Unit 4 Seminar – Title: DR Solutions Design and Review

Activity 1: DR Terms and Concepts

1. **What is the difference between Hot Standby, Warm Standby and Cold Standby? Frame your answers in terms of availability, RPO and RTO.**

Based on research by Alhazmi and Malaiya (2013), all organisations must have a cost-effective business continuity or disaster recovery plan (DRP). This plan should identify potential disasters, evaluate feasible recovery plans, and consider critical parameters such as initial cost, data transfer cost, and data storage cost. To facilitate the development and upkeep of the plan, it is recommended that a quantitative approach be used to evaluate costs.

Furthermore, in the event of a primary site malfunction, the transfer of incoming transactions to the backup site is known as a failover. This is termed a failback once the prior site's concerns have been addressed and reinstated.

According to Alhazmi and Malaiya (2013), there are three backup sites: cold, hot, and warm. Cold standby takes multiple days to recover, hot standby takes seconds or minutes, and warm standby combines both.

Moreover, RPO and RTO are important to Business Continuity Planning (BCP) and Disaster Recovery (DR) strategies. RPO determines data loss tolerance and how often data needs to be backed up. In contrast, RTO determines how quickly an organisation can recover after data loss due to a failure or disaster (Kerner, N.D.).

The differences between Hot Standby, Warm Standby, and Cold Standby in the context of availability, Recovery Point Objective (RPO), and Recovery Time Objective (RTO) are as follows:

Hot Standby is a redundancy method used to minimise system downtime. It provides a fail-safe against unexpected system outages, enhancing system reliability and operational continuity. In a Hot Standby scenario, there is a primary system and a secondary identical system that remains idle but ready to take over at any given moment. The secondary system is synchronised with the primary system in real-time, replicating every data modification. If the primary system fails, crashes, or needs maintenance, the hot standby server immediately activates to ensure uninterrupted service (DevX, 2023).

- **Availability:** In the event of a breakdown, a hot backup system is always up and prepared to take over. It entails keeping an identical set of resources (such servers, databases, or services) up to date and synchronised with the master system while in use.
- **RPO:** A hot standby's RPO is very close to zero. Any data loss is negligible since data is constantly copied.
- **RTO:** Usually, the RTO is less than a minute. A quick switchover to the standby system guarantees less interruption.

According to Damue (2023), the Warm Standby approach involves maintaining a partially provisioned environment ready to take over in a disaster, providing a faster recovery than starting from scratch. When a disaster occurs, the Warm Standby environment can be quickly scaled up by launching additional instances and activating the necessary services. A real-world scenario where the Warm Standby strategy is commonly used is in financial institutions requiring continuous critical system availability.

- **Availability:** While certain parts of a warm standby system are functioning, they are not actively handling user requests. To switch over, human assistance is needed.

- **RPO:** A warm standby's RPO is often expressed in minutes. Periodically, data synchronisation takes place (hourly backups, for example).

- **RTO:** The RTO is determined by how long it takes to turn on the standby system. It may take a few minutes or several hours.

Cold standby is a backup method in which a redundant system is kept offline until needed. It's cost-effective but requires more restore time than other standby methods (DevX, 2023).

- **Availability**: When a cold standby system is offline, it is inactive. Resources are available but not currently in use. Setup and personal intervention are needed for activation.

- **RPO**: The cold standby's RPO is ascertained based on the final backup before the failure. It can take a few hours or perhaps a few days.

- **RTO**: Because it requires starting up the standby system, recovering data, and configuring it, the RTO is quite lengthy. It might be hours or days.

2. **Does the technology deployed affect the options available? For example, can you create a high availability, hot standby solution between two on-premise data centres?**

Synchronous replication is necessary to ensure an organisational database cluster is resilient and highly available. For optimal fault tolerance, replicas should be distributed across various fault domains. In case of any failure, the functional replicas within the organisation should provide consistent readings and writings, as stated by Pachot (2023).

Organisations must develop disaster recovery plans that address all potential challenges, including cyberattacks, equipment malfunctions, natural catastrophes, and human error. They should explore different options, including managed cloud-based, on-premises, or a combination. IT managers should weigh the benefits and drawbacks of each option to choose the one that best suits their company's requirements, as pointed out by Lovett (2023).

