SECURE AI INFERENCING

IMDA PET SANDBOX – SINGAPORE GENERAL HOSPITAL CASE STUDY

INFOCOMM MEDIA DEVELOPMENT AUTHORITY

Contents

Use Case Background

Current process

1. Medical image analytics typically use computationally intensive neural network models and operate on large image sizes that can range from 1 to 5 Gigabytes and may exceed billions of pixels.
2. Therefore, leveraging cloud resources for data storage and computation can bring benefits such as **flexible and scalable performance and costs**.
3. However, medical images contain sensitive information about a patient and if cloud-based solutions are used, sharing medical data to and for use in the cloud could pose potential privacy and security risks.

Opportunity

4. The use of Privacy-Enhancing Technologies (PETs) such as TEEs could enable secure processing of sensitive data, protecting data in use, e.g. during secure inferencing.
5. Cloud-based TEE solutions could be relatively more cost-effective, as compared to deploying it within one's environment.

Use Case Details

Key POC Stakeholders:

6. **Singapore General Hospital (SGH)** – Evaluating value of conducting inferencing of an AI model situated in a cloud-based TEE
7. **Agency for Science, Technology and Research Institute of Advanced Intelligence and Computing (A*STAR IAIC)** – TEE Solution Provider

Proof of Concept and use of TEE

8. In this POC, SGH will be leveraging A*STAR IAIC's TEE solution, to address privacy and security related risks where medical data may need to be shared, for the purpose of conducting AI inferencing and using the results for clinical diagnosis.
9. A TEE is a type of PET which provides hardware-based, isolated execution, enabling sensitive data to be processed securely within a protected environment, mitigating risks associated with data in use, e.g. software breach, etc.

10. A*STAR IAIC's TEE solution is called Privacy Enhanced Guided Intelligence eNvironment (PENGUIN) and utilises AMD and Intel Confidential Computing solutions.

