Abstract

Outsourced data processing has become an attractive tool for cloud-oriented services to provide diverse functionalities on cloud-connected digital devices. State-of-the-art cloud infrastructure is the backbone of these services, which physically stores and processes users' data involving ordinary citizens to government/industrial establishments. Naturally, this brings attention to the strong privacy requirement of a client's data on the untrusted third-party cloud to protect it from unauthorised access. The use of plain data encryption to ensure users' data privacy trivially loses the ability to perform any computation/processing over the encrypted data. In this thesis, we explore an advanced cryptographic paradigm called Searchable Symmetric Encryption (SSE) which allows efficient processing of outsourced encrypted data without decryption. SSE outperforms other comparable cryptographic paradigms, including fully homomorphic encryption (FHE), functional encryption (FE) and oblivious random access machine (ORAM) in efficiency and scalability. However, despite being theoretically efficient, SSE often suffers from severe performance bottlenecks due to physical implementation-related issues. Additionally, state-of-the-art efficient SSE is restricted to single-client conjunctive query searches only. In this thesis, we study scalable and robust design approaches for fast, efficient SSE implementation system architectures and advanced SSE constructions covering complex queries, including multi-client searches with updates, for practical cloud applications.

In the first phase of the hardware-accelerated SSE implementation system design, we concentrate on developing primary hardware computing units for primitive crypto-operations. Typical software implementations suffer from significant execution latency and limited scope for parallelism, which quickly saturates a scheme's throughput for large real databases. We develop a lightweight programmable public-key cryptography processing core for complex cryptographic primitive operations, such as Elliptic Curve Cryptography, to achieve higher parallelism in hardware accelerator design and increased throughput for SSE execution. In the next phase, we design reconfigurable device-based accelerator architecture for SSE, which uses the processing core(s) developed in the first phase of the SSE implementation system development. The generic accelerator architecture takes advantage of the SSE-specific algorithmic structure and is exclusive to our custom hardware accelerator design. We attach these with the host system to realise the full hardware-software (HW-SW) co-design-based SSE implementation platform targeted for cloud applications. In the final phase of SSE implementation system development, we focus on addressing the performance bottleneck due to the costly storage access during SSE search execution. We develop a hardware-backed associative memory-based storage access mechanism for fast SSE search. We take advantage of SSE-specific query properties and locality information to efficiently *cache* data onto the associative memory subsystem for faster retrieval during repetitive accesses. The final end-to-end hardware-accelerated and memory-enhanced prototype SSE implementation system, called CAMiSE++, achieves greater than five-fold speed-up compared to naïve software implementations.

The aforementioned system-oriented developments focus on addressing multiple critical implementation-related challenges in SSE. This thesis's algorithmic contributions aim to support complex queries and dynamic update capability while allowing multiple clients to search in an SSE framework. While state-of-the-art SSE constructions have primarily focused on efficiently processing conjunctive Boolean queries, complex Boolean query processing is essential in real cloud applications. We present a storage-efficient SSE construction called TWINSSE that allows searching conjunctive and disjunctive Boolean queries over the encrypted database efficiently. The TWINSSE construction supports fast generic Boolean query search by decomposing the query into equivalent CNF/DNF expressions and is essential for practical search applications.

Finally, we present Nomos, the first efficient dynamic SSE construction in the literature supporting multi-keyword Boolean queries from *multiple* clients, to the best of our knowledge. While being an efficient multi-client construction, Nomos also mitigates a specific cross-term–based leakage present in multi-client constructions naïvely extended from single-client schemes, which can otherwise lead to a devastating attack breaking the scheme. In the end, both TWINSSE and Nomos constructions are designed to execute over the hardware-accelerated SSE implementation system CAMiSE++ for faster real-time performance in practical cloud applications.

Keywords. Searchable Symmetric Encryption, System Design, Algorithm Design, Hardware Architecture, Hardware-software co-design, Cloud, Reconfigurable Hardware