Lovett (2023) lists several benefits of cloud disaster recovery, including time-saving, scheduled testing, flexibility, quick failover, easy deployment, and reduced upfront expenses. However, on-premises disaster recovery provides reduced latency, bespoke application compatibility, better control, and custom infrastructure. On-premises disaster recovery requires high upfront expenditures, real estate, complicated setup, challenging scaling, fluctuating uptime and recovery times, and a deficiency of round-the-clock personnel.

Srivastava (2023) emphasises that when considering high-availability solutions between on-premise data centres or in the cloud, it's essential to carefully evaluate the technology being deployed. Choosing wisely based on specific requirements is critical to ensure that it can achieve the level of availability and reliability required by the business or organisation. Factors that should be considered include the scalability, security, and compatibility of the technology used and the level of support and maintenance required over time.

Understanding the technological differences between Disaster Recovery (DR) and High Availability (HA) is vital. Shuster & Krenn (2024) state that although DR is intended to restore service in the case of a catastrophic physical failure of the data centre or a substantial application-level failure, HA is intended to offer fault tolerance. While HA is usually included in design standards, DR incorporates automated or manual recovery techniques. Even though HA can encompass DR, it is advised to set up distinct Sites, Farms, and Server Groups for every data centre and to employ particular service accounts and authentication services across production and DR environments to reduce the possibility of single points of failure.

Furthermore, Shuster & Krenn (2024) clarify that to guarantee the effectiveness of a disaster recovery plan, the recovery time objective (RTO) and scope of recovery should be considered. To ascertain the architectural modifications necessary for DR, it is imperative to consider the use cases, capabilities, current disaster recovery, data criticality, and potential catastrophe types.

When implementing a DR strategy, the limitations on client network security and bandwidth use for Vital Data Asset (VDA) traffic must be considered. Planning for failback to propagate modifications made throughout the DR phase is also critical. Several expected recovery solutions are available, varying in cost and complexity from more straightforward and less expensive to more sophisticated and costly. "Always on" recovery methods are ideally suited to integrated technologies like NetScaler and single-image management solutions (Shuster & Krenn, 2024).

Activity 2: DR Solutions Design

1. **What are some of the leading vendor lock-in issues the authors identify? How would you mitigate them?**

Preventing customers from moving to another provider is known as vendor lock-in. Several things, including trade secrets, a lack of rivalry in the market, and legal limitations, may cause this. Frequently, vendors facilitate consumers' enrolment in their services while making it hard for them to move to a rival provider. Technological, monetary, and legal limitations that deter users from quitting are used to achieve this. Sometimes, these limitations are only made clear when a user's technological needs change or are buried in the vendor's Terms of Service fine print (Raza, 2020).

Cloud computing is a cost-effective and scalable solution for businesses to offer IT services. However, a common issue with cloud computing is vendor lock-in, which restricts its widespread usage. Although current cloud solutions are designed to be

vendor-locked, their compatibility with other cloud systems is limited (Opara-Martins et al., 2014).

To reduce the risk of cybersecurity threats, companies should follow four critical standards while moving their applications and systems to the cloud, per a report by Morrow et al. (2019). While these practices are mainly aimed at small and medium-sized businesses, any organisation can use them to enhance the security of their cloud usage. The four standards are due diligence, access control, data protection, and cybersecurity threat monitoring and defence. The report includes examples of actual cybersecurity issues from the real world, demonstrating the effectiveness of these practices. The report also highlights the application of these techniques to other cloud service models, including software as a service (SaaS), platform as a service (PaaS), and infrastructure as a service (IaaS).

Therefore, Opara-Martins et al. (2016) highlight that cloud computing is an appealing option for businesses, but significant challenges prevent its adoption. The primary concerns are security and the risk of being locked into one vendor. Cloud providers use different hardware and technologies, proprietary standards, and a lack of standard interfaces and open APIs. This makes it difficult for businesses to integrate, move their data, and ensure compliance and security.

Establishing standards for defining cloud applications and management parameters is essential to address these challenges. These standards can ensure interoperability, portability, compliance, trust, and security. Without standards, businesses may depend on a single cloud provider, stifling market competition and harming consumers. Proprietary standards can create anti-competitive environments, limiting cloud interoperability (Opara-Martins et al. 2016).

