Microsoft

Resiliency/Reliability:

Azure keeps your applications up and running and your data available

Trusted Cloud:
Microsoft Azure Security, Privacy, Compliance, Resiliency, and Protected IP

Author
Debra Shinder

Resiliency is not about avoiding failures but responding to failures. The objective is to respond to failure in a way that avoids downtime and data loss. Business continuity and data protection are critical issues for today's organizations, and business continuity is built on the foundation of resilient systems, applications, and data.

Reliability and resiliency are closely related. Reliability is defined as dependability and performing consistently well. Resiliency is defined as the capacity to recover quickly. Together, these two qualities are key to a trustworthy cloud service.

Despite your best efforts, disasters happen; they are inevitable but mostly unpredictable, and vary in type and magnitude. There is almost never a single root cause of a major issue. Instead, there are several contributing factors, which is the reason an issue is able to circumvent various layers of mitigations/defenses.

There is no way to always prevent bad things from happening. All we can do is add layers and minimize gaps.

Making systems reliable

Making systems reliable in the public cloud is not the same as in your own datacenter. The cloud is an ever-changing, constantly evolving platform, unlike the usual on-premises IT model where you can achieve greater availability by avoiding change. In the public cloud, change is both inevitable and beneficial, but you must plan for it.

Complex systems can fail in complex ways, and you need resilience to deliver reliability. Reliability is the goal, whereas resilience is the method by which you achieve that goal.

Azure was the first cloud platform to provide a built-in backup and disaster recovery solution.

Achieving resilience

Resilience begins with availability. Reducing the amount of downtime and the number of disruptions to service is important to the continued operation of your core business functions. Organizations today are dependent on their online presence for communications with vendors, sales to customers, financial transactions, and more. Downtime means lost revenue and can damage your business reputation.

High availability is about providing undisrupted continuity of operations whereas disaster recovery is about recovering from a natural or human-induced outage and providing continuity of operations. Disaster recovery usually involves some amount of downtime.

The direct and indirect monetary cost of down time for your organization depends on several factors, including your field of business, how you do business (percentage of sales made online), total revenues, time of day or day of week, and so forth. According to the Ponemon Institute, the average cost of a datacenter outage can be $9000 per minute.

An effective disaster recovery strategy has two parts: preparedness and recovery. The two are closely related but are not the same. Preparedness is the theoretical plan for the procedures you will follow in response to a catastrophic event, whether it's physical destruction of systems due to a natural disaster or a devastating cyberattack. Recovery is the actual implementation of the processes that make up that plan.

Your data is a valuable asset. According to multiple studies, the leading cause of data loss is human error, attributable to both users and IT professionals. Better training and tighter access controls can help reduce the incidence, but data loss can still occur. To implement a full recovery, you need a good backup system so that if your data is lost or corrupted, you can restore it.

Azure helps you to avoid many potential disasters and quickly recover if your organization does get hit by disaster. Azure offers resiliency for your cloud-based applications and data by providing for business continuity in the following ways:

- High availability
- Disaster recovery
- Backup
- Resilient app design best practices

Shared responsibility

Just as security is a responsibility that's shared between cloud provider and customer, building systems that will survive failure is also a shared responsibility. Microsoft builds and operates the resilient foundation, then you choose to enable relevant services to help with your resiliency needs. Your apps and workloads sit on top of both.

Azure resilient foundation

Everything is built on top of the resilient foundation, which is a requirement for any application to achieve resiliency. To achieve resilience—the application on top has to take advantage of the resilient services built on the foundation.

The three pillars of the Azure resilient foundation are:

- Design: How Microsoft designs its global fiber network, evolving datacenters, and storage protections built into the Azure platform.
- Operate: How Microsoft rolls out releases into the environment, performs maintenance (planned and unplanned), and uses machine learning to predict failures and protect customer workloads.
- Observe: How customers can observe what's happening in their environment(s), inform people and systems to make informed decisions before/during issues, and determine their own availability requirements.

IT systems are subject to failure. If all your systems and data are located on your premises, a fire, flood, tornado, or other natural disaster can bring your operations to a halt for weeks or months. When your applications and data are hosted on a cloud service, you have redundant, distributed implementations of your IT resources across physical locations.

Even in the cloud, however, hardware can fail, the network can have transient failures, and rarely, an entire service or region may experience a disruption. Thus, your cloud service provider's dedication to business continuity is vital. Azure provides a comprehensive set of native business continuity solutions that protect you against failures within datacenters and even the failure of entire datacenters.

