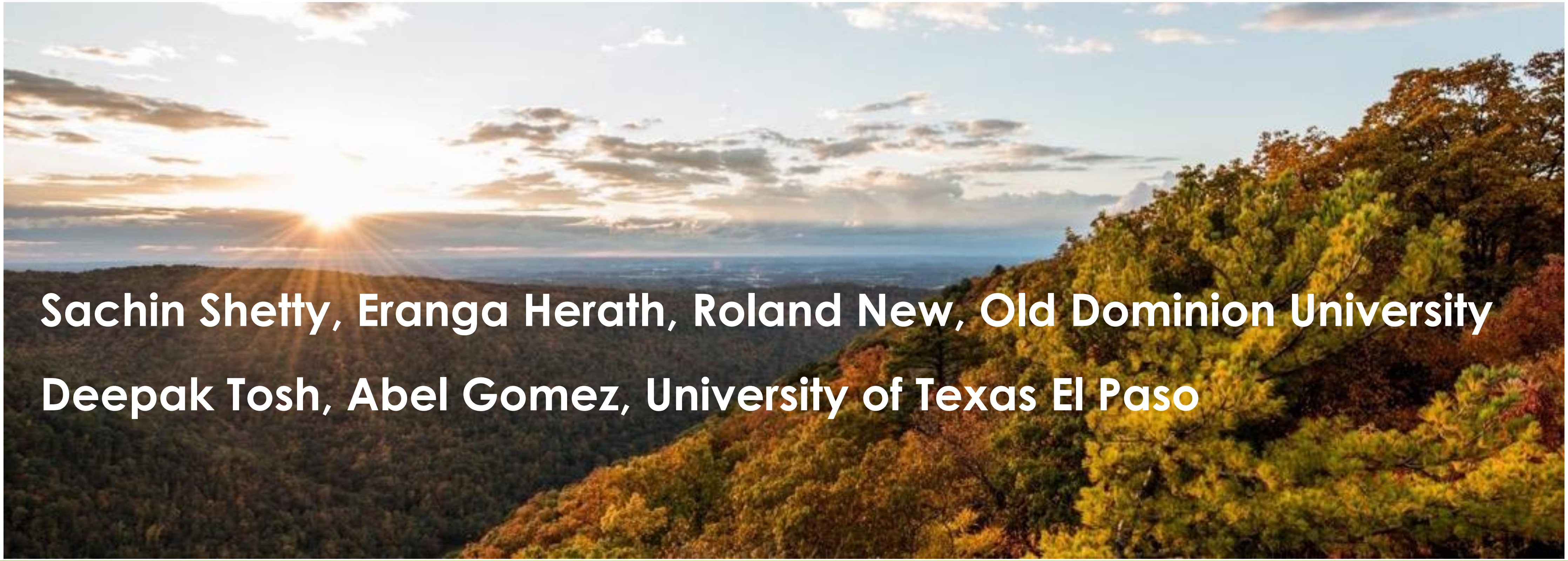
Blockchain Empowered Provenance Framework for Sensor Identity Management and Data Flow Security in Fossil-Based Power Plants

NATIONAL ENERGY TECHNOLOGY LABORATORY



Sachin Shetty, Eranga Herath, Roland New, Old Dominion University

Deepak Tosh, Abel Gomez, University of Texas El Paso

Project Description

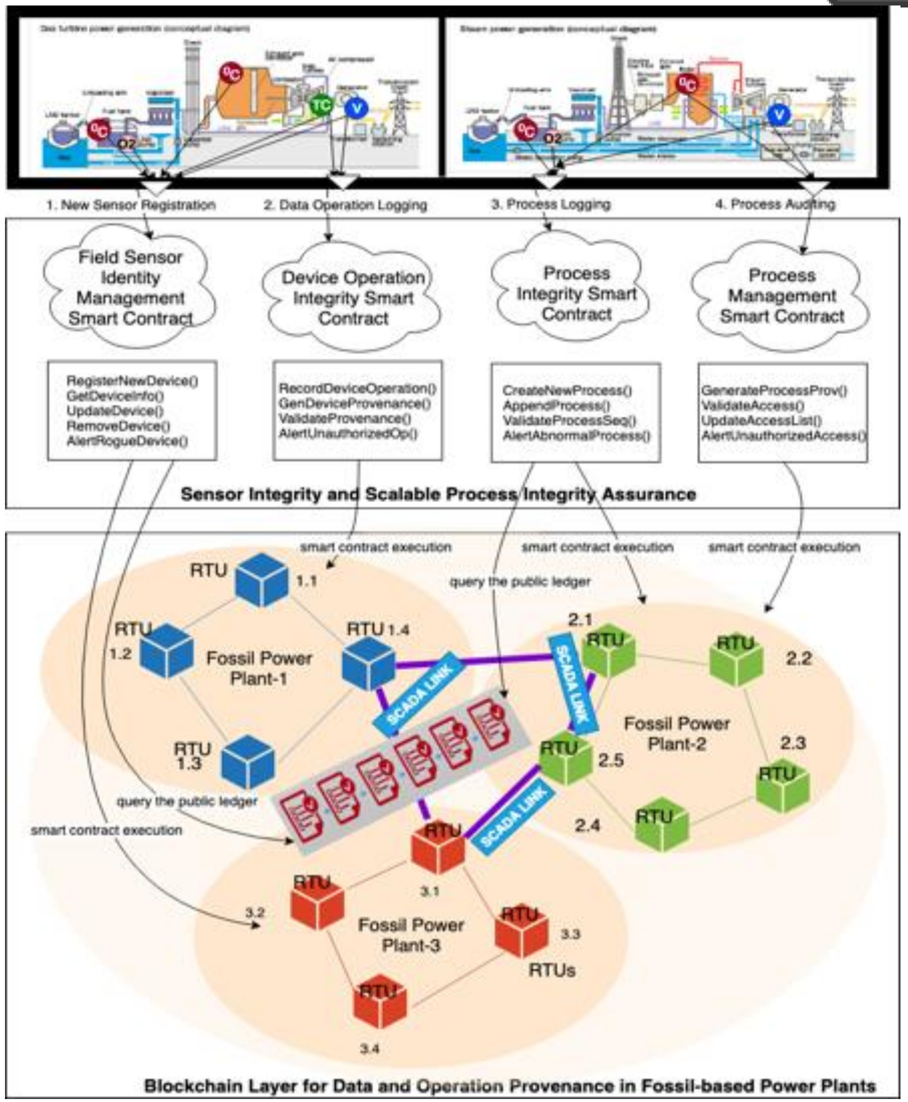
- **Project Goal** - Blockchain empowered provenance platform for identity management and process integrity for sensors in *Fossil-based Power Plants* (***FPP***).
- **Strategic alignment with DOE -** Improving electric grid reliability, resilience and availability
- **DOE-NETL** –Dr. Sydni Credle and Maria Reidpath
- **TEAM**
  - Old Dominion University – Virginia Modeling, Analysis and Simulation Center
  - University of Texas at El Paso – Computer Science
- **Partners**
  - Accenture, Argonne National Lab, ReliabilityFirst, Wood PLC
- **Contract**
  - October 1, 2019 – September 30, 2022

Project Objectives

Objective 1 - Sensor identity management via establishing a Peer-to-Peer (P2P) SCADA network

Objective-2: Networked Sensor Integrity and Scalable Process Integrity Assurance in FPPs

Objective-3: Prototype Development and Evaluation

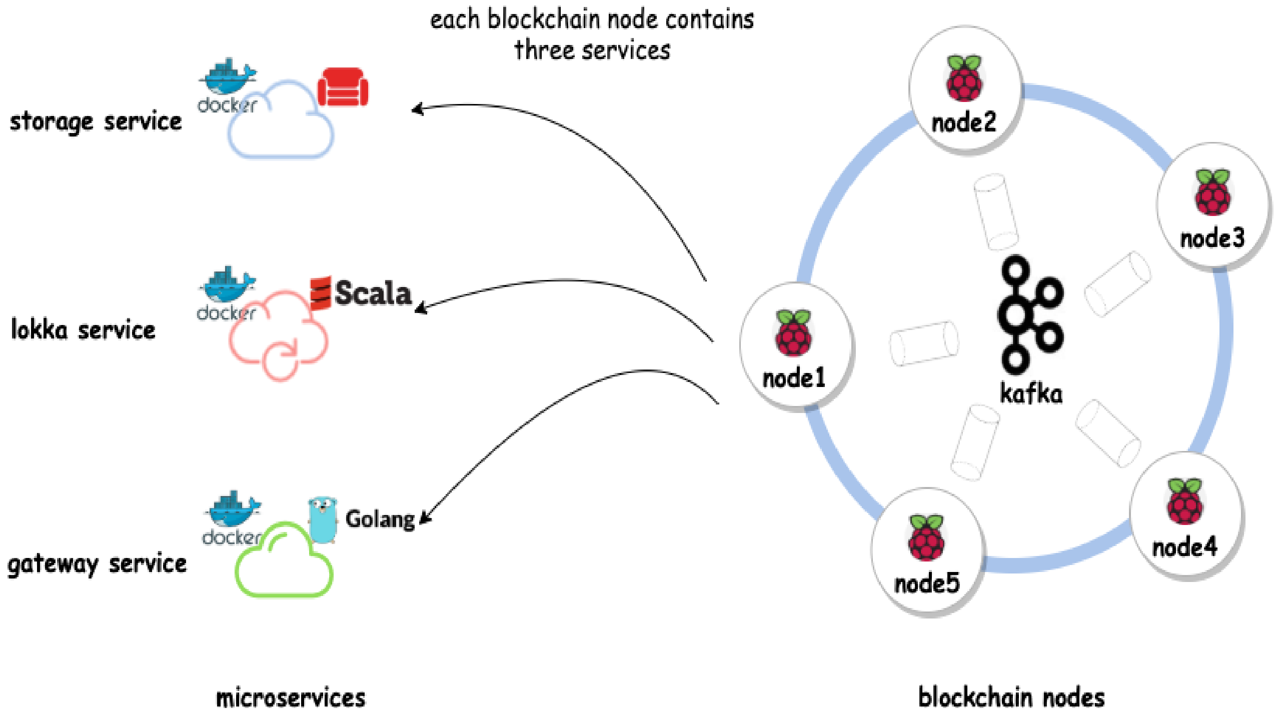


Project Accomplishments

- Developed Tikiri, a lightweight and scalable Blockchain platform
- Development of testbed using Raspberry Pis and LORAWAN to evaluate Tikiri platform
- Developed data authentication and integrity (SPAI) protocol is using SRAM-based Physical Unclonable Functions (PUF),
- Implemented the SPAI protocol in both Arduino and Raspberry Pi based testbed with DHT11 temperature sensors embedded into the boards
- Developed a lab based SCADA environment composed of HMI and PLCs for prototyping
- Discussions with partners, Accenture, WoodPLC and Argonne National Lab on platform development
- Webinar presentation to the EPRI Utility Blockchain Interest Group
- Submission of papers to IEEE SmartGridComm