Hence, to migrate per Raza's (2020) highlights, keep the following in mind to prevent vendor lock-in:

1. Determine intricate dependencies.
2. Look for similarities between the IT stack, cloud providers, and technical specifications.
3. Check to see if the apps work with various hardware.
4. Teach stakeholders to be more aware of the dangers of vendor lock-in.
5. Ensure that apps are open standards compliant and portable.
6. Make use of contemporary SDLC techniques like DevOps.
7. Once migrated, make sure it is portable.
8. Create a precise exit plan with workable conditions for both parties to agree to end the arrangement.
9. Take into account a multi-cloud approach.
10. Exercise caution.

2. What are some of the security concerns with the modern cloud? How can these be mitigated?

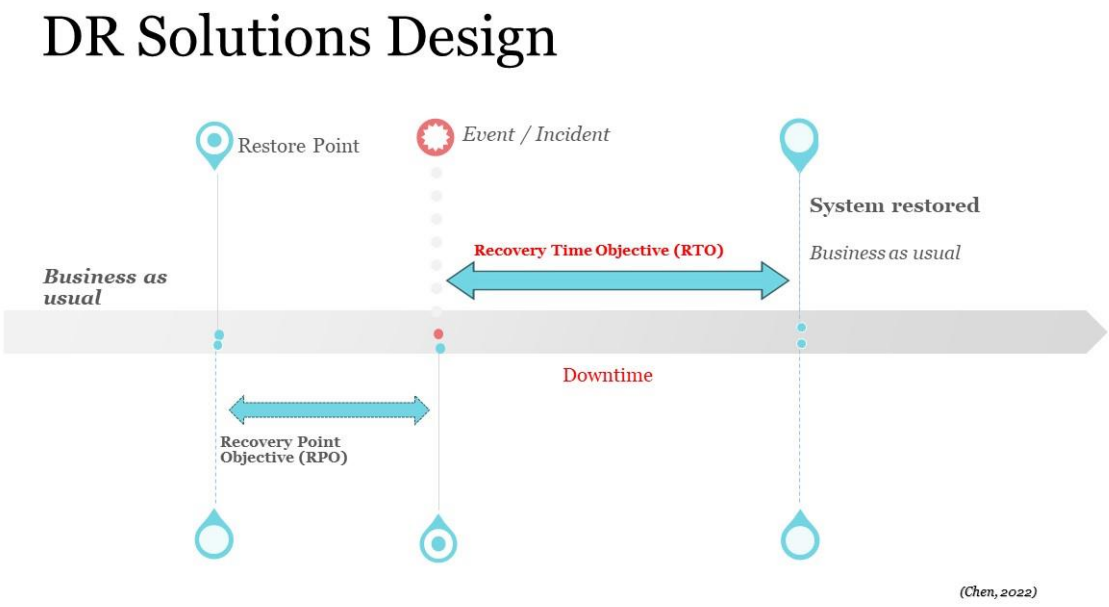
Protecting sensitive data in cloud environments is a top priority for organisations due to the high risk of threats such as misconfiguration, unauthorised access, insecure APIs, and account hijacking. Attackers can gain direct access to cloud resources if there is inadequate cloud security posture management. Therefore, it is crucial to have practical cloud-focused security tools and solutions to mitigate these risks. However, regulatory compliance can be challenging when moving data to the cloud, and traditional security tools may not be sufficient to secure cloud-based infrastructure. Concerns also include data loss or leakage, data privacy and confidentiality, accidental exposure of credentials, and incident response (CheckPoint, 2021).

Cloud infrastructure has become a critical component of today's corporate environment, and protecting it from cyber-attacks is imperative. Abbas (2024) emphasises that organisations must implement a robust and thorough cybersecurity strategy to protect cloud-based assets and operations. Preventing threats, identifying them, mitigating damage, encrypting data, controlling access, disaster recovery

planning, regular security audits, fostering security awareness and training, managing vendors, implementing a zero-trust security model, leveraging threat intelligence tools, automating processes, and engaging in DevSecOps are the best ways to strengthen cybersecurity posture. These steps are crucial for locating security holes in the system, guaranteeing compliance, and protecting against intrusions and breaches.

