

# Big Data and Future of Block Chain Technology

Mrs. Sukeshini Satish Gawai<sup>1</sup> and Mrs. Archana Vinod Bansod<sup>2</sup>

Lecturer, Y B Patil Polytechnic, Akurdi, Pune<sup>1</sup>

Lecturer, Y B Patil Polytechnic, Akurdi, Pune<sup>2</sup>

**Abstract:** Block chain is a type of database, but it differs entirely from traditional databases in how it stores and manages information. Instead of storing data in rows, columns, tables and files as traditional databases do, blockchain stores data in blocks that are digitally chained together. Moreover, a blockchain is a decentralized database managed by computers belonging to a peer-to-peer network instead of a central computer like in traditional databases and it is an immutable and secure digital ledger that records transactions permanently and verifiably. Bitcoin and Ethereum are examples of such blockchains. Blockchain and Big Data are two most advanced technologies that have been significantly transforming various industries. Block chain & big data integration can increase security, data integrity, and privacy efficiently than the traditional systems.

**Keywords:** Block chain, Big data, Cryptographic technique, Distributed Denial-of-Service (DDoS)

## I. FUTURE OF BLOCKCHAIN TECHNOLOGY

Worldwide organizations are using the blockchain technology to create digital currencies, it increases the transparency of data and lower the chances of potential fraud. It even allows customers to gain access to cloud-based solutions to host, build, and use blockchain apps.

Use Cases for Blockchain and Big Data Integration

### 1. Healthcare:

Blockchain can provide secure, decentralized storage for patient records, ensuring data integrity and privacy. Big Data analytics can then be applied to this data to derive insights into patient care, disease trends, and more.

### 2. Supply Chain Management:

Blockchain can track products as they move through the supply chain, ensuring the data collected about each product is accurate, transparent, and immutable. The Big Data system can analyze this data in real-time to optimize supply chain operations.

### 3. Finance and Banking:

Blockchain can ensure secure, transparent financial transactions, while Big Data can analyze customer behavior, market trends, and risk profiles. Together, they enable better decision-making, fraud detection, and compliance.

### 4. Energy Management:

Blockchain can securely track and manage energy transactions (such as from solar panels or wind turbines), while Big Data can be used to analyze energy consumption patterns, predict demand, and optimize energy distribution.

Example: Power Ledger uses blockchain to track renewable energy trading, while Big Data analytics optimizes the grid and predicts energy consumption.

### 5. IoT (Internet of Things):

IoT devices generate vast amounts of data, which can be securely stored on the blockchain to ensure data integrity.

Example: In smart cities, blockchain can manage data from IoT devices like traffic sensors or environmental monitors, while Big Data helps improve urban planning, traffic flow, and resource allocation

## II. BLOCKCHAIN AND BIG DATA INTEGRATION ARCHITECTURE

To integrate blockchain with big data effectively, you need an architecture that balances the strengths of both technologies while addressing their limitations.



**1. Data Ingestion & Collection Layer**

Data from various sources (IoT sensors, social media, web logs, etc.) is ingested into the Big Data platform through tools like Apache Kafka, Apache Flume, or AWS Kinesis.

The data can be processed and analyzed in real-time or batch mode using tools like Apache Spark or Apache Flink.

**2. Blockchain Integration Layer**

In this layer, blockchain is used to provide tamper-proof records of the data processed in the Big Data system. Instead of storing entire datasets, the blockchain stores metadata, transaction hashes, or pointers to the off-chain data.

Smart contracts can be deployed here to enforce business logic for data transactions and automate workflows.

**3. Big Data Storage Layer**

Big Data is typically stored in distributed storage systems like HDFS, NoSQL databases (e.g., MongoDB, Cassandra), or cloud-based storage like AWS S3 or Google Cloud Storage.

Large datasets that need to be queried and analysed are kept off-chain to avoid blockchain storage constraints.

**4. Data Processing & Analytics Layer**

Big Data processing and analytics tools (e.g., Apache Spark, Hadoop, TensorFlow) analyze the ingested data to derive insights, predictions, and actionable business intelligence. These tools can be integrated with blockchain to ensure that any insights derived from the data are linked to tamper-proof records for auditability and trust.

**5. Visualization and Reporting Layer**

Tools like Tableau, Power BI, or Apache Superset allow users to visualize insights from the Big Data processing layer. Blockchain metadata can be used to validate and verify the insights shown on the dashboards, ensuring that the data's provenance and integrity are auditable.

**III. OVERVIEW OF HOW BLOCKCHAIN CAN ENHANCE THE SECURITY OF BIG DATA**

Blockchain offer decentralized storage solutions for Big Data, where multiple copies of data exist across different nodes in the blockchain network. This reduces the risks associated with centralized storage, such as single points of failure and data breaches. Blockchain ensures that once data is written, it cannot be tampered with. Each block in the blockchain contains a cryptographic hash of the previous block, forming an immutable chain. This makes it perfect for securing Big Data, where integrity and traceability are critical. For instance, in financial transactions or healthcare records, Big Data systems benefit from blockchain's ability to ensure that records are unalterable, providing transparent data access and audit. Blockchain leverages advanced cryptographic techniques, such as public-private key pairs and digital signatures, to ensure the privacy and confidentiality of data stored and shared in Big Data systems. Users can retain control over their own data by giving permissions for specific data access without exposing everything to third parties. This is particularly valuable in sectors such as healthcare or finance, where data privacy regulations are stringent. Distributed Denial-of-Service (DDoS) attacks and other cyber threats pose a significant risk to centralized Big Data systems. By decentralizing data management and storage, blockchain makes it much harder for attackers to target and disrupt the system.