Overall POC Process Flow

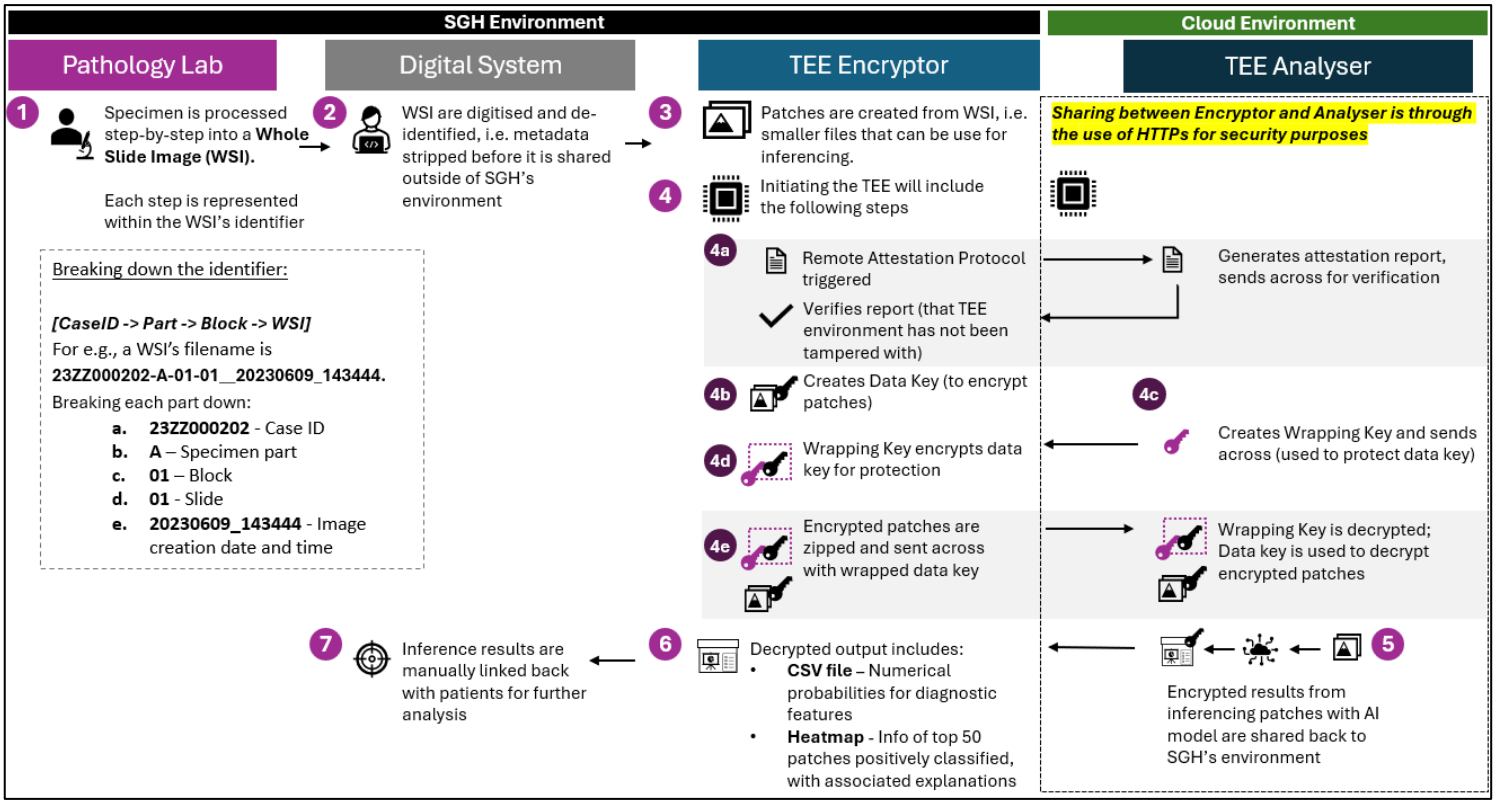

Diagram 1: POC Process Flow

POC Process Flow

11. Step 1:

a. To conduct pathological testing, patient specimens are submitted to SGH's laboratory. If the specimen comes from SGH, only required patient information will be shared through SGH's electronic medical record (EMR)[1] system to the Laboratory Information System (LIS). Only necessary and required patient demographic details for verification purposes would be sent alongside the specimen to the laboratory.

b. A case number ('case ID') is then generated and tagged to the patient, and recorded in the LIS, e.g. 23ZZ000202.

c. The specimen is then processed into multiple tissue blocks, and each block is further sectioned into multiple sections mounted on glass slides, and the glass slides with the tissue sections can be digitised into whole slide images (WSI).

[1] The EMR is access controlled, and there is no active exchange of messages or between the EMR and the LIS; only necessary information required would be shared.

d. The WSI can be identified via a string of numbers that includes the earlier generated case ID and references against the specimen, block, etc. For e.g., a WSI filename could be: **23ZZ000202-A-01-01__20230609_143444**, where in sequence:
    i. **23ZZ000202** represents the case ID listed above in 5b
    ii. A represents the part number,
    iii. 01 represents the block number
    i. 01 represents the slide number
    ii. 20230609_143444 represents the image creation date and time

12. Step 2:

a. The digitised WSI is uploaded into the internal digital pathology system. Each WSI possesses metadata containing patient-identifying information, e.g. patient's name, and this information is used to verify against the LIS to ensure the WSI corresponds to the correct patient.
b. If the WSI needs to be shared outside of the SGH environment, it undergoes a de-identification process, where patient metadata is removed. All details are stripped except the slide identifier, which is maintained in the name of the WSI file, e.g. **23ZZ000202-A-01-01__20230609_143444.**

13. Step 3:

a. By itself, the file size of the WSI is too large to be processed. To conduct inferencing, it is further split into smaller patch files, which are microscopy images of the pathology specimen. *[Note: the split patch files also retain the slide identifier.]* After the patches are created, the TEE process begins.

14. Step 4

a. The TEE solution consists of two key components:
    i. **TEE Encryptor (within SGH)**. The encryptor acts as a secure gateway for encrypting outbound data and decrypting returned results. *[Note: The **Key Management System** leverages **Federal Information Processing Standard 140-2 (FIPS 140-2)**[2] cryptographic algorithms to securely manage*[3] *encryption keys.]*
    ii. **TEE Analyser (within cloud environment)**: Primary secure computing environment in the cloud.

15. Step 4a

a. Since secure inferencing happens in the TEE Analyser, the TEE Encryptor would have to share the data key with the Analyzer to decrypt the patch files for

[2] FIPS 140-2 is a computer security standard that validates cryptographic modules, developed by the US National Institute of Standards and Technology (NIST).
[3] To securely manage the encryption keys, OpenSSL, which is a robust, commercial-grade, full featured open-sourced toolkit for transport layer security, is used.

inferencing. The data key is an encryption key and if intercepted, could be used to expose the sensitive data, it therefore needs to be protected in transit; the TEE Analyser generates a wrapping key that is used to encrypt and protect the data key. ***[Note: The wrapping key utilises asymmetric encryption, i.e. Public Key Infrastructure (PKI), which includes a pair of public and private keys. Public version of wrapping key is sent to and used by TEE encryptor to encrypt Data key.]***

16. Step 4b

a. To verify that the environment in the TEE Analyser has not been tampered with and that its integrity is intact, the remote attestation protocol is initiated by the TEE Encryptor. The TEE Analyser generates an attestation report with a hash of the wrapping key embedded within to bind the wrapping key to the attestation process.
b. The attestation report and the wrapping key are shared to the Encryptor where the report is verified and the wrapping key validated with the hash embedded in the report.

