UC San Diego
SCHOOL OF MEDICINE

Department of
BioMedical Informatics

Jeffrey Tellew • Dr. Lucila Ohno-Machado • Dr. Tsung-Ting Kuo

Contract-based On-chain Training Certificate Logging

This research was supported by grant T15LM011271

About Me

- 3rd-year student at UC Santa Barbara
- Computer Science major
- Originally from San Diego
- I like surfing, snowboarding, playing guitar, eating California burritos, and traveling

This research was supported by grant T15LM011271

Presentation Outline

1. Why training certificates?
2. Blockchain background
3. Project overview
4. Web app
5. Experiment results
6. Conclusion

This research was supported by grant T15LM011271

Why Training Certificates?

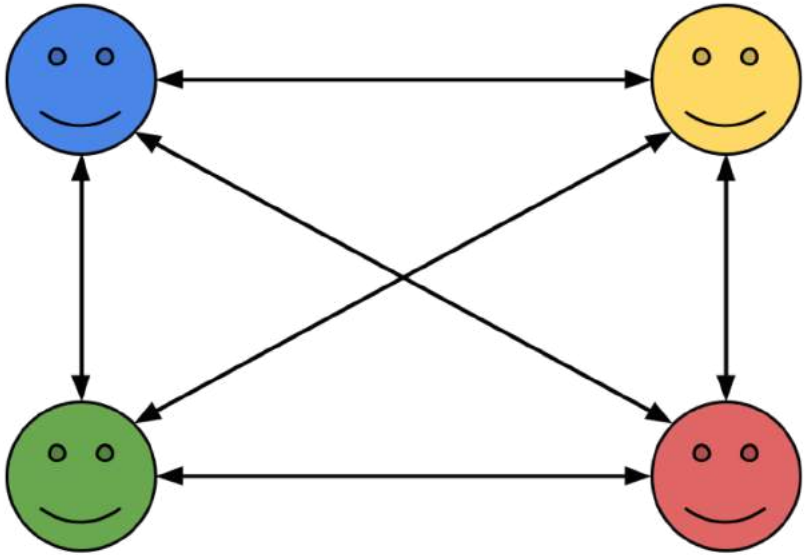
- Having a valid training certificate is essential to many different aspects of research
  - Access to data sets
  - Training in general
  - Grant funding
- Managing certificates can help avoid problems
  - Delaying grant funding
  - Unnecessary slowdowns in research and subsequent publications

Source: *citiprogram.org*

This research was supported by grant T15LM011271

The Current System: Researchers email certificates

- Slow and inefficient
- Prone to human error
  - Certificates could be mailed to the wrong person
  - Certificates might be expired



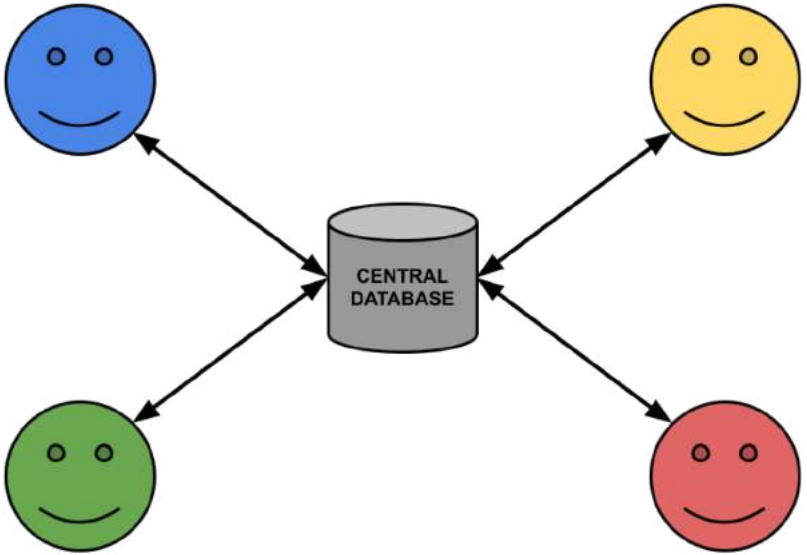
With the current system, each researcher has to email their certificates to whoever needs it

This research was supported by grant T15LM011271

A Potential Solution:

Researchers use a central database

- Vulnerable to attack
  - Entire system can be compromised with a single point of failure
- Difficult to get started

