DISASTER RECOVERY WITH IBM CLOUD SERVERS

S.No	Title	Page No
1.	Project Objective	2
2.	Design Thinking	3
3.	Innovation	4
4.	IBM Cloud Backup	7
5.	Solution	8
6.	Development Phase I	9
7.	Development Phase II	12

1. Project Objective:

Today's enterprises cannot afford planned or unplanned system outages. Even a few minutes of application downtime can result in considerable financial losses, eroded customer confidence, damage to brand image, and public relations problems. To better control and manage their IT infrastructure, enterprises have concentrated their IT operations into large (and on-demand) data centers. These data centers must be resilient (and flexible) enough to handle the ups and downs of the global market. They also must manage changes and threats with consistent availability, security, and privacy, both around the clock and the world. Most of the solutions are based on an integration of operating system(OS) clustering software, storage, and networking. How a system, server, or environment handles failures is characterized as its RAS. In today's world of e-business, the RAS of an OS and the hardware on which it runs have assumed great importance. Today's business require that IT systems be self-monitoring, self-healing, maintained without outages. More IT systems are meeting this requirement through techniques such as redundancy and error correction to achieve a high level of RAS. To minimize interruptions to normal operations. To limit the extent of disruption and damage. To minimize the economic impact of the interruption. To establish alternative means of operation in advance. Ensure uninterrupted operation of essential services, applications, and data to minimize downtime and maintain business continuity, even in the face of catastrophic events. Implement robust data protection measures, including backups, encryption, and access controls, to safeguard sensitive information from loss, corruption, or unauthorized access during and after a disaster. Optimize resource allocation to balance cost considerations with performance requirements, and design the DR solution to scale seamlessly with the evolving needs of the organization. Implement robust monitoring tools and procedures to continuously

assess the health and performance of the primary and secondary environments, with automated alerts for any deviations from predefined thresholds. Recovery point objective (RPO) is the point in time relative to the failure to which you need preservation of data. Data changes preceding the failure or disaster by at least this time period are preserved by recovery processing. Zero is a valid value and is equivalent to a "zero data loss" requirement.

2. Design Thinking:

- A Hybrid cloud application is a mix of on-premises, private or public cloud platforms with orchestration between these distributed platforms and workloads to perform as a single business service.
- The hybrid cloud applications that are built on IBM Power servers are known for their high performance.
- All HADR solutions must be methodically tests regularly. It is better to find a problem during planned testing.

Empathize:

Understand Stakeholder Needs: Engage with key stakeholders including IT teams, business leaders, and end-users to gather insights on their disaster recovery requirements, pain points, and priorities.

Ideate:

Brainstorm Solutions: Facilitate brainstorming sessions with cross-functional teams to generate a wide range of creative solutions. Encourage out-of-the-box thinking.

IBM Cloud Features and Capabilities: Leverage knowledge of IBM Cloud services, such as Virtual Servers, Object Storage, and Cloud Databases, to ideate solutions that make the most of the platform.

Prototype:

Conceptualize DR Architecture: Create visual representations of the proposed disaster recovery architecture using tools like diagrams or cloud modeling software. This could include primary and secondary site configurations, data replication methods, and failover mechanisms.

Simulation Tools: Utilize simulation tools or sandboxes provided by IBM Cloud to create prototypes for testing various aspects of the DR plan.

Test:

Scenario-Based Testing: Conduct simulated disaster scenarios to test the proposed architecture's effectiveness in real-world situations. Evaluate factors like failover time, data integrity, and application functionality.

Iterative Testing and Feedback: Continuously refine and retest the prototypes based on feedback from stakeholders and lessons learned from each iteration.

Feedback and Iterate:

Gather Stakeholder Feedback: Engage stakeholders for feedback on the prototypes, and use their input to refine the disaster recovery plan further.

3. INNOVATION:

• IBM cloud is a packaged software offering which is used to setup a private cloud on the IaaS of the users choosing. Here we focus on a Disaster Recovery use case where ICP is used to setup a Kubernetes based Private cloud on VMWare as described in ICP backup.