Business continuity is based on the ability to perform essential business functions during and after adverse conditions, such as a natural disaster or a downed service. Azure is the first hyperscale cloud provider to be certified under ISO-22301, the first international standard to demonstrate the ability to prevent, mitigate, respond to, and recover from disruptive incidents.

The reliability and performance of cloud services are determined in part by the network and (in addition to having more datacenter regions than any of our competitors) the Microsoft network is also one of the largest in the world. Unlike with many other public cloud providers, data that traverses between Azure datacenters and regions doesn't go through the public internet—it stays in the Microsoft network.

High availability

A key aspect of a resilient foundation is availability. High availability is all about maintaining acceptable continuous performance despite temporary failures in services, hardware, or datacenters, or fluctuations in load. Highly available systems are consistently operational over long periods of time.

Azure uptime, expressed as a rolling 12 month average to June 2019, was 99.996%, or approximately 26 minutes of downtime per year. Availability can never be 100% because hardware and software failures happen, and human error occurs. But the Service Level Agreement (SLA) describes our commitment for uptime and connectivity. Microsoft provides SLAs that define the guaranteed availability levels for each Azure service. Microsoft also provides support for high availability at the virtual machine, datacenter, and regional levels, through a number of features and functions across the categories of compute, storage, and networking.

Availability is often expressed as percentage of uptime, using a "table of nines." For example, if the level of availability over a year is 99.99%, it is said to be "four nines." This translates to average downtime of 1.01 minutes per week, 4.32 minutes per month, or a total of 52.56 minutes per year.

Traffic Manager and Load Balancer can be used individually, or you can use them together or in combination with Azure Application Gateway to create a deployment that is geographically redundant.

Azure Availability Zones. This provides redundancy at the regional level. It is a high-availability offering that protects your applications and data from datacenter failures. Availability Zones are unique physical locations within an Azure region. The physical separation of Availability Zones within a region protects applications and data from datacenter failures. With Availability Zones, Azure offers industry best 99.99% virtual machine uptime SLA.

Learn more about Azure Availability Zones.

Availability Sets. This provides redundancy at the datacenter level. An Availability Set is a logical grouping capability that you can use in Azure to ensure that the VM resources you place within it are isolated from each other when they are deployed within an Azure datacenter. If a hardware or Azure software failure occurs, only a subset of your VMs are impacted, and your overall application stays up and continues to be available to your customers.

Learn more about deploying highly available VMs by creating an Availability Set.

Data residency boundary. This provides redundancy across two regions that share the same regulatory requirements for data replication and storage. Your data is protected from loss of an entire region with geo-redundant storage and Azure Site Recovery.

Learn more about Azure Regions.

Azure Load Balancer. Load Balancer can scale your applications and create high availability for your services. Load Balancer automatically scales with increasing application traffic, and you can use the internal load balancer for traffic between virtual machines inside your private network.

Learn more about Azure Load Balancer.

Azure Traffic Manager. This is a DNS-based traffic load balancer that enables you to distribute traffic optimally to services across global Azure regions, while providing high availability and responsiveness.

Learn more about choosing the correct Azure load balancing solution.

Azure compute resiliency solutions. You can apply autoscaling to virtual machines for high availability and easily spread your workloads across the virtual machines in your virtual machine scale set.

Network support for high availability. You can deploy VPN and ExpressRoute gateways in Azure Availability Zones. This brings resiliency, scalability, and higher availability to virtual network gateways. Deploying gateways in Azure Availability Zones physically and logically separates gateways within a region, while protecting your on-premises network connectivity to Azure from zone-level failures.

Highly available storage options. The data in your Azure storage account is always replicated to ensure durability and high availability. Azure Storage replication copies your data so that it is protected from planned and unplanned events ranging from transient hardware failures, to network or power outages, to massive natural disasters, and so on. You can choose to replicate your data within the same datacenter, across zonal datacenters within the same region, and even across regions.

When you create a storage account, you can select one of the following replication options:

- Locally redundant storage (LRS) replicates your data within a storage scale unit that is hosted in a datacenter in the region in which you created your storage account.
- Zone-redundant storage (ZRS) replicates your data synchronously across three storage clusters in a single region, where each storage cluster is physically separated from the others and resides in its own availability zone.