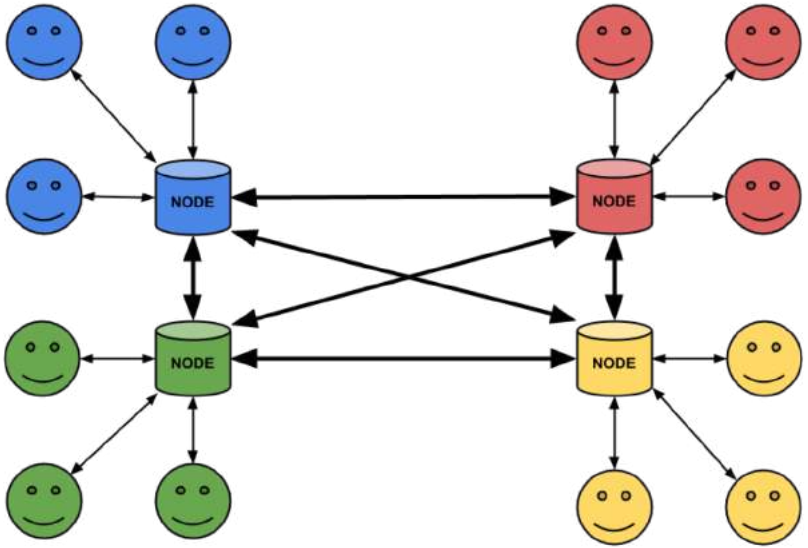


In this potential solution, researchers would use a central database to store their certificates

This research was supported by grant T15LM011271

A Better Solution: Use a decentralized system instead

- No single point of failure
- Maintenance costs do not fall on one entity

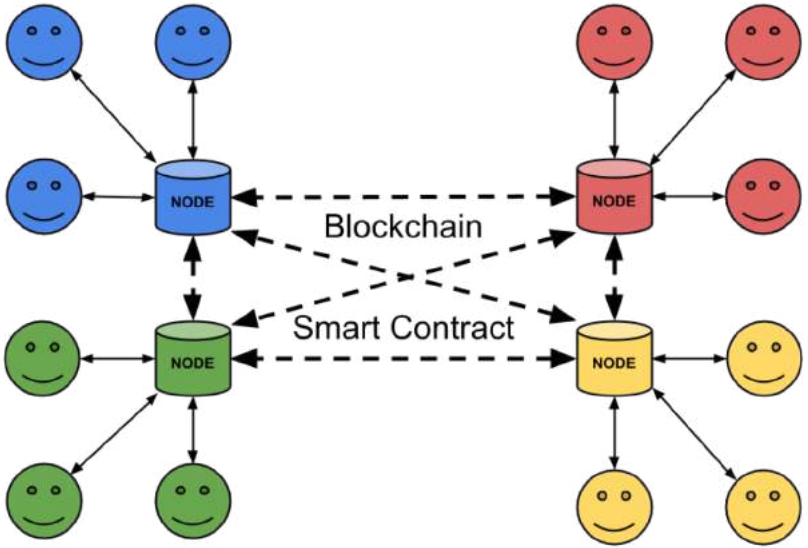


In this potential solution, researchers would use a central database to store their certificates

This research was supported by grant T15LM011271

Our Proposal: Use a blockchain system

- Blockchain is an immutable ledger that allows us to store the certificate data and share it
- Smart contracts are programs that allow you to save and query data



In this potential solution, researchers would use a central database to store their certificates

This research was supported by grant T15LM011271

What are Blockchain and Smart Contracts?

"Picture a spreadsheet that is duplicated thousands of times across a network of computers. Then imagine that this network is designed to regularly update this spreadsheet…"

- blockgeeks.com

Source: *blockgeeks.com/guides/what-is-blockchain-technology*

- **Blockchain** is a immutable, decentralized, ledger system
- In other words:
  - It cannot be changed
  - It is distributed throughout a network
  - It records transactions of some type
- **Smart Contracts** are contracts written in code that are distributed throughout the network

This research was supported by grant T15LM011271

Comparison: Centralized vs Blockchain

Metric	Centralized	Blockchain
Single Point of Failure	Yes	No
Setup Difficulty	Medium - Hard	Medium - Very hard
Maintained by Single Site	Yes	No
Speed	Very fast	Medium
Scalability	Good	Unclear

This research was supported by grant T15LM011271

Comparison: Centralized vs Blockchain

Metric	Centralized	Blockchain
Single Point of Failure	Yes	No
Setup Difficulty	Medium - Hard	Medium - Very hard
Maintained by Single Site	Yes	No
Speed	Very fast	Medium
Scalability	Good	Unclear

This research was supported by grant T15LM011271

Project Overview

Goals

- Design a training certificate management system based on blockchain technology
- Store the actual PDFs on the chain
- Test the scalability of the proposal