17. Step 4c

a. Next, the TEE Encryptor generates a data key, which is used to encrypt the patch files and decrypt encrypted results when it is shared back at the end of the securing inferencing process.

18. Step 4d

a. After attesting that environment has not been tampered with and encrypting the data key, the TEE Encryptor sends two items across to the Analyzer
    i. Data key – Encrypted using the wrapping key
    ii. Patch files – Encrypted using the data key

19. Step 5

a. The encrypted data key is decrypted using the corresponding wrapping key owned by TEE Analyser. ***[Note: Private version of wrapping key owned by TEE Analyser is used to decrypt data key encrypted using the public version of the wrapping key.]***
b. The decrypted data key is used to decrypt patch files, which would then be ready for inferencing with the AI model within the TEE environment.
[Note: Processing data in a TEE within the cloud adds an additional layer of security in the processing of the data, such that even if the cloud environment

was compromised, the sensitive medical data would remain encrypted within the memory of the TEE and is unintelligible due to the robust encryption measures in place.]

c. Patch files are then used for inferencing and two outputs are produced:
   i. **CSV file** contains numerical probabilities for specific diagnostic features
   ii. **PNG file** contains information of where the model identified top 50 patches positively classified with each medical condition

d. The output files are encrypted with the data key and sent back to TEE Encryptor.

20. Step 6 and 7

a. TEE Encryptor decrypts the encrypted results using the data key. Patches and results are then manually linked back via the slide identifier key that is only internal to SGH with specific patient case within the Laboratory Information System.

Regulatory Learnings

21. SGH sought Practical Guidance (Guidance) from the Personal Data Protection Commission (PDPC) on the following:

a. Are microscopy images of the pathology specimens considered personal data (PD)? If so, would sharing it to the cloud-based environment for processing constitute disclosure?

b. What responsibilities does SGH have under the PDPA when processing microscopy images of pathology specimens in the cloud-based TEE?

PDPC's assessment

Are microscopy images of the pathology specimens considered personal data? If so, would sharing it to the cloud-based environment for processing constitute disclosure?

22. Personal data is defined in section 2 of the PDPA to refer to data, whether true or not, about an individual who can be identified (a) from that data; or (b) from that data and other information to which the organisation has or is likely to have access.

23. PDPC notes that in this case, SGH, A*STAR IAIC and BDH each process different versions of the "microscopy image" which must be assessed separately.

24. For SGH, the pathology specimens are processed into digital pathology images (i.e. the patch files), and any direct patient identifiers associated with the patch files are removed for sharing outside the SGH environment. However, **these images retain the corresponding unique slide identifiers in the metadata and/or file name.** The processed patch files that result from these images also retain the corresponding slide identifiers. These slide identifiers contain the case ID, which allows for the identification of the corresponding patients and demographic details by SGH. **Hence, when SGH sends the patch files to the TEE Analyser, SGH will be considered to have disclosed personal data**

25. On the other hand, A*STAR IAIC and BDH would only process the encrypted patch files after transfer to the TEE environment, which still contain the unique slide identifier containing the case ID of each microscopy image. **These encrypted patch files are not personal data to A*STAR IAIC and BDH as they would not be able to identify an individual from the patch files or any other information, they are likely to have access to.** To elaborate:

    a. The **patch files are encrypted when received**, and neither A*STAR IAIC nor BDH would ordinarily have any means outside of the TEE Analyser to decrypt the files.

    b. The encrypted patch files are decrypted only within the cloud-based TEE Analyser which neither A*STAR IAIC nor BDH can access since **remote access will be removed after the solution is implemented**, notwithstanding that the TEE Analyser is located in BDH's cloud-environment. Without remote access to the TEE Analyser and since the TEE Encryptor is deployed on-premise, only SGH, and neither A*STAR IAIC or BDH, would have access to the encryption keys to decrypt the patch files and the subsequent inference results.

    c. Once the encrypted patch files are transferred to the cloud-based TEE environment, the **TEE Analyser ensures that the encrypted patch files are processed in a secure computing environment** where only codes and algorithms authorised by SGH can be run, and this is programmatically verified within the TEE.
    d. Unlike SGH, **neither A*STAR IAIC nor BDH has the information to identify the corresponding patient** through the unique slide identifiers on the patch files.

