Data Center High Availability

- Introduction
- Data center infrastructure Tiers
- N - Base requirement
- Concurrent maintainability and testing capability
- Capacity and scalability
- Isolation
- Data center tiering
- Secure Operation

Introduction

In general Tier 3 and Tier-4 datacenter will have redundancy to provide the business continuity with high availability to continue their function under unplanned or adverse conditions that would otherwise interrupt the data center’s telecommunications service.
The consistency of datacenter operation is dependent on the tier level of datacenter that was designed. Tier 3 and Tier 4 datacenter will have redundant cross-connect area and pathway that are physically
separated.

The datacenter needs to have multiple access providers to provide services, redundant network equipment for network level redundancy.

Minimal response time of technical support required in performing repairs to achieve the reliability of equipment to get high uptime.

The datacenter needs to have multiple entrance pathways from the property line to the entrance room to avoid a single point of failure; the maintenance hole and entrance pathways should be on opposite sides of the building and be a minimum 20m apart.

Both access providers are required to install two entrance cables in the datacenter with two entrance rooms.
One must go to primary entrance room and another one goes to a secondary entrance room and both the primary and secondary entrance room must have conduits from each other to provide flexibility.

Ensure that there are multiple access providers with multiple diverse pathways from the access provider to the datacenter for the business continuity.

The datacenter team must ensure that its services are provisioned for from different access provider central offices and the pathways to these central offices are diversely routed and this route pathways should be physically separated by at least 20m at all points along their route.

The datacenter admin team should make sure that access providers install circuit provisioning equipment in both entrance rooms so that circuits of all required types can be provisioned from either room.

The access provider provisioning equipment on both entrances must be the same and one room’s equipment should be up if other room’s equipment goes down.

The distance between two entrance rooms must be 20meter and it must be in different fire protection zone.

A secondary main distribution area (MDA) provides additional redundancy, but at the cost of complicating administration. Core routers and switches should be distributed between the two

MDAs. Circuits should also be distributed between the two spaces.

Main and secondary distribution area must be placed in different fire protection zone and both of them have to get power from different power distribution units and from different cooling equipment. Redundant backbone cabling protects against an outage caused by damage to backbone cabling.

Redundant backbone cabling may be provided in several ways depending on the degree of protection desired.

Backbone cabling between two spaces, for example, an HDA and an MDA, can be provided by running two cables between these spaces, preferably along different routes.

If the data center has redundant MDAs or redundant IDAs, redundant backbone cabling to the HDA from each higher-level distributor (IDA or MDA) is not necessary.

The routing of cables from the HDA to the redundant IDAs or MDAs should follow different routes. Horizontal cabling to critical systems can be diversely routed to improved redundancy. There should be enough attention not to exceed maximum horizontal cable lengths when selecting paths.

Critical systems can be supported by two different HDAs if maximum cable length restrictions are not exceeded. The two HDAs should be in different fire protection zones for this degree of redundancy to provide maximum benefit.

Data center infrastructure Tiers

Single point of failure should be eliminated to improve redundancy and reliability, both within the data center and support infrastructure as well as in the external services and utility supplies.

This Standard includes four tiers relating to various level of resiliency of the data center facility infrastructure. The tier ratings correspond to the industry data center tier ratings as defined by the uptime institute.

This Standard includes four tiers relating to various levels of resiliency of the data center facility infrastructure.

A data center may have different tier ratings for different portions of its infrastructure. For example, a data center may be rated Tier 3 for electrical, but tier 2 for mechanical.

For the sake of simplicity, a data center that is rated the same for all subsystems (telecommunications, architectural and structural, electrical and mechanical) can be called out by its tier overall (e.g. a tier 2 data center would have a tier 2 rating in all subsystems).

All portions of the infrastructure are at the same level, the tiering should be called out specifically. For example, a data center may be a tier rating of T2 E3 A1 M2 where:

- telecommunications are tier 2 (T2)
- electrical is Tier 3 (E3)
- architectural infrastructure is tier 1 (A1)
- Mechanical infrastructure is tier 2 (M2)

Although typically a data center’s overall rating is based on its weakest component, there may be mitigating circumstances relative to that facility’s specific risk profile, operational requirements or other factors that justify the lower rating in one or more subsystems.

Different areas within a data center may also be built and or used at different tier levels dependent on operational needs.

In such cases care should be given to describe these differences, for example, an area of a data center that has a tier 2 risk avoidance profile because it has T2, E2, A2 M2 services within a facility that may be Tier 3.

Care should be taken to maintain mechanical and electrical system capacity to the correct tier level as the data center load increases over time. For example, a data center may be degraded from Tier 3 or tier 4 to tier 1 or tier 2 as redundant capacity is utilized to support new computer and telecommunications equipment.