Activity 3: DR Solutions Design and Review

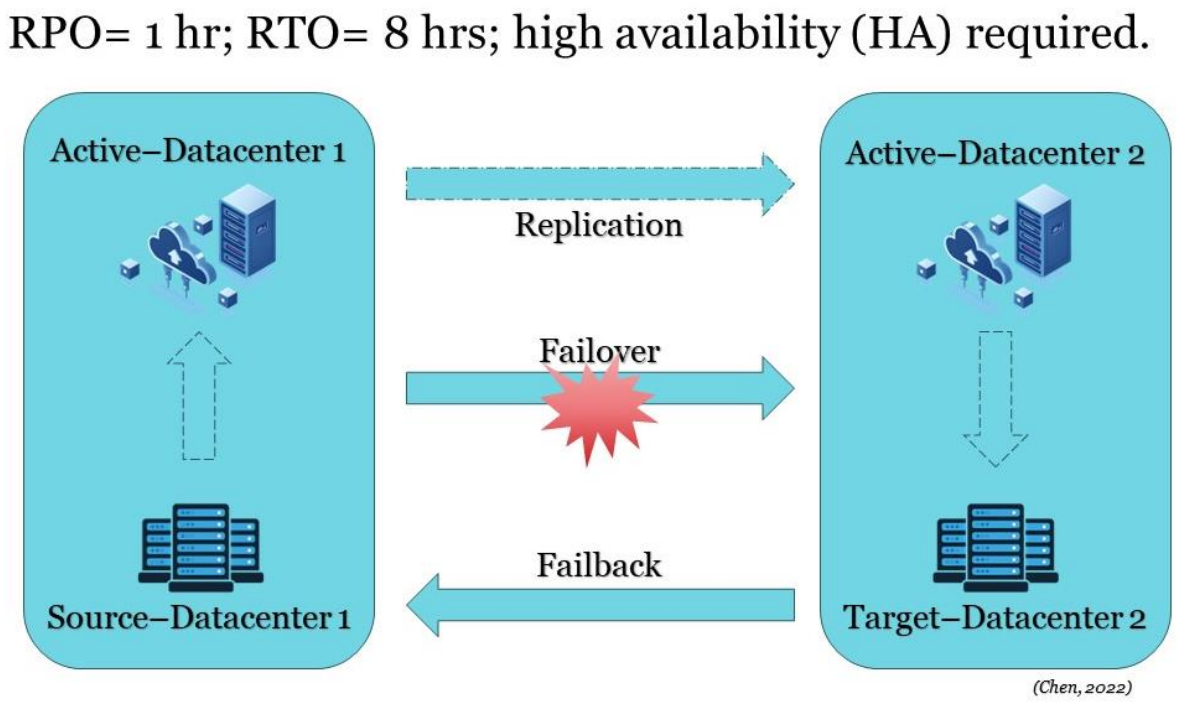
Any organisation that wants to select the best cloud backup and disaster recovery strategy must understand Recovery Point Objective (RPO) and Recovery Time Objective (RTO). The RTO is the most extended period a business process may recover following a disaster. Simultaneously, the RPO is the most extended period that might elapse during an interruption before the amount of data lost exceeds the upper limit permitted. Maximum allowable data loss, data storage choices, and the expense of putting disaster recovery plans into place are some variables that affect RPO. Disaster and business disruption drills are essential for highlighting the discrepancy between goals and reality (Singh, 2021).



1. **RPO= 1 hr; RTO= 8 hrs; high availability (HA) required.**

This solution ensures quick recovery within 8 hours and minimal data loss of 1 hour. High availability means a system is always accessible during maintenance or component failures (PureStorage, N.D). The diagram solution consists of the following elements based on the ideas of Chen (2022):

- Active-Active Data Centres: A pair of active data centres located separately. Real-time synchronous data replication occurs between the primary and secondary data centres.
- Rear-time replication: This element distributes data across operational data centres.
- Data Centre with Automatic Failover Mechanism: In the event of a failure detection, this element automatically switches to the backup data centre.



