Report from Dagstuhl Seminar 19261

Distributed Computing with Permissioned Blockchains and Databases

Edited by

C. Mohan[1], Beng Chin Ooi[2], and Gottfried Vossen[3]

1 **IBM Almaden Center – San Jose, US,** cmohan@us.ibm.com
2 **National University of Singapore, SG,** ooibc@comp.nus.edu.sg
3 **Universität Münster, DE,** vossen@wi.uni-muenster.de

Abstract

This seminar report contains the motivation, abstracts, and findings of Dagstuhl Seminar 19261 *Distributed Computing with Permissioned Blockchains and Databases* which took place in late June 2019. It brought together a very good mix of people from academia and industry as well as from databases and related areas for which blockchain is a current topic and who are either users or developers in that field.

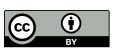
Seminar June 23–28, 2019 – http://www.dagstuhl.de/19261
2012 ACM Subject Classification Information systems → Data management systems
Keywords and phrases Distributed database, blockchain, permissioned
Digital Object Identifier 10.4230/DagRep.9.6.69

1 Executive Summary

C. Mohan
Beng Chin Ooi
Gottfried Vossen

The topic of blockchains, and in particular that of permissioned blockchains, has rapidly gained interest in both the industrial and the research communities in recent years. It particularly pertains to situations where trust among several parties that are about to do business together is difficult to establish (e.g., due to organizational, financial, or timing reasons) or impossible to establish at all. A blockchain is a decentralized, distributed ledger that consists of immutable blocks containing transactions that can be accessed by any party, and that provides trust via replication over all nodes and an agreed-upon execution order of the transactions. Of particular interest are permissioned blockchains in which the associated parties are known and authenticated, yet still do not fully trust each other.

Many applications have shown interest in the concept of blockchains, since the situation just described applies to many real-world scenarios, including (global) supply chains, the Internet of Things, connected cars, manufacturing, banking, and healthcare. As a consequence, a number of players in the IT industry work on a development of the technology, and several consortia have been formed to advance the technology across industries, among them Hyperledger and R3. Moreover, a number of companies have released Blockchain-as-a-Service (BaaS) platforms, including IBM, Oracle, Amazon, Baidu, and Alibaba.



Distributed Computing with Permissioned Blockchains and Databases, *Dagstuhl Reports*, Vol. 9, Issue 6, pp. 69–94
Editors: C. Mohan, Beng Chin Ooi, and Gottfried Vossen
Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

The technology has many links into the database community; however, the situation is basically like it was in the database area many years ago, when only a few systems had been released but users were on their own to figure out how to use them effectively. As the seminar has shown, many interesting issues remain to be solved, and there is a wide variety of aspects and research issues currently under investigation. Of these, the following were discussed:

- Blockchain scalability w.r.t. transaction throughput, one of the main roadblocks to business adoption
- Transaction ordering and endorsement, consensus of transaction commit
- Adjustments to the Proof of Work (PoW) consensus mechanism, other optimizations to consensus algorithms (e.g., Byzantine consensus) in the presence of transaction failures and in light of scalability
- Block validation
- Languages for smart-contract specification (e.g., Sandcastle SQL and Solidity)
- Amendments to Hyperledger Fabric, such as channels
- Cross-chain swaps using hashed timelocks
- Energy efficiency of blockchain applications

In addition, several participants reported on various working applications of blockchain technology.

2 Table of Contents

3 Motivation

A new era is emerging in the world of distributed computing with the growing popularity of blockchains. Traditionally, the Internet allows to exchange only data or information directly between two parties or users; a transaction, say, involving the purchase of an item requires a third party which can be trusted. Third parties often come in the form of a digital marketplace, a bank, or a trusted intermediary. Blockchains can eliminate third parties, since they are characterized by transparency, i.e., the blockchain content is visible to each participant, and being append-only, which is crucial for updating a blockchain. Conceptually, a blockchain is a decentralized and distributed digital ledger that consists of records representing transactions; since these records are tied to their history using hash values no existing record can be altered retroactively. The only kind of update allowed is to extend a given blockchain by additional records, which, assuming that the majority of participants does not pursue a dishonest intention, results in a stable view on transactions (which implies that not every party or node maintaining the blockchain needs to trust everybody else). The participants can verify and audit transactions, which results in a trustable workflow where participants' uncertainty regarding data security is marginal. The use of a blockchain also eliminates infinite reproducibility of digital assets; it confirms that each unit of value was transferred only once.

By storing data across its network, the blockchain eliminates the risks that come with data being held centrally, yet opens up for an application of distributed technology that was previously developed in other contexts. Blockchains come in two flavors: An open, permissionless, or public, blockchain network does not require any guarding against bad actors, and no access control is needed; anybody can join and leave. Hence applications can be added to the network without the approval or trust of others, using the blockchain as a transport layer. Permissioned (private) blockchains are emerging as open source protocols where openness and collaboration are encouraged among authenticated participants. They can hence restrict who can participate in the consensus processes as well as who can transact.

From a database point of view, a blockchain can be considered as a log of ordered transactions, since nodes keep replicas of the data and agree on an execution order of the transactions. A key property is the assumption that nodes behave in an arbitrary or Byzantine fashion. By being able to tolerate Byzantine failure by design, a blockchain offers stronger security than a database system. Although enterprise-grade database systems support applications like security trading and settlement, asset and finance management, or banking and insurance, blockchain technology has the potential to disrupt the status quo since they incur lower costs of infrastructure and human labor. In particular the immutability and transparency of a blockchain reduce human error as well as the need for manual intervention due to conflicting data.

While there is currently no standard in the blockchain space, all the ongoing efforts involve some combination of database, transaction, encryption, consensus and other distributed systems technologies. Some of the application areas in which blockchain pilots are being carried out are: smart contracts, supply chain management, know your customer, derivatives processing and provenance management. The seminar has surveyed some of the ongoing blockchain projects with respect to their architectures in general and their approaches to some specific technical areas. Its focus has been on how the functionality of traditional and modern data stores are being utilized or not utilized in different blockchain projects.

4 Topic Areas Discussed

This section lists the abstracts of talks given, ordered by the topic areas to which they belong.

Goals and Current State-of-the-Art of Blockchain Technology and Systems

4.1 State of Public and Private Blockchains: Myths and Reality

C. Mohan (IBM Almaden Center – San Jose, US)

It has been a decade since the concept of blockchain was invented as the underlying core data structure of the permissionless or public Bitcoin cryptocurrency network. Since then, several cryptocurrencies, tokens and ICOs have emerged. After much speculation and hype, a significant number of them have become problematic or worthless! The public blockchain system Ethereum emerged by generalizing the use of blockchains to manage any kind of asset, be it physical or purely digital, with the introduction of Smart Contracts. Over the years, numerous myths have developed with respect to the purported utility and the need for public blockchains. The adoption and adaptation of blockchains and smart contracts for use in the permissioned or private environments is what I consider to be useful and of practical consequence. Hence, the technical aspects of only private blockchains will be the focus of my talk. Along the way, I will bust many myths associated with public blockchains. I will also compare traditional database techniques with blockchain systems' features and identify desirable future research topics.

