HPChain: An MPI-Based Blockchain Framework for Data Fidelity in High-Performance Computing Systems

Abdullah Al-Mamun[1,2], **Tonglin Li**[2], **Mohammad Sadoghi**[3], **Linhua Jiang**[4], **Haoting Shen**[1], **and Dongfang Zhao**[1,3]

[1]University of Nevada, Reno [2]Lawrence Berkeley National Laboratory [3]University of California, Davis [4]Fudan University

Abstract

Data fidelity is of prominent importance for scientific experiments and simulations, as the data upon which scientific discovery rests must be trustworthy and retain its veracity at every point in the scientific workflow. The state-of-the-art mechanism to ensure data fidelity is through data provenance, which keeps track of the data changes and allows for auditing and reproducing scientific discoveries. However, the provenance data itself may as well exhibit unintentional human errors and malicious data manipulation. To enable a trustworthy and reliable data fidelity service, we advocate achieving the immutability and decentralization of scientific data provenance through blockchains. The challenges of leveraging blockchains in high-performance computing (HPC) are two folds. Firstly, the HPC infrastructure exhibits incompatible characteristics to the targeting platform of existing blockchain systems; Secondly, HPC's programming model MPI alone cannot meet the reliability requirements expected by blockchains. To this end, we propose HPChain, a new blockchain framework specially designed for HPC systems. HPChain employs a new consensus protocol compatible with and optimized for HPC systems. Furthermore, HPChain was implemented with MPI and integrated with an off-chain distributed provenance service to tolerate the failures caused by faulty MPI ranks. The HPChain prototype system has been deployed to 500 cores at the University of Nevada's HPC center and demonstrated strong resilience and scalability while outperforming state-of-the-art blockchains by orders of magnitude; we are working on deploying HPChain to the Cori supercomputer hosted at the Lawrence Berkeley National Laboratory.

Motivation

1. Lack of data fidelity in conventional scientific data provenance.
2. Leverage blockchains to ensure the immutability and reliability.
3. Existing consensus protocols are either based on
   - Intensive computation (e.g., proof-of-work, or POW)
   - Intensive network communication (e.g., practical Byzantine fault tolerance, PBFT)

State-of-the-Art

- Permissioned blockchains: Hyperledger [1], Inkchain [2].
- Public blockchains: Bitcoin [3], Ethereum [4].
- In-memory blockchains: IMB [5].
- Blockchain simulators: Vibes [6], BlockLite [7].

Open Challenges

1. Feasible consensus protocols for HPC blockchains.
2. Parallel block processing.
3. Resilient distributed ledgers implemented with MPI.

Design Objectives

1. A set of parallel consensus protocols with double layers of block validation.
2. A parallel mechanism for processing the data blocks through MPI.
3. A resilience mechanism.
4. A system prototype with OpenMP and MPI.

Architecture

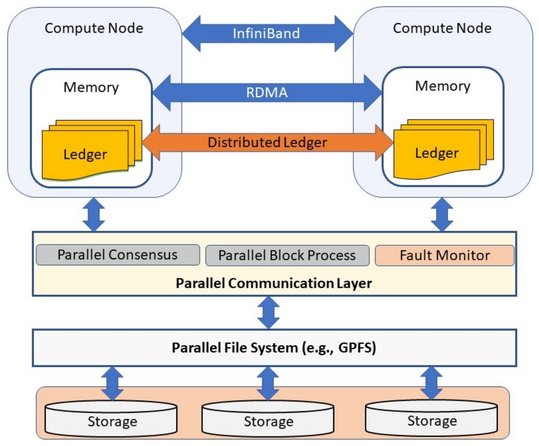

Figure 1: Architecture of HPChain in HPC systems.

