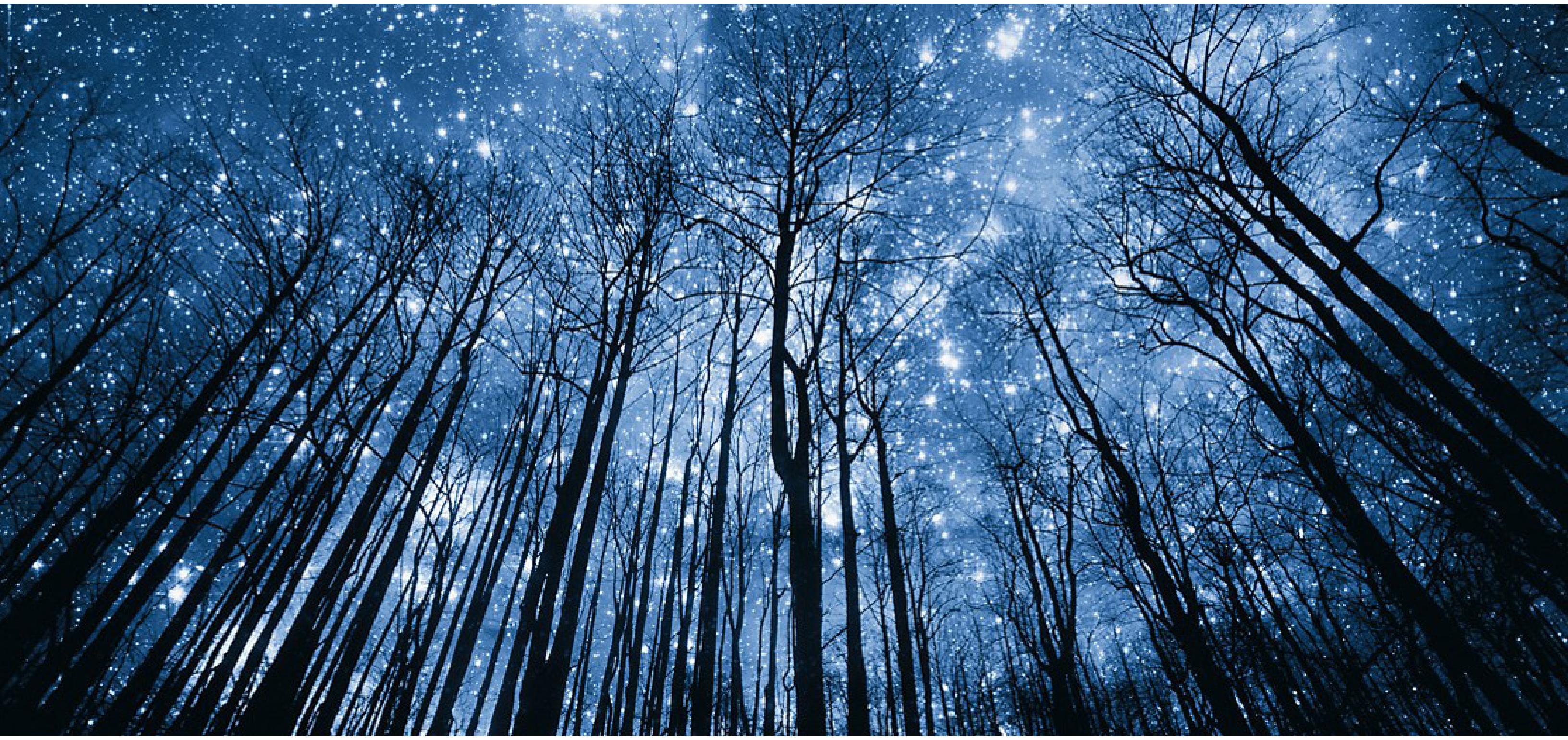
PROFICIENT STORAGE FOR BLOCKCHAIN 4.0



Saiganesh Venkataraman

Sagar Sajeev

Shruthi Krishnan
Specialist 2, Inside Product
Dell Technologies

Knowledge Sharing Article
© 2023 Dell Inc. or its subsidiaries.

The Dell Technologies Proven Professional Certification program validates a wide range of skills and competencies across multiple technologies and products.

From Associate, entry-level courses to Expert-level, experience-based exams, all professionals in or looking to begin a career in IT benefit from industry-leading training and certification paths from one of the world's most trusted technology partners.

Proven Professional certifications include:

- Cloud
- Converged and Hyperconverged Infrastructure
- Data Protection
- Data Science
- Networking
- Security
- Servers
- Storage

Courses are offered to meet different learning styles and schedules, including self-paced On Demand, remote-based Virtual Instructor-Led and in-person classrooms.

Whether you are an experienced IT professional or just getting started, Dell Technologies Proven Professional certifications are designed to clearly signal proficiency to colleagues and employers.

Learn more at www.dell.com/certification

Contents

Abstract 4
Introduction 5
History of Blockchain 5
Ownership in Blockchain Technology 7
Data Process Flow of a Blockchain Node Owner 8
Progression of Blockchain Scheme 10
Basic Hardware Design 11
Expansive Infrastructure Implementation 12
The Good: Virtuous Aspects of Blockchain Technology 14
The Bad & The Ugly: Deficient Aspects of Blockchain Technology 16
Proposed Environmental Blueprint 17
Primary Storage Strategy 18
Data Protection Strategy 19
Conclusion: Blockchain Transformation in Future 22
Bibliography 23

Abstract

Blockchain has deftly crafted its relevance in multiple verticals and will soon play a vital role in ensuring accountability across the internet. It has advanced from Blockchain 1.0 to 3.0 by evolving from being a simple distributive ledger technology to providing a platform for the development of Decentralized Application (dApps). Blockchain 4.0 aims to address the drawbacks of the previous versions by allowing Blockchain to be integrated with other technologies like Artificial Intelligence. As the shift to Blockchain 4.0 is underway, theoretical use cases of the revolutionary technology will be incorporated into mainstream business platforms under a unified umbrella of several enterprise blockchain models. With Web 3.0 aiming to provide a platform that is an intelligent internet, open, autonomous and will rely on decentralized protocols, Blockchain 4.0 becomes the best solution as it provides better accessibility to users and promises scalability. Blockchain 4.0 ecosystems are forecasted to provide an Integrated Development Framework for the development of dApps. Blockchain-as-a-Service providers help provide the environment for the development of dApps by helping manage and maintain the IT ecosystem for enterprises keen on developing dApps.

Blockchains rely on distributed networks which require component 'nodes' being connected furthermore, responsive and performant. Each of this Transactional information is stored on nodes that form the distributed network. The data itself can be stored on the user's hard drive in a Bare metal configuration; however, we can leverage a virtual server on a Cloud computing network if considering a Cloud infrastructure model. Data on these nodes increase over time as more Blockchain transactions are completed. Hence resulting in a rising concern for storage requirements not being met and this requires a scalable solution. With more benefactors investing Blockchain, it has become pivotal to inculcate better storage infrastructure to maintain the high availability and consistent service of Blockchain architecture.

Blockchain Technology is a mix of several current prominent technologies. The Current Blockchain Technology can be implemented over Bare Metal infrastructure or Cloud infrastructure as well. However, the technology also lacks disaster recovery measures that ensure a successful failover of the ledger data and lack cyber resilience. All these concerns can be addressed through the infrastructure storage portfolios across various enterprises in the industry C the vital role that infrastructure storage providers can play in architecting the environment for BaaS providers.

Introduction

Blockchain technology has risen over the last few decades, paving the path for concepts such as decentralized accountability and peer-owned networks to the forefront. With the rampant growth, we have observed the technology parse through three generations that have defined the future for contractual obligations, finance, and the logistical sector. However, the hardware requirements that the technology needs have retarded its potential growth. The future is now heading in the direction of Decentralized Applications, and infrastructure providers for Blockchains called, Blockchain-as-a-Service. We must architect better infrastructure than the components that make up mining rigs currently. Hence to ensure that the technology maintains its momentum, infrastructure storage groups across various enterprises must strive to solution products that are purpose-built for the Blockchain industry. With Blockchain 4.0 around the corner, it is essential to stay ahead of the curve by welcoming the new generation with hardware components that are powerful enough to run Blockchain services.