References

1 C. Mohan. State of Public and Private Blockchains: Myths and Reality. Proc. ACM SIGMOD International Conference on Management of Data, Amsterdam, July 2019. `http://bit.ly/sigBcP`
2 C. Mohan. Permissioned/Private Blockchains and Databases. `http://bit.ly/CMbcDB`

4.2 Introduction to Hyperledger

Hart Montgomery (Fujitsu Labs of America Inc. – Sunnyvale, US)

In this talk, I introduced Hyperledger, explained its structure and governance, and showed how to participate and contribute. Hyperledger is a "greenhouse" under the Linux foundation for permissioned blockchains. It is currently the largest and most popular permissioned blockchain project.

4.3 Usages of Blockchain Technologies for Data Stores

Bernhard Mitschang (Universität Stuttgart, DE)

Currently, blockchain technologies are seen as the foundation of a new business world: it will change the way the economy runs and thus will change the way we act and work, all triggered by means of some new ways to organize the relevant application data, e.g., in the areas of supply chain, health, and event storage. Blockchain technologies and systems are still in constant change and development. Hence, it is difficult to exactly define its ingredients and properties of the underlying technologies.

After having identified these technologies and associated characteristics (like transparency, provenance, fault tolerance, immutability, and authenticity), it is important to isolate and separate them into components that are subsequently used to enhance existing data stores as needed. Important questions that arise in this context are:

- How do certain Blockchain technologies and applications match?
- How to identify and separate Blockchain technologies?
- How to "append/integrate" Blockchain technologies to/with existing data stores?

Cryptography and Blockchain

4.4 Privacy, Confidentiality, Cryptography, and Security Modelling in Permissioned Blockchains

Hart Montgomery (Fujitsu Labs of America Inc. – Sunnyvale, US)

Achieving desired privacy and confidentiality properties on a blockchain can be quite difficult. This is especially true on permissioned blockchains, where it may be more difficult to hide or anonymize identities than on a public blockchain. In this talk, I explained some of the challenges that are commonly faced when attempting to achieve privacy and confidentiality on permissioned blockchains and how to go about using existing tools to achieve these properties.

One of the most important things when designing secure permissioned blockchains is the need for security modelling. Many people today pick cryptographic tools, apply them to blockchains, and then try to analyze the security properties that they get (if they even do that). This isn't a good idea for many reasons, but, in particular, it often means that blockchains do not provide the security guarantees that people want on a blockchain. For instance, even if transactions on the blockchain are encrypted or hashed, it could be the case that traffic analysis completely reveals the participants in a transaction or even information about the contents of transactions [5]. Intuitively, one might expect encryption might prevent such leakage, but it turns out that other "side channel" information on the blockchain nullifies some of the security properties of encryption.

Another very important discussion point was the notion of privacy and anonymity, and the fact that the two aren't equivalent. Many blockchain practitioners (both of the public blockchain and permissioned blockchain kind) frequently equate the two, and many disastrous

consequences can happen from this. As an example, I showed my credit card history from a week last year, which, if extended further, would easily deanonymize me. Solving this issue on a blockchain is a difficult task, and blockchain builders may not want to provide perfect privacy to their users (in some cases, functionality even demands imperfect privacy, like when KYC regulations apply). However, blockchain implementers certainly need to take into account privacy (and anonymity) into their security models when building blockchains.

We next discussed cryptographic tools that can be useful for obtaining various privacy and confidentiality properties in blockchain. The first topic was threshold signatures, which allow a cryptographic signing key to be split into n different shares such that any t out of the n shares are required to create a valid signature (and any t – 1 shares cannot "do anything") [1]. We went through the security game of threshold signatures in detail, which illustrated how one should look at a security game for a blockchain. We also briefly defined functional encryption [3] and explained why it would be very useful for a blockchain.

The next technique discussed was "channels." Channels, developed in Hyperledger Fabric [2], are a tool intended to enable private transactions on blockchains. The idea is that each channel acts as a private "sub-blockchain" for a limited number of participants, and that people without permission do not have visibility into what is going on inside the channel. While channels are a useful tool, they have not quite reached their full potential, so if they are to be used to achieve strong privacy requirements, more development on top of them is generally required.

The final technique we discussed was trusted execution environments (TEEs). Although they have been much maligned recently in terms of their security properties [4], TEEs such as Intel's SGX offer many potential benefits for secure and private blockchains. It is possible to essentially run blockchain nodes inside TEEs such that (assuming the TEEs are secure), even the blockchain node hosts cannot see what the blockchain is doing. TEEs could potentially give us very strong privacy and confidentiality options on blockchain if they can, in fact, be built securely.

The talk ended with many questions. Overall, the goal was to expose what was mostly an audience of researchers focused on databases to some of the privacy, security, and confidentiality challenges present on blockchain today.

References

1 Boldyreva, Alexandra: Threshold signatures, multisignatures and blind signatures based on the gap-Diffie-Hellman-group signature scheme; International Workshop on Public Key Cryptography. Springer, Berlin, Heidelberg, 2003.
2 Androulaki, Elli, et al: Hyperledger fabric: a distributed operating system for permissioned blockchains; Proceedings of the Thirteenth EuroSys Conference. ACM, 2018.
3 Boneh, Dan, Amit Sahai, and Brent Waters: Functional encryption: Definitions and challenges; Theory of Cryptography Conference. Springer, Berlin, Heidelberg, 2011.
4 Kocher, Paul, et al.: Spectre attacks: Exploiting speculative execution; arXiv preprint arXiv:1801.01203 (2018).
5 Murdoch, Steven J., and George Danezis: Low-cost traffic analysis of Tor; 2005 IEEE Symposium on Security and Privacy (S & P'05). IEEE, 2005.

4.5 A Hybrid Blockchain Architecture for Enhancing Privacy and Accountability

Murat Kantarcioglu (University of Texas – Dallas, US)

Joint work of Murat Kantarcioglu, Harsh Desai, Lalana Kagal

Unfortunately, existing public blockchains and smart contracts deployed on them may disclose sensitive information. Although there is some ongoing work that leverage advanced cryptography to address some of these sensitive information leakage issues, they require significant changes to existing and popular blockchains such as Ethereum and are usually computationally expensive. On the other hand, private blockchains have been proposed to allow more efficient and privacy-preserving data sharing among pre-approved group of nodes/participants. Although private blockchains address some of the privacy challenges by allowing sensitive data to be only seen by the select group of participants, they do not allow public accountability of transactions since transactions are approved by known set of users, and cannot be accessed publicly. Given these observations, one natural question that arise is, can we leverage both public and private blockchain infrastructures to enable efficient, privacy enhancing and accountable applications ? In this talk, we try to address this challenge in the context of digital auctions.