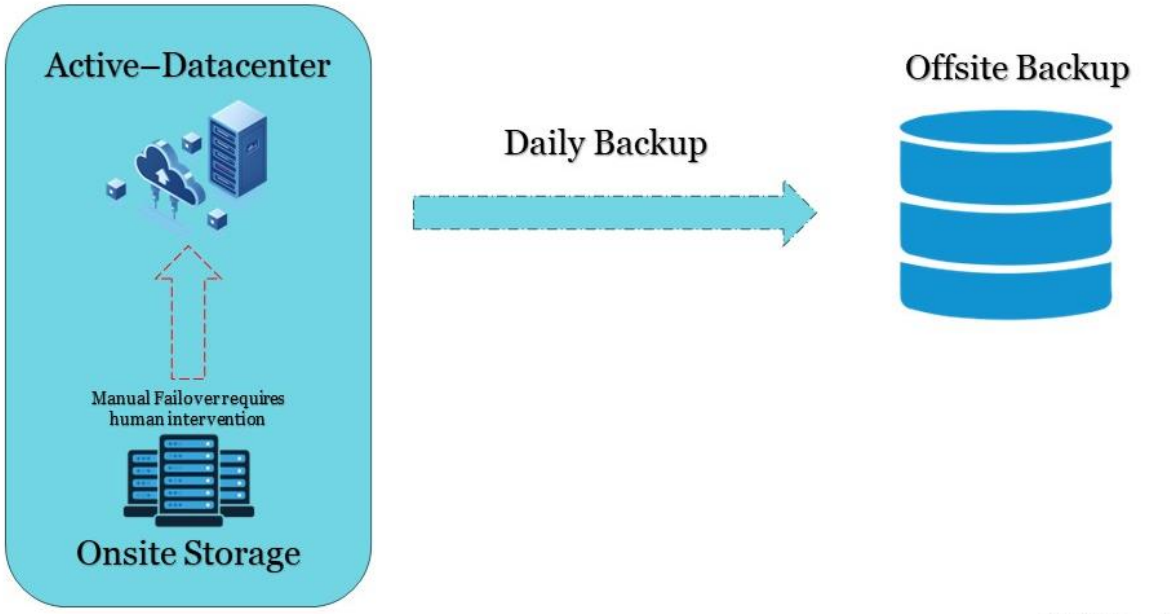
2. **RPO= 24 hrs; RTO = 72 hrs; HA NOT required.**

This solution prioritises data preservation with a Recovery Point Objective (RPO) of 24 hours and a Recovery Time Objective (RTO) of 72 hours. High availability is not the primary concern. Yasar et al. (N.D) note that this DR solution has cold sites with basic infrastructure but no pre-installed systems. Although businesses have data backups,

recovery requires manual work and hardware configuration, which increases recovery times. The diagram solution is based on ideas of the NAKIVO (2024) and includes:

- Offsite Backup is a separate location for backing up data.
- Scheduled Data Replication involves asynchronous replication (e.g., daily backups).
- Manual Failover requires human intervention to switch to the backup data centre.

RPO= 24 hrs; RTO = 72 hrs; HA NOT required.

