**Week 4: Workshop Session**

**Title: Data Availability**

**Important Note!**

The mid-module evaluations are live **for all students**. The mid-module evaluations are an online survey that students should be encouraged to complete within workshop sessions.

The link to the survey is here <https://cis.cardiffmet.ac.uk/ModuleEvaluation/>

There is a guidance video for students to support them in completing their module evaluations and why we do them. It would be helpful if you could play this in your workshop sessions and provide the link here:

<https://cardiffmet.cloud.panopto.eu/Panopto/Pages/Viewer.aspx?id=7b774e06-41ab-4458-86fa-b09a00d6787d>

**Expected Learning Outcomes:**

Understand the concept of data availability.

Understand how data availability is implemented using different methods such as:

* Redundancy techniques
* Threshold cryptography, example Shamir Secret Sharing
* Cache technologies

Know how to use Shamir Secret Sharing to demonstrate data availability.

**Practice exercise One: Answer all questions (10 marks each, 60% in all).**

Q1: Explain in your own terms the concept of data availability.

As-per NIST (National Institute of Standards and Technology) Availability ensures that users have reliable and proper access to data and services without an extended delay (Keller, 2019).

For example, a popular online e-commerce site would have to ensure high availability, to make sure that their customers can use the website to browse and purchase products. If they had a loss in availability (e.g., the site went down), they would likely have a loss in revenue as well.

The website could improve availability in a number of ways, the e-commerce website could move their infrastructure to a cloud provider with a higher availability (e.g. 99 – 99.99%).

They could also introduce load-balancing, which exposes a single server to users on the frontend, but on the backend it proxies the requests to multiple servers, to ensure that there isn’t a single point of failure. If one server goes down, then the other servers will still be available.

Q2: What are in your own words is redundancy technique?

Redundancy is essentially the practice of duplicating data in case of corruption or manipulation, to avoid a single point of failure (Li, Reiher and Popek, 1999).

There are a number of redundancy techniques, for example a RAID (Redundant Array of Independent Disks) is used within centralised storage devices, or a SAN (Storage Area Network). Depending on the RAID level used, data may be fully duplicated across multiple disks, to ensure that if one of the disks fail, the data is not lost.

Q3. What are in your own words is threshold cryptography?

As per (Brandao, Davidson and Vassilev, 2020) threshold cryptography is a method of sharing a secret key across multiple “parties”.

This is done by splitting the secret into an arbitrary number of subcomponents via an algorithm, which defines the number of subcomponents the secret should be split into (referred to as **n**) and the “threshold”, which defines number of subcomponents required to reform the original secret (referred to as **t**). If the number of parties is less than the threshold (**t**)thenthe original secret cannot be retrieved.

According to (NIST, 2018) threshold cryptography can be used on **any** cryptographic primitive (which includes signing, encryption and decryption).

The reason it is used, is to enhance availability and distribute trust across multiple parties, without compromising the confidentiality of the secret.

An example of this method being used for improving security is, if a secret (for example a private key) is stored on a single server, and that server is compromised and the key stolen, the entire cryptosystem would be at risk.

Threshold cryptography would allow this secret to be distributed across multiple servers, meaning an attacker would have to compromise each server (until the threshold is reached) to regain the original secret.

This method also improves availability, consider an example where a secret key is split into 5 subcomponents (**n** = 5) with a threshold of 3, (**t** = 3). The subcomponents are split across 5 separate servers, even if the data across 2 of the servers is completely lost, the secret key can still be derived, ensuring that the secret is still available.

Q4: What are in your own words is cache technique?

Caching is a technique to improve availability, by storing commonly used data (typically static data) in an area that provides faster access when compared to the original method of retrieving the data.

For example, when clients retrieve data from webservers like static content, images, JavaScript and CSS this takes time and resources. Caching allows the client to save a temporary copy of the data to retrieve it when it’s required, rather than making an HTTP request, as defined by RFC9111 (Fielding, Nottingham and Reschke, 2022).

Servers or Content Distribution Networks may also cache frequently accessed data, by saving a copy of it to memory which gives faster access when compared to reading from disk,

Caching is not only seen in webservers, databases also cache data that’s frequently accessed by storing that data in-memory improving the speed at which it’s accessed when compared to reading it from disk.

This reduces the amount of time required for the server to respond to clients, and the amount of requests that the server receives, which overall reduces the pressure on the servers and the overall network.

Q5. Compare and contrast three methods used in implementing data availability.

Q6. Find out other methods/tools that can be used in implementing data availability.

**Note: For these practice exercises, provide references and citations of any secondary material used in Harvard standard.**

**Practice exercise Two: Immersive labs (20 marks in all)**

1. Backups
2. Cyber fundamentals
3. Technical Fundamentals
4. Secure Fundamentals: Least Privileges
5. Secure Fundamentals: Defense in Depth
6. Secure Fundamentals: Authorization
7. Secure Fundamentals: Authentication

**Workshop Reflective Writing (20 marks in all)**

Reflect on your learning this week and put them down across these subheadings:

1. Depth and breadth of understanding. This is all about putting down some details of what you learnt and did not understand well in this week’s session. What will you do going forward on areas you did not comprehend during the session? **(10%)**
2. Critical information, contextualising the session’s learning components to assignment development. **(5%)**
3. Personal thoughts, subjective analysis of personal development during this week’s session. **(5%)**

**Submit your practices exercises on or before Monday 23/10/23 in Moodle to receive your FORMATIVE feedback.**

**Citations**

Brandao, L.T.A.N., Davidson, M. and Vassilev, A. (2020). NIST roadmap toward criteria for threshold schemes for cryptographic primitives. [online] doi:https://doi.org/10.6028/nist.ir.8214a.

Fielding, R., Nottingham, M. and Reschke, J. (2022). *RFC9111*. [online] IETF HTTP Working Group Specifications. Available at: <https://httpwg.org/specs/rfc9111.html>.

Li, J., Reiher, P. and Popek, G.J. (1999). Securing information transmission by redundancy. doi:https://doi.org/10.1145/335169.335205.

NIST. (2019). *Glossary*. [online] NIST. Available at: <https://www.nist.gov/itl/smallbusinesscyber/cybersecurity-basics/glossary>.

NIST. (2018). Multi-Party Threshold Cryptography | CSRC | CSRC. [online] CSRC | NIST. Available at: https://csrc.nist.gov/projects/threshold-cryptography [Accessed 20 Oct. 2023].

Shamir, A. (1979). How to share a secret. *Communications of the ACM*, 22(11), pp.612–613. doi:https://doi.org/10.1145/359168.359176.

‌

‌