Mainly, we discuss a novel hybrid blockchain architecture [1] that combines private and public blockchains to allow sensitive bids to be opened on a private blockchain so that only the auctioneer can learn the bids, and no one else. At the same time, we leverage public blockchains to make the auction winner announcement, and payments accountable [2]. Furthermore, using smart contracts deployed on public blockchain, we show how to incentivize truthful behavior among the auction participants. Our extensive empirical results show that this architecture is more efficient in terms of run time and monetary cost compared to pure public blockchain based auction implementations.

References

1 Harsh Bimal Desai, Murat Kantarcioglu, and Lalana Kagal. A hybrid blockchain architecture for privacy-enabled and accountable auctions. In *The 2019 IEEE International Conference on Blockchain (Blockchain-2019)*, 2019.

2 Aravind Ramachandran and Murat Kantarcioglu. Smartprovenance: A distributed, blockchain based dataprovenance system. In *Proceedings of the Eighth ACM Conference on Data and Application Security and Privacy, CODASPY 2018, Tempe, AZ, USA, March 19-21, 2018*, pages 35–42, 2018.

Data Models

4.6 Sandcastle: a SQL Ethereum Smart Contract Language

Shahan Khatchadourian (ConsenSys – Toronto, CA)

Enterprises rely on data management frameworks in order to serve their customer. However, enterprises face challenges when integrating blockchains with existing enterprise stacks in a way that makes it easy to query, understand, and transact across systems. Challenges arise

due to the complex composition of database and blockchain paradigms. As well, developers face the challenge of writing smart contracts in a low-level language like Solidity, with a need to understand concepts like decentralization, smart contracts, consensus and identity. This leads to developers building adhoc, incongruent solutions at application or protocol layers.

We propose Sandcastle, a SQL Ethereum smart contract language that integrates enterprise data management. Sandcastle works on all Ethereum blockchains (without modification or configuration), including public, private, permissioned, and permissionless networks. We showcase Sandcastle's relational features such as aggregation, triggers, functions, indexes, and row-based semantics in finance, electronic medical records, and governance use cases. We give architectural details, including the translation of Sandcastle SQL to Solidity. The Sandcastle roadmap includes optimization in performance, cost, and security. Sandcastle aims to help traditional enterprises build scalable, data-oriented blockchain platforms that span databases, Ethereum 1.0, and Ethereum 2.0 stacks and networks.

4.7 Blockchained Event Store

Dennis Przytarski (Universität Stuttgart, DE)

Consider different scenarios such as transportation/trucking and supply chain integrity. All these scenarios have one commonality: different parties generate events that need to be shared among themselves in an immutable and tamper-resistant manner. Because the stored events are used for forecasts, reports, or further process optimizations, powerful querying capabilities both on current and historical states are needed.

In general, the blockchain technology is suitable for these scenarios because it offers data immutability and tamper-resistance. For typical blockchain systems that assume transferable assets (i.e., transfer ownership of an object from one person to another person), the key-value data model and a simple query engine to answer queries such as "is a particular transaction included in a particular block" are sufficient enough.

As soon as either the data model or the query requirements increase, this basic blockchain setup does not suffice anymore. Instead, powerful query and data model capabilities are needed with immutability and tamper-resistance guaranteed. Therefore, I propose using triples (entity, attribute, value) as simple but powerful and flexible data model.

I am currently working on embedding the triple data model into a blockchain architecture with a powerful query engine. This will lead to an immutable, shared, tamper-resistant, and queryable data store for events. I am currently facing the following challenges: Data Model:

- How are changes to the data model done?
- How is the data in the triple data model (tamper-resistant) stored?

Query Language and Processing:

- How could the query language look like when there is a history to query?
- How to process a query on immutable data in the triple data model?

Consensus Protocols and Blockchain

4.8 ExpoDB Fabric: Efficient Transaction Processing in Byzantine Fault Tolerant Environments

Mohammad Sadoghi Hamedani (University of California – Davis, US)

The byzantine fault-tolerance model, studied in ExpoDB Fabric [6, 4, 5, 3, 2, 1], captures a wide-range of failures–common in real-world scenarios–such as ones due to malicious attacks and arbitrary software/hardware errors. We propose Blockplane [2], a middleware that enables making existing benign systems tolerate byzantine failures. This is done by making the existing system use Blockplane for durability and as a communication infrastructure. Blockplane proposes the following: (1) A middleware and communication infrastructure to make an entire benign protocol byzantine fault-tolerant, (2) Ahierarchical locality-aware design to minimize the number of wide-area messages, (3) A separation of fault-tolerance concerns to enable designs with higher performance.

We further investigate a protocol-agnostic approach to improve the design of primary-backup consensus protocols. At the core of our approach is a novel wait-free design of running several instances of the underlying consensus protocol in parallel [3]. To yield a high-performance parallelized design, we present coordination-free techniques to order operations across parallel instances, deal with instance failures, and assign clients to specific instances. Consequently, the design we present is able to reduce the load on individual instances and primaries, while also reducing the adverse effects of any malicious replicas. Our design is fine-tuned such that the instances coordinated by non-faulty replicas are wait-free: they can continuously make consensus decisions, independent of the behavior of any other instances.

We further develop DeltaBFT, a novel consensus protocol in which all algorithms necessary for normal-case operation only require linear communication costs, even if replicas fail [5]. At the center of our design is the delayed-replication algorithm, an algorithm we propose to reliably broadcast consensus decisions made by some non-faulty replicas to all replicas without any coordination and with low communication cost for all replicas involved. The delayed-replication algorithm is supported by our partial consensus algorithm, which uses threshold signatures to efficiently make consensus decisions.

The development of fault-tolerant distributed systems that can tolerate Byzantine behavior has traditionally been focused on consensus protocols, which support fully-replicated designs. For the development of more sophisticated high-performance Byzantine distributed systems, more specialized fault-tolerant communication primitives are necessary. As a result, we identify an essential communication primitive and study it in depth. In specifics, we formalize the cluster-sending problem [4], the problem of sending a message from one Byzantine cluster to another Byzantine cluster in a reliable manner. We not only formalize this fundamental problem, but also establish lower bounds on the complexity of this problem under crash failures and Byzantine failures. Furthermore, we develop practical cluster-sending protocols that meet these lower bounds and, hence, have optimal complexity. As such, our work provides a strong foundation for the further exploration of novel designs that address challenges encountered in fault-tolerant distributed systems.

References

1 S. Gupta and M. Sadoghi. *Blockchain Transaction Processing*. 2018.
2 F. Nawab and M. Sadoghi. Blockplane: A global-scale byzantizing middleware. In *2019 IEEE 35th International Conference on Data Engineering (ICDE)*, pages 124–135, April 2019.
3 S. Gupta, J. Hellings, and M. Sadoghi. Brief Announcement: revisiting consensus protocols through wait-free parallelization. In *33rd International Symposium on Distributed Computing (DISC 2019)*.
4 J. Hellings, and M. Sadoghi. The fault-tolerant cluster-sending problem. In *33rd International Symposium on Distributed Computing (DISC 2019)*.
5 J. Hellings, and M. Sadoghi. Brief Announcement: Byzantine fault-tolerant consensus with almost linear communication complexity. In *CoRR. 2019*.
6 M. Sadoghi, S. Blanas. Transaction Processing on Modern Hardware. In *Synthesis Lectures on Data Management, Morgan & Claypool Publishers 2019*.