26. Without remote access to the TEE Analyser, prevailing protection measures in place against authorised access to the TEE Encryptor deployed on SGH's premise, and with verification of authorised codes and algorithms conducted programmatically, i.e. without the need for human intervention, **A*STAR IAIC and BDH would not have access to the TEE Encryptor or Analyser, and the encryption keys or codes and algorithms within.**

27. Hence, A*STAR IAIC and BDH are not collecting, using, or disclosing personal data.

28. In the event that any of the above circumstances change such that A*STAR IAIC or BDH are able to identify individuals or correlate the information to identifiable individuals,

they may be considered to be processing personal data and should promptly notify SGH for a reassessment of the case.

What responsibilities does SGH have under the PDPA when processing microscopy images of pathology specimens in the cloud-based TEE?

29. Uploading of patch files into the cloud-based TEE would constitute disclosure of personal data, and SGH must comply with all obligations under the PDPA e.g. consent, purpose limitation, notification, protection etc. For more information on how to comply with data protection obligations under the PDPA, please refer to the Advisory Guidelines on Key Concepts in the Personal Data Protection Act.

30. PDPC notes that the implementation of TEEs is a good technical data protection practice to ensure sensitive medical data can be securely processed in a cloud environment, while enabling use of data for insights and data innovation (e.g. training or conducting inferencing of AI models)

Additional safeguards that can be implemented

31. PDPC notes that BDH's data infrastructure is ISO 27001[4] certified, aligning with international best practices and standards. As added security measures to reduce the risk of re-identification, the following safeguards should be considered:

TEE Hardware and Firmware Supply Chain Security

32. Ensure that TEE hardware and firmware meet verifiable provenance and integrity standards, including evidence of trusted manufacturing origin, up-to-date firmware, and an attestation mechanism that can be independently verified.

TEE Access Controls

33. Implement strict role-based access controls that clearly separates administrative, operational, and audit roles within the BDH environment, to ensure no single administrator should have unilateral access to both the TEE infrastructure and the key management system simultaneously.

34. Privileged access to TEE host infrastructure should be governed by a Privileged Access Management (PAM) solution with just-in-time provisioning, ensuring elevated access is granted only for specific, time-bounded tasks and automatically revoked thereafter.

[4] ISO 27001 is an international standard for Information Security Management Systems and provides a framework for organizations of any size to systematically manage, monitor, and continually improve information security—protecting the confidentiality, integrity, and availability of sensitive data.

TEE Attestation

35. The current use of AMD SEV-SNP (Secure Encrypted Virtualization - Secure Nested Paging) framework with AMD's Key Distribution Service (KDS) as the root of trust is a recognised industry approach for TEE hardware-based attestation. As an area for improvement, consider providing SGH with a human-readable attestation report at initial deployment, covering key integrity indicators such as enclave configuration, memory encryption status, and source code measurements validated against the VCEK certificate from AMD KDS. It is also recommended that attestation evidence be made available to SGH for independent re-verification, particularly in the context of significant platform changes or security incidents.

Key Management

36. Schedule automatic key rotation and revocation based on predefined policies to reduce risk of human error.

Encryption of Model Weights

37. Encrypt model weights at rest, instead of storing in plain text. Decryption key should be managed within the TEE to ensure weights are only decrypted inside the secure boundary for inference;

Periodic Reviews and Audit

38. Implement comprehensive audit trails for all TEE operations, including attestation events, data provisioning, and computation requests to ensure detailed security monitoring and forensic capability;

39. Conduct periodic penetration testing and security audits specifically focused on TEE implementations, including side-channel analysis and attestation verification procedures.

40. While TEE solutions offer secure infrastructure, **organisations that engage third party TEE solution providers remain responsible under the PDPA for the data processing within the TEE.** SGH may wish to conduct due diligence to assess and validate if the data security afforded by TEE solution, and the data governance policies and practices of the solution provider are sufficient and appropriate for their business needs.

Results and Next Steps

41. SGH had concluded that the deployment of a TEE in a cloud-based environment for secure AI inferencing can potentially lead to greater capacity and more secure analysis, compared to an on-premise deployment of the TEE. Key benefits include:

    a. More inference loads can be processed in parallel in the cloud;

b. Improved security in processing of sensitive medical data in the cloud with the use of TEE

42. However, the use of TEE had also increased:

a. The amount of time it takes to conduct inferencing, e.g. around 10% more than a scenario without a TEE.
b. The overall cost, e.g. TEE infrastructure, with one encryptor and analyser each, costing around SGD 4,000 per month.

43. Nevertheless, the benefits of using a cloud-based TEE, e.g. increase in analysis capacity, outweigh the cons, e.g. increase in overheads associated with the use of a TEE, and can potentially be considered for other use cases within the hospital. The POC results have been presented to internal management and next steps, e.g. whether to scale or deploy in other parts of SGH, are being deliberated.