Consensus Protocol Workflow

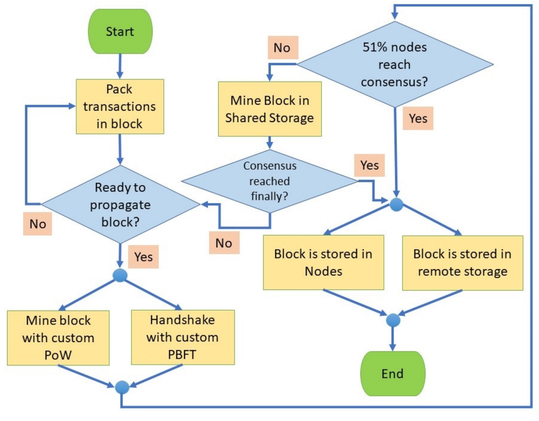

Figure 2: HPChain Protocol

Preliminary Results

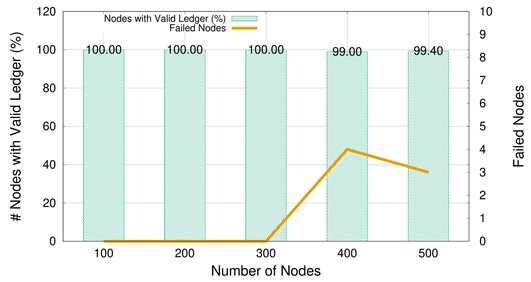

Figure 3: Ledger Validity in HPChain.

- Figure 3 shows at all scales at least 99% nodes keep the valid ledger.
- 100% nodes hold valid ledgers on up to 300 nodes.
- Reminder: A blockchain can work correctly as long as at least 51% of the nodes are not compromised. (HPChain raises the bar ensuring 99% reliable nodes on up to 500 nodes)

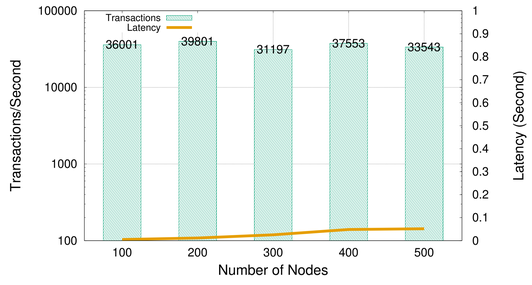

Figure 4: HPChain Performance.

- Figure 4 shows performance, in terms of throughput and latency, at different scales ranging from 100 to 500 nodes.
- HPChain exhibits insignificant gap in throughput and incurs negligible latency degradation (i.e., below 0.1 second on up to 500 nodes).
- Recall that scalable performance remains one of the most challenging issues in blockchain systems [8]. HPChain seems scalable on up to 500 nodes.

Conclusion

- HPChain represents the first blockchain system tailed to HPC.
- HPChain employs a set of new consensus protocols compatible with HPC architecture.
- HPChain delivers scalable performance on up to 500 nodes.

Future Work

1. Evaluate HPChain with real-world scientific applications at extreme scales.
2. Develop off-line distributed provenance for HPChian to tolerate MPI's faulty rank(s).

References

[1] Hyperledger. https://www.hyperledger.org/, Accessed 2018.

[2] Inkchain. https://github.com/inklabsfoundation/inkchain, Accessed 2018.

[3] Bitcoin. `https://bitcoin.org/bitcoin.pdf`, Accessed 2018.

[4] Ethereum. https://www.ethereum.org/, Accessed 2018.

[5] Abdullah Al-Mamun, Tonglin Li, Mohammad Sadoghi, and Dongfang Zhao. In-memory blockchain: Toward efficient and trustworthy data provenance for hpc systems. In *2018 IEEE International Conference on Big Data (Big Data)*, pages 3808–3813. IEEE, 2018.

[6] Lyubomir Stoykov, Kaiwen Zhang, and Hans-Arno Jacobsen. Vibes: Fast blockchain simulations for large-scale peer-to-peer networks: Demo. In *Proceedings of the 18th ACM/IFIP/USENIX Middleware Conference (Middleware)*, 2017.

[7] Xinying Wang, Abdullah Al-Mamun, Feng Yan, and Dongfang Zhao. Toward accurate and efficient emulation of public blockchains in the cloud. In *12nd International Conference on Cloud Computing (CLOUD)*, 2019.

[8] Kaiwen Zhang and Hans-Arno Jacobsen. Towards dependable, scalable, and pervasive distributed ledgers with blockchains. In *38th IEEE International Conference on Distributed Computing Systems (ICDCS)*, 2018.

Acknowledgement

This work is in part supported by DOE, NSFC, NASA, and Google.

Contact Information

Email: dzhao@unr.edu