History of Blockchain

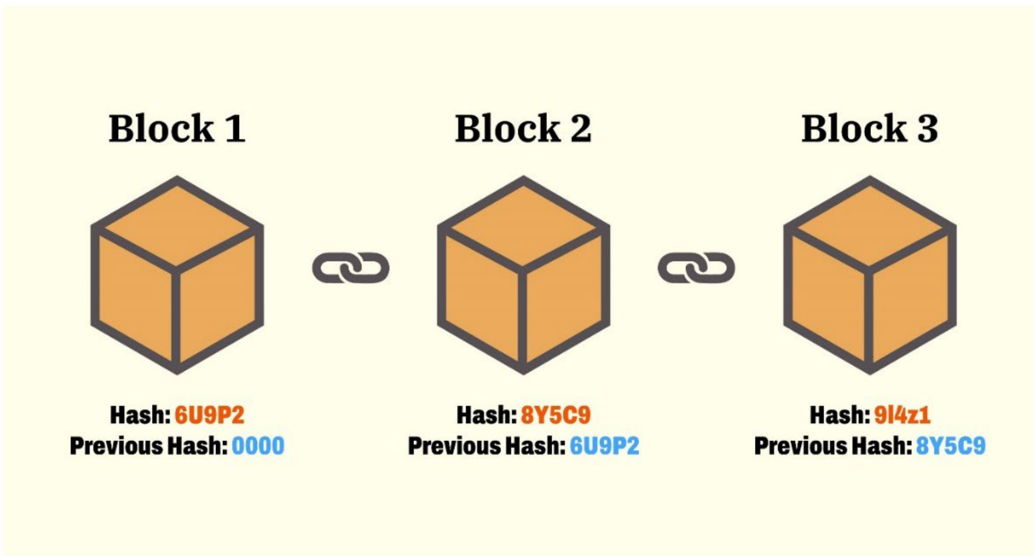


Fig 3.1. Hashed Blocks of Transactional Data

Even though blockchain gained traction only recently, its history dates back to the 1990s. Since its popularity, various applications have been born, highlighting the impact it would have in the current and future industries. Blockchain started with the vision of Stuart Haber and W. Scott Stornetta in the year 1991. They first worked on a chain of blocks that was cryptographically secure, something which no one could tamper. Stornetta and Haber were cited many times in Satoshi Nakamoto's original whitepaper called 'Bitcoin' years later. In 1992 they worked on renewing their system and upgrading it to include Merkle trees.

Merkle trees are also known as 'Binary hash trees,' a type of data structure in computer science. They are also used to encrypt blockchain data more efficiently. These Merkle trees also store all the transactions in any block by creating a digital footprint of the entire set of transactions. While this happened, blockchain gained more attention only in 2008 because of Satoshi Nakamoto [1]. He is often called the brains behind blockchain technology. Although there is a debate if it was a singular person who led this popular obsession with blockchain or a particular group of people under the name, it is clear that from there, the first application of the digital ledger technology was in place. Satoshi Nakamoto was the one who conceptualized the first blockchain, from where the technology expanded and found its way into many applications and not just cryptocurrencies. An infamous whitepaper was also released the following year [2], where details of the technology being suitable and appropriate to enhance trust digitally were explained, giving credit to the decentralized aspect of blockchain. This meant that it was clear that no singular body could have authority over the nodal transactions and ledger data.

There is a common misconception that bitcoin and blockchain are the same. But this is not the case: one is the underlying technology, and the other is an actual application based on blockchain. The term bitcoin also existed in 2008 as blockchain's first application. It found its mention in the whitepaper as an 'electronic peer-peer system. After the success of the idea surrounding bitcoin, other applications came into existence that aimed to leverage the capabilities of the digital ledger technology. Hence, it is safe to say that the first evolution following the introduction of blockchain technology was the emergence of bitcoin.

Ethereum is one of the most popular blockchain platforms in the market, and many would be familiar with it. The second evolution was triggered by the founder of this open source blockchain platform and co-founder of the bitcoin magazine: Vitalik Buterin [3]. He felt the need to add Bitcoin because it did not fully exploit the capabilities of blockchain technology. His mind set on its limitations, he worked on developing the technology further, which gave it the ability to perform additional functions apart from being a peer-peer network. And that is how Ethereum was born in the year 2013. In addition, it came with added advantages, allowing users to record other assets such as contracts. This made it a suitable platform to develop decentralized applications and not just stay limited to being a cryptocurrency.

Once the ability to develop applications grew, many projects were undertaken in the recent past that involved leveraging the blockchain technology as we know it today. A few notable technologies that played an essential part in developing it further include the Internet of Things. In 2015, Hyperledger [4] became an umbrella project of the Linux foundation. It is an open-source project created to champion the development of distributed ledgers in the blockchain. Various well knows tech enterprises have contributed and collaborated with its development ever since it was launched. Hyperledger promotes the usage of blockchain in supporting a critical vertical: global business transactions and banks.

Ownership in Blockchain Technology

Centralization is known to introduce synchronization in how things work. In this model, the control lies with a single authority at the center. In Blockchain, however, the core of its functioning depends on decentralization. To define it in simpler terms, decentralization is shifting power from one central entity to a more distributed system [5]. A few features of such a system include members having equal authority, shared governance, and eliminating a central authority.

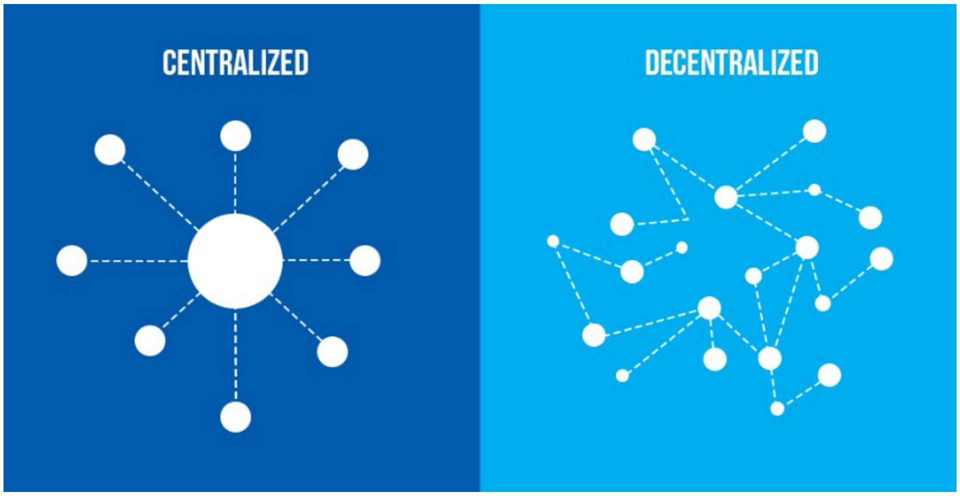


Figure 4.1. Centralized vs Decentralized Network

Blockchain is a decentralized peer-to-peer network; it provides an efficient way to work with unknown members without trusting each other, forming a more secure pathway to share information. On an architectural level, there is no single point of failure. By spreading the data held in the database among several network nodes at various locations, redundancy is created within a blockchain network. If an alteration is attempted at a record at any database instance, the other nodes will not be altered. In cases where a user tampers with the record of transactions in Bitcoin, all the other nodes can cross-reference each other and detect the node with the wrong information. This system establishes a proper and transparent order of events. Thus, the most significant advantage of a decentralized system is that within a network, no single node can alter the information held inside of it.

