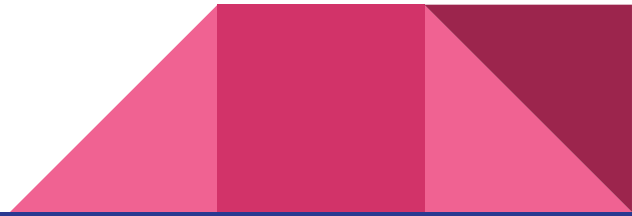
LogSafe: Secure and Scalable Data Logger for IoT

Hung Nguyen, Radoslav Ivanov, Linh T.X. Phan, Oleg Sokolsky, James Weimer

Presented by: Rachell Kim

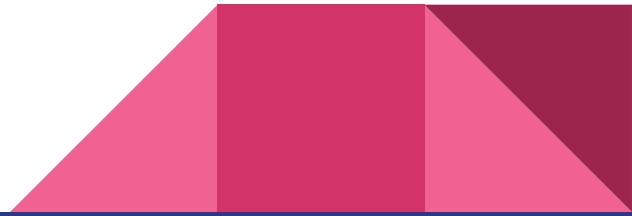
Introduction: Vulnerabilities in IoT

- IoT devices collect large amounts of personal data
  - Opens up security and privacy concerns
  - Devices become vulnerable to both physical and network attacks
- Securing all IoT devices is an impossible task
  - 20.4 billion IoT devices predicted to exist by 2020 [1]



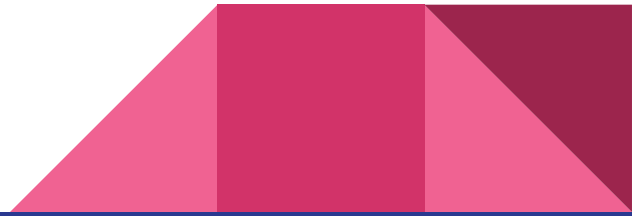
Introduction: Surfaces of Attack

- Cloud services in IoT
  - Yahoo, Ashley Madison, Equifax, etc.
- Replay, injection, eavesdropping, side-channel attacks
- Cyber-physical attacks
- DoS/DDoS



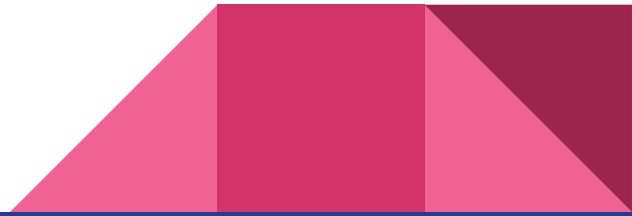
Introduction: Data Logging

- Need to secure and store data in a safe manner:
  - Methods of data collection
  - End-to-end security guarantees
- Communication protocols and data storage
- Previous works:
  - Not fault tolerant
  - Very slow
  - Scaling issues



Problem Definition

- Logging system must be:
  - Tamper evident
  - Fault tolerant
  - Scalable
- Should defend against network attacks such as:
  - Replay
  - Injection
  - Eavesdropping
  - Side-channel
- Must satisfy CIA properties



Problem Definition

- Logging system must be:
  - Tamper evident
  - Fault tolerant
  - Scalable
- Should defend against network attacks such as:
  - Replay
  - Injection
  - Eavesdropping
  - Side-channel
- Must satisfy CIA properties

