and VRRP

NOTE 1: Restarts can occur for a number of reasons, for example service, maintenance, upgrades, software recovery, network reconfiguration and network outages.

NOTE 2: As of 3300 ICP Release 8.0, 90 percent of the phones will fail-over within 40 seconds. If there are three restarts per year, the total time per year for 90 percent of the phones to fail-over will be 120 seconds or two minutes.

Figure 10: High Availability

12.4 Alternate High Availability - 5 ‘9s: Resilient Controllers, Redundant LAN, and Survivable Remote Gateway

In this example, a company has 1,000 phones in total. There are 900 phones located at the headquarters and 100 phones located at the branch office. The company has two MXe Controllers with subsystem redundancy located at the headquarters and these controllers are configured for resilient operation. The headquarters LAN is redundant and headquarters and the branch office are interconnected with a non-redundant WAN link.

A CX Controller / CXi Controller is located at the branch office and acts as a local survivable gateway. This provides phones at the branch office with local PSTN access and allows phones at the branch office to continue to operate locally should there be a WAN link failure or a PBX failure at the headquarters.
Solutions Availability

This resilient configuration with rapid phone registration provides 5 '9s of availability.

Solutions Availability = 99.9992%

Configuration

- Resilient controllers used in a load-sharing fashion so that 500 phones are registered on each controller
- Two MXe Controllers with subsystem redundancy
- Redundant LAN configuration based on RSTP
- Registration loading is shared between the headquarters and the branch office
- Local gateway at the branch office continues to provide service to the branch office in the event of a WAN failure.

Assumptions

- Three restarts per year with controller outage of 10 minutes
- Forty seconds to fail-over 90 percent of phones (See Notes 1 & 2)
- WAN SLA of 99.95 percent availability

Availability

- 0.99999999 Hardware for controller pair
- 0.999992 Registration time for phones
- 0.999999 Redundant LAN
- 0.9995 Single WAN link
- 0.99995 Hardware for branch office controller

Note 1: Restarts can occur for a number of reasons, for example service, maintenance, upgrades, software recovery, network reconfiguration and network outages.

Note 2: As of 9300 ICP Release 8.0, 90 percent of the phones will fail-over within 40 seconds. If there are three restarts per year, the total time per year for 90 percent of the phones to fail-over will be 120 seconds or two minutes.
12.5 Geographically Distributed Call Control

In this configuration local call control is provided at each of the remote locations. The sites and controllers are networked together to produce a single virtual system. Failure of call control at one location is offset by other locations taking over control of the phones at the remote site. In this situation the WAN link is no longer central to the operation but rather is part of the backup solution. It therefore only becomes critical during a failure condition.

Solution Availability

\[
\text{Solution Availability} = 99.9991\%
\]

**Configuration**

As this is a complex configuration, each of the sites needs to be evaluated in turn to determine the individual impact each site has on the overall availability. As additional backup sites are included, the limiting factor on the level of backup becomes the availability of the common WAN link. If business availability cannot be achieved with this single link, then an additional link can be added to improve this. It must therefore be assumed that the backup sites have minimal impact on the result and that the WAN link is the overriding factor.

The configuration is therefore:

- Local Call Control, in parallel with serial combination of
- A number of backup sites in parallel
- Connected to the local site through a common WAN link
Assumptions

The big assumption is that the backup sites through the WAN link will provide sufficient local site backup and availability. This can be tested. Other assumptions include:

- Only a single site fails
- Backup is only critical during a local outage
- Later calculations can show that 3 backup sites are needed to this local site
- All sites have equal system and geographic availability
- All sites back each other up

Can the backup sites through the WAN link provide the necessary level of availability? To test this, the parallel combination of local call control and the WAN link must be greater than the required business availability. Thus:

- 99.98 % Local call control availability
- 98.33% Local geographic site availability
- 99.95% in parallel with WAN availability
- 99.9992% Combined Availability – This meets the criteria. (Note that this result assumes infinite number of backup sites. In practice this is not the case, so actual result may be smaller – see below)

Availability

- 98.31% Local Controller availability, including geographic availability
- 99.95% WAN availability to this site
- 98.31% Availability of individual backup site, assuming all sites are equal
- 3 backup sites, 1 local
- 99.9995% for three backups sites in parallel
- 99.9991% System Combined

This example takes into consideration layers 1 to 5 of the business continuity model, taking full account of the geographic availability across multiple sites.

Three backup sites are needed to achieve the desired level of business availability of five ‘9’s.
13 Glossary

AC  Alternating Current: the form in which electrical power is delivered to end users over the power distribution grid.

AFR  Annualized Failure Rate: reciprocal of the MTBF. This indicates the percentage of failures to be expected over the period of one year.

Availability  The percentage of time that a system is in an operational state.

CO  Central Office: The physical building that houses inside plant equipment including the telephone switch. Also known as a local telephone exchange.

CPU  Central Processing Unit: The processor in a digital computer or the processor in a piece of equipment that contains an embedded computer.

HSRP  Hot Standby Routing Protocol: A Cisco proprietary protocol used to increase availability of default gateways used by end hosts.

ICP  IP Communications Platform: Includes gateway function, call control, plus a number of other features, such as voice mail.

IP  Internet Protocol: An encapsulation protocol that allows data to be passed from one end user to another. Typically this was over the Internet, but the same protocol is now used within businesses.

LED  Light Emitting Diode: A semiconductor device that emits light when it conducts current, used as visible indicators on electronic equipment.

