香港浸會大學理學院
HKBU Faculty of Science

DEPARTMENT OF COMPUTER SCIENCE

PhD Degree Oral Presentation

PhD Candidate:	Mr. WANG Canhui
Date	15 June 2023 (Thursday)
Time:	4:00 pm – 6:00 pm (35 mins presentation and 15 mins Q & A)
Venue:	ZOOM (Meeting ID: 945 6834 4535) (The password and direct link will only be provided to registrants)
Registration:	https://bit.ly/bucs-reg (Deadline: 12:00 nn, 15 June 2023)

Measurement, Analysis, and Modeling of Bitcoin and Hyperledger Fabric Blockchains

Abstract

Blockchain is a Distributed Ledger Technology (DLT) originally proposed to solve the double-spending attack on the Bitcoin cryptocurrency. A blockchain is a chain of blocks that is replicated among participants on a Peer-to-Peer (P2P) network. An important property of blockchain is that it does not rely on any third party for transaction clearing; instead, it stores and updates all valid historical transactions as a future reference, with which whether a new coming transaction is valid can therefore be determined. However, blockchain suffers from various performance challenges, including openness, energy consumption, scalability, throughput, and latency.

In this thesis, we study the performance characteristics of blockchain from two types of blockchains: permissionless and permissioned blockchains. Specifically, a permissioned blockchain requires prior authorization with which only authenticated users have access permission to access specific blockchain data sets, while a permissionless blockchain does not require that. In practice, the permissionless blockchain is popular for data sharing, of which a killer application is the Bitcoin cryptocurrency. The permissioned blockchain is prevalent in many fields of industry where data security and privacy is the most significant concern, and a popular application is Hyperledger Fabric. In the following, we introduce our studies on Bitcoin and Hyperledger Fabric blockchains.

Bitcoin network is one of the most popular permissionless blockchain systems. Mining pools are the main components of the Bitcoin network that invest a large amount of computing power to maximize their mining payoffs, which in turn guarantees the security of the Bitcoin network. Although there are many existing works about mining pools on the Bitcoin network, the long-term evolution of mining pools and their effects on both the Bitcoin system and its end-users remain to be investigated. To fill this gap, we trace over 2.54 hundred thousand blocks from Feb 2016 to Nov 2020 and collect over 12 million unconfirmed transactions from Mar 2018 to Nov 2020. We then conduct a broad range of analyses on these data, including the pool evolution, labeled transactions, and label blocks. We make the following observations from our measured data: 1) A few top mining pools control most of the P2P network's computing power. 2) The long-term computing power of top mining pools grows exponentially while its continuous-time mining strategy decreases linearly. 3) The computing power of the Bitcoin network converges to the Nash equilibrium. We then propose game-based strategies for mining activity analysis, i.e., the best-response strategies for mining pools when the mining revenue increases or decreases sharply. Moreover, we study the transaction fee dilemma of mining pools and the transaction fee strategies for

end-users. Overall, our models and analysis can help to understand and improve the Bitcoin system quantitatively.

Hyperledger Fabric is one of the most popular permissioned blockchain platforms. Although many existing works on the overall system performance of Hyperledger Fabric are available, a phase decomposition analysis of Hyperledger Fabric still remains to be explored. Admittedly, an overall system performance of Hyperledger Fabric might provide an end-user with satisfied performance information when invoking a transaction; however, it is far from informative when deploying a complex distributed system with specific performance goals, except for understanding each component and each phase in Hyperledger Fabric. Besides, it is challenging to analyze the performance of a distributed system with many dependent phases like Hyperledger Fabric, where an output of a phase becomes the input of the next phase, and each phase's output might not follow a Poisson distribution. In this thesis, we develop a measurement framework to characterize each phase's transaction and block data in Hyperledger Fabric based on the Fabric SDK Nodejs, where we thoroughly analyze and open-source the implementation details of the measurement framework. We evaluate the performance of Hyperledger Fabric and have some interesting observations: 1) The number of CPU cores has a linear impact on the throughput of an endorsing peer. 2) The Raft-based ordering service shows good scalability with the number of ordering service nodes. 3) The communication latencies of both the client and the service sides in Hyperledger Fabric are significant. We then identify each phase's dominant resource and latency in Hyperledger Fabric via primitive operation analysis and propose a stochastic computation model for performance analysis. Specifically, an endorsing peer in the execute phase is modeled as a M/D/c queue, an OSN leader in the order phase is modeled as a G/G/1 queue, and a committing peer in the validate phase is modeled as a G/G/1 queue. We also apply the Alpha-Beta communication model to analyze the corresponding communication latency. Finally, we validate the accuracy of the proposed performance model, i.e., the stochastic computation model and the alpha-beta communication model, on both local and cloud clusters. Overall, the experiment results and the performance model can help guide the deployment of the Hyperledger Fabric service.

***** ALL INTERESTED ARE WELCOME *****