4.9 CUB, a Consensus Unit-based Storage Scheme for Blockchain System

Lei Chen (HKUST – Kowloon, HK)

Joint work of Zihuan Xu, and Siyuan Han
Main reference Zihuan Xu, Siyuan Han, Lei Chen: "CUB, a Consensus Unit-Based Storage Scheme for Blockchain System", in Proc. of the 34th IEEE International Conference on Data Engineering, ICDE 2018, Paris, France, April 16-19, 2018, pp. 173–184, IEEE Computer Society, 2018.
URL https://doi.org/10.1109/ICDE.2018.00025

Recently, Blockchain becomes a hot research topic due to the success of Blockchain in many applications, such as cryptocurrency, smart contract, digital assets, distributed cloud storage and so on. The power of Blockchain is that it can achieve the consensus of an ordered set of transactions among nodes which do not trust each other, even with the existence of malicious nodes. However, compared to traditional databases, the current Blockchain technology still cannot handle a massive number of transactions, which is caused by many factors, such as the consensus protocol, structure of the blocks and storage challenge. Among them, the high storage requirement is a key factor that prevents the wide usage of Blockchain on various devices such as mobile phones or low-end PCs.

In this talk, I will discuss a novel concept called Consensus Unit (CU), which organizes different nodes into one unit and lets them to store at least one copy of Blockchain data in the system together. Based on this idea, the Blocks Assignment Optimization (BAO) problem is defined to determine the optimal assignment of blocks such that the storage space is fully used and the query cost is minimized. The problem is NP-hard. Thus, three efficient heuristic algorithms are presented to solve the static assignment problem. Furthermore, the dynamic scenarios are discussed when new blocks arrive or nodes join or depart from the CU. At the end of this talk, I will highlight some future research directions on Block chain systems.

4.10 Beyond Consensus in Permissioned Ledgers: Experiences in Using BFT Replication on DLTs

Alysson Neves Bessani (University of Lisbon, PT)

Permissioned Blockchains such as Hyperledger Fabric and R3 Corda rely on modular consensus-as-a-service components for ordering transactions. In this talk I explained how these components can be implemented using "traditional" consensus protocols (such as PBFT) and argued that such protocols are only the first step in building a robust and efficient service for these blockchains. I also showed how the required features were implemented in BFT-SMART, a replication library used for implementing Byzantine-resilient consensus-as-a-service components for both Fabric and Corda.

4.11 Red Belly Blockchain: Byzantine Consensus Is Back but Is It the Same?

Vincent Gramoli (The University of Sydney, AU)

Byzantine Consensus was proposed in the early 80's for multiple machines to reach agreement on a unique value. A practical solution, called PBFT, used a leader for implementing a network file system in a local area network in 1999. Today, with the advent of Blockchain, various companies are now trying to avoid double spending by having a large number of machines reach an agreement upon a block at any given index of the blockchain. Most companies take off-the-shelf leader-based Byzantine consensus protocols, inspired by PBFT, to solve this old consensus problem.

The issue is that the Blockchain Consensus is different from this classic Byzantine Consensus problem because the number of machines that should agree is large. Our recent design of the Democratic Practical Byzantine Fault Tolerant (DBFT) consensus algorithm solves a variant of the Byzantine Consensus problem that allows to scale by leveraging the cryptographic primitive of the blockchain to decide whether a proposal is valid. It contrasts with off-the-shelf solutions in that it is fully decentralised and does not rely on a leader to avoid bottlenecks.

The blockchain we built using DBFT, called Red Belly Blockchain, is a community blockchain whose set of consensus participants changes over time. Red Belly Blockchain uses the ECDSA public-key cryptosystem, it verifies all cryptographically signed transactions in an efficient way to avoid CPU wastage, it involves all participants to collaboratively solves this Blockchain Consensus instead of relying on a leader bottleneck. It resolves conflicts between transactions, never forks and is provably starvation-free. Our experiments show that Red Belly scales to 1000 replicas spread across 4 different continents with an average latency of 3 seconds and its peak throughput exceeds 660,000 TPS.

Performance

4.12 Enhancing Performance, Scalability, and Confidentiality of Permissioned Blockchains

Divyakant Agrawal (University of California – Santa Barbara, US)

Joint work of Divyakant Agrawal, Mohammad Java Amiri, Amr El Abbadi

Blockchains have unique features, such as transparency, provenance, fault tolerance, and authenticity, which appeal to a wide range of distributed applications, e.g., supply chain management and healthcare. However blockchain systems suffer from *performance*, *scalability*, and *confidentiality* limitations.

Existing blockchains mostly utilize an order-execute architecture where nodes agree on a total order of the blocks of transactions using a consensus protocol and then the transactions are executed in the same order on all nodes sequentially. The sequential execution of transactions on all nodes, however, reduces the blockchain performance in terms of throughput and latency. While Hyperledger Fabric increases the performance of blockchains by switching the order of the execution and ordering phases and executing the transactions in parallel, it performs poorly on workloads with high-contention, i.e., many *conflicting transactions* in a block, due to its high abort rate. To address this problem, we introduce a permissioned blockchain system *ParBlockchain* [1]. ParBlockchain is mainly introduced to support distributed applications processing workloads with *some degree of contention*. ParBlockchain consists of *orderers* and *agent* nodes. Orderers establish agreement on the order of the transactions of different applications, construct the blocks of transactions, and generate a *dependency graph* for the transactions within a block. A dependency graph enables higher concurrency by allowing the parallel execution of non-conflicting transactions. The agents of each application are then responsible for executing the transactions of that application following the generated dependency graph.

Scalability is one of the main roadblocks to business adoption of blockchain systems. Despite recent intensive research on using sharding techniques to enhance the scalability of blockchain systems, existing solutions do not efficiently address cross-shard transactions. We introduce a permissioned blockchain system, *SharPer* [2], that enhances the scalability of blockchain systems by clustering (partitioning) the nodes and assigning different data shards to different clusters. SharPer supports both intra-shard and cross-shard transactions and processes intra-shard transactions of different clusters as well as cross-shard transactions with no overlapping clusters simultaneously. In SharPer, the blockchain ledger is formed as a directed acyclic graph where each cluster maintains *only* a view of the ledger. SharPer also incorporates a protocol to establish consensus on the order of cross-shard transactions among *only* the involved clusters.

Many distributed applications need to collaborate with each other following service level agreements to provide different services. Distributed applications are often designed and implemented in different blockchain systems. In this case, inter-application collaboration could be performed as an atomic cross-chain swap, however, such an operation could negatively affect the performance of the blockchain. Furthermore, while collaboration between applications, e.g., cross-application transactions, should be *visible* to all applications, the internal data of each application, e.g, internal transactions, might be *confidential*. To support both internal and cross-application transactions of collaborating distributed applications, a permissioned blockchain system, *CAPER* [3], is introduced. In CAPER, the blockchain ledger is formed

as a *directed acyclic graph* where each application accesses and maintains only its own *view* of the ledger including its internal and all cross-application transactions. CAPER also introduces three consensus protocols to globally order cross-application transactions between applications.