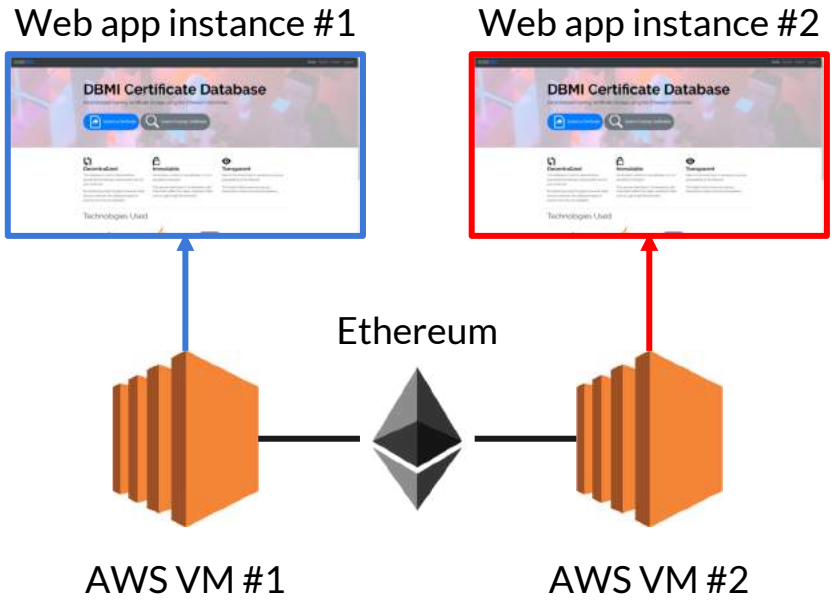
Technology

- Remix, go-ethereum (geth), web3j, Spring Boot, Bootstrap
- Solidity, Java, Bash, HTML, CSS, JavaScript, R

This research was supported by grant T15LM011271

System Architecture

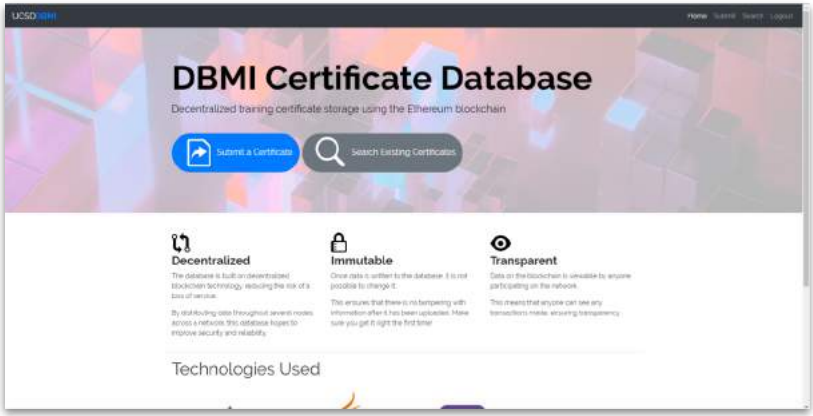
- Two Ubuntu Virtual machines on UCSD AWS
- Machines are connected via a private Ethereum blockchain
- Each machine hosts an instance of the web app