Tikiri - Lightweight and scalable blockchain

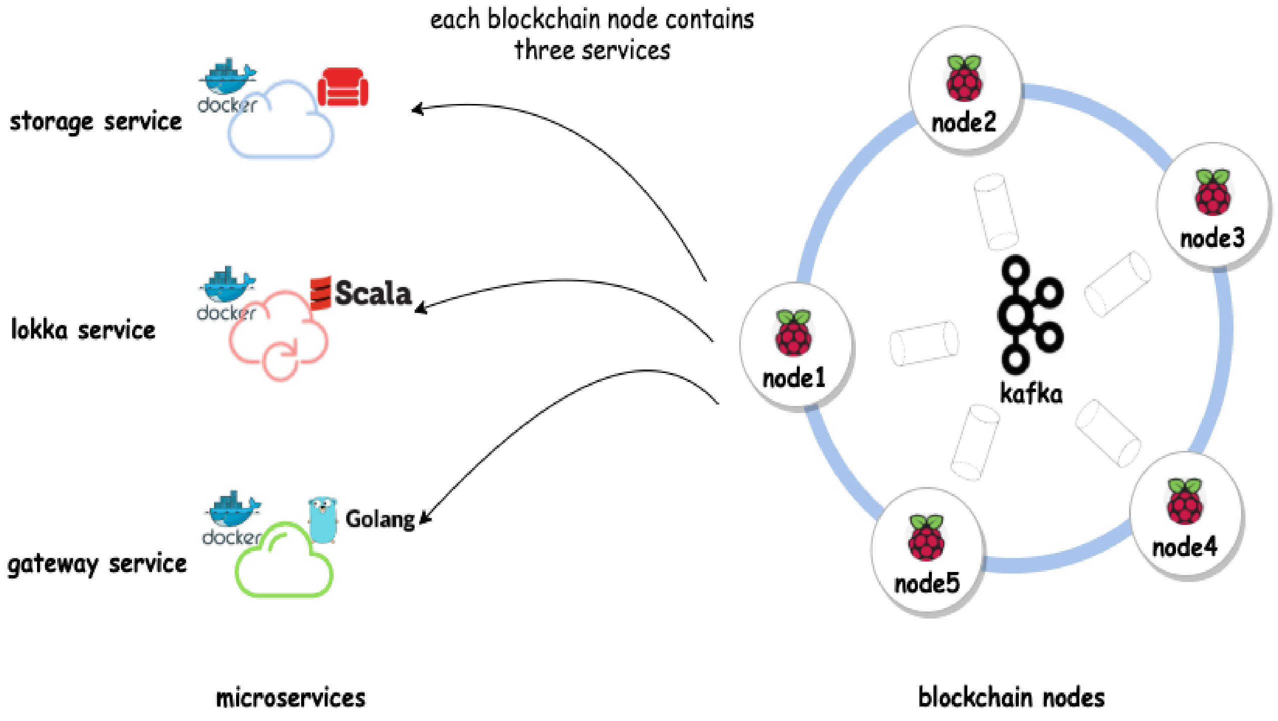
- Support real-time transaction
- Concurrent execution of blockchain transactions
- Support sharding based data replication to reduce the communication overhead
- Apache kafka based consensus to increase the scalability and throughput



Tikiri - Lightweight and scalable blockchain

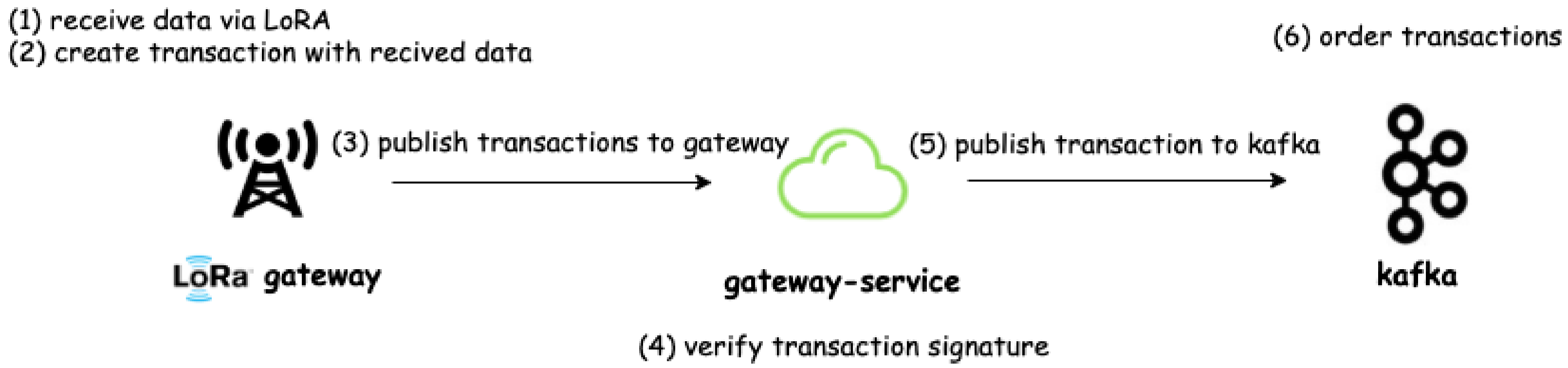
- Microservices based distributed system architecture for lightweight blockchain

- Reactive-streaming based back-pressure operation for scalabilty



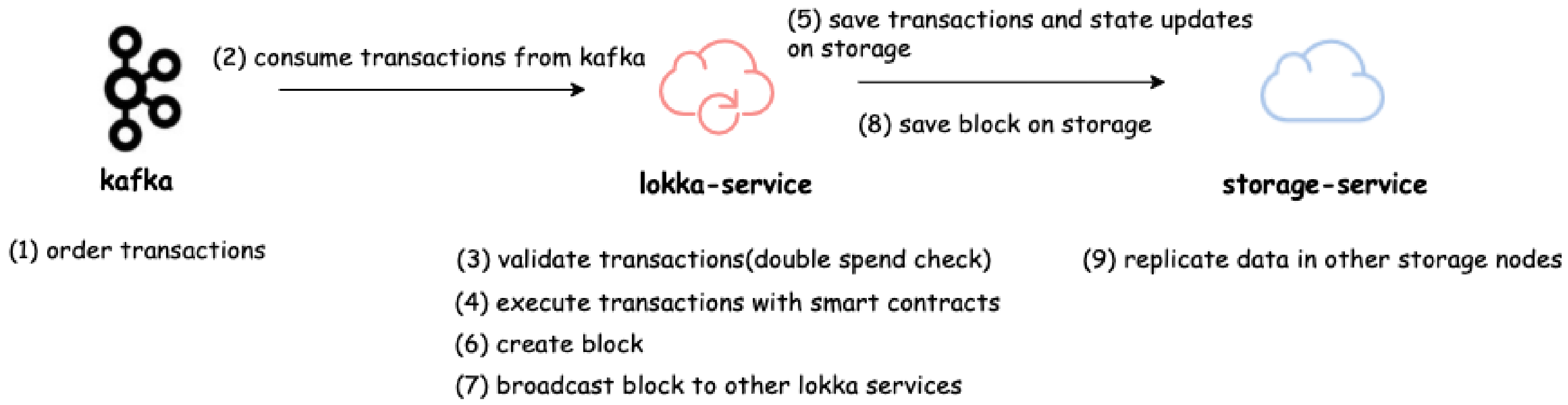
Tikiri Architecture

- Blockchain client(e.g LoRa gateway) connected to gateway service and publish transactions.
- Gateway service publishes the transactions to kafka.



Tikiri Architecture

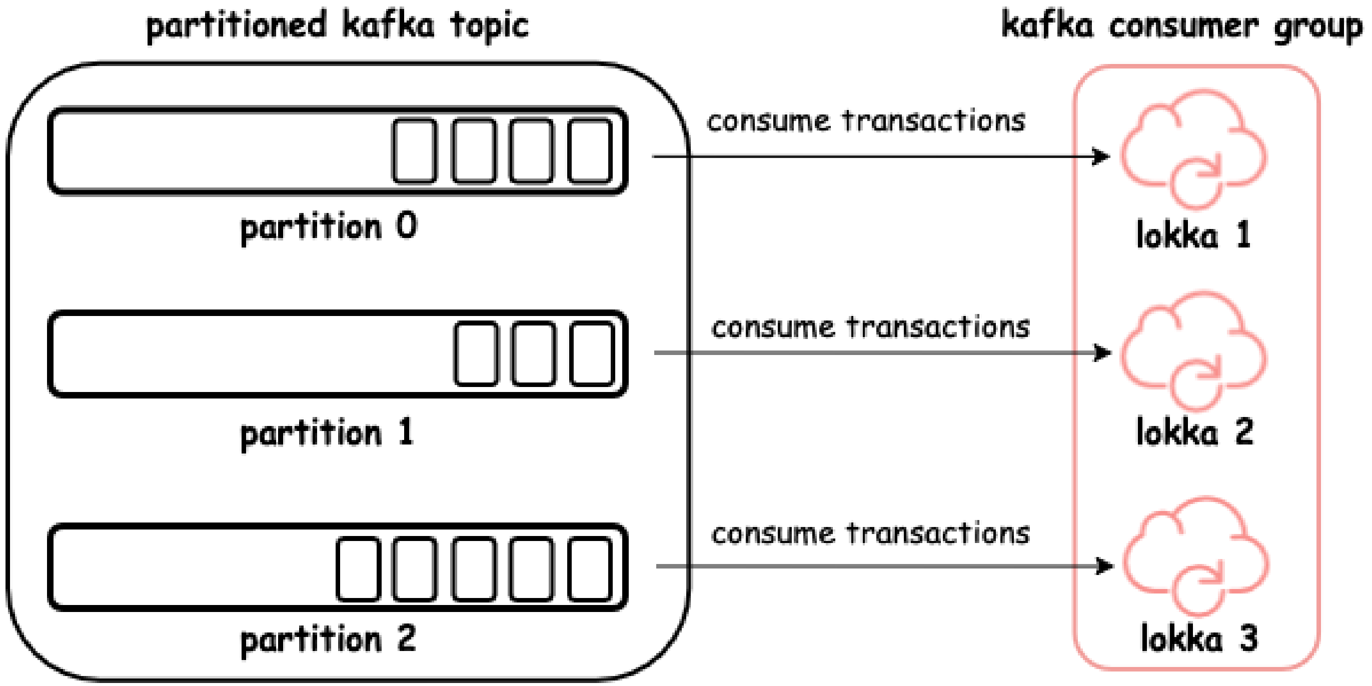
- Lokka services consume transactions via kafka, execute them and create blocks.
- Multiple Lokka services can work in parallel with kafka partitioned topics.