References

1 Mohammad Javad Amiri, Divyakant Agrawal, and Amr El Abbadi. *ParBlockchain: Leveraging Transaction Parallelism in Permissioned Blockchain Systems*. In Proceedings of the 39th IEEE International Conference on Distributed Computing Systems (ICDCS'19), pp. 1337-1347, Dallas, 2019.

2 Mohammad Javad Amiri, Divyakant Agrawal, and Amr El Abbadi. *On Sharding Permissioned Blockchains*. In Proceedings of the Second IEEE International Conference on Blockchain, pp. 282-285, Atlanta, 2019.

3 Mohammad Javad Amiri, Divyakant Agrawal, and Amr El Abbadi. *CAPER: A Cross-Application Permissioned Blockchain*. The 45th International Conference on Very Large Data Bases (VLDB'19), PVLDB 12(11), pp. 1385-1398, Los Angeles, 2019.

4.13 Hyperledger Fabric's Read-Set Conflicts and Conflict-Free Replicated Datatypes

Hans-Arno Jacobsen (TUM, DE & Univ. Toronto, CA) and Pezhman Nasirifard (TU München, DE)

Permissioned blockchains such as Hyperledger Fabric provide a robust ecosystem for developing enterprise and production-grade decentralized applications. However, the additional latency between the execution and committing the transactions, due to Fabric's adapted transaction lifecycle, is a potential scalability bottleneck. This latency can increase the probability of the occurrence of conflicting transactions, leading to the failure of a high number of transactions, which increases the application development complexity and decreases the Fabric's throughput and availability. We study an approach for integrating Conflict-Free Replicated Datatypes (CRDTs) to Hyperledger Fabric, to understand how CRDTs can improve the Fabric's availability and scalability. CRDTs are abstract data types that can resolve conflicts automatically in the presence of concurrent updates without coordination.

4.14 FastFabric: Scaling Hyperledger Fabric to 20,000 Transactions per Second

Srinivasan Keshav (University of Waterloo, CA)

Joint work of Christian Gorenflo, Stephen Lee, Lukasz Golab, Srinivasan Keshav
Main reference Christian Gorenflo, Stephen Lee, Lukasz Golab, Srinivasan Keshav: "FastFabric: Scaling Hyperledger Fabric to 20, 000 Transactions per Second", CoRR, Vol. abs/1901.00910, 2019.
URL https://arxiv.org/abs/1901.00910

Blockchain technologies are expected to make a significant impact on a variety of industries. However, one issue holding them back is their limited transaction throughput, especially compared to modern enterprise database systems. We have re-architected Hyperledger

Fabric to increase transaction throughput to 20,000 transactions per second. We focus on performance bottlenecks beyond consensus, proposing architectural changes that reduce computation and I/O overhead during transaction ordering and validation. Notably, our optimizations are fully plug-and-play and do not require any changes to Hyperledger Fabric.

4.15 Blockchains and Distributed Databases, a Twin Study

PingCheng Ruan (National University of Singapore, SG) and Beng Chin Ooi (National University of Singapore, SG)

Since the rise of Bitcoin, the public are stirring tremendous hype on its underlying blockchain technology. Over the years, the scope of blockchains has long been limited to cryptocurrency. Since the introduction of the smart contract, blockchains start to support general transactional workload, as RDBMs do. Further considering about their distributed nature, a proliferation of literature start to draw a parallel between blockchains and distributed databases. However, they mainly focus on their distinct properties to applications, but fail to identify their common technical aspects.

In this paper, we perform a joint study on blockchains and distributed databases and show that both are a twin of distributed transactional systems, with the former focusing on security while the latter on efficiency. On this common basis, we abstract out four technical aspects, replication, sharding, transaction management and storage, to lay out our comparison. Throughout, we show how the security–efficiency trade-off implicate their design goals and architectural choices. Next, we conduct an extensive performance study on two blockchains, QUORUM and FABRIC, with three distributed databases, Cockroach DB, TiDB and etcd. Our results indicate that even though the performance of blockchain is still far behind distributed databases, blockchains may still outperform them in some specific transactional workloads.

4.16 Optical Chips

Yong Tang (Univ. of Electronic Science &Technology – Chengdu, CN)

The current hardware solutions for PoW are CMOS based ASIC chips, which are slow and energy-consuming. Considering that the CMOS has met the limits of physics, it's hard to further improve speed. Moreover, it's hard to avoid energy wasting. In this talk, I introduce a design of an optical chip for PoW computations. Most of the PoW can be implemented using optical components. It's promising and attractive to do PoW with optical chips which might save energy and enjoy high speed. The possibility of doing PoW with innovative solutions such as optical chips might lead to reconsiderations of PoW and the design of cryptocurrencies.

4.17 Minimizing Transaction Failures in Permissioned Blockchains

Jeeta Ann Chacko (TU München, DE) and Hans-Arno Jacobsen (TUM, DE & Univ. Toronto, CA)

Permissioned blockchains have generally two models, namely the order-execute model and the execute-order-validate model. The order-execute model orders the incoming transactions based on a consensus algorithm and then executes the transactions on every peer in the blockchain network. Quorom, Tendermint and Ripple are examples of permissioned blockchains that follow this model. Hyperledger Fabric, on the other hand, uses the execute-order-validate model. Here the transactions are initially executed on specific peers known as endorsers which endorse these transactions. Endorsed transactions are then ordered based on a distributed consensus algorithm. The ordered transactions are then validated and committed by every peer. Both the permissioned blockchain models can be compared to database systems in certain aspects. The distributed consensus algorithms used to order the transactions are used also in replicated databases to reach consensus. Also, the order-execute-validate model is similar to the optimistic concurrency control model which has been used in various database systems. Given these parallels, it is a fruitful research direction to integrate existing database optimization strategies to improve permissioned blockchains. Our research goal is to minimize the transaction failures in permissioned blockchains. We are currently focusing on the Hyperledger Fabric implementation. The main types of of transaction failures in Hyperledger Fabric is MVCC read conflicts (inter block and intra block), phantom reads and endorsement failures. The first research area we are exploring is to use transaction reordering to reduce the number of transaction abortions. Transaction reordering is a well-known database optimization technique for databases that use optimistic concurrency control. We first create a conflict graph to find the transaction dependencies, then the minimum feedback vertex set is detected and finally the transactions are topologically sorted to minimize transaction abortion. A similar approach has been successfully used in [1] with good results. Our work differs from [1] in one aspect. We used an exact algorithm that has an exponential complexity to detect the minimum feedback vertex set. This resulted in high latency during the ordering phase resulting in more inter block MVCC read conflicts. Therefore, we were not able to show total reduction of transaction failures even though the intra block MVCC read conflicts were reduced. Currently our focus in on early commit of independent transactions and immediate re-endorsement of dependent transactions to counter the latency in the ordering phase.