The unique ability of this method enables Blockchain to be used as a record for a variety of other information like state identifications or even legal contracts. An interesting concept concerning the same technology is called Proof of Work (PoW) [6]. It is a decentralized consensus mechanism that requires the network members to solve a mathematical puzzle to prevent unwanted break-ins. It is widely used in cryptocurrency mining for the validation of transactions.

What are some major benefits of a decentralized system? [7] The first would be transparency. Due to the decentralized nature of Blockchain, all transactions can be viewed by having a node in the network. Since each node has its copy of the chain, bitcoin can be tracked wherever it goes. The copies also get updated as new blocks are added.

Another critical aspect of blockchain technology is that it achieves trust and security in many ways. New blocks are always stored chronologically; they are always added to the end of the Blockchain. Once this is done, it is quite challenging to alter the contents of the block unless everyone on the network gives a consensus to do so. Each block also contains its hash apart from the previous block's hash and the time stamp, forming a 'chain'. If this information is changed, the hash code also changes.

Due to this, hackers will have to spend considerable time, money, and effort as they would need to work backward on the changed blocks again as they would have different hash codes and time stamps. Furthermore, since cryptocurrency is growing at such a fast pace, it would be really difficult to attempt hacking it.

Data Process Flow of a Blockchain Node Owner

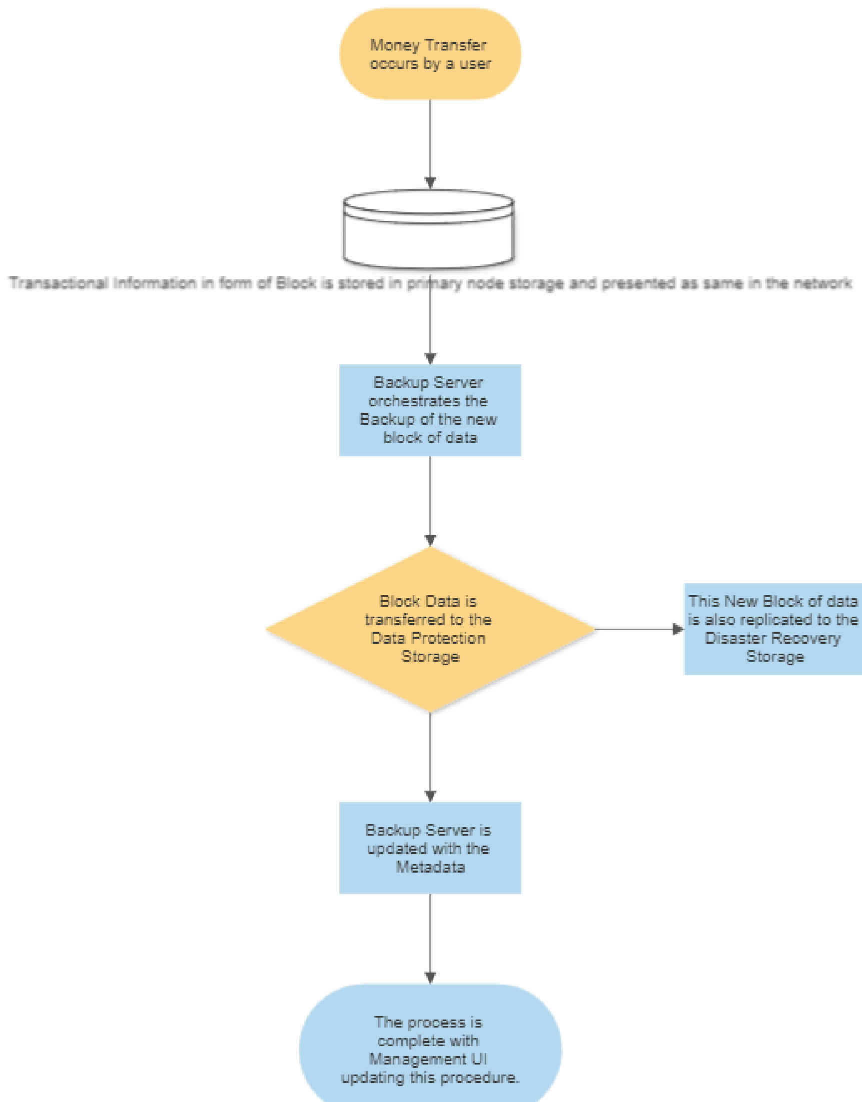


Fig 5.1. Data Flow Chart

From the previous section, we understood the decentralized ownership aspect of the Blockchain technology. However, to maintain discretion and privacy of data that is generated during monetary transactions, the "Miners" need to highlight their efficiency in decrypting the trust algorithm while being in network and with other miners who intend to become a successful node user - this activity is known as the proof of stake or proof of work. Miners participating in proof of stake can edit and change the Transactional ledger being in the Blockchain network. This problem-solving capability has to increase over a period of time as the complexity of the algorithms increases with an increasing number of potential node users and transactional value. This activity is performed on every transaction basis and requires computing power of the nodal system intending to be in the Blockchain network.

Moreover, the Transactional data that the miner once accesses after successfully decrypting the algorithm or "puzzle" can be edited and updated in the local ledger, where this particular data is updated to in the Blockchain network. The transactional data itself is unreadable to the miner and maintains the privacy of the original transactor of the money on the Blockchain network. Once the Transactional data is procured and updated in the ledger, it is locally stored to the miner's compute node, which contains dedicated storage attached to the compute node in terms of low-scale miner but, it can be a dedicated storage system or a network of storage systems to store this local transactional data in terms of large-scale organizations.

Furthermore, Backup Server is the system in the above figure 5.1 that orchestrates the Backup process from the miner's storage system to the Backup storage dedicated and unattached to the primary storage. A Backup server should be deployable virtually and physically as a dedicated computer with attached minor storage. Several methodologies or transfer protocols can be leveraged for this localized yet significant transfer of data. In an optimal backup strategy, newly generated data is protected daily when the production level activity, i.e., in our terms, "mining," is not performed. Utilizing this Backup Strategy, new data blocks are transferred efficiently to Backup storage. Efficient strategies to transmit this data can include deduplication, compression, and other novel data transfer concepts. As a result, efficient Backup architectures can perform backups in a faster manner under a lesser time period with least overhead to the network when compared to a generic data transfer. The Backup program automatically initiates data transfer from production, which is overseen and orchestrated by the Backup Server. After placing blocks of new data in backup storage, the Backup server updates the metadata about the respective backup activity.

Users will also leverage replication storage as a further addition to better protection and high availability strategy. The replication storage will hold copies of the new blocks of backup data. Again, the transfer of replicated data can be efficiently managed using protocols or intelligent protocols mentioned in the earlier portion of this paragraph. The replication job can be performed concurrently, or an automated strategy can be positioned considering time management or quantity of data. The activities in the figure 5.1 are culminated by finally updating the particular activities through an efficient visual and feedback-based user management system. The user management activities can include:

- Forced termination.
- Overriding strategic protection plans.
- Other features that increment trust and control over the protected data.