N - Base requirement

System meets base requirements and has no redundancy. N+1 redundancy
N+1 redundancy provides one additional unit, module, path, or system in addition to the minimum required to satisfy the base requirement. The failure or maintenance of any single unit, module, or path will not disrupt operations.

N+2 redundancy

N+2 redundancy provides two additional units, modules, paths, or systems in addition to the minimum required to satisfy the base requirement. The failure or maintenance of any two single units, modules, or paths will not disrupt operations.
2N redundancy provides two complete units, modules, paths, or systems for every one required for a base system. Failure or maintenance of one entire unit, module, path, or system will not disrupt operations.

2(N+1) redundancy

2 (N+1) redundancy provides two complete (N+1) units, modules, paths, or systems. Even in the event of failure or maintenance of one unit, module, path, or system, some redundancy will be provided and operations will not be disrupted.

Concurrent maintainability and testing capability

The facilities should be capable of being maintained, upgraded, and tested without interruption of operations.

Capacity and scalability

Data centers and support infrastructure should be designed to accommodate future growth with little or no disruption to services.

Isolation

Data centers should be (where practical) used solely for the purposes for which they were intended and should be isolated from non-essential operations.

Data center tiering

Tier I Data Center: Basic

A Tier I data center is susceptible to disruptions from both planned and unplanned activity. If it has UPS or generators, they are single-module systems and have many single points of failure.

The infrastructure should be completely shut down on an annual basis to perform preventive maintenance and repair work. Urgent situations may require more frequent shutdowns. Operation errors or spontaneous failures of site infrastructure components will cause a data center disruption.

Tier II Data Center: Redundant Components

Tier II facilities with redundant components are slightly less susceptible to disruptions from both planned and unplanned activity than a basic data center. They have UPS, and engine generators, but their capacity design is “Need plus One” (N+1), which has a single-threaded distribution path throughout.

Maintenance of the critical power path and other parts of the site infrastructure will require a processing shutdown.

Tier III Data Center: Concurrently Maintainable

Tier III level capability allows for any planned site infrastructure activity without disrupting the computer hardware operation in any way.

Planned activities include preventive and programmable maintenance, repair and replacement of components, addition or removal of capacity components, testing of components and systems, and more.

Sufficient capacity and distribution must be available to simultaneously carry the load on one path while performing maintenance or testing on the other path.

Unplanned activities such as errors in operation or spontaneous failures of facility infrastructure components may still cause a data center disruption.

Tier IV Data Center: Fault Tolerant

Tier IV provides site infrastructure capacity and capability to permit any planned activity without disruption to the critical load. Fault-tolerant functionality also provides the ability of the site infrastructure to sustain at least one worst-case unplanned failure or event with no critical load impact.

This requires simultaneously active distribution paths, typically in a System + System configuration.

Tier 3

The data center should be served by at least two access providers. Service should be provided from at least two different access provider central offices or points-of-presences.

Access provider cabling from their central offices or points-of-presences should be separated by at least
20 m (66 ft.) along their entire route for the routes to be considered diversely routed.

The data center should have two entrance rooms preferably at opposite ends of the data center but a minimum of 20 m (66 ft.) physical separation between the two rooms.

Do not share access provider provisioning equipment, fire protection zones, power distribution units, and air conditioning equipment between the two entrance rooms. The access provider provisioning equipment in each entrance room should be able to continue operating if the equipment in the other entrance room fails.

The data center should have redundant backbone pathways between the entrance rooms, MDA, intermediate distribution areas (IDAs), and HDAs.

Intra-data center LAN and SAN backbone cabling from switches to backbone switches should have redundant fiber or wire pairs within the overall star configuration. The redundant connections should be in diversely routed cable sheathes.

There should be a “hot” standby backup for all critical telecommunications equipment, access provider
provisioning equipment, core layer production routers and core layer production LAN/SAN switches.

All cabling, cross-connects and patch cords should be documented using software systems or automated infrastructure management systems as described in the ANSI/TIA-606-B.

Some potential single points of failure of a tier 3 facility are:

- Any catastrophic event within the MDA may disrupt all telecommunications services to the data center; and any catastrophic event within a HDA may disrupt all services to the area it servers.

A tier 3 data center should have protection against most physical events, intentional or accidental, natural or manmade, which could cause the data center to fail.

All systems of a tier 3 facility should be provided with at least N+1 redundancy at the module, pathway, and system level, including the generator and UPS systems, the distribution system, and all distribution feeders.

The configuration of mechanical systems should be considered when designing the electrical system to ensure that N+1 redundancy is provided in the combined electrical-mechanical system.

This level of redundancy can be obtained by either furnishing two sources of power to each air conditioning unit or dividing the air conditioning equipment among multiple sources of power.