• The Disaster Recovery site is being simulated using another VMware vCenter Server on IBM Cloud (VCS) with similar hardware. VMware Site Recovery Manager (SRM) is used to manage the Disaster Recovery of the VMs. SRM is expected to recover:

- Network
- VMs used as Nodes of the cluster
- Storage Volumes

• In addition to recovery of the VMs, the following Kubernetes state may need to be recovered if the distributed state gets corrupted. Therefore it is a good practice to backup this state and restore it in the recovered VMs if needed. The backup and restore process is described in ICP Component Backup

- Image Registry
- Cloudant DB
- MariaDB

High availability disaster recovery concepts

The following concepts are used in this chapter:

Split-brain or split-cluster: A cluster split-brain can occur when a subset of nodes in a cluster cannot communicate with the remaining nodes. Although it is possible for this situation to occur within thedata center, it is far more likely to happen to a cluster across data centers due to the greater exposure of the interconnecting networks to potential risk.

Tie breaker or third site: In HADR clusters, it is a best practice to use a tie breaker or a third site to prevent a split-brain situation. Although it isstill important to avoid this situation for clusters within a single data center, it is far less likely because multiple communication paths connect all nodes in the cluster, which is a less common situation between sites.

Split policy: When a split-brain situation occurs, each partition attemptsto acquire the tie breaker by placing a lock on the tie-breaker disk or on the NFS file. The partition that holdsthe lock on the SCSI disk or reserves the NFS file wins, and the other loses.

Synchronous replication: Writes are committed at the remote storage before an acknowledgment can be returned to the application. This delay degrades the application performance and limits thedistance between the application and the remote storage to around 80 - 120 km.

Asynchronous replication: Writes are cached locally in some form of non-volatile storage and an acknowledgment is returned to the application. Later, the write is committed to the remotestorage, and then the record is removed from the localcache.

3.2 IBM Power Virtual Server offering:

The IBM Power Virtual Server offering provides a secure and scalable server virtualization environment that is built on the IBM Cloud platform for on-demand provisioning. The IBM Power Virtual Servers are in IBM data centers, which are distinct from the IBM Cloud servers,with separate networks and direct-attached storage. The environment is in its own pod, and the internal networks are fenced but offer connectivity options to meet customer requirements. This infrastructure design enables IBM Power Virtual Server to maintain key enterprise software certification and support because the IBM Power Virtual Server architecture is identical to the certified on-premises infrastructure. The virtual servers, also known as logical partitions (LPARs), run on IBM Power hardware with the PowerVM hypervisor.

3.3 IBM Cloud Disaster Recovery Solutions:

IBM Cloud offers built-in capabilities and services for business continuity, resiliency, and security. IBM Cloud Disaster Recovery Solutions are categorized into three major areas:

› Management: Improve the management of infrastructure, apps, processes, and entire cloud environments.

› Migration: Move existing applications and data to the cloud with a portfolio of disasterrecovery (DR) focused migration tools and services.

› Storage: Scale capacity without interruption and deploy globally to achieve higherapplication performance.

3.4 IBM Backup as a Service:

IBM Backup as a Service (BUaaS) from IBM offers fully managed, end-to-end data protectionand data backup in a security-rich environment. Its benefits include:

› Reliable data protection that complies with government and industry regulations.

› Scalability based on your business needs.

› Remote management and operation.

› Monitoring solutions to ensure the health of data protection.

3.5 IBM Resiliency Services:

IBM offers a full range of readily deployable services,solutions, and technologiesfor data protection and recovery:

› Security & Resiliency Consulting Services

› Disaster Recovery as a Service (DRaaS) for hybrid platform recovery

› Data Protection with BUaaS

› Cybersecurity and recovery

› Data center services

3.6 IBM Resiliency Disaster Recovery as a Service:

IBM Resiliency DRaaS offers continuous business resiliency of applications, infrastructure, data, and cloud systems with health monitoring and comprehensive DR services. Its benefits include:

› A less expensive operating expenses (OpEx) based solution compared to a self-managed on-premises model

› Reliable DR orchestration with automation

› Risk-based approach to protect critical IT services

› Data-driven service environment for testing DR, patches, and upgrades. This section provides information about a few migration solutions options.