Progression of Blockchain Scheme

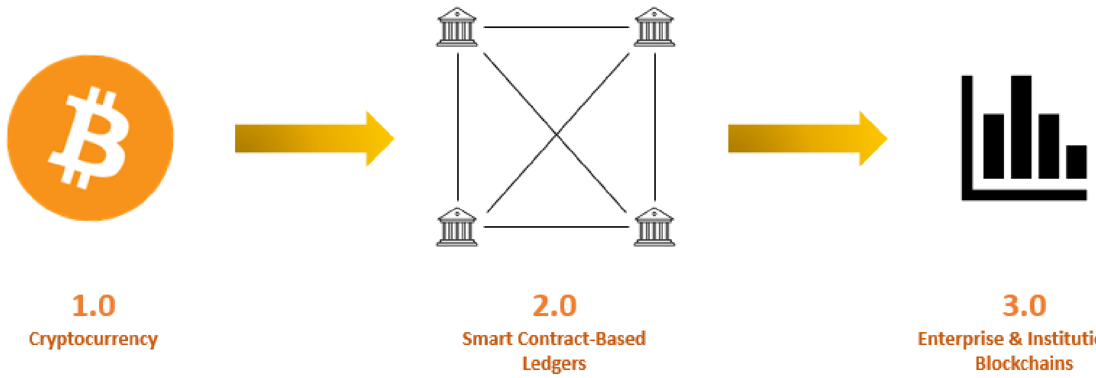


Fig 6.1. Blockchain Evolution Diagram

The evolution of Blockchain ever since its inception has been rapid and well-planned. Around the 1980s and 1990s, Electronic Cash (ECash) was on the rise. The idea behind it was to have a secure, peer-to-peer transaction of a currency that everyone in the network owned. In other words, it was decentralized. This was Blockchain's inception, which was soon considered the first version of its kind [8]. Blockchain 1.0 technology includes components such as a Blockchain Core, respective to the cryptocurrency, wallet software, mining rigs, and mining software. This stage, however, has yet to transform into its next version but has been running parallel to 2.0 and evolving and branching into the world of cryptocurrencies, with many new ones coming up.

With the rise of many cryptocurrencies, the concept of "smart contracts" started to kick in. Smart Contracts are conditional lines of code that tend to execute themselves and are strictly handled by only the buyer and seller themselves through the crypto network. This proved an advantage for many, as it eliminates the requirement for an intermediary to be anywhere near the picture. This drastically reduced the possibility of security breaches and data leakage. The 2.0 stage paved the way for smart contracts, smart property, Decentralized Applications (dApps), DAOs (Decentralized Autonomous Organizations), and DACs (Decentralized Autonomous Corporations).

Blockchain 1.0 & 2.0 have transcended technology to the next level [9]. However, the two versions have a severe drawback, a lack of scalability and time taken to confirm transactions. Blockchain 3.0 aims to make cryptocurrencies globally viable, and to do so; this generation directed its efforts toward the growth of dApps. Decentralized Applications are digital programs run on Blockchain networks, hence not under any central authority's purview. Blockchain 3.0 also utilizes Proof of Authority and Proof of Stake mechanisms to achieve user consensus. The entire generation is based on the "FFM" concept, which stands for Fast, Feeless, and Miner-less networks that can be used to develop decentralized applications. With the rampant increase in the dApps concept, the need for infrastructure to develop it is also rising. With that, we see Blockchain-as-a-

Service (BaaS) providers coming up to meet the developers' infrastructure requirements, creating a new market in the process.

This brings us to the newest rendition of the technology, Blockchain 4.0. This generation aims to deliver the technology as a platform to develop and run applications, hence making this technology mainstream for the public. This holds paramount importance for the future, with the concept of decentralization on the rise. Therefore, the technology enabled populous must lean into this foresight, become proactive and start developing infrastructure that will catalyse the growth of Blockchain 4.0. Furthermore, with the technology taking centre stage over the past decade, any infrastructure aimed at enabling the fourth generation will help the transaction speed to be well over one million transactions per second, which is currently impossible with the infrastructure we possess.

Basic Hardware Design

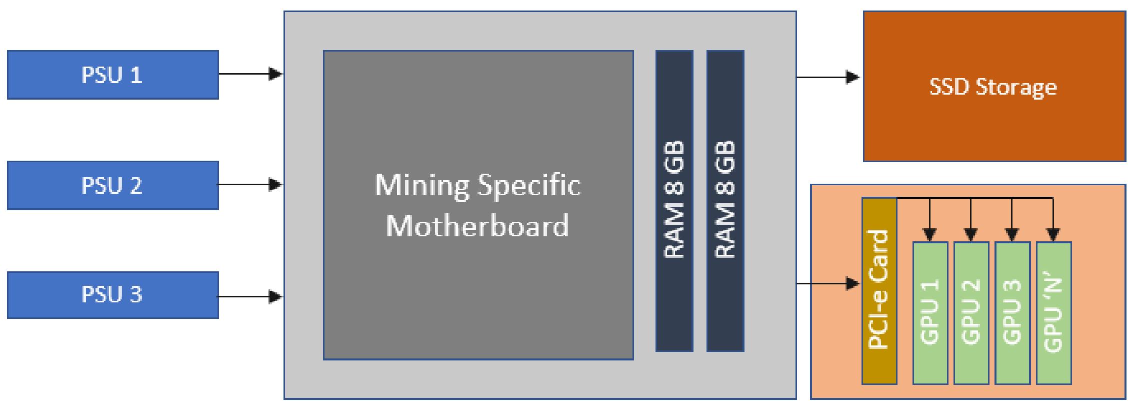


Fig 7.1. Basic Hardware Layout

If one is entering the world of crypto intending to mine their place in it, a few basic requirements must be fulfilled to build a small rig. A rig is nothing but a PC customized to meet a specialized need. However, a mining rig is deviant from the usual norm of PCs. When mining cryptocurrency, the brain on the unit will often be the GPU, as compared to the CPU [10]. This necessitates a powerful GPU that can accommodate the load, and there is a high chance that more than one GPU may need to be purchased. The objective of including this section within the article is to give you a fair understanding of what is required by a singular miner and the complexities and limitations that come in the way when trying to kick-start a large-scale mining operation.

There are a few basic requirements that a miner can quickly scavenge or purchase without any difficulty. However, they must remember that this is not an economical endeavour. Do refer to Fig X to understand what the basic layout will look like. To begin with, one will require a motherboard that has been proposed only for crypto mining. They do not come cheap but are necessary as they can accommodate nearly nineteen graphic cards. Next, we will need a 14nm processor, consisting of four cores and RAMs that surmount to around 16 GB for a grass-root level operation.

The number of PSUs is directly proportional to the number of GPUs that are being installed, and that number may vary from one to nineteen based on the mining requirements. Therefore, it is strongly recommended that one connects the graphic cards through PCIe cards which the motherboard is specifically designed to accommodate.

The final and most significant requirement to be met is storage. With a substantial amount that is needed to store off-chain structured/unstructured data and offline ledgers, proficient storage that promises lightning-quick fetching is required. This aspect completely rules out HDDs from being used, making a miner rely on efficient but expensive SSD storage.

Overall, these are the minimum requirements a new entrant needs in the mining field, but it is an expensive investment. Food for thought, now imagines how much a full-blown, 24x7, completely operational enterprise-level mining rig will have to spend.

Expansive Infrastructure Implementation

In the previous section, we had an insight into the hardware requirements for the low-scale user to begin their foray into Blockchain mining. However, as enterprises grow their infrastructure, utilizing this infrastructure is necessary to run the business continuously. Therefore, several large-scale corporations decided to enter the Blockchain market to provide the miners with the required hardware infrastructure either in an on-prem fashion or "as-a-Service" approach, where most of the Blockchain infrastructure is provided by the cloud model and cloud network. According to the referenced article [11], miners that intend to participate in the proof of stake activity, which is explained previously, must attain proper technical expertise, time, resources, and, most importantly, a resilient infrastructure that can scale and be trusted with stable internet connection for communication and be part of the blockchain network. This paradigm shift in the evolution of Blockchain significantly impacted the type of Network deployment of Blockchain, such as the shift from a Peer-to-Peer network to Global Wide Area access for consortium-type implementation.

Additionally, another important aspect of the large-scale implementation of Blockchain is the network security and hardware security. According to the article by zeeve [12], the Bitcoin core gets affected due to any viruses on the network. Even when an antivirus program is installed on the nodal computer, windows computers are affected while self-deploying the node, a process undertaken by low-scale or standalone miners to establish their presence in the Blockchain network. Physical damages and network threats that can prove costly to a standalone or low-scale miner can be easily averted by big corporations and their infrastructural security measures. Many such reasons spreading across the factors of scalability, security, privacy, and management brought up the demand of bringing a professional and proficient Blockchain infrastructure to many miners interested in contributing to Blockchain network activities.