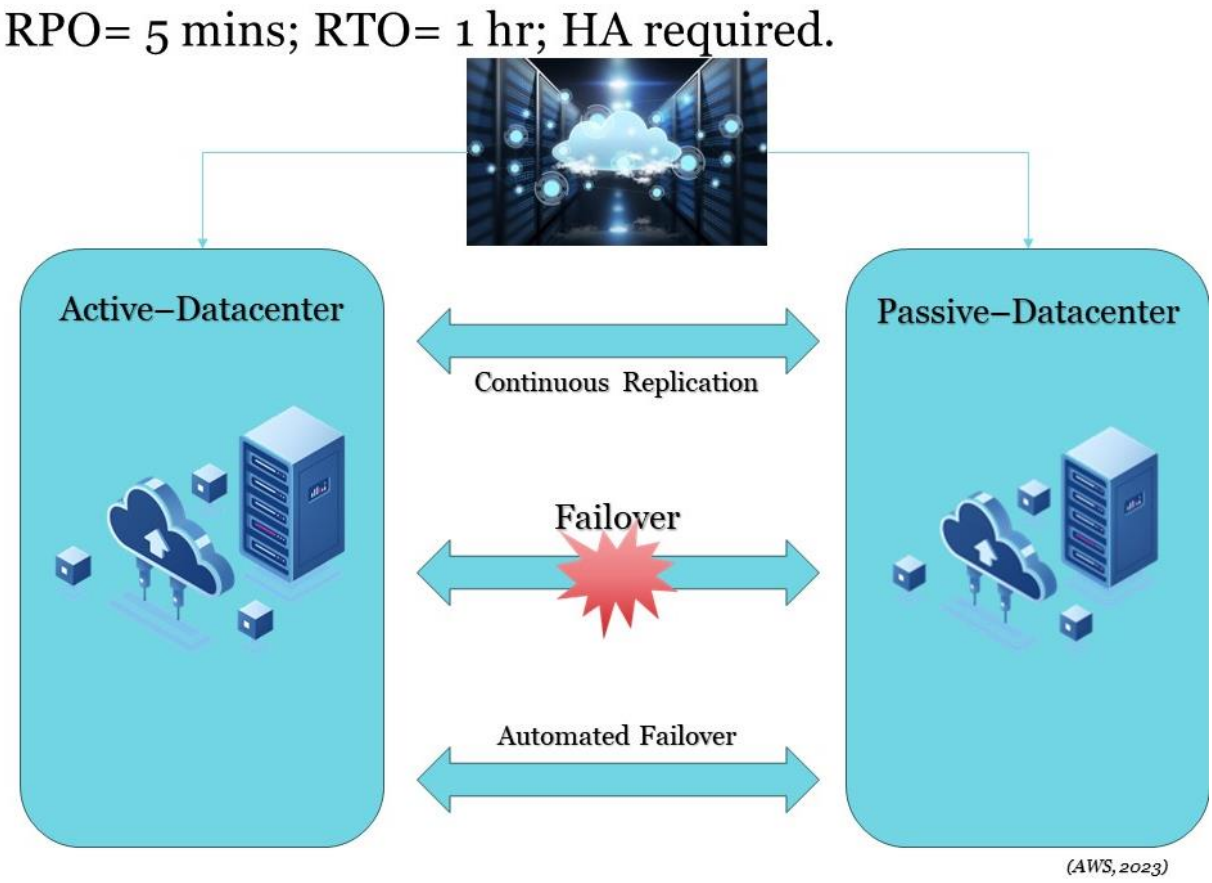


(NAKIVO 2024)

3. **RPO= 5 mins; RTO= 1 hr; HA required.**

This DR solution offers High Availability with a 5-minute RPO and 1-hour RTO. This DR solution uses multi-region active/passive architecture, with a dynamic site for hosting and traffic and an inactive site for recovery. The data is continuously replicated and backed up for protection against disasters (AWS, 2023). It is designed to provide real-time data protection while ensuring quick recovery in case of a system failure. The solution is comprised of the following components:

- Active-Passive Data Centres: Two are primary (active) and one secondary (passive). The primary data centre is where the system is currently running, while the secondary data centre is a standby location that can be activated in case of a failure.
- Continuous Data Replication: The system replicates data to the secondary data centre in real-time. This ensures that the secondary data centre always has an up-to-date copy of the data from the primary data centre.
- Automated Failover: In case of a failure at the primary data centre, the system will automatically switch to the secondary data centre. This ensures that there is no interruption to the service and that the system can continue to function seamlessly.



References:

- Alhazmi, O.H. & Malaiya, Y.K. (2013). *Evaluating disaster recovery plans using the cloud*. [online] IEEE Xplore. doi:https://doi.org/10.1109/RAMS.2013.6517700.
- Kerner, S. M. (N.D.). *What Is A Recovery Point Objective (RPO) And How Does It Work?* [online] Available at: https://www.techtarget.com/whatis/definition/recovery-point-objective-RPO.
- DevX. (2023). *Hot Standby*. [online] Available at: https://www.devx.com/terms/hot-standby/ [Accessed 1 Mar. 2024].
- Damue, B. (2023). *Ensuring Business Continuity: A Guide to Choosing the Right Disaster Recovery Strategy on AWS*. [online] Medium. Available at: https://medium.com/@dbrandonbawe/ensuring-business-continuity-a-guide-to-choosing-the-right-disaster-recovery-strategy-on-aws-728de6ab853a [Accessed 26 Feb. 2024].
- DevX. (n.d.). *Cold Standby: Definition, Examples*. [online] Available at: https://www.devx.com/terms/cold-standby/ [Accessed 1 Mar. 2024].
- Pachot, F. (2023). *How to Achieve High Availability and Disaster Recovery with Two Data Centers*. [online] Yugabyte. Available at: https://www.yugabyte.com/blog/high-availability-disaster-recovery-two-data-centers/.
- Lovett, C. (2023). *Disaster Recovery Cloud vs On-Premises | TierPoint*. [online] TierPoint, LLC. Available at: https://www.tierpoint.com/blog/disaster-recovery-cloud-vs-on-premise/.
- Srivastava, S. (2023). *On-Premise vs. Cloud: A Detailed Analysis*. [online] Appinventiv. Available at: https://appinventiv.com/blog/on-premise-vs-cloud/.
- Shuster, M. & Krenn, G. (2024). *Design Decision: Disaster Recovery Planning*. [online] Available at: https://docs.citrix.com/en-us/tech-zone/design/design-decisions/cvad-disaster-recovery.html.
- Raza, M. (2020). *10 Best Practices to Avoid Cloud Vendor Lock-In*. [online] BMC Blogs. Available at: https://www.bmc.com/blogs/vendor-lock-in/.
- Opara-Martins, J., Sahandi, R. & Tian, F. (2014). Critical review of vendor lock-in and its impact on adoption of cloud computing. *International Conference on Information Society (i-Society 2014)*. doi:https://doi.org/10.1109/i-society.2014.7009018.
- Morrow, T., Lapiana, V., Faatz, D., Hueca, A. & Richmond, N. (2019). *Cloud Security Best Practices Derived from Mission Thread Analysis*. [online] doi:https://doi.org/10.1184/R1/12363563.V1.
- Opara-Martins, J., Sahandi, R. & Tian, F. (2016). Critical analysis of vendor lock-in and its impact on cloud computing migration: a business perspective. *Journal of Cloud Computing*, [online] 5(1). doi:https://doi.org/10.1186/s13677-016-0054-z.
- CheckPoint (2021). *Top Cloud Security Issues, Threats and Concerns*. [online] Check Point Software. Available at: https://www.checkpoint.com/cyber-hub/cloud-security/what-is-cloud-security/top-cloud-security-issues-threats-and-concerns/.

- Abbas, M. (2024). *Top 11 Strategies to Mitigate Cloud Security Threats*. [online] Available at: https://www.educative.io/blog/mitigate-cloud-security-threats.
- Singh, J. (2021). *Understanding the Difference Between RPO and RTO | Druva*. [online] www.druva.com. Available at: https://www.druva.com/blog/understanding-rpo-and-rto.
- Chen, S. (2022). *Achieve Near-Zero RPO & RTO with Orchestrated Application Recovery*. [online] Available at: https://www.rubrik.com/blog/architecture/22/5/achieve-near-zero-rpo-and-rto-with-orchestrated-application-recovery [Accessed 3 Mar. 2024].
- PureStorage (N.D.). *What Is Oracle High Availability? | Pure Storage*. [online] Available at: https://www.purestorage.com/au/knowledge/what-is-oracle-high-availability.html [Accessed 3 Mar. 2024].
- Yasar, K., Sullivan, E., Crocetti, P. (N.D.). *What is Disaster Recovery (DR)?* [online] Available at: https://www.techtarget.com/searchdisasterrecovery/definition/disaster-recovery.
- NAKIVO (2024). *Storage Tiering Guide for Data Archival*. [online] Available at: https://www.nakivo.com/blog/storage-tiering/ [Accessed 3 Mar. 2024].
- AWS (2023). *Disaster recovery options in the cloud - Disaster Recovery of Workloads on AWS: Recovery in the Cloud*. [online] docs.aws.amazon.com. Available at: https://docs.aws.amazon.com/whitepapers/latest/disaster-recovery-workloads-on-aws/disaster-recovery-options-in-the-cloud.html.