3.7 IBM Spectrum Protect Plus:

IBM Spectrum Protect Plus is a modern data resilience solution that provides recovery, replication, retention, and reuse for virtual machines(VMs), databases, applications, file systems, software as a service (SaaS) workloads, and containers in hybrid cloud environment.

3.8 Storage:

Actifio GO on IBM Cloud - Actifio GO on IBM Cloud is the next-generation, multi-cloud Copy Data Management SaaS solution that enables customers to back up enterprise workloads directly to IBM Cloud while being ableto instantly access the backup images within their data center.

4. IBM Cloud Backup:

IBM Cloud Backup is a full-featured, automated, and agent-based backup and recovery system that is managed through the IBM Cloud Backup WebCC browser utility. Its benefits include:

› Implement and monitor backup policies from anywhere by using a web-based GUI.

› You can choose an IBM data center or keep the backup outside the network.

› Recover from more than one facility by using multi-vaulting capabilities.

› Scheduled backup with intelligent compression of data.

› End-to-end encryption with Deltapro Deduplication.

› Restoration options from a previous backup or available multiple other recovery points.

IBM Cloud Object Storage:

IBM Cloud Object Storage is a flexible, cost-effective, and scalable cloud storage for unstructured data. Its benefits include:

› Less expensive because you can save costs that are related to server, power, and data center space requirements.

› Streamlined storage environment for increased agility and reduced downtime.

› Supports exponential data growth and built-in high-speed file transfer capabilities.

› Enhanced data security with role-based policies and access permissions.

5. Solutions:

Critical applications are no longer frequently hosted on the same frame (server) or, in many cases, in the same data centre, thanks to the evolution of IT operations over the past ten years. But rather than being fueled by a thorough analysis, this development frequently occurs in more fragmented ways. A thorough analysis looks at the HADR application needs and then compares those requirements to the available solutions throughout the entire infrastructure. For workloads running on systems powered by IBM POWER® processors and meeting the availability needs of crucial enterprise applications, IBM has long been recognised as a leader in HADR solutions. In recent years, the portfolio has grown to now cover data centre application protection for "less critical" workloads. These are the applications that don't have as strict criteria for data loss or can tolerate a somewhat longer outage. However, IBM currently offers a variety of LPAR restart alternatives if you're searching for a less expensive and complex HADR solution (for more details, see "LPAR and virtual machine restart option. Taking into account that mission-critical business workloads have reportedly increased by an average of 15%–36% over the past three years, the ITIC 2020 Reliability poll2 found that 87% of respondents consider 99.99% (52.56 minutes) of unplanned per server/per annum downtime to be the minimum acceptable level of reliability for those servers and applications. Even if it is not taken into account here, the same study, which deals with the projected costs of outages. It is a good idea to assess what is available, what has changed, and how these options match your application availability requirements now that IBM has a more extensive portfolio of

HADR solutions. Although HADR solutions' main goal is to avoid infrastructure failures, these tools can also be used to manage maintenance and upgrade processes. For instance, PowerHA SystemMirror on AIX comes with a solution to handle Service Packs and interim patches throughout the cluster. PowerHA SystemMirror development has recently put a strong emphasis on usability, thereby dispelling the long-held (and frequently false) misconception that the software is challenging to use. Organisations typically have a variety of apps with linked (but distinct) SLAs. In order to meet your various SLAs and the various OSs that can be running in your IBM Power system.

6. Development Phase – I

A Disaster Recovery Plan (DRP) is a documented strategy outlining the procedures an organization follows to recover and protect its IT infrastructure in the event of a disaster. This plan is essential for minimizing downtime, data loss, and financial impact during unforeseen events.

Disaster Recovery Strategy:

1. RTO (Recovery Time Objective): Define the acceptable downtime for the services. For example, if RTO is set to 4 hours, it means that services should be restored within 4 hours after a disaster occurs. **2. RPO (Recovery Point Objective):** Determine the maximum amount of data loss that is acceptable. For instance, if RPO is set to 1 hour, it means that no more than 1 hour's worth of data can be lost in the event of a disaster.