Hence, this demand gave rise to Blockchain-as-a-Service for many clients across the globe. The figure 8.1 below [13] highlights the involvement of Cloud providers, Clients, and miners who update the Ledger in case of cryptocurrency transactions. The image also shows intelligent contracts govern the complete automated actions amongst the Blockchain network of nodal devices.

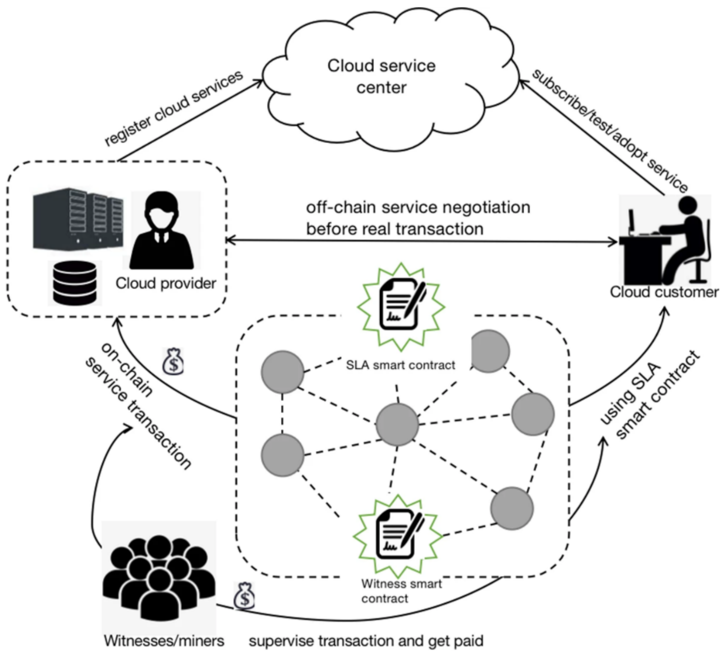


Fig 8.1. Miners contained Cloud deployment methodology.

Although several big game players in the IT industry have established their presence in this market, several others are slowly leveling up their attributes to be sound infrastructure providers in this model. With the advent of Blockchain 4.0, Large Infrastructure farms are configured to keep up with the demands of better power, high-performance levels that can process data blocks quickly, and more robust cooling to keep the equipment that participates in the process from overheating during intense computations. To put into perspective, the largest mining infrastructure facility in the City of Bratsk, Russia, uses 100 Megawatt of power per year to be operational [14]. In addition, the plant uses several Application specific integrated circuits (ASIC) boards and devices that are rented out to Miners for their activities. The City of Bratsk is friendly for such a large-scale implementation because of the Siberian climatic conditions, which remain cold and has long winters. The site currently hosts 20,000 mining devices, and we can see a further add-up to 67,000 mining units in the future.

Furthermore, the Blockchain-as-a-Service model is contemporarily used by organizations to innovate their Supply Chain Management, Resource maintenance, and other critical business functionalities. At a hardware level, Infrastructure providers intend to assign a portion of their compute and storage infrastructure to tenants who leverage them for production. At an abstraction level, multiple tenants are managed efficiently by Trained Engineers and Logistics management tools to maintain discretion and privacy amongst intra-miners. This model is currently available with AWS, Oracle, and IBM vendors with adequate professional support to their clients. After this section, it is clear that as Decentralized platforms such as Blockchain is becoming more prevalent for business and revenue generation amongst a wide spectrum of clients, the need for expansive infrastructure providers has become a crucial juncture for the further evolution of this technology.

The Good: Virtuous Aspects of Blockchain Technology

From the previous sections, we understood the importance of Blockchain and the different ways Blockchain is implemented in the contemporary world. This section will portray the additional advantages of Blockchain 4.0, the latest evolution period in Blockchain Technology. We will also discuss Decentralized Applications that have become more prevalent as part of evolution.

Distributed Ledger Technology is a prime and disruptive technology in our current era; with this technology coming to the attention of various industries and corporations, it has gained mass investments and researchers' eyes in developing a Blockchain-based safe ecosystem. According to Leeway Hertz [15], over 100 private-permissioned blockchains and 50+ Layer 1 Blockchain protocols exist. The respective market is expected to be worth 60 million dollars within three years. As Blockchain evolved from Cryptocurrencies to Smart contracts, which in turn gave birth to Decentralized applications, and finally, in our current era of technologies like Web 3.0, Metaverse, and Industry 4.0, we have seen huge transitions and several advantages these transitions have created.

Firstly, **Decentralization** is a significant trait of Blockchain or Distributed Ledger Technology. This novel concept made equal accessibility and efficient data management in today's world. Decentralization ensures there is no single governance body, decreasing the chance of data corruption, targeted or distributed attacks on systems, and data loss. Furthermore, because of the high redundancy persisting in the Blockchain network, data loss is almost improbable, and the chance of system unavailability is minimal.

Secondly, The Blockchain network emphasizes **Security** amongst its connected systems. Whether implemented over a permissioned-network, public network, or consortium-based implementation, the data blocks containing critical information about transactions in the case of cryptocurrencies are always encrypted and unreadable for a nodal user participating in Ledger updating strategy. The transactional information, which is segmented and hashed into several data blocks, is unrecognizable to the miner and can only be confirmed by the users who participated in the transactional money activity.

Thirdly, **Smart contracts** came into the picture from the advent of Blockchain 2.0, where the necessity to automate the actions or activities when specific terms of the Ledger updating agreement are met was required. It is a computer program that works on the if-then logic of basic computer programming. When required conditions amongst a host of nodal systems are agreed upon, the smart contracts are executed amongst the systems that have completed the proof-of-stake. This piece of the program brought a significant amount of governance amongst systems and neutrality in data management and workflow.

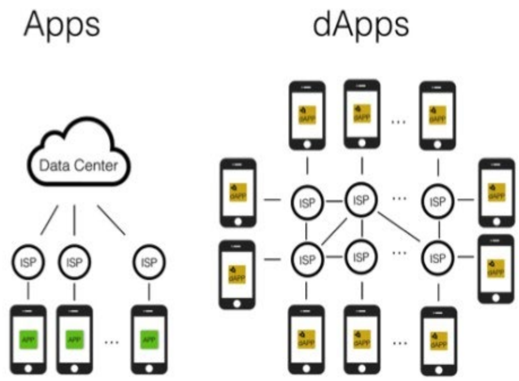


Fig 9.1. Centralized Application vs Decentralized Application Model

Also, one of the recent innovations of Blockchain Technology was **Decentralized Applications**. In a traditional system application, where the hardware and software abstraction layer cannot be shared between multiple systems was improved and refined to a decentralized application. According to the referenced article [16], Decentralized Applications are applications that run on computers that are part of the P2P network. Multiple participants are consuming content, feeding, or seeding content, or simultaneously performing both functions. The figure 9.1 above [17] gives the idea regarding the differentiation between traditional centralized application deployment and decentralized application deployment. They are developed to decentralize and distribute multiple functions and applications. They are built on platforms like Ethereum that provide developers with the necessary infrastructure to develop better applications. These also allow users to open-source, create transparency and censorship resistance, and remove the intermediaries from any ratification activities. These applications have made Blockchain technology much more lightweight and faster. Decentralized applications are leveraged in banking, real estate, gaming, social media, and e-commerce.

Finally, **Blockchain-as-a-Service** comes into the foray for large-scale deployments and infrastructure renting for a user who wants to leverage Blockchain for production purposes. In the previous section, we discussed BaaS and the Cloud deployment model. Several advantages of BaaS include the maximum utilization of hardware infrastructure, secure and efficient data management, better support and service capabilities, scalability and integration, dedicated project management for organizations or multiple users, and great accessibility to solutions that can be deployed as well as used for mining, supply chain or other production activities. Furthermore, according to the article [18], BaaS provides beneficial cost savings since it usually resides in a shared or multi-tenant environment, where the hardware and software costs are low compared with the traditional model. Considering all the advantages above, it is evident that the onset of the Blockchain industry has transformed from just being a money transfer and decentralized governing medium. With innovative projections for the future, we can be confident that the Blockchain Industry will boom and being part of this roaring opportunity can see huge investments from Technological corporations.