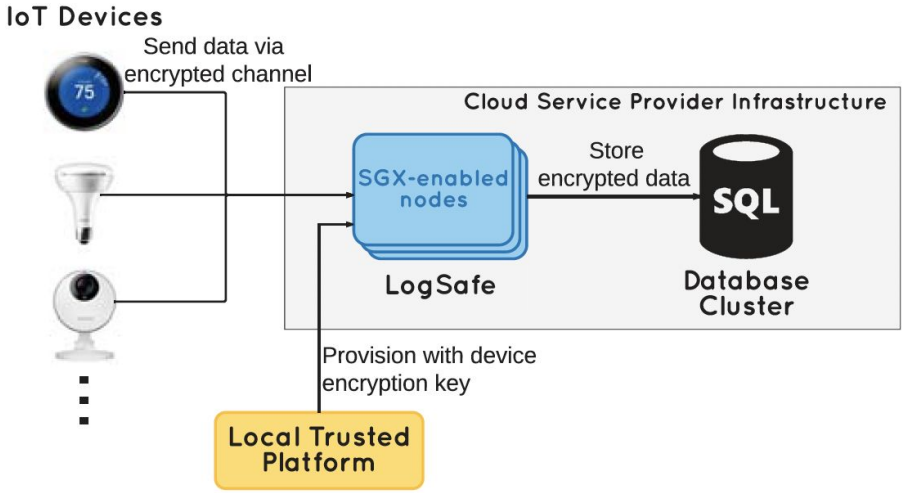
How do we build that?

Solution: LogSafe

- Decentralized logging architecture
  - Logger nodes built in a distributed fashion on the cloud
- Logger nodes are placed in a ring-like structure to back up other nodes
- Uses Transport Layer Security (TLS)
- Employs Intel Software Guard Extensions (SGX)
  - Allows for the creation of an isolated execution environment on the cloud
  - SGX-enabled nodes are decentralized to minimize performance hits
- Other contributions → Snapshot algorithm

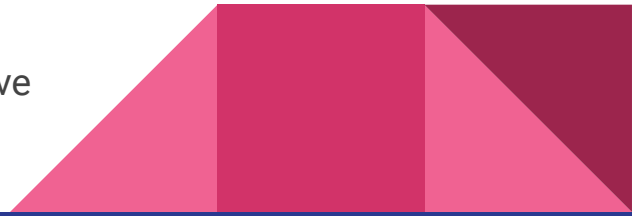
LogSafe Architecture I

- Includes:
  - SGX-enabled nodes
  - The cloud
  - IoT devices
  - Method to store data



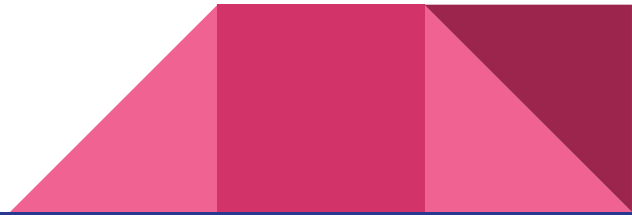
SGX

- Software Guard Extensions (SGX)
  - Designed by Intel
  - Enables safe execution of code in untrusted environments
- Enclave
  - Secure, encrypted region for code and data
  - Instantiated by SGX
  - Provides encrypted memory, but not for disk (i.e. not for I/O operations, etc.)
- Remote attestation
  - Allows other parties to verify the trustworthiness of the enclave