**IV. CONCLUSION**

The concept of blockchain aid in the management of enormous amounts of data and efficiently secure very large quantities of data, which can be highly beneficial for government organization. Centralized servers having a high risk of data hacking, loss, or human errors. Cloud storage can be made more secure and robust against hacking with the implementation of Blockchain technology, just like its application in cybersecurity. The use of blockchain can reduce time delays and human errors and monitor employment, costs, and releases at each step of the supply chain. Through traceability, Blockchain can also ensure the fair-trade status and legitimacy of products. Digital advertising faced challenges including bot traffic, lack of transparency, domain fraud, inefficient payment models, etc, Blockchain has been found to resolve such issues in the supply chain. Advertisement-related transactions can be better handle with by employing this technology. Blockchain technology was firstly developed for cryptocurrencies like Bitcoin, but its applications have extended to various industries, including finance, supply chain, healthcare, and more. Its decentralized and secure nature makes it attractive for scenarios where trust, transparency, and immutability of records



are very important. Blockchain technology will soon attract all business sectors and worldwide organizations to invest more in it.

## REFERENCES

- [1]. S 4. Wu J, Ota K, Dong M, et al. Big data analysis-based security situational awareness for smart grid. IEEE Transactions on Big Data 2018; 4(3): 408–417. doi: 10.1109/TBDATA.2016.2616146 5. Tole AA.
- [2]. Big data challenges. Database Systems Journal 2013; 4: 31–40. 6. Liu CH, Lin Q, Wen S. Blockchain-enabled data collection and sharing for industrial IoT with deep reinforcement learning. IEEE Transactions on Industrial Informatics 2019; 15(6): 3516–3526. doi: 10.1109/TII.2018.2890203
- [3]. Liu G, Dong H, Yan X, et al. B4SDC: A blockchain system for security data collection in MANETs. IEEE Transactions on Big Data 2022; 8(3): 739–752. doi: 10.1109/TBDATA.2020.2981438
- [4]. Nakamoto S. Bitcoin: A peer-to-peer electronic cash system. In: 2018 Annual National Seminar; 30 May–1 June 2018.
- [5]. Bernabe JB, Canovas JL, Hernandez-Ramos JL, et al. Privacy-preserving solutions for blockchain: Review and challenges. IEEE Access 2019; 7: 164908–164940. doi: 10.1109/ACCESS.2019.2950872
- [6]. Parkin J. The senatorial governance of bitcoin: Making (de) centralized money. Economy and Society 2019; 48(4): 463–487. doi: doi.org/10.1080/03085147.2019.1678262
- [7]. Casino F, Dasaklis YK, Patsakis C. A systematic literature review of blockchain-based applications: current status, classification and open issues. Telematics and Informatics 2019; 36: 55–81. doi: 10.1016/j.tele.2018.11.006
- [8]. Zhang J, Zhong S, Wang T, et al. Blockchain based systems and applications: A survey. Journal of Internet Technology 2020; 21: 1–14. doi: 10.3966/160792642020012101001
- [9]. Shalini S, Santhi H. A survey on various attacks in bitcoin and cryptocurrency. In: 2019 International Conference on Communication and Signal Processing (ICCSP); 4–6 April 2019; Chennai, India. pp. 0220–0224.
- [10]. Huang H, Kong W, Zhou C, et al. A survey of state-of-the-art on blockchains: Theories, modelings, and tools. arXiv 2020; arXiv:2007.03520. doi: 10.48550/arXiv.2007.03520
- [11]. Dey AK, Akcora CG, Gel YR, Kantarcioglu M. On the role of local blockchain network features in cryptocurrency price formation. The Canadian Journal of Statistics 2020; 48(3): 561–581. doi: 10.1002/cjs.11547
- [12]. Angelis J, da Silva ER. Blockchain adoption: A value driver perspective. Business Horizons 2019; 62(3): 307–314. doi: 10.1016/j.bushor.2018.12.001
- [13]. 17. Chen Y, Bellavitis C. Blockchain disruption and decentralized finance: The rise of decentralized business models. Journal of Business Venturing Insights 2020; 13: e00151. doi: 10.1016/j.jbvi.2019.e0015
- [14]. . Moin S, Karim A, Safdar Z, et al. Securing IoTs in distributed blockchain: Analysis, requirements and open issues. Future Generation Computer Systems 2019; 100: 325–343. doi: 10.1016/j.future.2019.05.023
- [15]. . Zhang R, Xue R, Liu L. Security and privacy on blockchain. ACM Computing Surveys 2019; 52(3): 1–34. doi: 10.1145/3316481
- [16]. Leonardos S, Reijbergen D, Piliouras G. PREStO: A systematic framework for blockchain consensus protocols. IEEE Transactions on Engineering Management 2020; 67(4): 1028–1044. doi: 10.1109/TEM.2020.2981286