This research was supported by grant T15LM011271

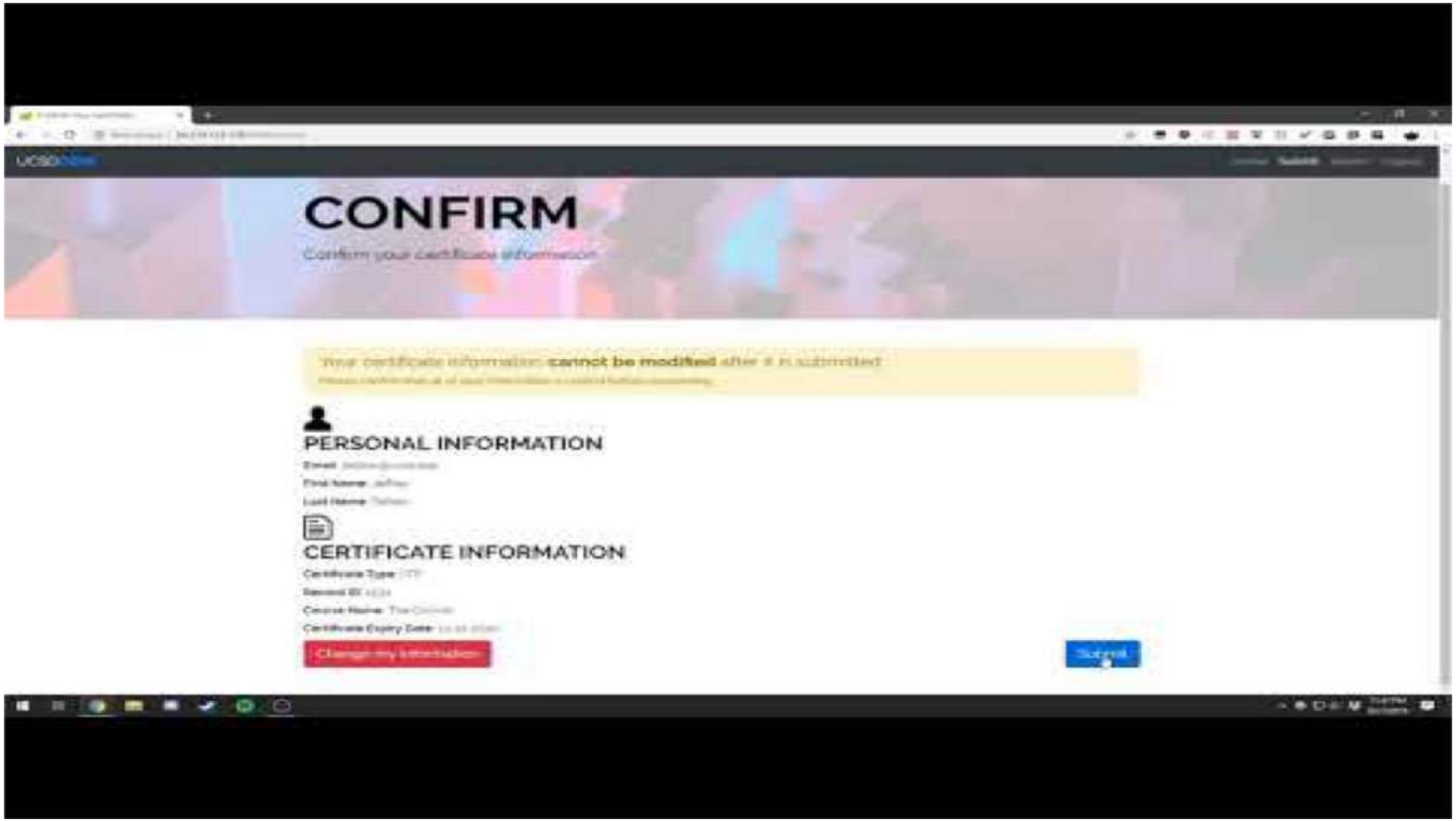
Web App:
General Overview

- Created a web app to make it easy to use the system
- Features:
  - Basic login
  - Submitting new certificates
  - Searching existing certificates



The main menu of the web app

This research was supported by grant T15LM011271



A short video demo of the web app

This research was supported by grant T15LM011271

Web App: Initial Results

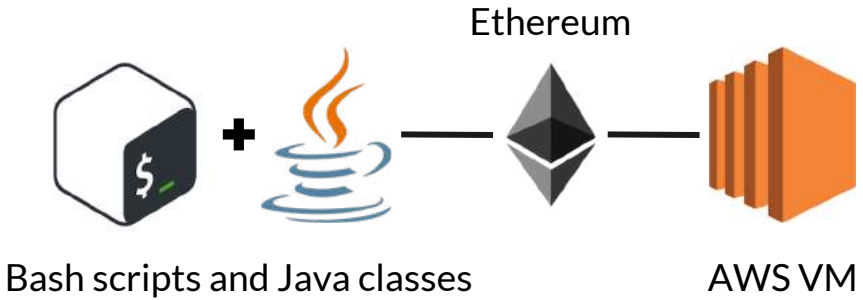
- Collected PDFs from faculty and other interns
- System was able to handle the data
- Will invite more faculty, staff, and students to test the system

	People	CITI	HIPAA	Certificates
Interns	14	13	2	15
Staff	2	2	0	2
Faculty	1	2	1	3
Total	17	17	3	20