3. Priority of Virtual Machines: Categorize virtual machines based on their criticality to business operations. For example: Priority 1 (High): These are mission-critical systems with little tolerance for downtime. Priority 2 (Medium): These are important systems but can tolerate a bit of downtime. Priority 3 (Low): These systems are less critical and can handle extended downtime. RTO (Recovery Time Objective) and RPO (Recovery Point Objective) are critical parameters in any disaster recovery strategy. They define the acceptable levels of downtime and data loss in the event of a disaster, helping organizations prioritize their recovery efforts.

Recovery Time Objective (RTO):

RTO refers to the maximum amount of time that a system or service can be down before it starts causing significant harm to the business. It represents the acceptable downtime for a service or application. Example: If a business has an RTO of 4 hours, it means that in the event of a disaster, the affected service or application must be restored and operational within 4 hours. Factors influencing RTO:

- Criticality of the system or service.
- Complexity of recovery processes.
- Availability of resources and technology.
- Business impact of prolonged downtime.

Recovery Point Objective (RPO):

RPO is the maximum tolerable amount of data loss measured in time. It signifies the point in time to which data must be recovered after a disaster. It essentially defines how much data the organization can afford to lose. Example: If an organization has an RPO of 1 hour, it means that in case of a disaster, they can't afford to lose more than 1 hour's worth of data.

Factors influencing RPO:

- Frequency of data backups or snapshots.
- Rate of data change within applications.
- Importance of real-time data updates.
- Relationship between RTO and RPO:

Alignment:

RTO and RPO should be closely aligned. In many cases, RPO will influence the RTO. If a very low RPO is required (meaning minimal data loss is acceptable), it may dictate a shorter RTO because more frequent backups or replication will be necessary. Balancing Act: Achieving

a very low RTO and RPO can be costly. There's a trade-off between the cost of implementing high-availability solutions and the potential loss incurred from extended downtime or data loss.

Application Dependency:

Different applications may have different RTO and RPO requirements. Mission-critical applications may require near-zero RTO and very low RPO, while less critical services may have more relaxed objectives.

Technology and Infrastructure:

The choice of technology and infrastructure (like cloud services, backup solutions, etc.) greatly influences the ability to meet RTO and RPO targets.

Testing and Validation:

Regular testing and validation are essential to ensure that the set RTO and RPO objectives can be met in practice.

Setting Up Regular Backups:

IBM Cloud Backup Service:

Utilize IBM Cloud's native backup service to automate backups of virtual machines. This service provides options for scheduling, retention policies, and easy recovery.

Custom Scripts for On-Premises Backups:

Develop custom scripts to perform regular backups of on-premises virtual machines. These scripts can leverage tools like Veeam, Acronis, or native OS backup utilities.

Backup Frequency:

Define the frequency of backups based on the RPO. For example, if RPO is 1 hour, backups should be taken at least every hour.

Retention Policies:

Determine how long backups will be retained. This should align with compliance requirements and business needs. For example, keep daily backups for 7 days, weekly backups for 4 weeks, and monthly backups for 6 months.

Offsite Storage:

Ensure that backups are stored in an offsite location, either in a different physical location or in a cloud storage solution. This safeguards against disasters that might affect the primary site.

Regular Testing and Validation:

Periodically perform recovery tests to ensure that backups can be successfully restored. This helps identify and address any issues in the backup process.

Additional Considerations:

Failover and Failback Procedures: Develop detailed procedures for failing over services to the IBM Cloud Virtual Servers in the event of a disaster. Also, outline steps for failing back to on-premises infrastructure once the disaster is resolved.

Monitoring and Alerting:

Implement robust monitoring solutions to keep track of the health and availability of both on-premises and cloud-based resources. Set up alerts for any anomalies or outages. Documentation and Communication: Document the entire disaster recovery plan, including procedures, contacts, and configurations. Ensure that all stakeholders are aware of the plan and their roles in the event of a disaster. Compliance and Regulatory Considerations: Ensure that the disaster recovery plan aligns with any industry-specific compliance requirements (e.g., HIPAA, GDPR) that may apply to your organization.