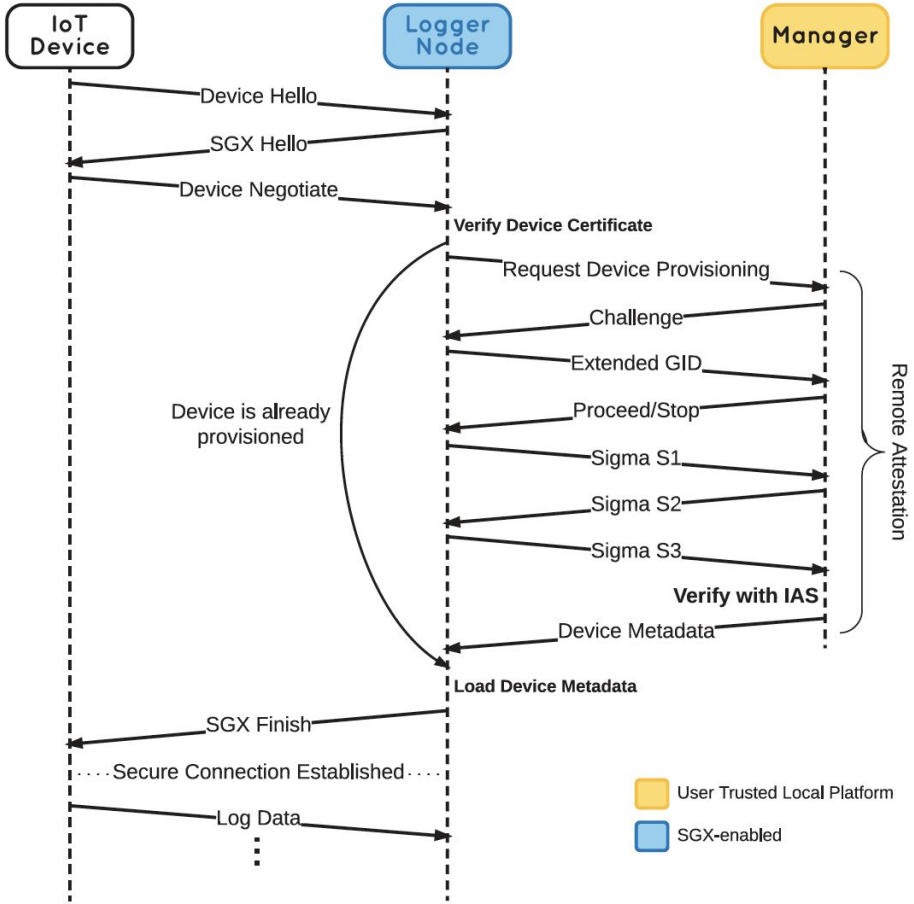


LogSafe Architecture II

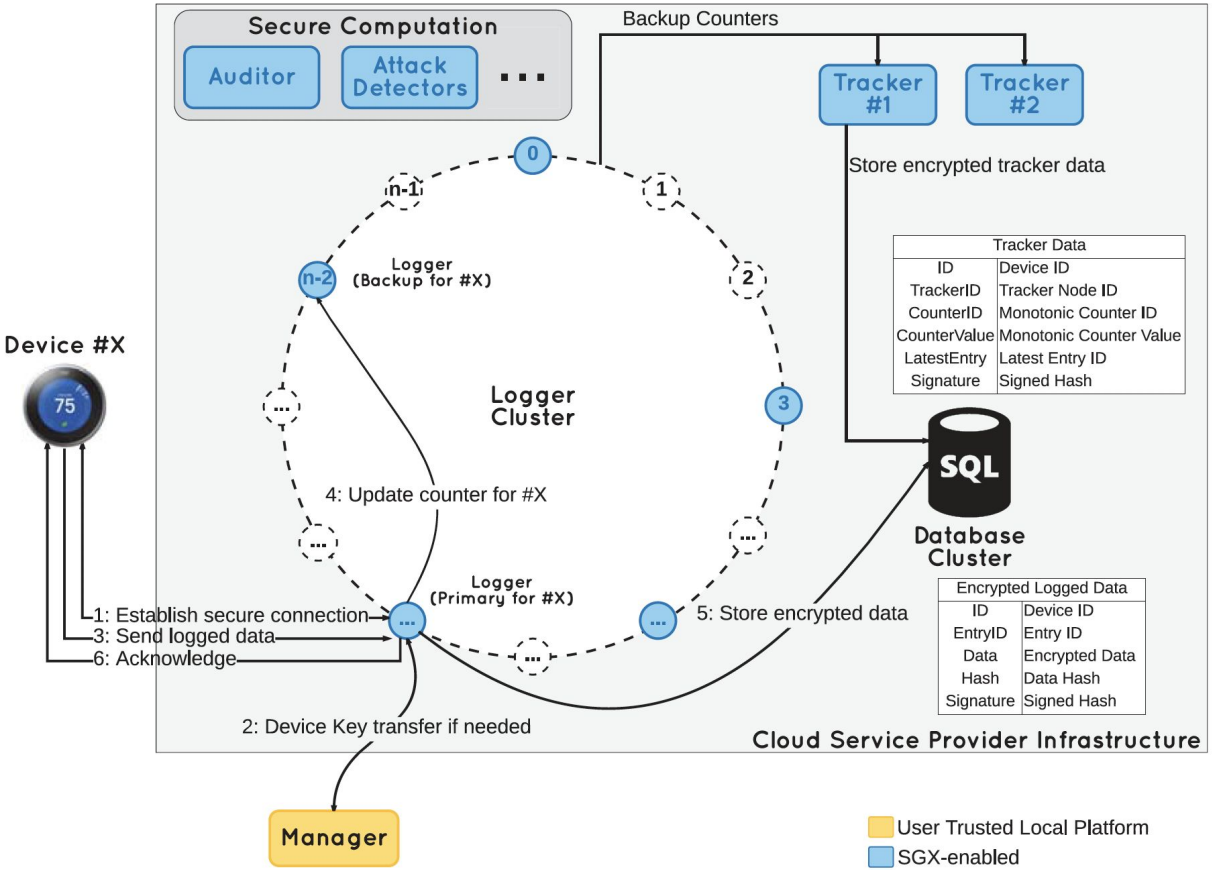
- Main modules of the LogSafe architecture:
  - IoT Device
  - Logger — Establishes connection to the IoT device and listens for incoming messages
  - Tracker — Takes a “snapshot” of the latest logged data of a given IoT device for verification purposes
  - Manager — Provisions both IoT devices and Logger nodes; assists in communication protocols (i.e. remote attestation)
- Written in C++