References

1 Ankur Sharma, Felix Martin Schuhknecht, Divya Agrawal, and Jens Dittrich. *Blurring the Lines between Blockchains and Database Systems: the Case of Hyperledger Fabric*. In Proceedings of the 2019 International Conference on Management of Data (SIGMOD '19). ACM, New York, NY, USA, 105-122.

4.18 Blockchain Goes Green? A Time Energy Performance Study of Blockchain on Low Power Systems

Dumitrel Loghin (National University of Singapore, SG)

Permissionless blockchains are well-known to be energy inefficient, mainly because of their compute-intensive Proof-of-Work consensus protocols. On the other hand, the energy profile of permissioned blockchains is less studied. With the increasing performance of low-power, wimpy devices based on ARM or x86/64 CPUs, our goal is to analyze their time-energy performance when running blockchain applications, in comparison with traditional, brawny servers. In this work, we select three wimpy systems with power profiles in the range 5-25W, namely, (i) an Intel NUC with Intel Core i3 CPU, (ii) a Jetson TX2 with 64-bit ARM CPU and (iii) a Raspberry Pi 3 with 32-bit ARM software stack. We run BLOCKBENCH on three blockchains, namely, Hyperledger Fabric v0.6, Ethereum and Parity, in a private, permissioned setup. We show that low-end wimpy nodes, such as Raspberry Pi 3, are struggling to run full-fledged blockchains due to their small memory size and low I/O bandwidth. However, higher-performance wimpy nodes, such as Jetson TX2, achieve around 80% and 30% of the throughput of Xeon servers for Parity and Hyperledger, respectively, while using 18x and 23x less energy.

Applications

4.19 Blockchain and New Economies

Feida Zhu (SMU – Singapore, SG)

URL https://symphonyprotocol.com/

Despite its most successful and well-known application for cryptocurrencies, it is our belief that the true power of blockchain technology is to unleash the great potential of a whole class of virtual assets, whose value are long known but not yet well established. Such assets include data, influence, social network, credit, to name a few. As an example, I will demonstrate in this talk how blockchain technology can be used to establish individual data as an emerging asset class to solve the bottleneck in today's data-driven economy. We will examine the key issues we face today from both the perspectives of the businesses and the individual users, and explore how blockchain-based platform could provide both the "trust" and "incentive" necessary to foster a self-growing data ecosystem. We introduce "Symphony Protocol", which is a blockchain-based protocol to create an ecosystem that unlocks personal data for democratized and personalized intelligence, with privacy by design.

4.20 Blockchain-based Cross-Site Genomic Dataset Access Audit

Li Xiong (Emory University – Atlanta, US)

Genomic data have been collected by different institutions and companies and need to be shared for broader use. In a cross-site genomic data sharing system, a secure and transparent access control audit module plays an essential role in ensuring the accountability. The goal of the iDASH competition 2018 first track is to develop blockchain-based ledgering solutions to log and query the user activities of accessing genomic datasets across multiple sites. We designed a Multichain-based log system which can provide a light-weight and widely compatible module for existing blockchain platforms. The submitted solution won the third place of the competition. Our method introduces an on-chain indexing data structure which can be easily adapted to any blockchains that use key-value database as their local storage.

4.21 Leveraging Decentralized, Secure and Governed Exchange of Confidential Information with Permissioned Blockchain

Gabriela Ruberg (Banco Central do Brasil – Rio de Janeiro, BR)

Joint work of Gabriela Ruberg, Marcus Vinicius Cursino, Jose Deodoro, Rafael Sarres, Aristides Cavalcante-Neto

The impressive popularity of blockchain applications, such as Bitcoin, has fostered the emergence of a variety of software tools to develop decentralized P2P systems. This has opened up the way for several new possibilities to explore blockchain technology beyond cryptocurrencies and financial services. In particular, permissioned blockchain networks (that is, when participants are identified and previously authorized) allow benefiting from relevant blockchain properties, especially tamper-proof data and non-repudiation, with better performance.

Sharing confidential data among autonomous entities in a secure and governed environment remains a challenge that can benefit from this new blockchain perspective. In practice, canonical solutions involving either centralized databases or traditional information integration are not sufficient nor sustainable. They usually require significant up-front efforts and cannot easily support updates with new datasets and views. Also, they present long time-to-data (namely, the time for new information to be available), require frequent (and expensive!) data transfers and lack trustful data governance. In many cases, choosing trusted third parties is not trivial. Moreover, recent regulation on data protection has further highlighted the disadvantages of siloed-data solutions.

This problem is relevant, for instance, in the context of public agencies and regulators, which need to frequently exchange protected data in order to perform due diligence processes and to provide integrated public services.

To tackle these issues, at the Central Bank of Brazil we developed a blockchain platform, called PIER, to enable entities to share, integrate and exchange sensitive data in a flexible, secure and governed environment. The PIER platform runs a permissioned blockchain network where participant nodes can easily discover and publish datasets from off-chain

data sources, and then share metadata on the available datasets using Open API standards. Moreover, PIER nodes can create request models, which are views defined on the available datasets, possibly joining them.

PIER nodes rely on a powerful and agile oracle (namely, a component of the blockchain system that can read data stored externally), called Olinda, to create data services based on the OData protocol. Nonetheless, PIER nodes can recognize and import any dataset description that is Open-API compliant, as well as they can access any data service that supports the corresponding dataset RESTful API. By running configurable smart contracts, PIER nodes execute request models to retrieve data, and register all data requests (that is, the executions of the request models) in the distributed ledger, along with their responses.

In summary, in the PIER platform, blockchain ledgers are used to store: a decentralized catalog of datasets and request models; and an audit trail of all the data requests. Each participant is concerned only with the maintenance of its datasets and request models, which are automatically combined by the platform to compose the full catalog in the ledger. The PIER platform uses both public and private ledgers to enable flexible privacy control of the shared information. It explores the concept of dataspaces [1], such that the PIER platform provides information integration in a pay-as-you-go approach.

We developed the PIER platform using the JPMorgan Quorum software, and we are running a pilot in production since September of 2018 with the Brazilian financial regulators to support due diligence in authorization processes.

Currently, we are investigating further developments in the PIER platform, such as integrating off-chain credentials and datasets versioning in the decentralized catalog. Also, we are interested in exploring natural language processing and machine learning to classify and match datasets and their embedded data entities, as well as to automatically generate request models from datasets based on high-level user specifications.

References

1 Michael J. Franklin, Alon Halevy and David Maier. *From databases to dataspaces: a new abstraction for information management.* SIGMOD Record, vol. 34, n. 4, pp. 27-33, 2005.

