FHE Use-Cases and Benchmarking

Breakout Session

7th HE Standardization Workshop

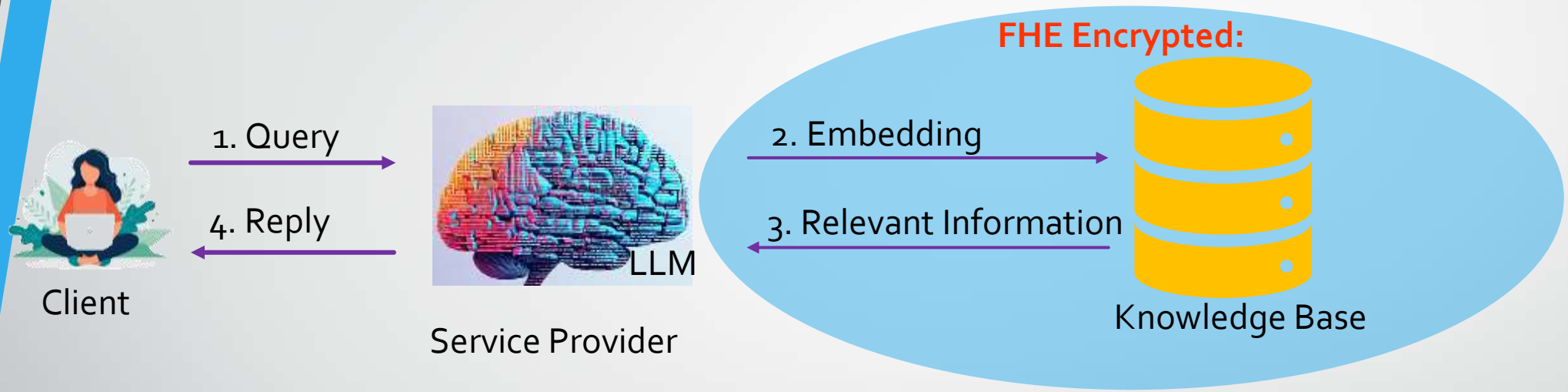
Salt-Lake City, UT, USA, October 13, 2024

Goal: "Standard" FHE Benchmarking Suite(s)

- From a technology-consumer point of view
- Help businesses make decisions:
  - Which of my workloads could feasibly use FHE?
  - Which implementation(s) of what variant(s) should I use?
- Method:
  - Identify use-cases of interest
  - Describe workloads that shed light on the performance characteristics for these use-cases
  - Collect workloads into (one or more) benchmarking suites

Example: RAG with External Knowledge Base

RAG = Retrieval Augmented Generation



- Enables loose coupling of ML and knowledge bases
  - LLM from OrgA can call knowledge-base from orgB
- Representative workload: Compute cosine similarity, return closest match

Example: Private Key-Value Fetch

Customer

1. queried key(s)

2. returned value(s)

FHE Encrypted:

Key	Value
vteK	uEGL tMSe xv4UK8N Z ZtT3t MPaqc9z 5M
G9JV	Fw 9cIEy vWN1 S4Vb4 xiUW dJU G9w6 pXE
PJ4Q	dV6 Rpzno LQw UK6N TshCuo 62To3 AeLy
CXYP	ctp3 YJQ Wfon EJ6 7qdS jBkz Qsu HD450
27eG	ct p3YJ QW fonE J67q dSjB kzQ Su HD450
JrZ	M Ci MzL Vjy2 h3T 5kBe Wj7 SvTK Kes JWL
BezR	ID6 WKc byZ8 qEfu bjq5 QRXYSU gbc VDi
AjWM	Mt Nvm Syx 6 2z8 TK Egw w4 5z7 EoN NIC dm
…	…

- Commercial customers access to noSQL databases elsewhere
  - Maybe on behalf of end users
- Representative workload: PIR by keyword

Example in the web3 space

- Use Case examples from Remi's excellent talk
- Representative workloads?
  - Many concurrent heterogenic "simple" operations?
    - 64-bit comparisons, additions, thresholding
    - Small-table lookups
    - …
- Who are the technology-consumers here?
  - Blockchain providers that was to adopt FHE?
  - Blockchain users that was to deploy smart contracts?

What Makes Good Benchmarks?

- Relevant: Benchmark numbers provide useful input for making FHE-deployment decisions
  - Not just marketing value to the implementors
- Well Specified
  - Data: Database size? Datatypes? Precision? Dynamicity?
  - Process: Preconditions? Function? Postconditions?
  - Performance: Time? Space? Single-threaded? Multi-threaded? Accelerated?

Breakout Session Blurb

- This breakout session focuses on both applications where FHE provides real-world value, and benchmarks that allow solution providers and users to asses the performance of solutions. Example use-cases may include ML/AI applications where the model or the query must remain secret, data-access with private queries such as PIR or PSI, and simple general-purpose program such as smart contacts in web3 architectures. A major goal of this breakout session is to continue community efforts to develop a benchmarking suite for FHE, with concrete workloads relevant to real-world applications, and their performance characteristics.