LogSafe Data Flow



LogSafe Data Flow



Snapshot Algorithm

- Use cases:
  - Logger cluster shutdown
  - IoT device is inactive for long period of time
- Creates a hash chain and signature without encrypting data
  - Uses the SGX monotonic counter
- Verifies in-memory Logger counter and authenticity of logged data

Solution Summary: LogSafe

- Decentralized logging → Availability, Fault-tolerance
- Cloud infrastructure → Scalability
- SGX implementation → Confidentiality, Integrity
- TLS + Hash Chaining → Integrity
- Attacks:
  - Replay and injection → physical monotonic counter
  - Eavesdropping and side channel → protection from the enclave
  - DoS/DDoS → *decentralized logger nodes

Evaluation

- Biggest issues:
  - Computation overhead of setup time
  - Scalability
- Setup time experiment
  - Measured the time needed to successfully establish a TLS connection when:
    - The device is being set up for the first time
    - The device is re-connecting to the Logger node

Evaluation

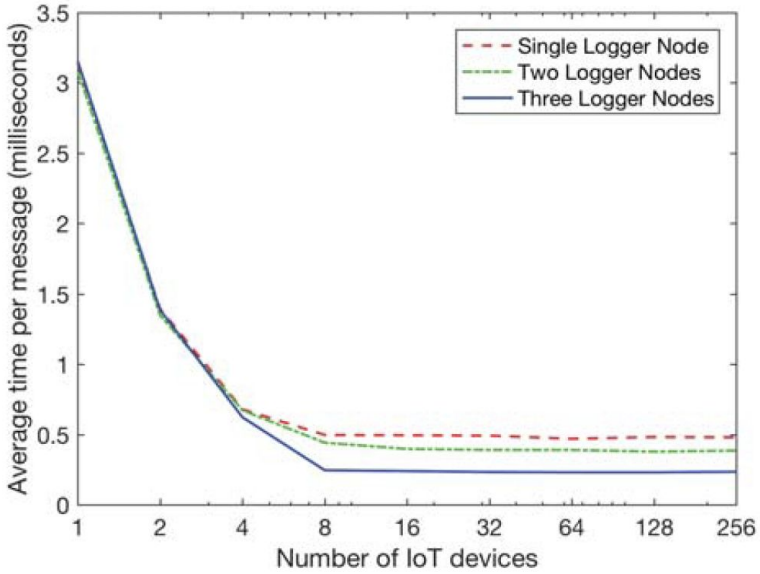
- Results:
  - Set up time with remote attestation tends to be significantly higher