4.22 Blockchain Empowered Drug Development Financing

Yong Tang (Univ. of Electronic Science & Technology – Chengdu, CN)

The drug developments are very risky business with very high failure rates and require massive investment. The procedure can last for near ten years before a successful drug is finally approved by FDA. The high risk makes the drug developers and investors less interested to invest in the early stages. To encourage investors to fund the underinvested stages requires innovative business model and platforms. In this talk, I'd like to introduce a blockchain empowered megafund for drug development financing. Using a blockchain-based special purpose vehicle (SPV), we get all stakeholders involved in drug development such as developers, SPV, regulators, institutional investors, retail investors, credit rating agencies, credit enhancers onto a platform. The data are shared, and the procedures are executed as smart contracts. All parties can enjoy better data sharing and enhanced services. More importantly, expensive management costs can be saved to allow better investment returns.

4.23 Turning a Vehicle Into an Economic Platform

Michael Huth (Imperial College London, GB)

Joint work of Kwok Cheung, Michael Huth, Laurence Kirk, Leif-Nissen Lundbæk, Rodolphe Marques, Jan Petsche
Main reference Kwok Cheung, Michael Huth, Laurence Kirk, Leif-Nissen Lundbæk, Rodolphe Marques, Jan Petsche: "Owner-Centric Sharing of Physical Resources, Data, and Data-Driven Insights in Digital Ecosystems", in Proc. of the 24th ACM Symposium on Access Control Models and Technologies, SACMAT 2019, Toronto, ON, Canada, June 03-06, 2019, pp. 73–81, 2019.
URL https://doi.org/10.1145/3322431.3326326

Consumer expectations and fierce market competition have led to margins becoming increasingly thinner for manufacturers of consumer and commercial vehicles. These actors realize now more than ever, that the value of their goods no longer rests on the basic functions they provide, but rather on the types and qualities of user experiences they can offer: extra horse-power on demand, ability to share usage, selling data streams to third parties – to name a few.

Increasingly, manufacturers are exploring ways to capture this value by turning a vehicle into a mini-economic platform that facilitates value exchange. Usage of that platform must be controlled so that value creation and consumption are neither impeded, nor corrupted, for the tenants that interact on it.

Our R & D in policy-based access control, distributed ledger technology, and embedded systems has led to the development of FROST Technology for fully programmable sharing ecosystems and flexible usage control on a vehicle's compute systems. FROST can thus provide consumers with novel, on-demand services whilst enabling manufacturers to tap into additional revenue streams.

4.24 Distributed Blockchain Systems across Distributed Data Centers

Dilip Krishnaswamy (Reliance Jio Infocomm Ltd., IN)

Emerging 5G and future networks will be realized leveraging programmable infrastructure that utilizes VMs and containers across hierarchical / distributed data centers. For transaction processing in such distributed deployments, distributed blockchain systems will need to be supported with consideration for data communication latency and bandwidth availability across these data centers. It would be interesting if a distributed blockchain system can be designed with lazy decoupling of blockchain ledgers that has a transaction throughput performance (tps) closer to a performance that is achievable in a local data center, while meeting the end-to-end latency constraint requirements across the distributed data centers over which the blockchain system is deployed. In particular, as edge data centers get deployed to provide support for latency sensitive applications (such as video streaming, Virtual and Augmented Reality applications, healthcare services, data privacy related services, financial applications, retail services, telecommunications services, etc.) at the edge of the network, transaction data will be produced at the edge where such transaction data will need closure in short time scales. Therefore a distributed producer-consumer blockchain framework with a scalable microservices-based approach is suggested in [1], where a producer blockchain sub-network commits transaction data locally, and then eventually commits the data to

subscribing consumer remote ledgers with additional latency. Thus remote ledgers are only eventually consistent in such cases. For applications that are not latency-constrained, one can continue processing blockchain transaction data over a wider-area-network with reduced throughput. In general, based on the latency and throughput constraints that need to be met, one can choose to utilize such edge distributed ledger systems that synchronize lazily with remote blockchain ledgers, if desired.

References

1 D. Krishnaswamy at al, "A Microservices-based Virtualized Blockchain Framework for Emerging 5G Data Networks", IEEE Global Communications Conference (Globecom) 2019, Hawaii, December 2019.

4.25 Atomic Cross Chain Swaps

Eric Lo (The Chinese University of Hong Kong, HK)

Joint work of Eric Lo, Ziliang Lai, Lucian Ng, Sherman Chow, Yongyun Zhao

Since the birth of Bitcoins, thousand of new blockchains emerges. Allowing exchanges of digital currency and goods between blockchains helps users to enjoy benefits from different blockchains and improves the liquidity. To this end, we need a mechanism where multiple untrusted parties can exchange assets on different blockchains in an all-or-nothing manner, i.e., atomic cross-chain swaps. However, reaching consensus across different blockchains is challenging. Two outstanding issues are how to ensure all the blockchains 1) agree on swapping on not and 2) faithfully execute the swap protocol.

A native and common approach is running an exchange center to provide such service. However, such an approach violates the decentralized nature of blockchains since it places trust in the service provider. A common solution is to use smart contracts to escrow assets. Combining with hashed timelocks, a party holding a secret can decide swapping assets or not. However, hashed timelocks require synchronous clocks on different blockchains, which is missing in most blockchains. In this seminar, I introduced several solutions to attack this problem.

Collaborators: Lucien Ng, Sherman Chow, Yongjun Zhao, ZiLiang Lai

4.26 Blockchain Analytics

Murat Kantarcioglu (University of Texas – Dallas, US)

Joint work of Murat Kantarcioglu, Yulia Gel, Cuneyt Akcora
Main reference Cuneyt Gurcan Akcora, Yitao Li, Yulia R. Gel, Murat Kantarcioglu: "BitcoinHeist: Topological Data Analysis for Ransomware Detection on the Bitcoin Blockchain", CoRR, Vol. abs/1906.07852, 2019.
URL http://arxiv.org/abs/1906.07852

In this talk, we give an overview of the blockchain data analytics [3] where transactions recorded on blockchains such as Bitcoin can be represented as a heterogeneous graph [2] and then different graph patterns named chainlets [1] can be mined for predicting cryptocurrency prices [1] to detecting ransomware activities [4]. In addition, we briefly discuss why

some of the existing graph analytics techniques could not be directly applied for blockchain transaction graphs.

References

1 Cuneyt Gurcan Akcora, Asim Kumer Dey, Yulia R. Gel, and Murat Kantarcioglu. Forecasting bitcoin price with graph chainlets. In *Advances in Knowledge Discovery and Data Mining – 22nd Pacific-Asia Conference, PAKDD 2018, Melbourne, VIC, Australia, June 3-6, 2018, Proceedings, Part III*, pages 765–776, 2018.

2 Cuneyt Gurcan Akcora, Yulia R. Gel, and Murat Kantarcioglu. Blockchain: A graph primer. *CoRR*, abs/1708.08749, 2017.

3 Cuneyt Gurcan Akcora, Murat Kantarcioglu, and Yulia R. Gel. Blockchain data analytics. In *IEEE International Conference on Data Mining, ICDM 2018, Singapore, November 17-20, 2018*, page 6, 2018.

4 Cuneyt Gurcan Akcora, Yitao Li, Yulia R. Gel, and Murat Kantarcioglu. Bitcoinheist: Topological data analysis for ransomware detection on the bitcoin blockchain. *CoRR*, abs/1906.07852, 2019.