- Geo-redundant storage (GRS) replicates your data to a secondary region that is hundreds of miles away from the primary region.
- Read-access geo-redundant storage (RA-GRS) provides read-only access to the data in the secondary location, in addition to geo-replication across two regions.

Learn more about these Azure storage replication options.

Disaster recovery

Your disaster recovery strategy is key to business continuity. Site recovery and data backup are elements of a disaster recovery plan. Organizations using the cloud tend to take the reliability of the public cloud for granted, not recognizing that they may be responsible for choosing and implementing backup and recovery mechanisms.

As a cloud customer, you will confront more opportunities to spend extra time and money on optional backup than you can ever take advantage of, so you need to make explicit and careful choices as to what you will and will not do.

Your disaster recovery plan should:

1. **Identify** and classify the threats and risks that may lead to disasters.
2. **Define** the resources and processes that ensure business continuity during the disaster.
3. **Define** the reconstitution mechanism to get the business back to normal from the disaster recovery state, after the effects of the disaster are mitigated.

An effective disaster recovery plan plays its role in all stages of operations and it is continuously improved by disaster recovery mock drills and feedback capture processes.

Disaster recovery happens in the following sequential phases:

1. **Activation Phase:** In this phase, the disaster effects are assessed and announced.
2. **Execution Phase:** In this phase, the actual procedures to recover each of the disaster-affected Azure services are executed. Business operations are restored into the Azure paired region.
3. **Reconstitution Phase:** In this phase the original Azure region hosted system/service is restored, and execution phase procedures are stopped.

Microsoft provides tools and services to help you implement and test your disaster recovery plan.

Azure Site Recovery is the Azure built-in disaster recovery as a service (DRaaS) solution that can help keep your applications up and running during an IT outage. You can ensure compliance by testing your disaster recovery plan without impacting production workloads or end users and keep applications available during outages with automated recovery from on-premises to Azure or Azure to another Azure region.

Azure helps you to reduce the cost of deploying, monitoring, patching, and maintaining on-premises disaster recovery infrastructure by eliminating the need for building or maintaining a costly secondary datacenter. You pay only for the compute resources you need to support your applications in Azure.

Several different types of disaster scenarios can affect a customer's current Azure infrastructure topology. Region-wide service disruptions are not the only cause of application-wide failures. Poor design and administrative errors can also lead to outages. It's important to consider the possible causes of a failure during both the design and testing phases of your disaster recovery plan. A good plan takes advantage of Azure features and augments them with application-specific strategies.

Learn more about Azure Site Recovery.

Backup

Azure Backup helps you reduce data restoration time and reliability challenges. It's built into the Azure platform, with seamless support for virtual machines running in Azure and on-premises. It's cost effective because it doesn't require any additional infrastructure. Multiple authentication layers help to keep your data safe and guard against loss from ransomware.

Data backup is a critical part of disaster recovery. If the stateless components of an application fail, you can always redeploy them. If data is lost, the system can't return to a stable state. Data must be backed up, ideally in a different region in case of a region-wide disaster.

Azure provides resiliency for your databases. Azure Backup automatically discovers if a selected virtual machine is running SQL and backs up your SQL database natively with support for fifteen-minute recovery time objective (RTO). Azure also provides data resiliency. You can back up important files natively in Azure, with item-level restore. Azure Backup supports full, differential, and incremental backup.

Backup is distinct from data replication. Data replication involves copying data in near real time, so that the system can fail over quickly to a replica. Data replication can reduce the length of time it takes to recover from an outage by ensuring that a replica of the data is always standing by. However, data replication won't protect against human error. If data is corrupted because of human error, the corrupted data just gets copied to the replicas.

Thus, you still need to include long-term backup in your disaster recovery strategy.

Learn more about Azure Backup.

Resilient app design best practices

Your mission-critical applications and data should be built for resiliency. One of the primary ways to make an application resilient is through redundancy. You need to plan for this redundancy when you design the application. The level of redundancy that you need depends on your business requirements; in general, there is a tradeoff between greater redundancy and reliability versus higher cost and complexity.

Azure has a number of features to make an application redundant at every level of failure, from an individual VM to an entire region. These include Availability Sets and Availability Zones as well as Azure Site Recovery and Azure Backup.

The Azure Architecture Center provides detailed information to get started quickly and build apps correctly the first time. This includes guidance on building for security, scalability, performance, cost, and manageability—including tested deployment scripts and verified recommendations for your production workloads.

Learn more about designing resilient applications for Azure.


Microsoft