This research was supported by grant T15LM011271

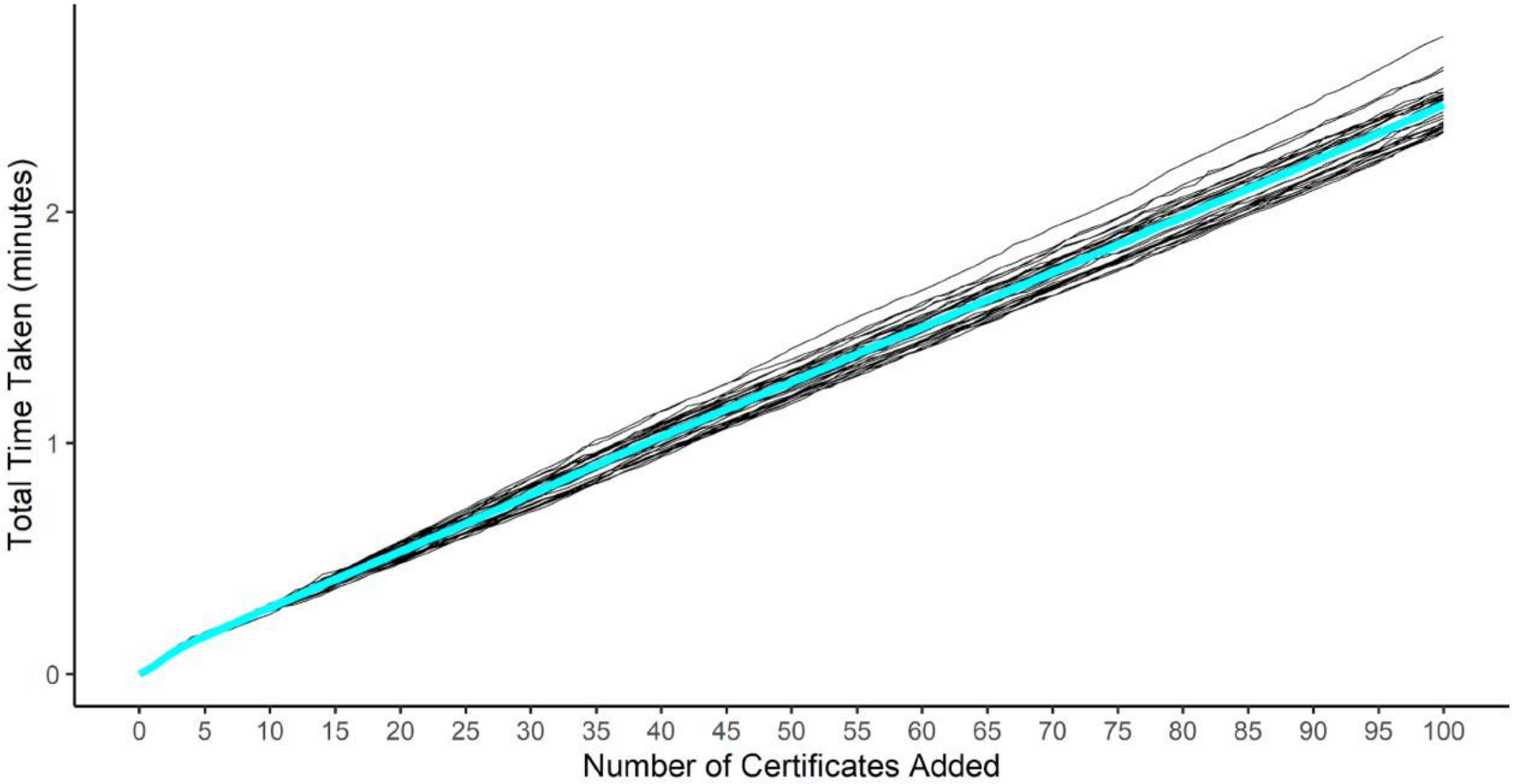
Experiment:
Setup and Methods

- Used script to simulate large volume of submissions from the web app
- Each experiment consisted of 30 trials, where 100 PDFs were added per trial
  - Compute average for the 30 trials
  - System was reset after every trial
- Settings
  - 1 node vs 2 nodes
  - With PDF vs without PDF
  - Strawperson method vs Parallel method