4.27 Blockchain and Open Source Governance

Juho Lindman (University of Gothenburg, SE)

Early public blockchain cryptocurrency projects (such as Bitcoin and Ethereum) are licensed under open source licenses and governed openly in developer communities using governance mechanisms, practices and tools inherited from the open source world. Early governance ideals of these blockchain projects followed closely the governance of OSS such as Linux operating system or Apache Web Server. In my research I am investigating whether earlier open source software (OSS) research can help us to explain blockchain-related phenomena. OSS research offers several insights that may be reusable in Blockchain context regarding how to solve different kinds of tension between voluntary (developer) communities and profit-seeking commercial companies. The openness of artifact is obviously an interesting point of departure, but more critical questions may be related to guaranteeing the incentives of the different actors and matching divergent interests and misaligned incentives. The re-emergence of governance foundations and forks – splits of the development community – also raise questions where OSS may provide some analytical tools to raise the analytical clarity.

5 Findings

5.1 General Conclusions

Blockchain technology has many connections to database technology, yet the current situation is comparable to that in which the database field was in the 1980s: There were some commercial systems already, but users had to figure out for themselves how they could efficiently and effectively be used. As the seminar has shown, a number of important questions still have to be answered for blockchains, including but not limited to the following

- Blockchain scalability and performance w.r.t. transaction throughput is one of the core hurdles enroute to a wider application of blockchains. While commercial systems, e.g., those run by credit card companies, meanwhile achieve a throughput of 25,000+ transactions per second, blockchain systems can currently offer 15 to 20 transactions per second, due to their high verification effort. As a consequence, participants of a permissioned blockchain have to wait a long time for a transaction commit and hence for progress in the execution of a smart contract.
- Further development of the Proof of Work (PoW) consensus mechanism as well as other optimizations of finding consensus in the presence of transaction failures and w.r.t. scalability. Establishing consensus is of central importance since participants have to reach an agreement on the execution order of running transactions; this has to be identical for every party. The challenge here lies in the fact that while the (potentially large) participants in a Permissioned Blockchain are known, there is not necessarily trust among them, and some nodes may even be faulty (i.e., Byzantine errors have to be tolerated). However, blockchains do not always assume a complete lack of trust of actions among the participants, and so some do not get built on Byzantine-tolerating protocols.
- Languages for the specification of smart contracts are needed for an integration of blockchains into an existing enterprise IT. In addition, it is currently necessary for developers of smart contracts to use languages like Solidity which are low-level.

There is a lot of differences among the way words and terms are used, and what assumptions are made by the players in this field. Also, there is a wide variation in what aspects people are focused on: improving the performance of the protocols, integration with other data and computational platforms, understanding the security and fault-tolerance properties, applications, organizational aspects of managing the platforms, etc. It is also interesting to note that even when a trusted party does exist, there could be organizational constraints (budgets, mandates etc.) that lead to adopting a blockchain as a good architecture in practice.

5.2 More Blockchain Analytics

We saw that blockchains are used to manage both physical and digital assets, e.g., in finance, shipping, or energy (where I work a lot). Blockchains have both strengths and weaknesses, most notably performance, compared to other technologies for managing distributed data and transactions, e.g., database systems. It is thus interesting to investigate the optimal technology mix for certain types of analytics applications. Specifically, it is very interesting to be able to analyze the large amounts of data on blockchains. Research questions include the following:

- How can blockchains be optimally combined with (existing) database and analytics technology for different types of analytical workloads?

- How can data in blockchains be analyzed in powerful and scalable ways, like for data in normal databases?
- Which new, specific types of analyses are needed for blockchain data?

5.3 Virtual Assets

Blockchain technology can be used to establish a whole class of virtual assets, such as individual data, by providing both the "trust" and "incentive" necessary to foster a self-growing value ecosystem. In particular, one can explore more on using blockchain to solve the bottleneck in today's data-driven economy – how we initiate and push along "Symphony Protocol" to create an ecosystem that unlocks personal data for democratized and personalized intelligence, with privacy by design. The domain presents a wealth of interesting research questions, such as data pricing and trading.

5.4 Areas for Future Work

Among the activities for future work in the area of (permissioned) blockchains, participants of the seminar suggested the following:

- Foundation of an Academic Research Special Interest Group: The goal of this group is to be a forum for academic researchers in Hyperledger. We want researchers that are interested in Hyperledger or Hyperledger-related topics to be able to interact and collaborate on problems. This might take the form of presentations, discussions, or collaborative working sessions. We also will incorporate bidirectional communication with developers and engineers.
- Writing of a "Blockchain Manifesto" which helps clarifying the terminology used in this area and sorts out as well as organizes the main directions of development by which the field is characterized. Several participants have expressed their interest in contributing to such an endeavor.

Acknowledgements We want to thank the Dagstuhl staff for providing an environment that truly encourages scientific exchange and discussions and that cares for the guests in an unmatched way.

Participants

- Divyakant Agrawal
University of California – Santa Barbara, US
- Alysson Neves Bessani
University of Lisbon, PT
- Jeeta Ann Chacko
TU München, DE
- Lei Chen
HKUST – Kowloon, HK
- Mariano P. Consens
University of Toronto, CA
- Tien Tuan Anh Dinh
Singapore University of Technology and Design, SG
- Alan Fekete
The University of Sydney, AU
- Michael J. Franklin
University of Chicago, US
- Vincent Gramoli
The University of Sydney, AU
- Krishna P. Gummadi
MPI-SWS – Saarbrücken, DE
- Michael Huth
Imperial College London, GB
- Hans-Arno Jacobsen
TUM, DE & Univ. Toronto, CA
- Murat Kantarcioglu
University of Texas – Dallas, US
- Srinivasan Keshav
University of Waterloo, CA
- Shahan Khatchadourian
Toronto, US
- Dilip Krishnaswamy
Reliance Jio Infocomm Ltd., IN
- Juho Lindman
University of Gothenburg, SE
- Eric Lo
The Chinese University of Hong Kong, HK
- Alexander Löser
Beuth Hochschule für Technik – Berlin, DE
- Dumitrel Loghin
National University of Singapore, SG
- Bernhard Mitschang
Universität Stuttgart, DE
- C. Mohan
IBM Almaden Center – San Jose, US
- Hart Montgomery
Fujitsu Labs of America Inc. – Sunnyvale, US
- Pezhman Nasirifard
TU München, DE
- Beng Chin Ooi
National University of Singapore, SG
- Torben Bach Pedersen
Aalborg University, DK
- Dennis Przytarski
Universität Stuttgart, DE
- PingCheng Ruan
National University of Singapore, SG
- Gabriela Ruberg
Banco Central do Brasil – Rio de Janeiro, BR
- Mohammad Sadoghi Hamedani
University of California – Davis, US
- Yong Tang
Univ. of Electronic Science & Technology – Chengdu, CN
- Gottfried Vossen
Universität Münster, DE
- Li Xiong
Emory University – Atlanta, US
- Feida Zhu
SMU – Singapore, SG