Layer 2  Also called the data link layer, this is the second layer of the seven-layer OSI model.

Layer 3  Also called the network layer, this is the third layer of the seven-layer OSI model.

LAN  Local Area Network: A communications network that interconnects computers in a small geographic area such as a business.

LACP  Link Aggregation Control Protocol: Described in the IEEE standard 802.3ad, it is a networking protocol that allows multiple physical links to be grouped or aggregated together to form a single logical communications link.

MTTF  Mean Time to Failure: This is the statistical length of time that a component will operate before a failure occurs.

MTTR  Mean Time to Repair: This is the length of time that will elapse from the time a component failure was detected to the component being restored to service.

MTBF  Mean Time between Failures: This is the statistical length of time elapsed between sequential component failures. This metric also includes the time it takes for the component to be restored to service. In other words, MTBF is equal to the MTTF plus the MTTR.

PBX  Private Branch Exchange: A telephone switch or exchange that serves a particular business.
PoE  Power over Ethernet: A method of transmitting power over unshielded twisted pair LAN wiring, so that end devices can be powered with the same wires that are used for data communications. The most prevalent form of PoE is described in IEEE standard 802.3af.

PSTN  Public Switched Telephone Network: The telephone network that provides local and long distance connections, eg., Bell, AT&T, BT.

RAID  Redundant Array of Independent Disks: It is an array of hard drives on which the information is duplicated. A controller manages the disks, switching automatically from the primary to the secondary in the event of the failure of the primary hard drive.

Reliability  Is the capability of a system to remain operational for a given period of time under well-defined conditions.

Redundancy  Refers to the practice of providing a component, subsystem or a system with a backup to increase system availability.

Resiliency  In networking and communications, it is the ability to maintain an acceptable level of availability under fault conditions.

RSTP  Rapid Spanning Tree Protocol: A means whereby the network can determine that there are multiple paths between two points and disconnect them to leave a single path, removing broadcast issues. RSTP is described in the IEEE standard 802.1w.

SLA  Service Level Agreement: In data communications, a contract between the service provider and the customer that defines the availability of the communication service being provided.

TDM  Time Division Multiplex: A means of combining a number of digitally encoded data or voice channels onto a common digital stream, eg., T1.

UPS  Uninterruptible Power Supply: A unit capable of providing output power for a period of time when the local mains supply fails. Usually relies on storage devices such as batteries.


VRRP  Virtual Router Redundancy Protocol: A protocol described in RFC 3768. Allows two or more routers to be addressed by a single IP address so that hosts can access redundant routers with a single default gateway IP address.

WAN  Wide Area Network: A communications network used to provide communications links across metropolitan, regional, or national boundaries.
14 Appendix A – Availability Formulas

This appendix provides details on how to calculate availability figures for components in series and components in parallel, and shows the benefits of parallel operation over series operation. The calculations showing how availability is impacted with service recovery and without service recovery are also provided.

14.1 Series Operation versus Parallel Operation

For purposes of this example a device which we refer to as a “unit” has an availability figure of 0.9995 and an unavailability figure of 0.0005. The unit’s availability figure equates to the probability of the unit working successfully, and the unit’s unavailability figure equates to the probability of the unit not working successfully.

Therefore the following definitions can be made:

- \( S \) = Probability (Success) = 0.9995
- \( F \) = Probability (Failure) = 0.0005

Two Units Operating in Series

When two units with the same probability of success are operating in series there is only one possible state of operation that will provide overall success:

- Both units operating successfully, this is expressed as \( S \times S \)

To determine the overall probability of success when two units with the same probability of success are operating in series the following formula is used:

\[ S \text{ (overall)} = S \text{ (unit)}^2 \]

Substituting values for \( S \text{ (unit)} \):

\[ S \text{ (overall)} = (0.9995)^2 = 0.99900025 \]

Therefore the probability of success will be 0.99900025, or the overall availability of this system will be 99.9 percent.

Two Units Operating in Parallel

When two units with the same probability of success are operating in parallel there are four possible states of operation:

- Both units operating successfully, this is expressed as \( S \times S \)
- One unit operating successfully and one unit in failure mode, this is expressed as \( S \times F \)
- One unit in failure mode and one unit operating successfully, this is expressed as \( F \times S \)
- Both units in failure mode, this is expressed as \( F \times F \)

However, when two units with the same probability of success are operating in parallel, there are only two modes of operation that will provide overall success:

- Both units operating successfully, this is expressed as \( S \times S \) or \( S^2 \)
- One of the two units operating successfully, this is expressed as \( SF + FS \) or \( 2SF \)
To determine the overall probability of success when two units with the same probability of success are operating in parallel, the following formula is used:

\[ S_{\text{overall}} = S_{\text{unit}}^2 + (2 \times S_{\text{unit}} \times F_{\text{unit}}) \]

Substituting values for \( S_{\text{unit}} \):

\[ S_{\text{overall}} = (0.9995)^2 + (2 \times 0.9995 \times 0.0005) = 0.99999975 \]

Therefore, the probability of success will be 0.99999975, or the overall availability of this system will be 99.9999 percent.

NOTE: These equations are a simplified version and rely on the serial and parallel parts being identical in reliability. More complex analysis is needed for dissimilar units.

Conclusion

As can be seen from the above examples, operating two units in series provides an overall availability of 99.9 and 3 '9s, but operating the same units in parallel provides an overall availability of 99.9999 percent or 6 '9s.