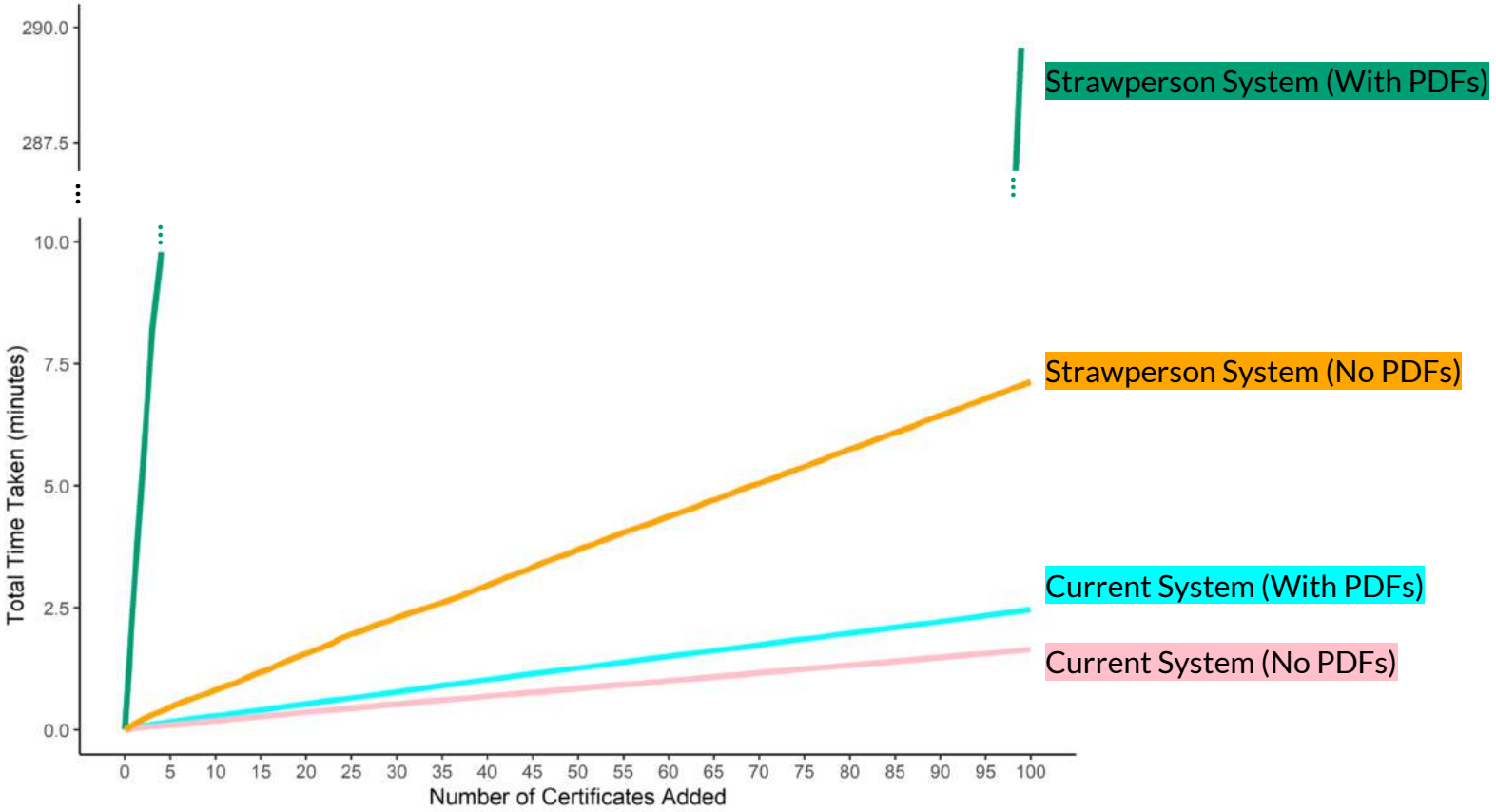
This research was supported by grant T15LM011271

Result: One-node System



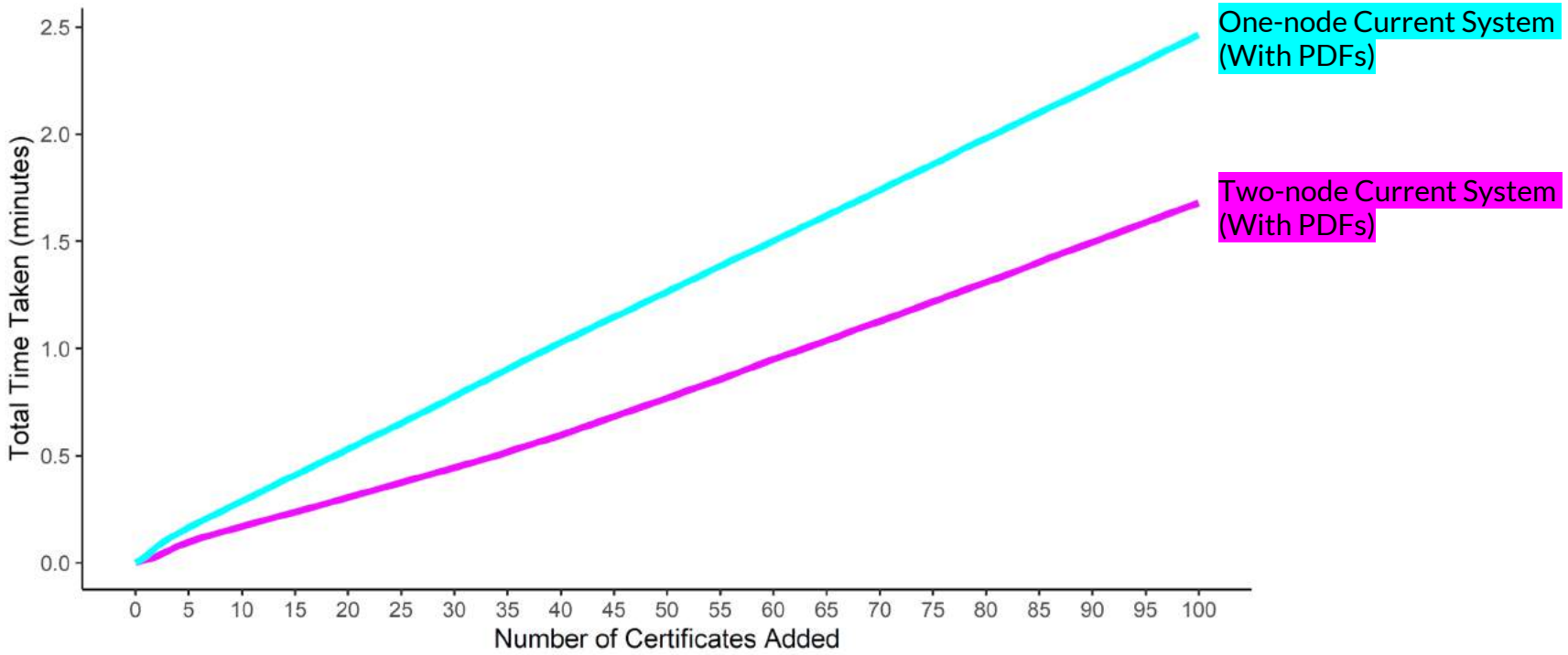
This research was supported by grant T15LM011271