Tikiri Architecture

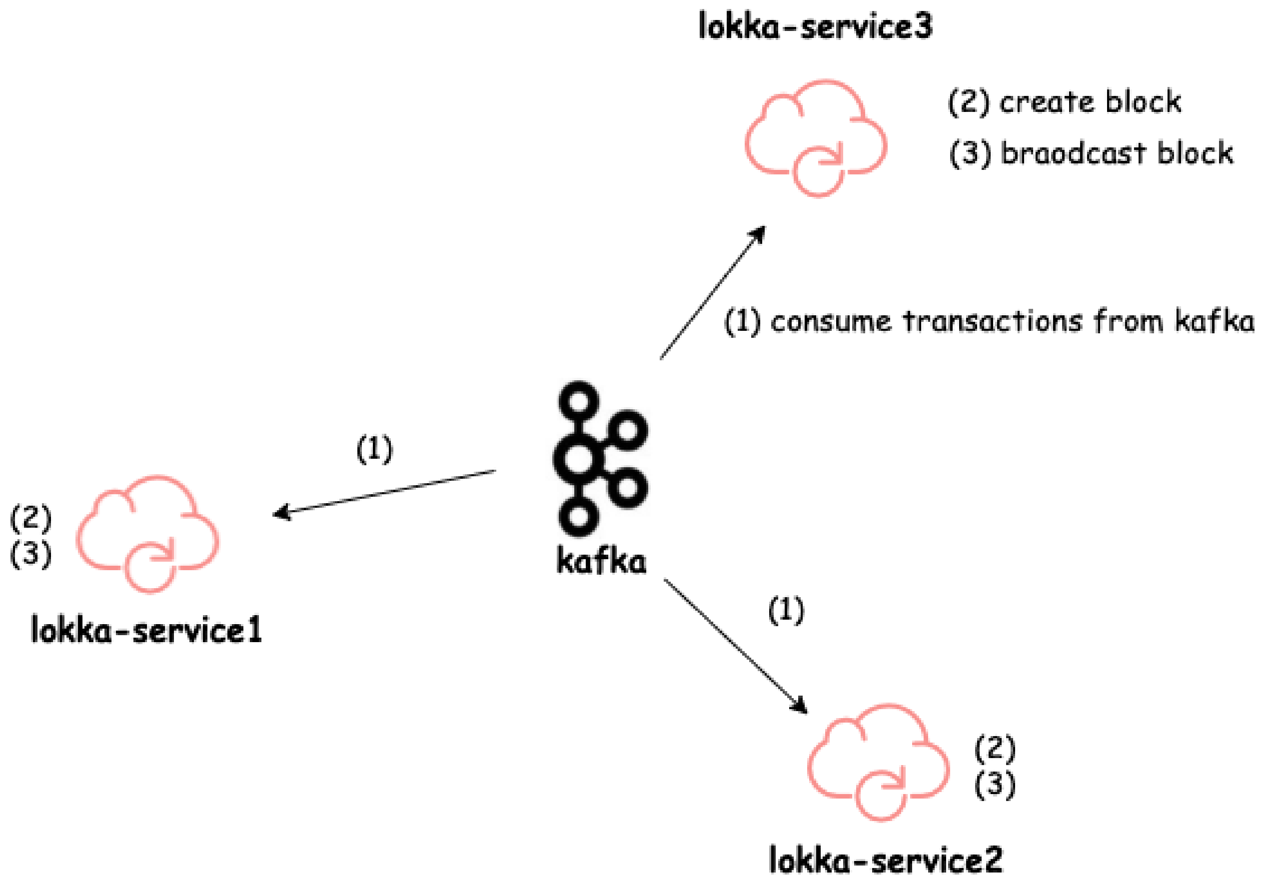
NATIONAL ENERGY TECHNOLOGY LABORATORY

- Kafka will handle message broadcasting between multiple Lokka services by guaranteeing total order



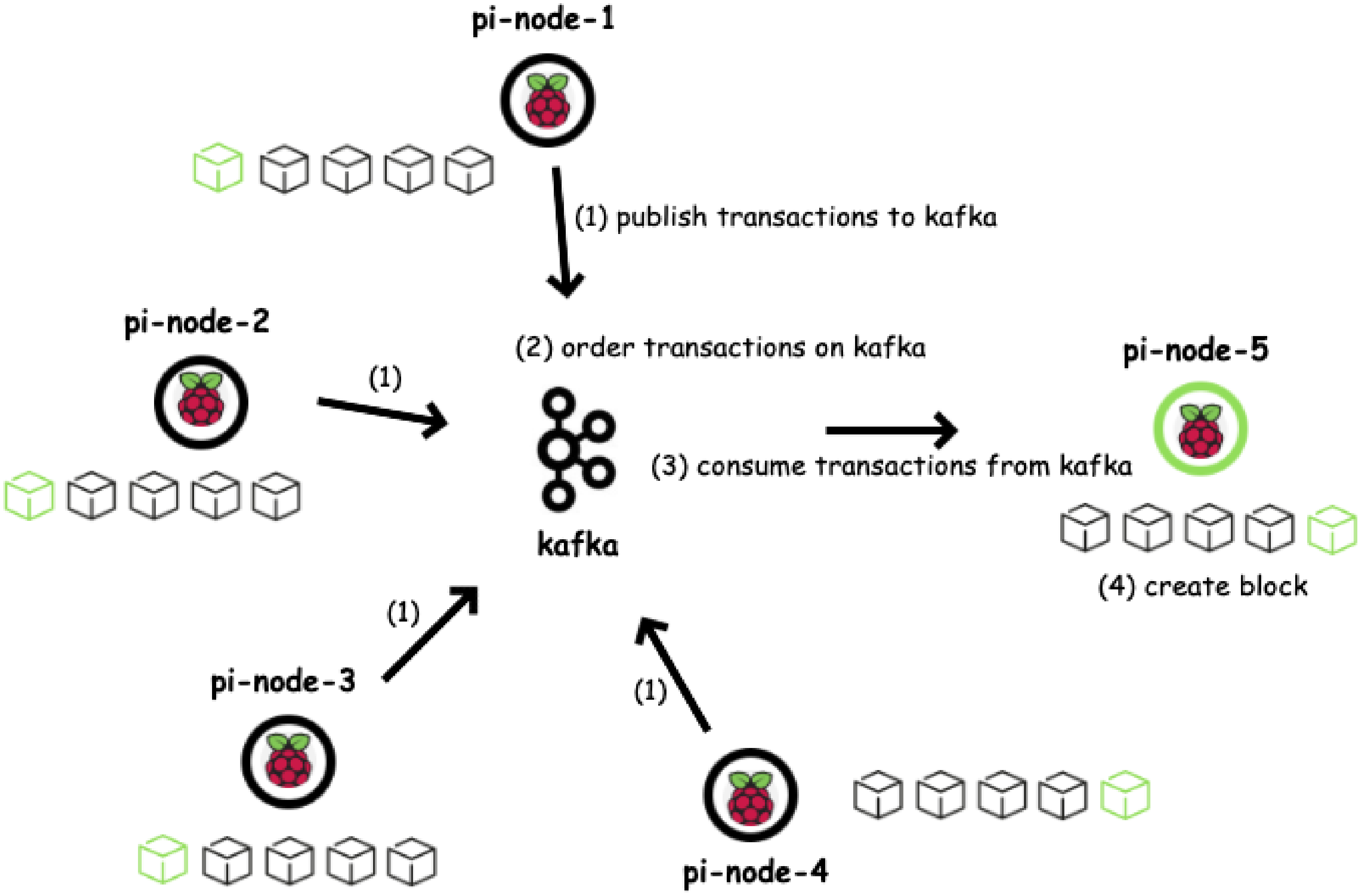
Tikiri Architecture

- Execute transactions and create blocks in parallel by multiple Lokka services in Tikiri blockchain.