The Bad & The Ugly: Deficient Aspects of Blockchain Technology

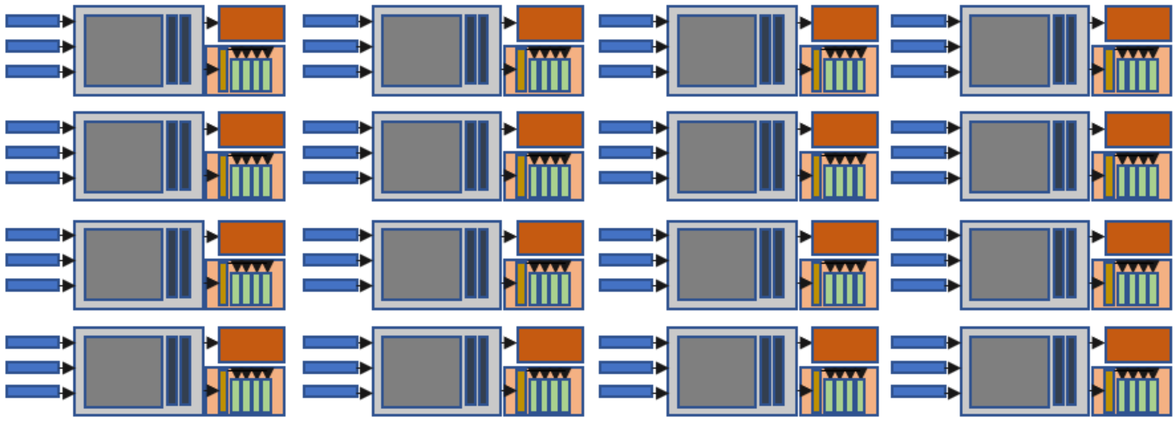


Fig 10.1. Large-Scale Entry-Level Mining Devices

With the rising cost of components, the initial investment for an entrant rig is becoming a distant dream for many. Now try to imagine over one hundred of these entrant-level gigs working simultaneously for 24 hours a day, seven days a week. Not only does the capital weigh down the investor, but the operational costs also pose a goliath-like threat that affects an immediate return on investments.

Expansive mining farms that farm up to 600 BTC monthly, consuming close to 110 terawatt-hours per year, have amounted to 0.55% of global energy consumption [19]. To be put in perspective, that is the amount of power Malaysia or Sweden consumes yearly. On the transaction side, with over 400,000 transactions happening daily worldwide, we are at a critical juncture. The requirement for platforms that can host the development of decentralized applications has significantly shot up in the last few years, with Blockchain 3.0 transitioning into 4.0. Given the large-scale requirements for mining infrastructure that can be directed toward the development of 4.0 in the foreseeable future, it is paramount that the world targets its resources toward building better infrastructure.

If we were to extrapolate the current infrastructure by multiplying the small setup, the power consumption and efficiency at which the compute and storage can function drastically drops or stays stagnant. This will retard the progress that Blockchain 4.0 proposes. Hence, enterprise-level infrastructure must be built specifically for mining to find its way to the market. Due to the lack of the same, the overall mining capacity has been hit purely because no enterprise level infrastructure has been repurposed to befit the needs of a mining farm. It is high time we focus on experimenting and working toward a better initiation toward R&D. Without such an initiative, the decentralized network will remain stagnant and may even regress.

Proposed Environmental Blueprint

The proposed environment consists of presently available enterprise-level infrastructure that can be used to architect a compute-heavy solution that prioritizes on storage as well. Concerning Fig 11.1 below, the production site has been connected to the Blockchain network through a Blockchain API. The API, however represented as one medium, may vary in number in real life based on the mining network the user wishes to mine data from.

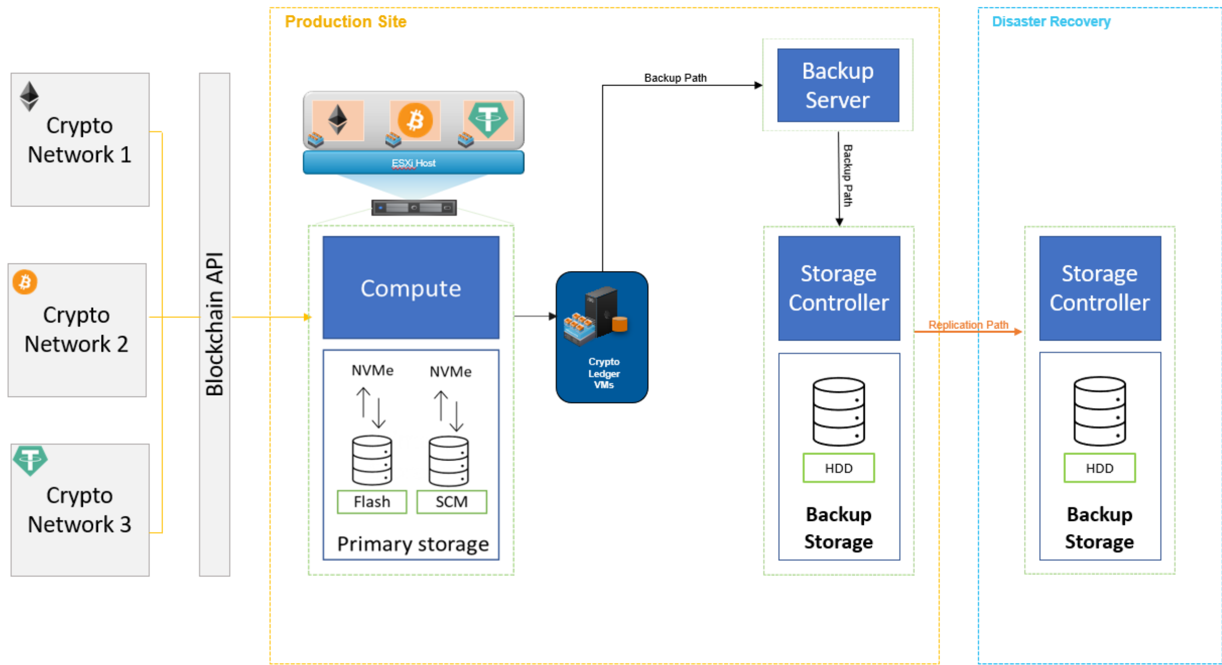


Fig 11.1. Proposed Environmental Architecture

The mid-business level computes infrastructure should do the job of performing high-level mathematical calculations that PoW requires. As mentioned earlier, the GPU acts in the place of a CPU in the data mining industry, hence making it an absolute requirement for the server to have GPU slots and PCIe extensions. To make data fetching more efficient and precise, NVMe storage must be used for the reasons mentioned later in the upcoming section. To accommodate a miner's intent to work on multiple cryptocurrencies at the same time, there is an increased importance to possess a hypervisor that helps virtualize and segment infrastructure resources according to the mining requirements for each currency. To stay on the safer side, we propose a Virtual Machine backup to a secondary storage device that helps store the offline copy of the ledger and any off-chain storage as well. The backup path shall traverse via a backup server to the secondary storage where the data may sit.

One of the Blockchain industry's main drawbacks is the need for suitable disaster recovery mechanisms. Hence, a replica of the secondary storage architecture from the production site is placed within the Disaster Recovery site to ensure a successful and accommodative fail-over in the case of a calamity affecting the production site. Both the secondary storage devices may use cheaper storage disks to store their data and need not fret when it comes to having quick fetching of data.

Primary Storage Strategy

Blockchain storage technology is transparent and tamper-proof. Since blockchain depends on a decentralized network to store and transmit data, the storage is responsible for breaking the data from the web into small manageable chunks. These chunks are then distributed across nodes. The system further encrypts each data shard on the local system to ensure that it is inaccessible by others on or off the network except the valid owner of the shard. Then the blockchain storage creates a unique hash and an encrypted output string based on the data chunks or encryption keys. Finally, the hash is added to the ledger and data shard (metadata) to link transactions to the stored data shards. Once this is done, the shard is replicated several times to ensure enough copies to provide redundancy and avoid data loss.