7. Development Phase – II

Database Replication:

a. Primary Site: Set up MySQL server as the primary database. Enable binary logging and configure server ID. Create a replication user with appropriate privileges.

b. Secondary Site: Set up MySQL server as the secondary database. Configure it as a replica of the primary server. Start the replication process.

c. Monitoring and Maintenance: Implement monitoring for replication status. Regularly monitor the replication lag.

Sample MySQL Replication Configuration:

Primary Server (my.cnf):

pythonCopy code:

[mysqld] server-id=1

log-bin=mysql-bin binlog-do-db=my_database

Secondary Server (my.cnf):

[mysqld] server-id=2 relay-log=mysql-relay-bin replicate-do-db=my_database

Create Replication User: CREATE USER 'replication_user'@'%' IDENTIFIED BY 'password'; GRANT REPLICATION SLAVE ON *.* TO 'replication_user'@'%'; FLUSH PRIVILEGES;

Testing Recovery Procedures:

a. Planned Failover: Simulate a controlled failover from primary to secondary. Stop the primary database. Promote the secondary to the primary role. Update application configurations to point to the new primary.

b. Unplanned Failover: Simulate an unexpected failure in the primary site. Manually trigger the failover process. Monitor the failover process to ensure it completes successfully.

c. Data Validation: Verify that data on the secondary site matches the primary. Use tools like checksums or manual spot-checking.

d. Failback (if applicable): After primary site is restored, set up replication back to the original primary.

Conduct recovery tests to ensure that the disaster recovery plan works as intended.

Simulate a disaster scenario and practice recovery procedures.

1. Preparation: Before conducting the recovery test, make sure to: Notify Stakeholders: Inform relevant stakeholders about the test, including the date, time, and expected duration. Document the Test Plan: Clearly outline the objectives, steps, and expected outcomes of the recovery test. Backup Data: Ensure that you have recent backups of critical data and configurations. Isolate the Test Environment: If possible, conduct the test in an isolated environment to avoid impacting production systems.

2. Simulate the Disaster Scenario: Choose a disaster scenario to simulate. For example, you could simulate: Hardware Failure: Simulate a critical server failure. Data Corruption: Introduce data corruption or deletion. Network Outage: Simulate a network failure.

3. Initiate the Recovery Process: Follow the steps outlined in your disaster recovery plan to initiate the recovery process based on the simulated disaster scenario. This may include: Failover Procedures: If you're switching from a primary to a secondary site, follow the failover procedures. Data Restoration: Restore data from backups if necessary. Configuration Updates: Update configurations as per the recovery plan.

4. Monitor and Validate: During the recovery process, monitor the progress and validate that each step is executed correctly. Ensure that: Data is restored accurately. Services and applications are brought back online. Network connectivity is re-established.

5. Functional Testing: Perform functional tests to verify that all critical systems and applications are functioning as expected. This may involve: Accessing key applications. Performing transactions or operations.

6. Performance Testing: If applicable, conduct performance testing to ensure that the recovered systems meet performance requirements.

7. Data Verification: Verify the integrity and consistency of the recovered data. This can include comparing checksums or performing spot checks on critical data.

8. Document the Test Results: Document the results of the recovery test, including any issues encountered, the time taken for recovery, and any deviations from the expected outcomes.

9. Post-Test Review: After the test, gather feedback from the team members involved. Discuss any challenges faced, areas for improvement, and lessons learned.

10. Update the Disaster Recovery Plan: Based on the findings from the test, update the disaster recovery plan as necessary. This could involve refining procedures, addressing gaps, or improving documentation.

11. Communicate the Results: Share the results of the recovery test with relevant stakeholders. Highlight any improvements made to the disaster recovery plan.

12. Regularly Repeat Tests: Schedule regular recovery tests to ensure that your disaster recovery plan remains effective and up-to-date