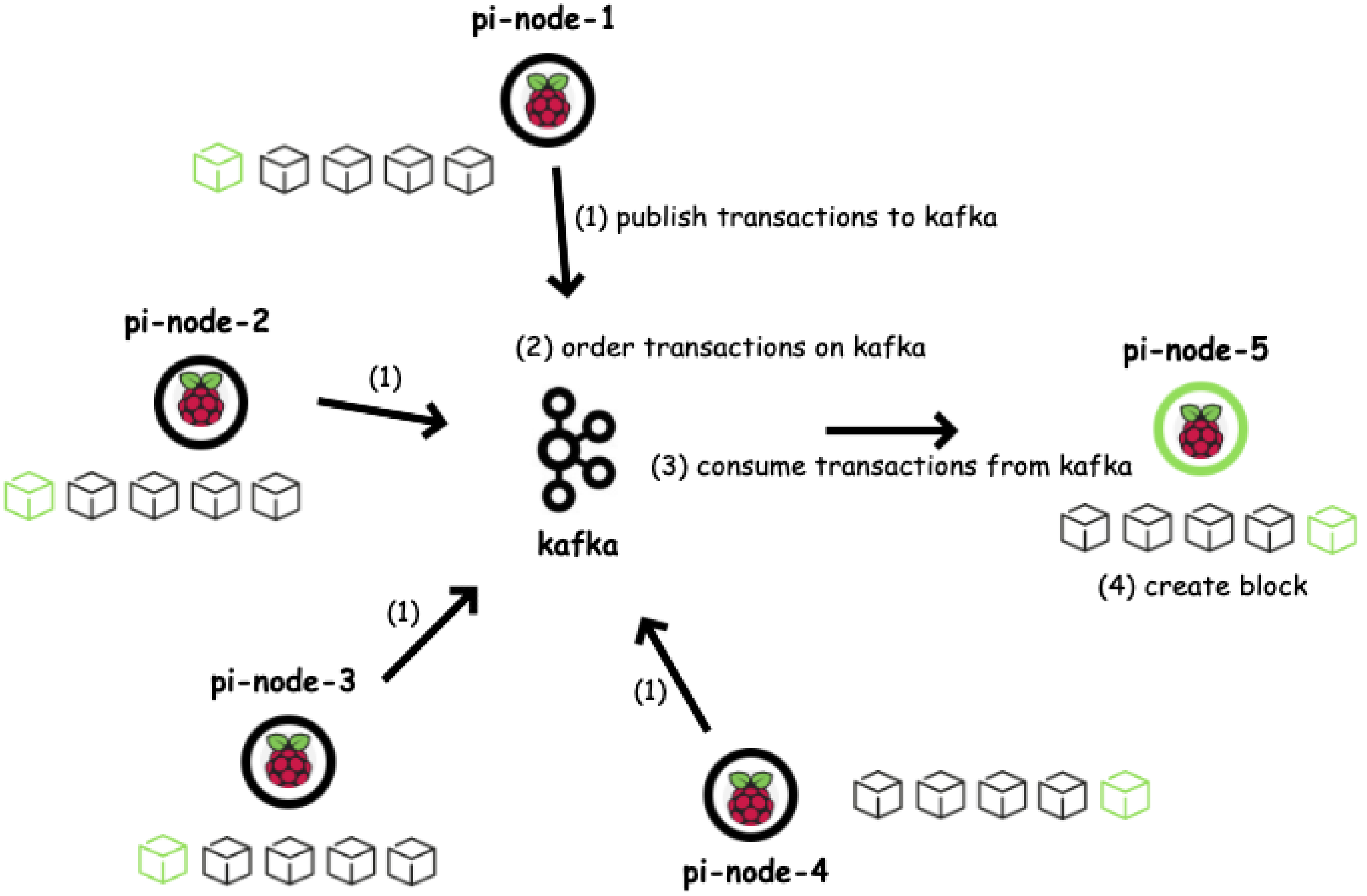
Tikiri Architecture

- Implemented on raspberry-pi cluster
- Apache kafka(zookeeper atomic broadcast protocol) used as the consensus mechanism
- Functional programming based concurrent smart contract platform

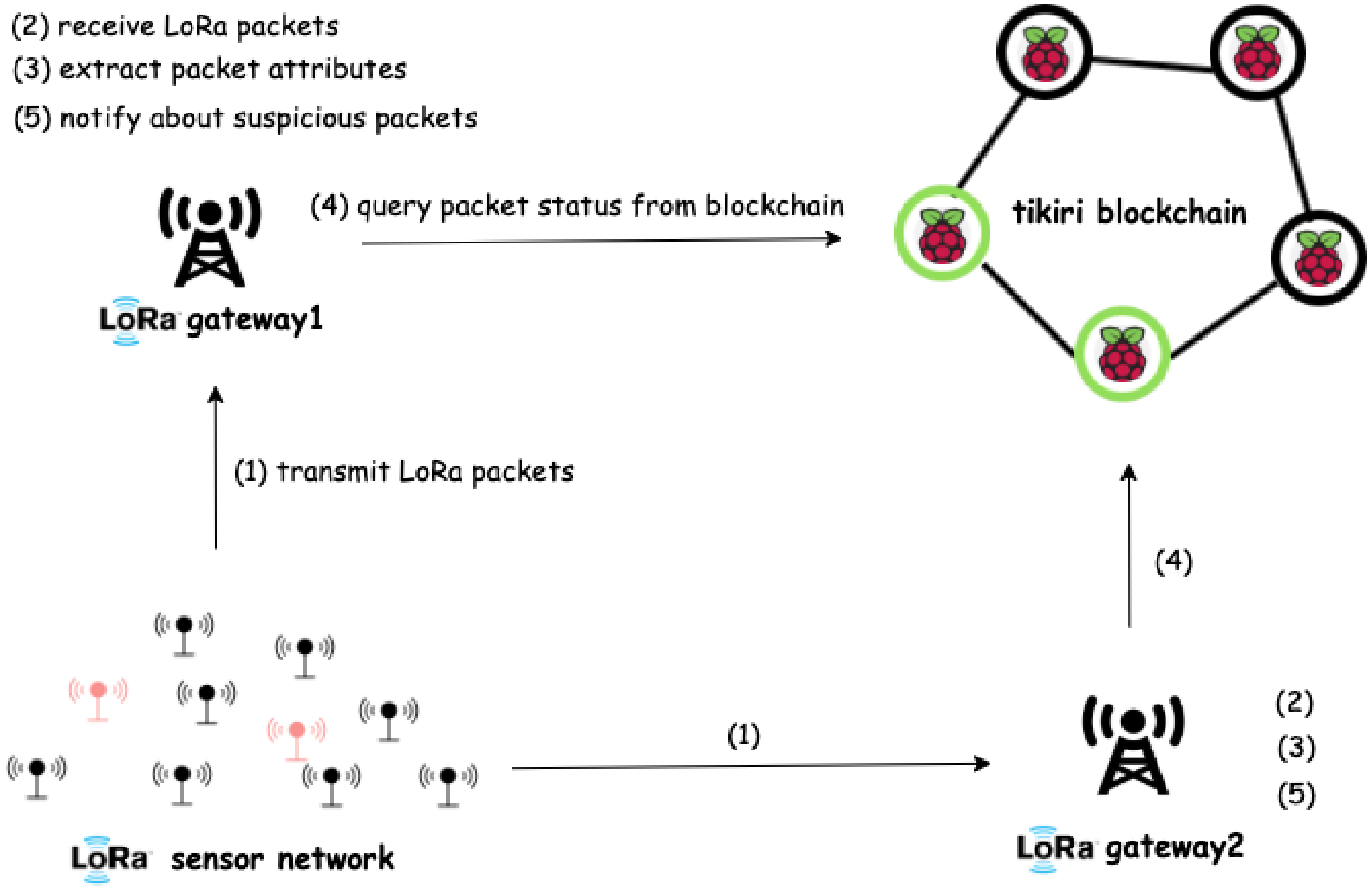


Tikiri Architecture

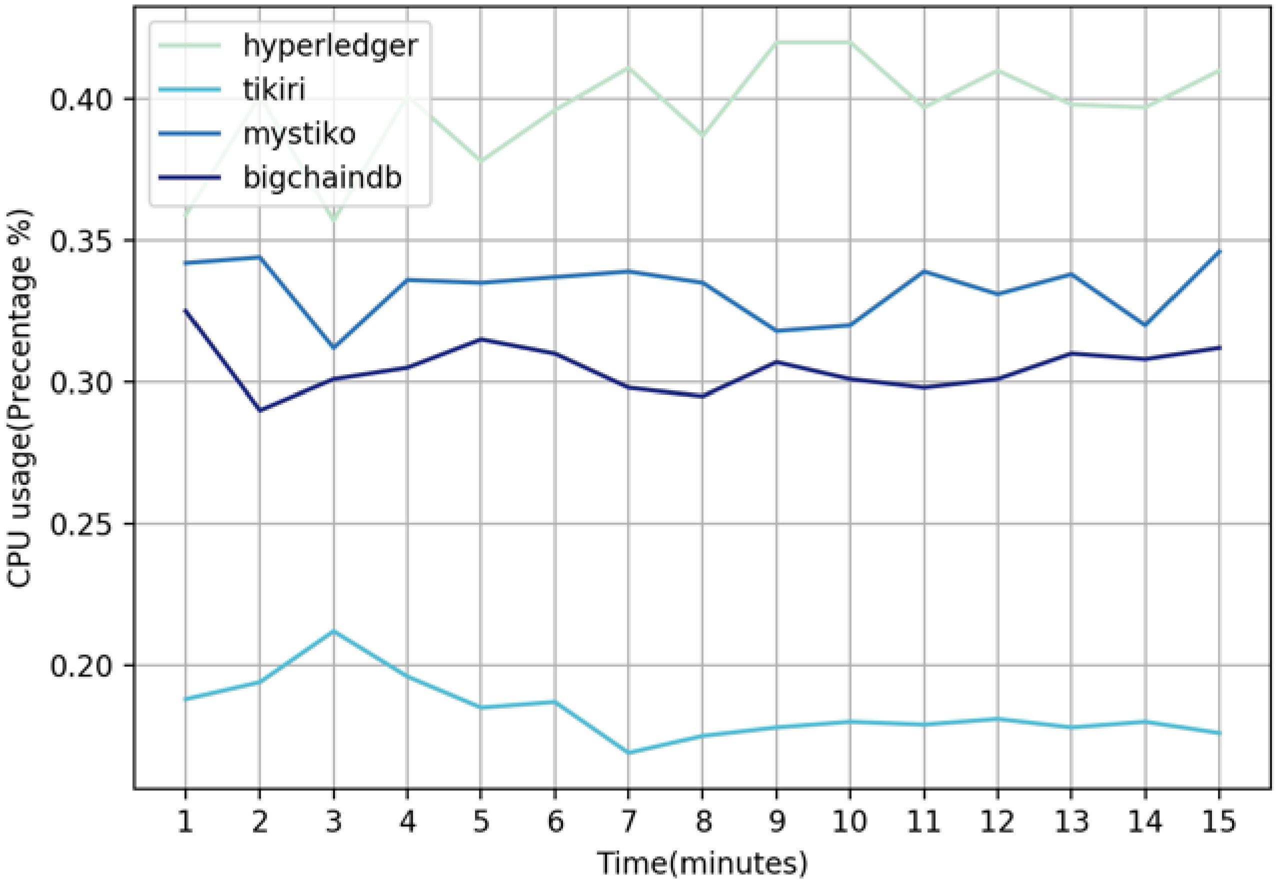
- Store power plant sensor device identities in Tikiri blockchain
- Sensor devices communicate with LoRa interface