Following that, a peer-to-peer network distributes the replicated data chunks to various dispersed storage nodes that are owned by individuals known as 'farmers. However, the data is still accessible only to the true content owner. Finally, the storage system records all these transaction details, like the shard location and hash in the blockchain ledger. This information is then synced across all nodes. The approach to blockchain differs based on the storage system and how the data is managed.

There are two blockchain data storage solutions approach: off-chain and on-chain. Below is a closer peek into the off-chain storage that answers the blockchain's scalability issue.

Off chain data is also called real-world data because it is external to the blockchain and can consist of various data, including but not limited to weather forecast data, financial information, and even game scores. Blockchains are isolated systems by nature. Connecting them to off-chain storage and off-chain data is like connecting a computer to an internet connection. So, it enables an isolated system to interact with the external environment. Decentralized applications have contributed to expanding the blockchain ecosystem, and storage systems with enough capacity to scale are required, along with the ability to compute faster. Off-chain storage help address the waste of storage and computation resources that is needed for every blockchain node.

Off-chain storage can help in unburdening the blockchains from storing large datasets. Still, adding drives or a storage system externally can impact performance as physical disks sometimes need to be faster. Hence NVMe technology can provide the desired storage considerations and implementation. NVMe is a high-performance, optimized, and scalable protocol connecting the host to the memory subsystem. The protocol is based on parallel and low-latency data access paths to the underlying media, like many high-performance processor architectures. Hence, NVMe storage systems would be ideal for storing off-chain data as they can help reduce the processor cycle and provide the necessary performance without too much overhead.

Along with SCM (storage class memory), these systems can ensure the data persists even after a failover and perform better than NAND flash drives. Another crucial aspect of incorporating storage infrastructure for blockchain plans is that it should be scalable. To complement the nature of blockchain data, primarily defined by its growth and efficiency, the storage system should be fast, highly scalable, and accommodative of different types of data workloads.

But it does not stop there; we need decentralized storage to make the blockchain applications secure to high standards. It refers to data being stored on the same network, but multiple systems can connect to this network and share data. Decentralization ensures security and immutability because there is no central authority involved.

Data Protection Strategy

In the previous section, we saw the Blockchain on-chain and off-chain storage, which from a Hardware Environment perspective, would mean primary storage where the Transactional blocks of data are stored. In this section, we will look into the Data Protection strategy to safeguard the data further from external threats, data losses, and production downtime.

Firstly, Backup is the most common data protection methodology leveraged by both small scaled and large scaled users in different industries and verticals. Backup ensures a safe, secure, and executable format of your most recent data that can be leveraged to protect your system from downtime or data loss. Additionally, one can create multiple copies of similar data throughout time can perform restores as required by the business to ensure production continuity and consistency.

Furthermore, the initial primary storage based on performance included NVMe protocol being leveraged for internal communications between the computer and the disks. The disk themselves accommodated faster speeds as the SCM-based drives are leveraged in the architecture. However, when discussing Backup Storage, it should necessarily run-on Hard disk drives using spinning disks. The disk architecture leveraging RAID methodology will allow better fault tolerance to the Backup protection system. The network cards should accommodate the primary environmental network. The storage should cater to Ethernet protocol and Optical connection for Backup activities.

Now, Disaster Recovery forms a pivotal part of the unlimited storage and data protection strategy. The Environmental explanation in the previous page covers the overview of block storage capability for the transactional information necessary to be over a Blockchain network. It also covered the protection storage ability of the intended architecture. Disaster Recovery, as the term suggests, was introduced in low-scale as well as enterprise-level environments to safeguard and employ high availability over a dedicated displaced location separate from the primary and backup environment.

Further, the intention is to have an additional copy of the critical data to tackle the environmental disaster that can be human-induced or natural. It is governed by policies, strategic requirements, and subsequent actions to fulfill desired intentions. In the architecture proposed in this paper, protection storage that can cater to the bandwidth of daily utilization up to an adjustable 100 MB/s at the Backup and Replication end is adjudged to be leveraged. This transfer can vary according to the Mining job the Node performs in the Blockchain network. For example, a node assisting in more transactions will have more blocks of data generated to update the Blockchain ledger. Specific mechanisms to efficiently increase the data transfer of replicated data from Backup storage to Replication storage can include transfer between similar types of partition, to name one. This can decrease the processing time of the controller computer used at the replication storage to position the data blocks in the storage disk. It can also be accommodated with similar transfer protocols tested and production purposes amongst the two storages.

Since the architecture requires the replication storage to be physically displaced from the primary and backup storage, the Blocks of data that needs to be transferred have to traverse the Global internet to fulfill the strategy. Several DR Topologies can be leveraged to expand the data protection plan that a miner wants to implement. This can include One to Many, One to One, Bi-directional topologies for better efficiency. For a better understanding of the DR Topologies, the example images of several Disaster Recovery strategies are displayed in the images below. Several configurational requirements are necessary to implement an excellent replication plan. The destination system or the replication system must have at least the same size as the most significant source directory of the original design. The backup storage and replication storage systems must be in an active, visible route through the IP network as the Blocks of data transfer through global networking. Hence, addresses of these systems are critical in such an implementation.

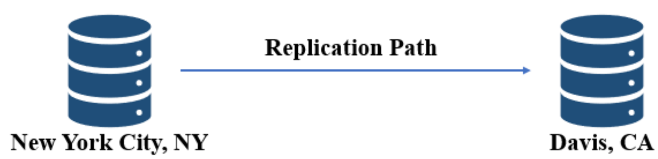


Figure 13.1. One to One Replication Topology

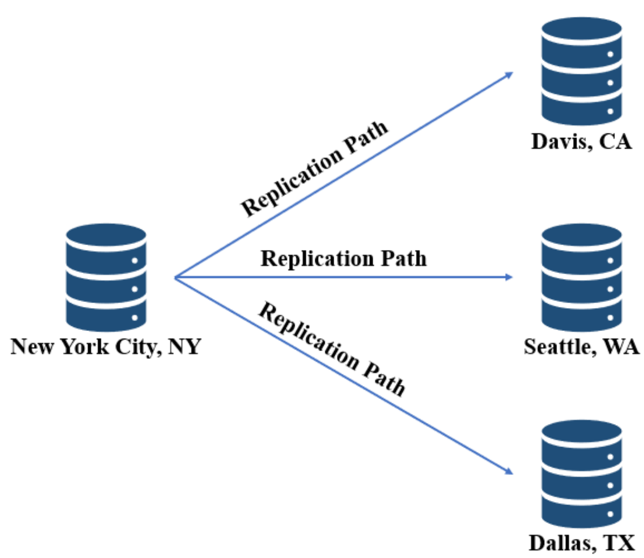


Figure 13.2. One to Many Replication Topology

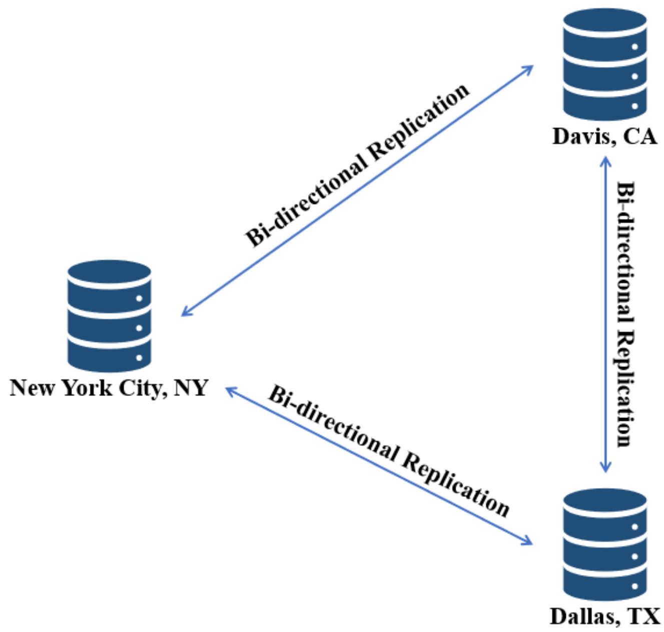


Figure 13.3. Three Way Bi-directional Replication Topology

Additionally, after the identification of the compatriot systems by each other, through test data packets can, the replication activity is initialized. These procedures are governed by particular policies and respective units of measure that can identify the amount of data transfer, bandwidth limitations, compression, and deduplication statistics with mathematical representation in the form of averages, maxima, and most recent activities. These policies are adjusted based on the miner's strategic requirements, which can vary according to monetary capability, computing ability, and fundamental infrastructure. Following these checklists, a miner can decide the type of replication strategy and efficiency of their data protection plan.