Intel Edison	Dell 5480	Task	Logger
$717\mu s$	$87\mu s$	Device Hello $\rightarrow$	
		$\leftarrow$ SGX Hello	$5{,}886\mu s$
$370{,}055\mu s$	$8{,}617\mu s$	Device Negotiate $\rightarrow$	
		(*) Remote Attestation	$1.038s$
		$\leftarrow$ SGX Finish	$5{,}935\ \mu s$
$1.420s$	$1.059s$	Total time (with remote attestation)	
$382.5ms$	$20.5ms$	Total time (without remote attestation)	

- Claim:
  - Logger provisioning is a one-time cost

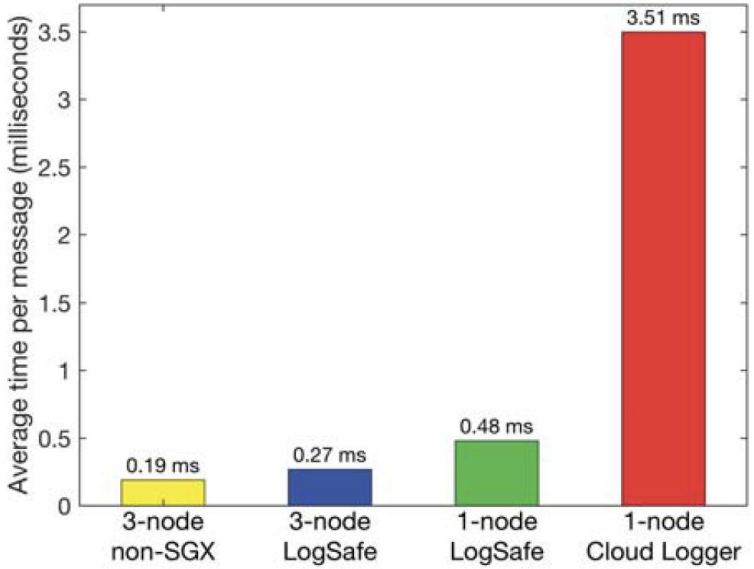
Evaluation

- Logging performance / Scalability results:
  - Average processing time for 3 nodes is much lower as the number of IoT devices increases
  - Multi-threading implementation benefits after 2+ devices joining the system



Evaluation

- Performance comparison with other implementations
  - Overhead of SGX
- Cloud Logger
  - Previous implementation using SGX
- Claim:
  - LogSafe provides more security with the price of slower performance



Critique

- Claims logger provisioning is a one-time cost, but also states it must reconnect once session time expires
  - What is the true cost of logger provisioning on long-term performance?
- Does not address methods to perform forensic analysis on logged data
  - Only that it can
- Evaluation did not include tests to challenge security of the LogSafe system

Conclusions

- Paper proposes a new, decentralized logging system using SGX-enabled notes to guarantee fault-tolerance and confidentiality of private information collected from IoT devices
  - Enhanced tamper detection
- Emphasizes the scalability of LogSafe
  - Incorporates cloud infrastructure to support better management of a large number of devices
- (Almost) Outperforms previous implementations with added layer of security through SGX features

Paper Feedback

- Is there a model that can determine the percentage of machines needed to be taken offline to disrupt a system of linked nodes sharing computation? (Niko Reveliotis, Comprehension)
- If enough nodes received sufficient I/O requests, could this potentially compromise the system as a whole? (Sam Frey, Comprehension)
- Is the number of snapshots created enough to maintain integrity? (Alvaro Albero, Critical)
- In Figure 4, why does performance seem to improve as more IoT devices push logs? (Sean McBride, Critical)
- How does a hash chaining work? (Sean McBride, Critical)

Works Cited

[1] - https://www.vxchnge.com/blog/iot-statistics