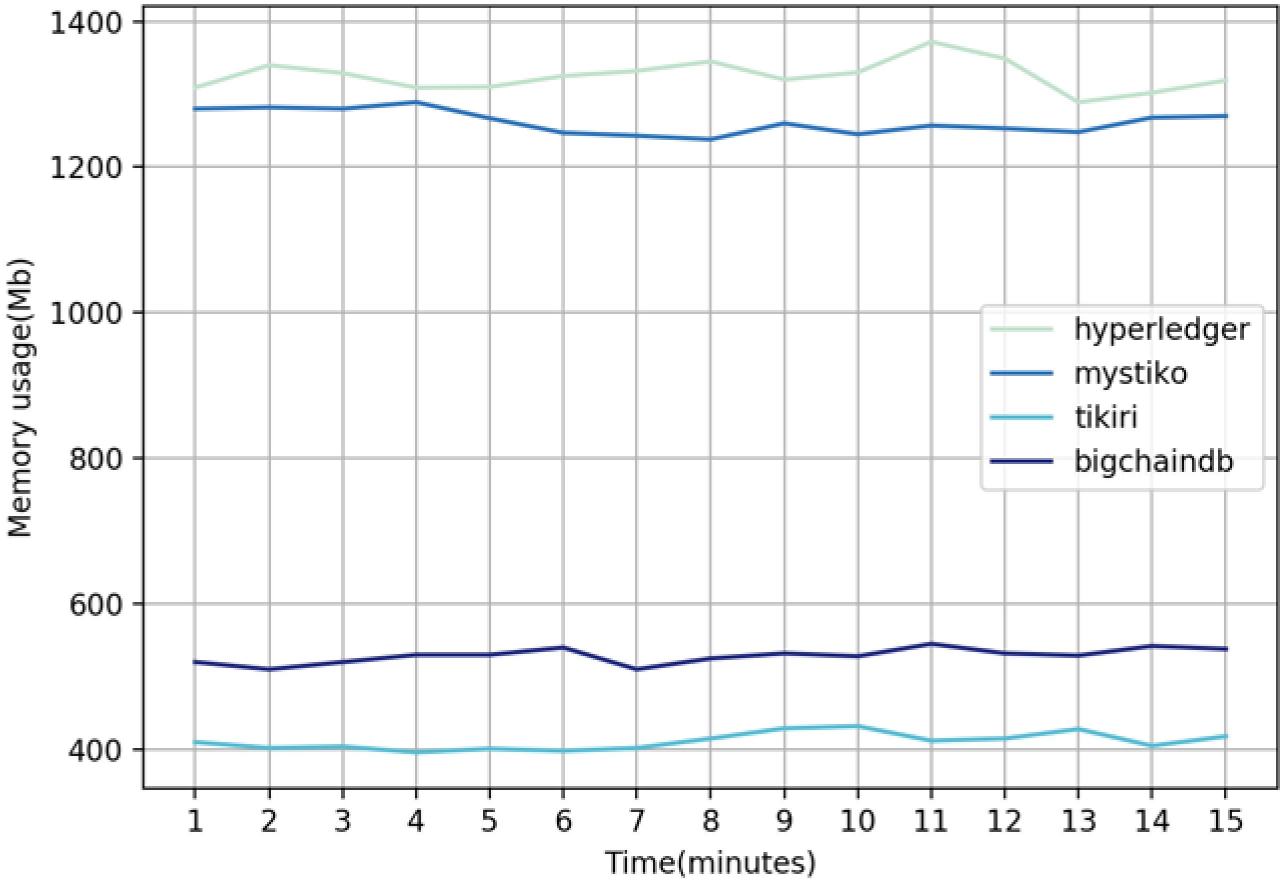
Tikiri Architecture – Information Flow



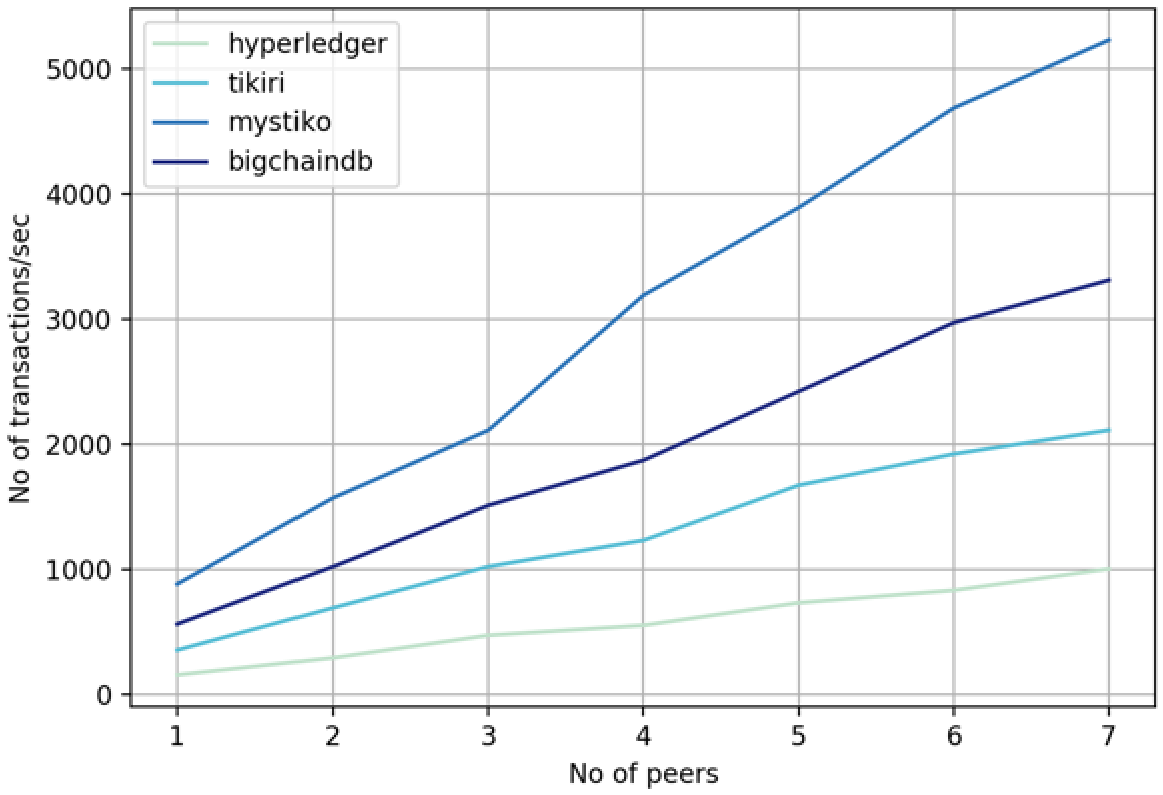
Tikiri Results – CPU Overhead

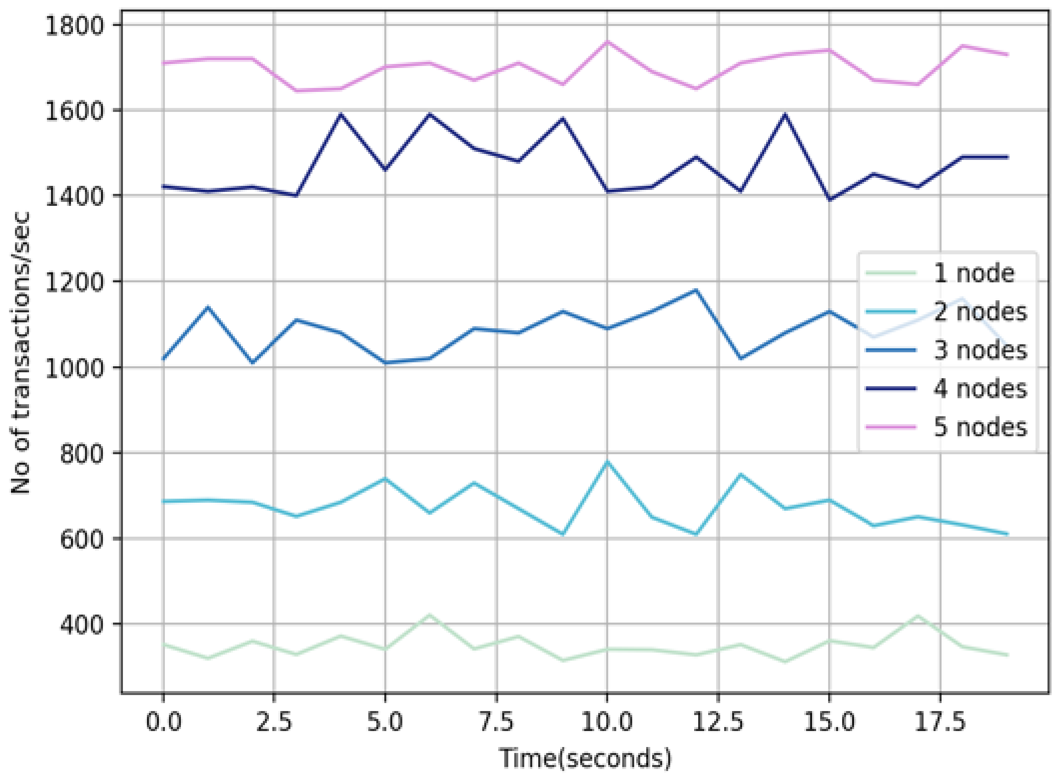


Tikiri Results – Memory Overhead

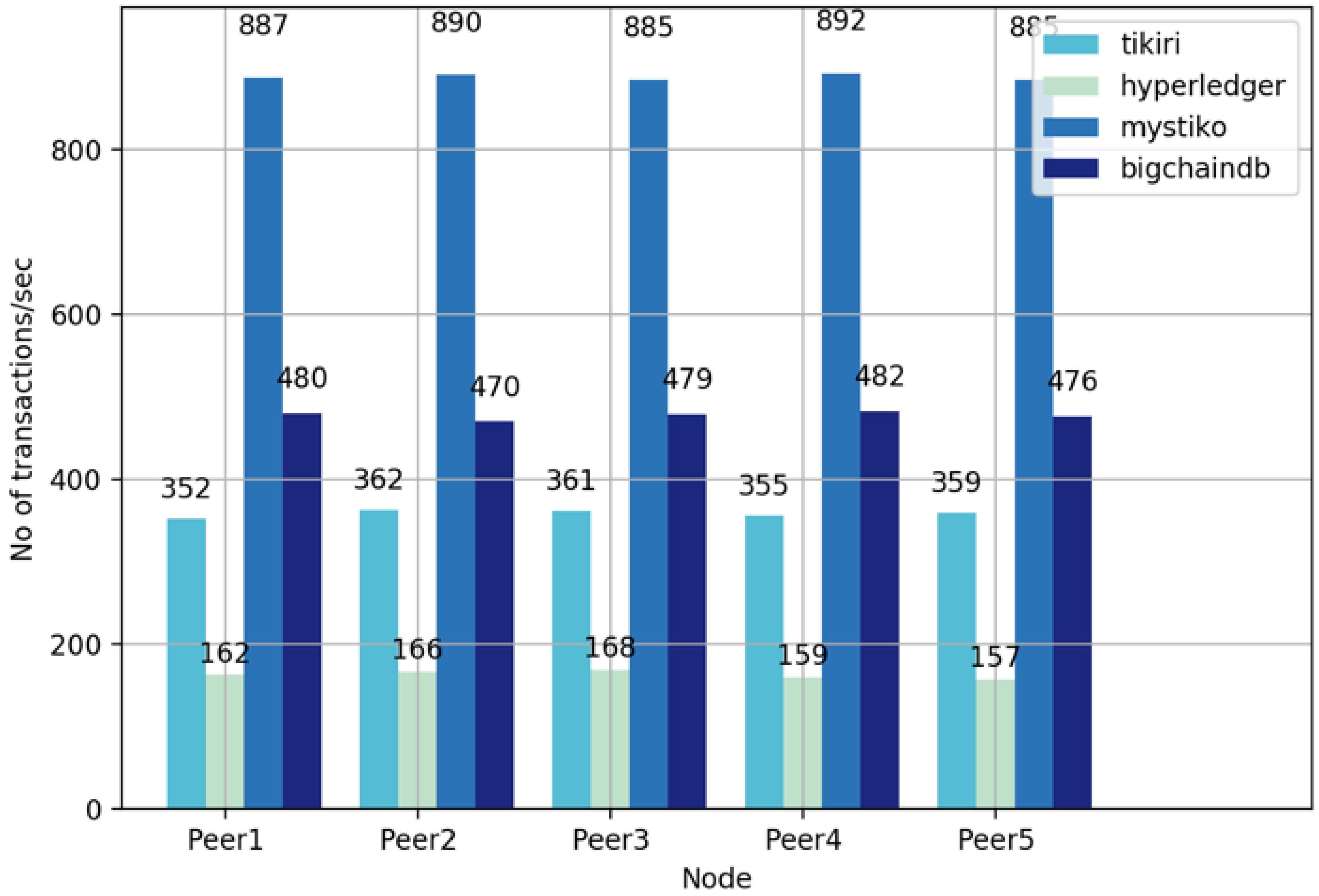


Tikiri Results – Scalability





Tikiri Results – Transaction Throughput



LoRaWAN Testbed – Sensor network traffic

- Represent Industrial IoT Sensor environment
- Provide realistic and reliable generation of end node data
- Transmit end node information to downstream LoRaWAN infrastructure
- Provide oracle services to the Tikiri blockchain

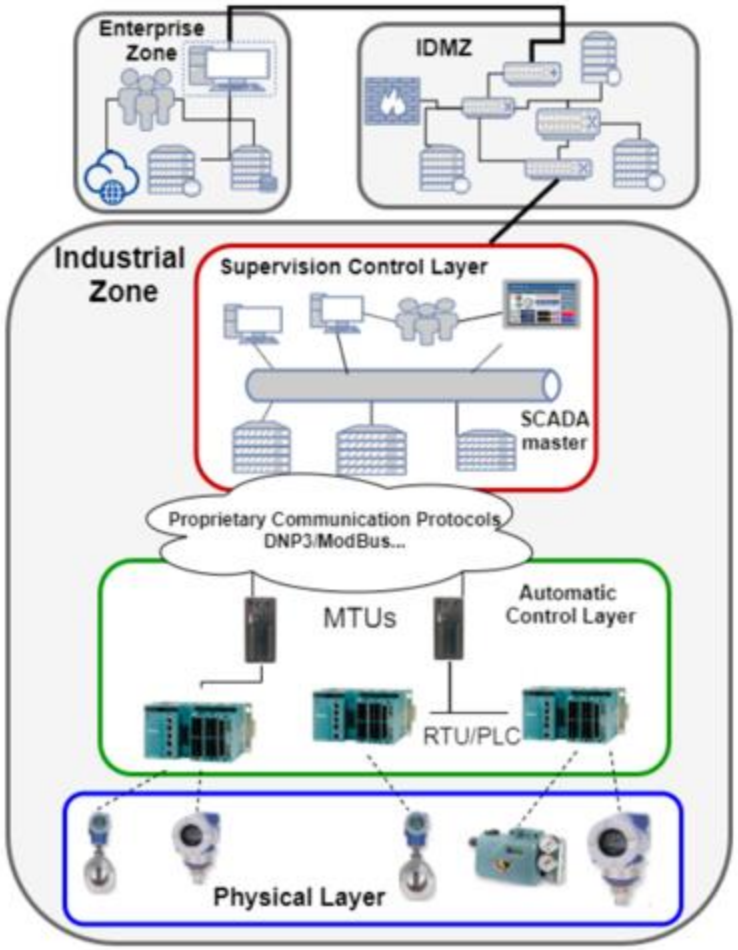
LoRaWAN Testbed

- Current Progress
  - Single Arduino Uno acts as end node and generates/transmits LoRaWAN traffic
  - Single Dragino LG01-P Gateway forwards LoRaWAN traffic to The Things Network (TTN)
  - Gateway packages LoRaWAN traffic and sends to Kafka broker

- Future Work
  - Integration with Tikiri
  - Add additional end nodes & gateways
  - Perform physical layout testing
  - Gather performance metrics
  - Future enhancements (ex: support additional Tikiri communication protocols)

Sensor Data Authentication using PUFs

- Industrial Control Systems (ICS) are integral components of national critical infrastructures
  - Example: Power plants, Water and gas distribution centers, transportation
- Commonly monitored by Supervisory Control and Data Acquisition (SCADA) systems



