RestAssured: Securing Cloud Analytics

Oshrit Feder
IBM Research - Haifa
oshritf@il.ibm.com

Gidon Gershinsky
IBM Research - Haifa
gidon@il.ibm.com

Eliad Tsfadia
IBM Research - Haifa
eliad.tsfadia@ibm.com

ABSTRACT

Protecting sensitive business and personal information is a cornerstone requirement when enterprises and organizations move to the cloud. Many aspects of this requirement are already handled at various levels. Data-at-rest can be secured in cloud stores by encrypting it before persisting the data to storage, while data-in-flight is transmitted using protected channels such as TLS and HTTPS. Data-in-use, processed in cloud compute nodes, is the most vulnerable link in the end-to-end information flow, since the process memory can be accessed by malicious privileged software or system administrators.

IBM Research - Haifa takes part in a European H2020 research project RestAssured [2] that aims to deliver end-to-end cloud architectures and methodologies for assuring secure data processing in the cloud. We build a trusted analytic platform based on a combination of hardware and software components, and collaborate with the RestAssured partners to implement cloud analytic use cases ranging from social care services to pay-as-you-drive insurance policies.

The platform uses the Intel SGX (Software Guard Extension) technology [4], available in Skylake and later processors, that allows to create memory regions (enclaves) protected with hardware encryption in the SoC (system on chip). The data resides unencrypted only inside the processor. It is encrypted in SoC before being written to main memory, and decrypted in SoC upon fetching from main memory. Our team has designed and developed a framework for trust management in SGX enclaves [3] that performs verification (remote attestation) of the enclave hardware and software components, and assists in trusted delivery of secrets (such as data encryption keys) to the enclaves.

Apache Spark SQL [1] is the analytic engine of the RestAssured platform. We use the Opaque [6] open source technology [5] from the Berkeley RISELab that integrates the Spark SQL with Intel SGX hardware, and offers data protection by running SQL transformations inside trusted enclaves. We have augmented Opaque with a few key mechanisms for secure data processing in SGX enclaves, by *integrating Opaque with our trust management framework to enable remote attestation and data encryption key delivery to Opaque enclaves*. We have also *developed a component that serves as a gateway between RestAssured use case applications and Opaque clusters*. The gateway supports a REST endpoint that accepts SQL query from applications, sends the query for governance verification and modification by a rule engine, and executes the modified query in Opaque. The results are serialized into a JSON object and sent back to the application on a secure REST channel.

ACKNOWLEDGMENTS

The research work has received funding from the European Community's Horizon 2020 Research and Innovation Programme under grant agreement number 731678

CCS CONCEPTS

• **Security and privacy** → *Security in hardware*; *Distributed systems security*; *Software security engineering*;

KEYWORDS

Secure computing, big data analytics

REFERENCES

[1] Apache Spark. 2018. Spark SQL and DataFrames. https://spark.apache.org/sql/. (2018).

[2] EU H2020 RestAssured. 2018. Secure Data Processing in the Cloud. https://restassuredh2020.eu/. (2018).

[3] Gidon Gershinsky, Danny Harnik, Eliad Tsfadia. 2018. Linux SGX Trust Management Framework. https://github.com/IBM/sgx-trust-management. (2018).

[4] Intel. 2018. Intel Software Guard Extensions. https://software.intel.com/en-us/sgx/. (2018).

[5] UC Berekeley RISELab. 2018. Opaque: A data analytics platform with strong security. https://github.com/ucbrise/opaque/. (2018).

[6] J. G. Beekman R. A. Popa J. E. Gonzalez W. Zheng, A. Dave and I. Stoica. 2017. Opaque: An oblivious and encrypted distributed analytics platform. In *Proceedings of the Network and Distributed System Security Symposium (NDSS)*,.

SYSTOR '18, June 4–7, 2018, HAIFA, Israel
© 2018 Association for Computing Machinery.
ACM ISBN 978-1-4503-5849-1/18/06…$15.00
https://doi.org/10.1145/3211890.3211908