Result: One-node System Comparison



This research was supported by grant T15LM011271

Result: Two-node VS One-node System



This research was supported by grant T15LM011271

Conclusion: Summary

- Blockchain vs Centralized Database
  - Benefits from no single point of failure and is not maintained by a single site
- Feasibility
  - Web app is usable
- Scalability
  - Found to be more scalable than expected (linear growth with latest system version)
  - Without PDF was much faster than with PDF initially, but the gap shrank with the new system
  - Two-node experiment was about twice as fast as one-node equivalent
  - Current system using parallel transactions vastly improved on strawperson system

This research was supported by grant T15LM011271

Conclusion: Next Steps

- Code is available on GitHub so someone else can continue with the project
- Integrate UCSD Active Directory with the web app login
- Fix bugs and add features to web app
- Run more experiments with two or more nodes

This research was supported by grant T15LM011271

Challenges Faced

- Blockchain is new and constantly evolving at a rapid pace
- First time diving deep into web apps
  - Spring/Spring Boot has a steep learning curve
- Experiments ran very slowly at first

This research was supported by grant T15LM011271

Things I Learned

- I learned a **lot** about blockchain technology, how it works, what it's being used for now, and things it may be used for in the future
- I really enjoyed learning about web apps and trying new things with Spring Boot
- Got to use R for the first time to make some graphs from data that I collected