Feeders and distribution boards are dual path, whereby a failure of or maintenance to a cable or panel will not cause interruption of operations.

Enough redundancy should be provided to enable isolation of any item of mechanical or electrical equipment as required for essential maintenance without affecting the services being provided with cooling.

By employing a distributed redundant configuration, single points of failure are virtually eliminated from the utility service entrance down to the mechanical equipment, and down to the PDU or computer equipment.

To increase the availability of power to the critical load, the distribution system is configured in a distributed isolated redundant (dual path) topology. This topology requires the use of automatic static transfer switches (ASTS) placed either on the primary or secondary side of the PDU transformer.

Automatic static transfer switches (ASTS) requirements are for single cord load only.

For dual cord (or more) load design, affording continuous operation with only one cord energized, no automatic static transfer switches (ASTS) is used, provided the cords are fed from different UPS sources. The automatic static transfer switches (ASTS) will have a bypass circuit and a single output circuit breaker.

A central power and environmental monitoring and control system (PEMCS) should be provided to monitor all major electrical equipment such as main switchgears, generator systems, UPS systems, automatic static transfer switches (ASTS), power distribution units, automatic transfer switches, motor control centers, transient voltage surge suppression systems, and mechanical systems.

A separate programmable logic control system should be provided, programmed to manage the mechanical system, optimize efficiency, cycle usage of equipment and indicate an alarming condition.

The HVAC system of a Tier 3 facility includes multiple air conditioning units with the combined cooling capacity to maintain a critical space temperature and relative humidity at design conditions, with enough redundant units to allow failure of or service to one electrical switchboard.

If these air conditioning units are served by a water-side heat rejection system, such as a chilled water or condenser water system, the components of these systems are likewise sized to maintain design conditions, with one electrical switchboard removed from service.

This level of redundancy can be obtained by either furnishing two sources of power to each air conditioning unit or dividing the air conditioning equipment among multiple sources of power.

The piping system or systems are dual path, whereby a failure of or maintenance to a section of pipe will not cause interruption of the air conditioning system.
Redundant computer room air conditioning (CRAC) units should be served from separate panels to provide electrical redundancy.

All computer room air conditioners (CRAC) units should be backed up by generator power.
Refrigeration equipment with N+1, N+2, 2N, or 2(N+1) redundancy should be dedicated to the data center.

Enough redundancy should be provided to enable isolation of any item of equipment as required for essential maintenance without affecting the services being provided with cooling.

Subject to the number of Precision Air Conditioners (PAC’s) installed, and consideration of the maintainability and redundancy factors, cooling circuits to the Precision Air Conditioners (PAC's) should be sub-divided.

If chilled water or water-cooled systems are used, each data center dedicated sub-circuit should have independent pumps supplied from a central water ring circuit.

A water loop should be located at the perimeter of the data center and be in a sub floor trough to contain water leaks to the trough area.

Leak detection sensors should be installed in the trough. Consideration should be given to fully isolated and redundant chilled water loops.

Tier 4

Data center backbone cabling and distributor locations should be redundant.

Cabling between two spaces should follow physically separate routes, with common paths only inside the two end spaces.

Backbone cabling should be protected by routing through a conduit or by use of cables with interlocking armor.

There should be an automatic backup for all critical telecommunications equipment, access provider provisioning equipment, core layer production routers and core layer production LAN/SAN switches. Sessions/connections should switch automatically to the backup equipment.

The data center should have redundant MDAs preferably at opposite ends of the data center, but a minimum of 20 m (66 ft.) physical separation between the two spaces.

Do not share fire protection zones, power distribution units, and air conditioning equipment between the redundant MDAs. The redundant MDA is optional, if the computer room is a single continuous space, as there is probably little to be gained by implementing two MDAs in this case.

The two MDAs should have separate pathways to each entrance room. There should also be a pathway between the MDAs.

The redundant routers and switches should be distributed between redundant distribution spaces (e.g. redundant MDAs, redundant pair of IDAs, or redundant pair of HDAs, or redundant pair of entrance rooms).

Each HDA should be provided with connectivity to two different IDAs or MDAs. Similarly, each IDA should be provided with connectivity to both MDAs.
Critical systems should have horizontal cabling to two HDAs. Some potential single points of failure of a tier 4 facility are at:

- The MDA (if the secondary distribution area is not implemented).
- The HDA and horizontal cabling (if redundant horizontal cabling is not installed).

A tier 4 data center must consider all potential physical events that could cause the data center to fail. A tier 4 data center must be provided with specific and in some cases redundant protections against such events.

Tier 4 data centers should consider the potential problems with natural disasters such as seismic events, floods, fire, hurricanes, and storms, as well as potential problems with terrorism and disgruntled employees.