- Integration of advanced sensors in power plants introduces security challenges:
  - How to ensure authenticity of sensor data?
  - Can the received data be trusted?
  - What lightweight mechanism can verify device identities in such Cyber-Physical Systems?

Sensor Data Authentication using PUFs

NATIONAL ENERGY TECHNOLOGY LABORATORY

Problem Statement:

Given resource constrained IoT nodes,

How can we uniquely identify them and perform continuous authentication in order to avoid maliciousness (node & data)?

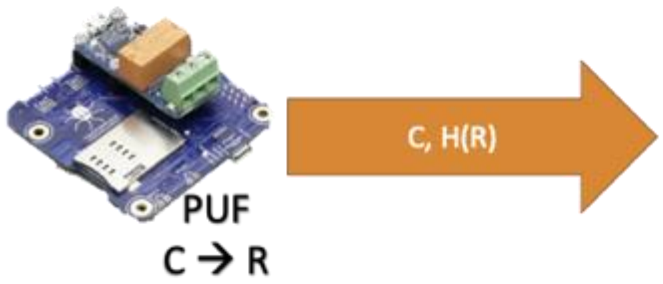
IDEA:

With **lightweight** hardware security primitive called Physical Unclonable Functions (PUF) to act as a **hardware fingerprint generator** and use it to dynamically authenticate sensor data

Physical Unclonable Functions (**PUFs**)

A hardware security primitive that maps challenges and responses

$$\gamma: \{0,1\}^n \rightarrow \{0,1\}^m$$



Node	Challenge	Response
Node 1	C1	R3
Node 2	C2	R2
...	...	...
Node N	CN	RN

Authentication server

Physical Unclonable Function (PUF)

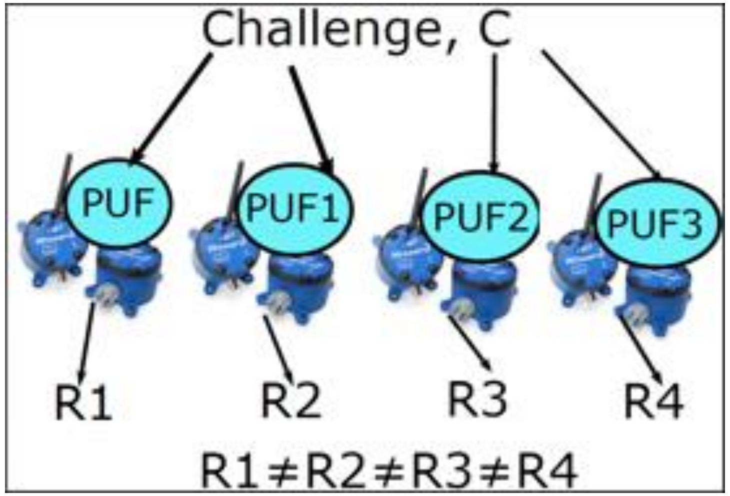
- Challenge-Response
- Low-cost fingerprint generator
- Generate unique identities for all devices
- Offload complex state-of-art crypto solutions
- Different types such as SRAM-based
  - High availability and performance [2]

Offloads the complexity of managing and storing keys on the IoT devices

Proposed Approach

- We exploit the fundamental property of embedded sensors to generate unique identities through PUFs and derive hash-key functions