This research was supported by grant T15LM011271

UC San Diego
SCHOOL OF MEDICINE

Department of
BioMedical Informatics

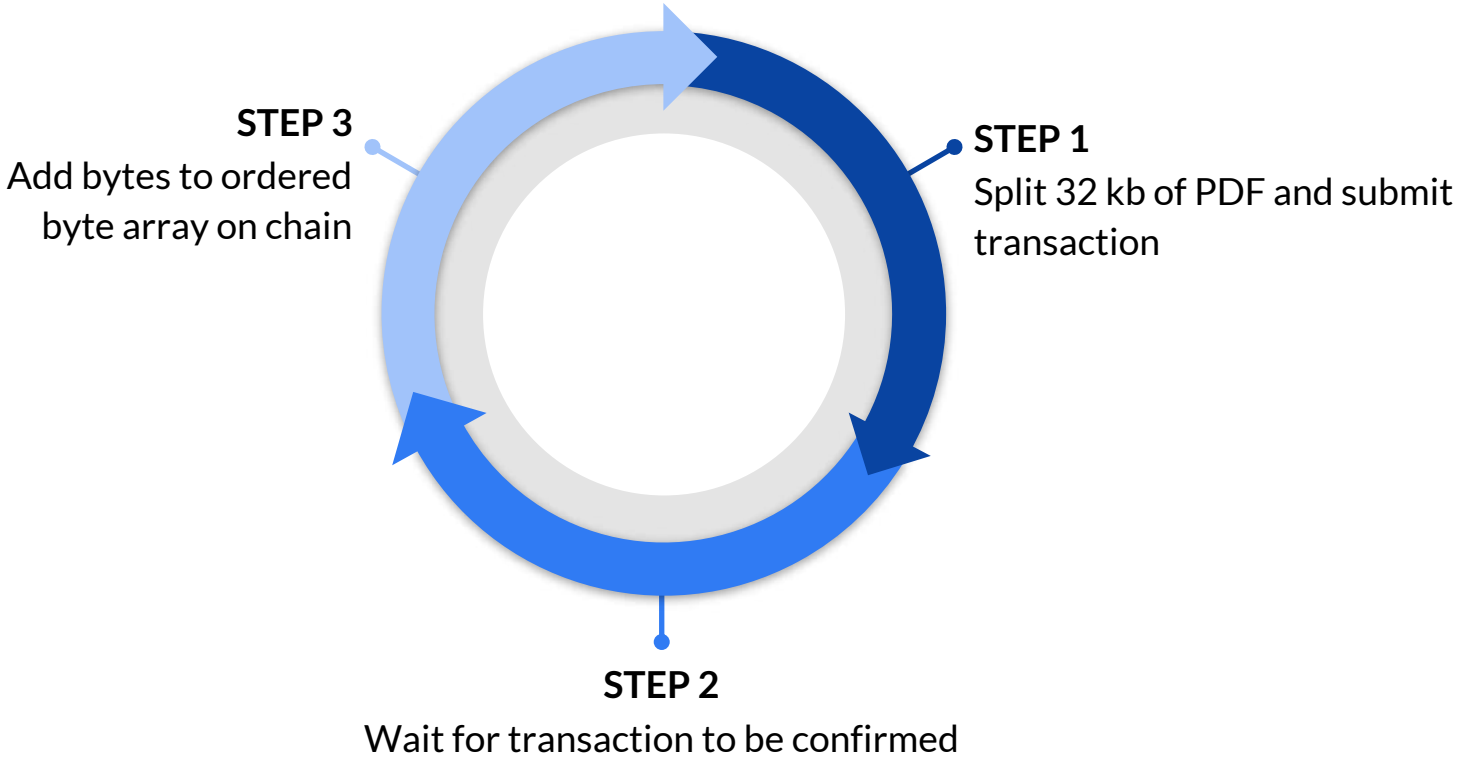
Acknowledgements

- Dr. Ohno-Machado, Dr. Kuo
- Nancy Herbst, Kai Post, and Tyler Bath
- Reid Otsuji and Stephanie Labou
- Dr. Koola and Elizabeth Santillanez

Questions?

This research was supported by grant T15LM011271

PDF Upload: Strawperson System



This research was supported by grant T15LM011271

PDF Upload: Parallel System

Split entire PDF into 32 kb slices labeled by position

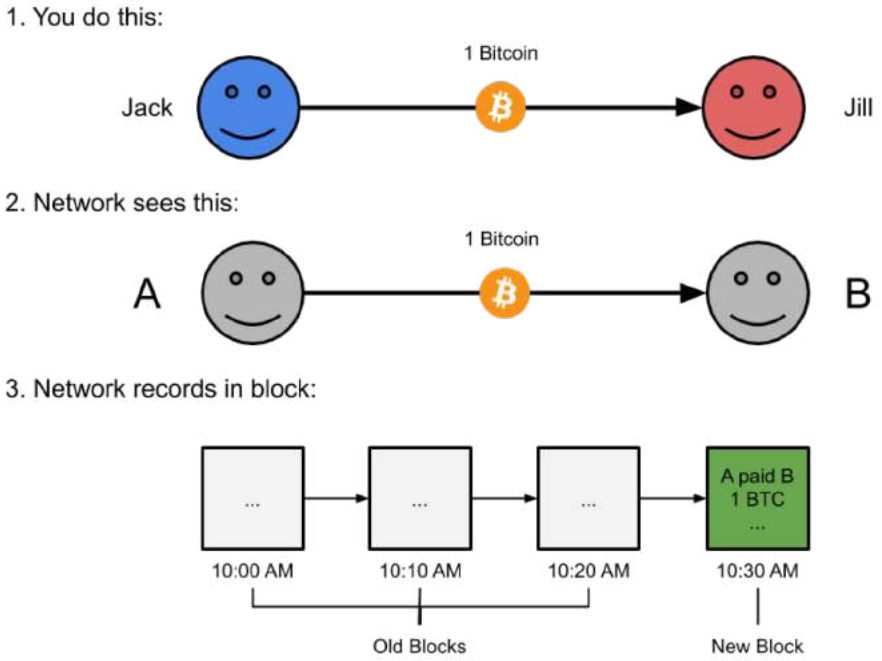
Submit all of the slices at once in unordered transactions

Store byte slices indexed by their position within the file

This research was supported by grant T15LM011271

Blockchain Basics

- Most typical use is for money
- Jack sends Jill some money, in this case, one Bitcoin
- The network sees that a transaction has occurred in which person A sent person B one Bitcoin
- This transaction is recorded in a block, which is added to the chain
- Everyone in the network updates their chain to reflect this new change



This research was supported by grant T15LM011271

Web App Access

Web App Instance #1

cutt.ly/dbmi-certificates-1

Web App Instance #2

cutt.ly/dbmi-certificates-2

This research was supported by grant T15LM011271