Tier 4 data centers should have control over all aspects of their facility. Tier 4 facilities should be designed
in a ‘2(N+1)’ configuration in all modules, systems, and pathways.

All feeders and equipment should be capable of manual bypass for maintenance or in the event of failure. Any failure should automatically transfer power to critical load from a failed system to the alternate system without disruption of power to the critical electronic loads.

A battery monitoring system capable of individually monitoring the impedance or resistance of each cell and temperature of each battery jar and alarming on impending battery failure should be provided to ensure adequate battery operation.

The utility service entrances should be dedicated to the data center and isolated from all noncritical facilities. The building should have at least two utility feeders from different utility substations for redundancy.

The HVAC system of a tier 4 facility includes multiple air conditioning units with the combined cooling capacity to maintain a critical space temperature and relative humidity at design conditions, with sufficient redundant units to allow failure of or service to one electrical switchboard.

If these air conditioning units are served by a water-side heat rejection system, such as a chilled water or condenser water system, the components of these systems are likewise sized to maintain design conditions, with one electrical switchboard removed from service.

This level of redundancy can be obtained by either furnishing two sources of power to each air conditioning unit or dividing the air conditioning equipment among multiple sources of power.

The piping system or systems are dual path, whereby a failure of or maintenance to a section of pipe will not cause interruption of the air conditioning system.

Alternative source for water storage is to be considered when evaporative systems are in place for a tier
4 system.

Secure Operation

Managing and operating a datacenter requires to follow tailored processes to reap expected results from the datacenter. While considering standard operating procedure (SOP); security in all the aspects is a most needed aspect. Datacenter SOP should be developed based on Standards like ISO 27001 and best practices like ITIL, which will provide the clear understanding on the control's requirements such as Administrative, Technical and Physical.

SOP should have minimum following aspects in it;

Security of Datacenter

The datacenter should take all the required security measures to guarantee the confidentiality, integrity, availability of their client's information, networks and services. Appropriate technical and organizational measures should be identified and put in place to ensure minimum level of security.

A comprehensive Information Security framework that includes the essential components such as but not limited to;

- Risk Assessment and Management;
- Configuration Management;
- Change Management;
- Incident Management;
- Secured application acquisition, development and maintenance;
- Business continuity plan and Disaster recovery plan;
- Vulnerability assessment and Audit;
- Internal and external penetration testing and
- Legal and Regulatory compliance identifying, maintaining and monitoring.

Data Centre Team is solely responsible for security of the Data Centre Infrastructure, Network and Communication Infrastructure and Servers and Applications.
The required minimum Service Levels and the management of Data center should include but not be limited to:

- Measurement and Reporting of Service Level achieved
- Service Level target for external Service Providers
- Data Centre commitment on what must be provided to Customers

IT Infrastructure Resources Management:
HVAC – Heating, Ventilation, Air conditioning (Cooling, Humidification, De-humidification)

- Operations Parameter for Data Centre Room
- Minimum required standby Spare Parts available on-site
- Regular monitoring Duties and Responsibilities

- Monitoring external Water Supply for HVAC Status and condition of UPS – Uninterrupted Power
  Supply
- Minimum required standby Spare Parts available on-site Racks
- Regular Duties and Responsibilities
- Required Spare Parts available on-site

Internal Network and Communication Infrastructure Management

- Core Switches
- Cabling

Network and Communications management

- Broadband connections

Servers and Applications management

- Servers Management life cycle
- Rack-Mounting of Servers and Cabling
- Installation of Operating System
- Installation of Monitoring Agents and connect to Monitoring System
- Installation of Agents for Backup-System and configuring Backup
- Regular Monitoring by Manual Checks and / or automatic Warnings and Alarm from Monitoring
  System
- Infrastructure Applications
- Domain Name Servers (DNS)
- Central Authentication Server

Common ICT Processes

- Day to day operations procedures
- Emergency Reboot of Servers as and when required
- Regular Reboot of Serves Scheduled / unscheduled
- Onsite Spare Parts / Reserve Parts for Servers planning and making available
- Replacing faulty parts / parts with limited life time in Servers
- Regular Restoration Tests of Servers and Services
- Access control Process for allowing customer / visitor Data Centre
- Processing access Requests to locked cages / locked racks
- Capacity Management for Data Centre
- New acquisition / Project triggered processes
- Adding / Modify new Hardware to the Data Centre
- removing of Hardware and other Equipment
- Software Life-Cycle Management of Data Centre related software
- Third party / vendor of Support Contracts
- Regular recurring processes

- Data Centre Capacity Planning / Management

Disaster Recovery Planning

- DR – Tests
- Fail-over Test from public Power Supply to UPS

Asset management
Resources:

- Human Resources – Roles and Responsibilities
- Shift / Rota Planning
- Staff Technical Training and Certifications
- Capacity Building