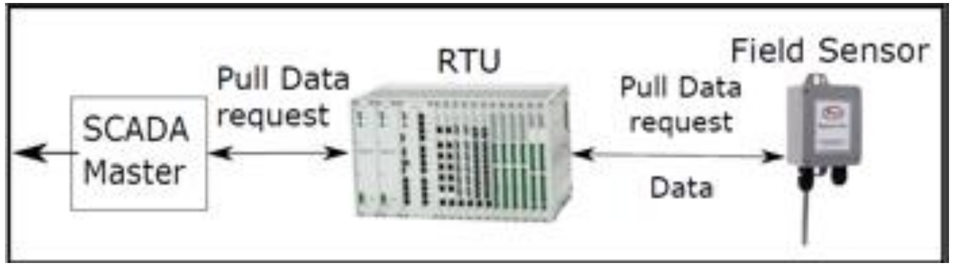
- Design a lightweight SRAM-based PUF Authentication and Integrity (**SPAI**) protocol
  - Ensures the integrity of data flow from field sensors
  - Eliminates rogue devices from the SCADA architecture



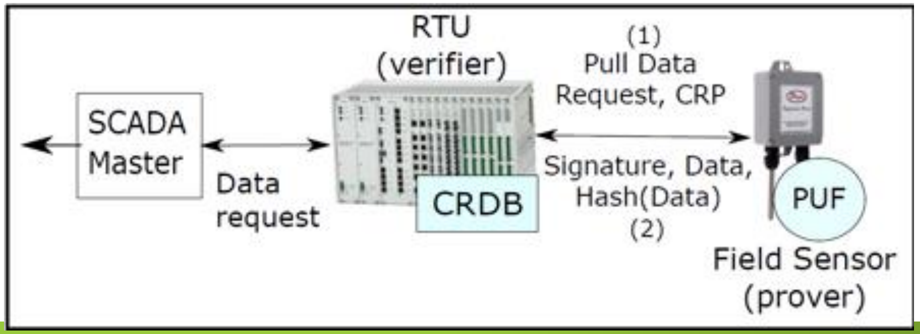
System Overview

NATIONAL ENERGY TECHNOLOGY LABORATORY

- State-of-art SCADA communication
  - RTU sends a pull request to a field sensor
  - The field sensor read the request and sends the environmental data
  - Operational and commands are sent in clear text without security



- SPAI Protocol Overview
  - RTU sends a pull request to a field sensor
  - It appends a challenge-response pair from a secure Challenge-Response Database (CRDB)
  - The field sensor (prover) uses the CRP to generate a unique response through the PUF
  - The prover appends the unique response while sending the sensor data
  - The verifier validates the identity of the sensor and integrity of data



SPAI Protocol

- **Three-phase protocol**
  - Profiling
  - Enrollment
  - Authentication

1. Profiling
   - Error Correction Code, necessary for PUFs introduce a significant overhead
   - We identify strongest cells in the SRAM through a Data Remanence Algorithm [4]

2. Enrollment: Generate and store CRP
   - The verifier gets a challenge and appends it to the CRDB
   - Once a challenge is generated, the verifier sends it through a secure channels to the prover
   - Using the addresses in the challenge, the prover reads the bit value [0, 1] that is stored in the SRAM, creating a response
   - The prover sends the response to the verifier
   - Finally, the verifier appends the response to the CRDB

SPAI Protocol: [3. Authentication]

- Ensures the authenticity of request and data flow integrity
- Communication during this phase can be over unsecured channels

Algorithm 1: setup

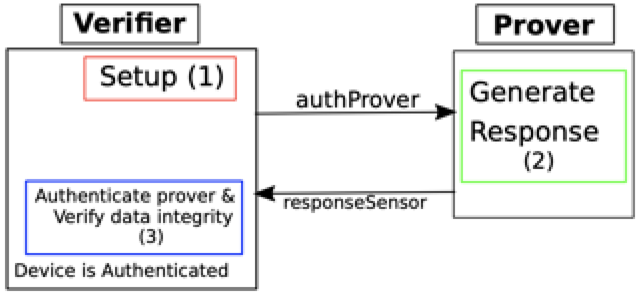
Require: Access to CRDB

$R_1, C_1 = getTuple(prover)$
$TS = time()$ `// keep track of active requests`
$H_1 = HMAC(R_1, C_1||TS)$
$TS' = TS \oplus hash(R_1)$
$respSensor = authProver(C_1, TS', H_1)$

Algorithm 2: generate response

Require: Access to PUF mechanism

$C_1, H_1, TS' = readRequest()$
$R_1 = getPUF(C_1)$
$TS = TS' \oplus hash(R_1)$
$H_{11} = HMAC(R_1, C_1||TS)$
if $H_{11}! = H_1$ **then** `// verify prover identity`
  exit
end if
$data = readSensor()$
$H_2 = HMAC(R_1, data)$ `// ensure data integrity`
$responseSensor =< data, H_2 >$
return $responseSensor$

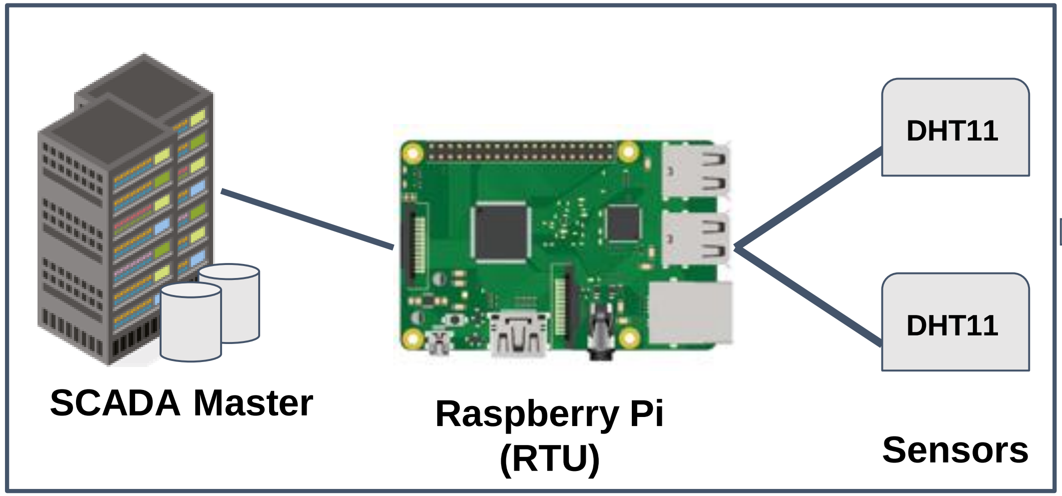


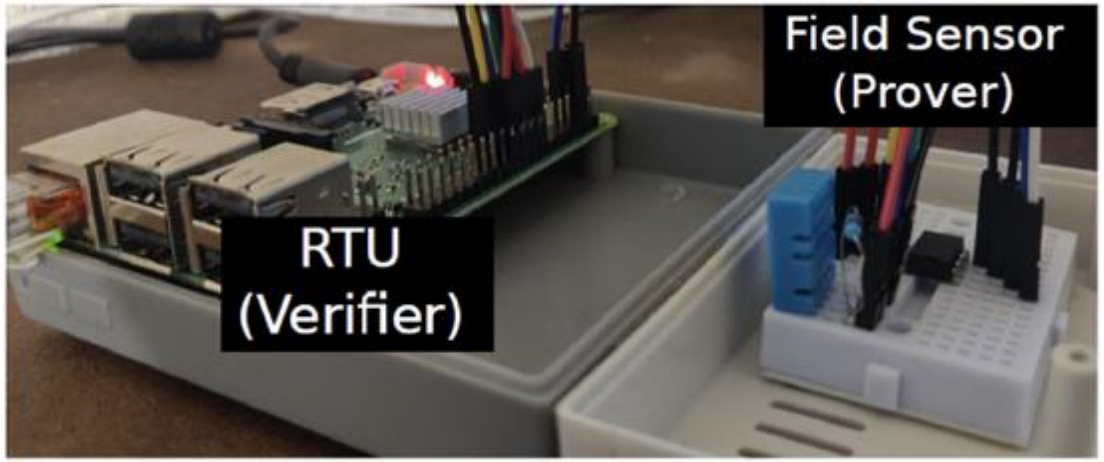
Algorithm 3: authenticate prover and verify data integrity

$data, H_2 = unpack(respSensor)$
$t_2 = time()$ `// keep track of active requests`
$H_{22} = HMAC(R_1, data)$
if $t_2 - TS > \Delta t$ **then** `// prevent brute force modeling attacks`
  exit
else if $H_2! = H_{22}$ **then** `// verify integrity of data and verifier identity`
  exit
end if
Device is Authenticated
$sendToSCADA(data)$

Testbed setup

- Emulated an RTU functions, in a Raspberry Pi 3 model B.
- Raspberry pi is connected to the external SRAM microchip 23LC1024, with a capacity of 128k bits.





- Evaluated the overhead of the SPAI protocol in a temperature and humidity sensor DHT11, the sensor reads and sends the data in every two seconds

Evaluation

Attack scenario

- Man-in-the-Middle attack
- An attacker has full access to the communication link and message

```
pi@raspberrypi:~/Desktop/sram-puf/Authentication $ ./master
Hash of response: 22093202986069e29585db37a3aa95086868151ab756dd8b5cd16d4740ed10df
Timestamp: 1588793221
Master hmac: ae75172e1c21c5615f7447041471d330f9b5eff427ab39f6c7f703587abe7477
Reading sensor...
Temperature read from sensor: 999
Failed to verified data integrity
```

Fig. 4. Unsuccessful Authentication of Field Sensor

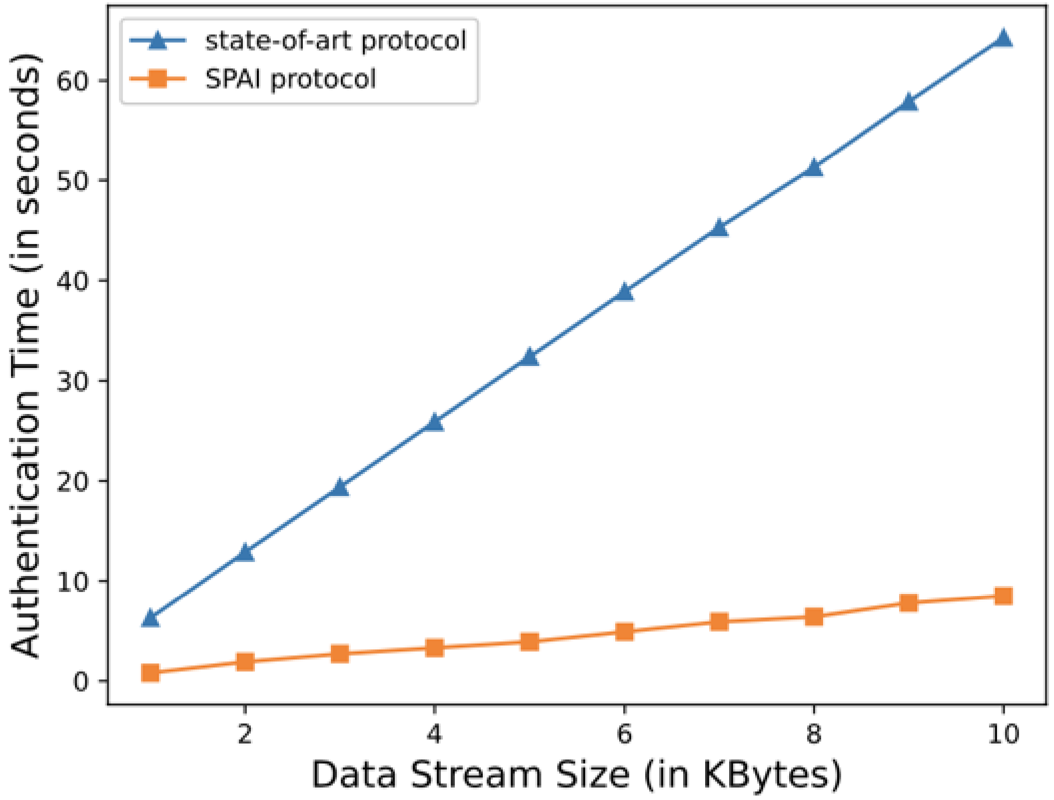
```
pi@raspberrypi:~/Desktop/sram-puf/Authentication $ ./master
Hash of response: 22093202986069e29585db37a3aa95086868151ab756dd8b5cd16d4740ed10df
Timestamp: 1588793262
Master hmac: ae75172e1c21c5615f7447041471d330f9b5eff427ab39f6c7f703587abe7477
Reading sensor...
Temperature read from sensor: 25
Data Integrity verified..
```

Fig. 5. Successful Authentication of Field Sensor

Evaluation

Performance analysis of the SPAI:

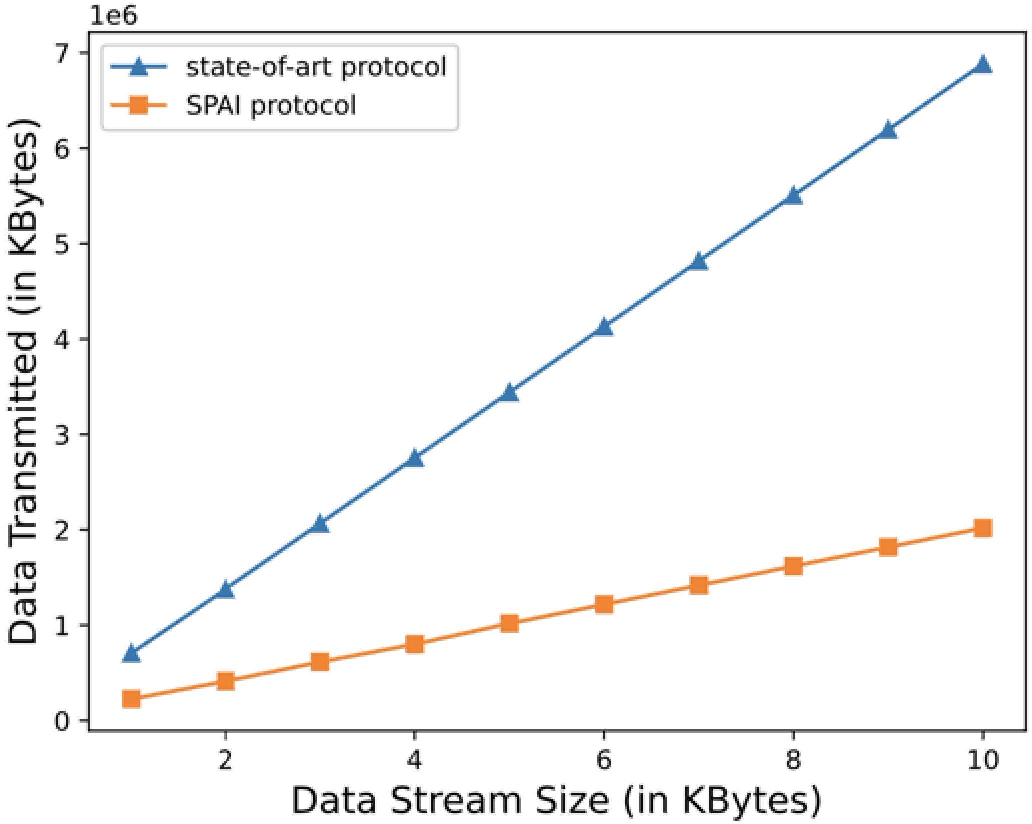
Time to complete authentication process for continuous data stream



Evaluation

Performance analysis of the SPAI:

Total data communicated over modbus protocol for continuous data stream



Evaluation

- Comparison of performance of different protocols when unit byte of data is communicated and authenticated.

Protocol	**Transmission Time**	**Transmitted Data**
Raw Modbus	20ms	857 Bytes
SPAI Protocol	**218ms**	**2803 Bytes**
State-of-art PUF protocol [9]	379ms	5483 Bytes

Benefits and Other Research Areas

- Scalable data and process integrity assurance in FPP would help plant managers to better maintain the components
    - Reduce operational cost over long-run

- Establishment of overlay Blockchain for SCADA environment can also be applicable for achieving **access control and accountability**
    - Large and multi-site energy companies have many independent contractors, whose access to the infrastructure must be vetted

- **Supply-chain provenance** in energy delivery systems is critical and the proposed platform has potential to enable this service

Preparing Project for Next Steps

Market Benefits/Assessment

- The project addresses the need for an infrastructure based identity management and provenance solution that can provide early detection of rogue devices.
- The proposed technology would realize a low cost security solution that would provide protection to large number of sensors in the power plant and lead to cost savings

Technology-to-Market Path

- The Blockchain platform will be integrated into state-of-practice security monitoring solutions
- Ensuring the ability to provide desired benefits at lower cost
- Integration with AI solutions to also provide trusted source of ground truth
- Collaborating with Accenture, ReliabilityFirst, WoodPLC

Concluding Remarks

- The technology developed by the project will address the following specific challenges in fossil energy
  - Identity management and provenance that would enhance the infrastructure cybersecurity.
  - Increasing system reliability due to early detection of attacks
  - Optimize utility efficiency by identifying and isolating faults
  - Enhanced security of monitoring technology by improving resilience to cyber attacks

Concluding Remarks

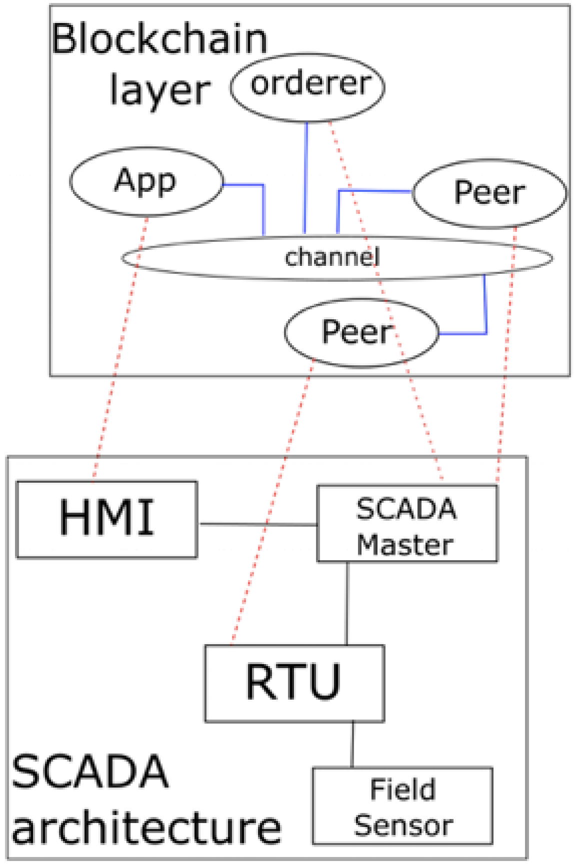
- We demonstrated that our proposed SPAI protocol is capable of preventing MITM attacks and **detect abnormality in data flow**
- The SPAI protocol **ensures the integrity of sensor data** through HMAC functions and unique PUF-responses. It implements **lightweight crypto** solutions required to improve the security of SCADA systems
- Through preliminary experiments, we found that our proposed SPAI protocol serves its purpose (data authentication and integrity) at a **minimal overhead** compared to existing PUF-based models

Concluding Remarks and Future Work

- Leverage Tikiri and Hyperledger Fabric Blockchain to store PUF-based sensor identity profiles and authenticate sensor data
  - Performance evaluation for bulk transactions in FPP

- Analyze the overhead induced by Blockchain in terms of storage and time to verify data integrity and improve resilience of SPAI protocol

- Transitioning SPAI to emulated SCADA Network with several OpenPLC controllers interfacing with multiple sensors and actuators



Concluding Remarks and Future Work

- Support MQTT based communication between sensor network and blockchain network
- Integration with distributed database (Apache Cassandra) to store transaction data
- Integrate off-chain storage to guarantee privacy of the data and address storage issues in the blockchain nodes
- Development of sharding-based consensus that handles the network partitions and failures handling scenarios where the voting nodes that are not reachable