Conclusion: Blockchain Transformation in Future

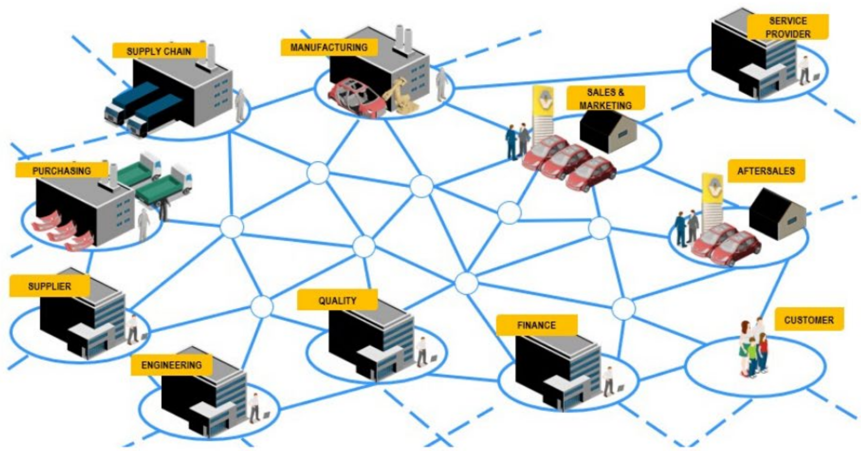


Fig 14.1. Blockchain Transformation in Future

The question does arise, where is this headed? With technology gaining prominence rapidly, how can one stay ahead of the curve? We are sure these questions are floating around in your mind. But let us not get carried away…...yet. The Blockchain wave will course through the industries such as finance, cloud storage, cyber security, and digital advertising [20].

With finance being the most prominent industry to be positively impacted, Blockchain will provide an efficient means to track financial properties and a clear view into the decentralized ledger that will provide access to the network users on the inflow and outflow of cash. This makes it a completely client-owned bank that needs no broker, clerk, or banker in the midst, and only the customer runs the show.

In the distant future, we can see cloud storage that all can own, built on a peer-to-peer network. A storage infrastructure that grows as the number of users grows is Blockchain helping answer its problems of scalability. This will be necessary as we step into Blockchain 4.0 with multiple users working on a BaaS platform.

To make things better, Blockchain will soon be integrated with AI, which in turn is going to make the use of smart contracts mainstream. The fourth generation of Blockchain also offers its users proof of integrity [9]. With AI-powered software powering verification and a decentralized ledger that the AI can use to cross-reference data, the technology is unstoppable. It will take on a more prominent role in multiple sectors.

It is now high time that the infrastructure storage industry spends its resources on developing industry-specific hardware that can be used to mine and take on high-intensity mathematical calculations efficiently. It will be a matter of time before theoretical and conceptual technology catches up with the real world. With no enterprise-level infrastructure available to support the Blockchain world and the industry, its time infrastructure storage providers aim to stay ahead of the curve.

Bibliography

[1] - https://101blockchains.com/who-is-satoshi-nakamoto-bitcoin/

[2] - https://101blockchains.com/history-of-blockchain-timeline/

[3] - https://101blockchains.com/who-is-vitalik-buterin-ethereum/

[4] - https://www.techtarget.com/searchcio/definition/Hyperledger

[5] - https://www.investopedia.com/terms/b/blockchain.asp

[6] - https://www.naukri.com/learning/articles/decentralization-in-blockchain/

[7] - https://www.blockchain-council.org/blockchain/what-is-decentralization-in-blockchain/

[8] - https://medium.com/the-capital/blockchains-evolution-1-0-2-0-3-0-4fdb2c5e52be

[9] - https://www.researchgate.net/publication/351263701_Blockchain_10_to_Blockchain_40-The_Evolutionary_Transformation_of_Blockchain_Technology?enrichId=rgreq-6124cc270470f3244472a36be61a0a91-XXX&enrichSource=Y292ZXJQYWdlOzM1MTI2MzcwMTtBUzoxMDYwODk0NTQ1MDMxMTc3QDE2Mjk5NDgzNzg3NjM%3D&el=1_x_2&_esc=publicationCoverPdf

[10] - https://www.zdnet.com/article/how-to-build-a-cryptomining-rig/

[11] - https://www.coinbase.com/cloud/discover/dev-foundations/blockchain-infrastructure

[12] - https://www.zeeve.io/blog/a-detailed-guide-on-blockchain-node-infrastructure/

[13] - https://www.google.com/url?sa=i&url=https%3A%2F%2Fjournalofcloudcomputing.springeropen.com%2Farticles%2F10.1186%2Fs13677-021-00247-5&psig=AOvVaw2HVweXtIOtfQra_Ri0RGYV&ust=1673967650526000&source=images&cd=vfe&ved=0CA0QjRxqFwoTCPCF_s6tzPwCFQAAAAAdAAAAABAQ

[14] - https://www.bloomberg.com/news/features/2019-11-24/seo-inside-russia-s-largest-bitcoin-mine

[15] - https://www.leewayhertz.com/blockchain-4-0/

[16] - https://www.investopedia.com/terms/d/decentralized-applications-dapps.asp#:~:text=What%20Are%20Decentralized%20Applications%20(dApps,control%20of%20a%20single%20authority

[17] - https://www.google.com/imgres?imgurl=https%3A%2F%2Fblogs.lse.ac.uk%2Fbusinessreview%2Ffiles%2F2017%2F10%2Fapps-dapps.jpg&imgrefurl=https%3A%2F%2Fblogs.lse.ac.uk%2Fbusinessreview%2F2017%2F10%2F02%2Ftraditional-firms-are-opening-up-to-blockchain-and-its-decentralised-apps%2F&tbnid=wNr4mFU9zFWgKM&vet=12ahUKEwiN5KbGpcz8AhVsDbcAHbv2AqcQMygDegUIARC8AQ..i&docid=QG4dXD7DKudIFM&w=670&h=335&q=Decentralized%20Apps%20Images&ved=2ahUKEwiN5KbGpcz8AhVsDbcAHbv2AqcQMygDegUIARC8AQ

[18] - https://www.ibm.com/cloud/blog/top-5-advantages-of-software-as-a-service

[19] - https://www.sunbirddcim.com/infographic/largest-bitcoin-mining-farms-world

[20] - https://timesofindia.indiatimes.com/readersblog/raghib-blogs/opportunities-and-the-future-of-blockchain-technology-43325/

Fig 3.1 - https://money.com/what-is-blockchain/

Fig 4.1 - https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.aceinfoway.com%2Fblog%2Fcentralized-vs-decentralized-apps&psig=AOvVaw2NyMvr81CATlBUc6v3HfwL&ust=1673989884418000&source=images&cd=vfe&ved=0CA0QjRxqFwoTCLiTgb2AzfwCFQAAAAAdAAAAABAD

Fig 14.1 - https://www.renaultgroup.com/en/news-on-air/news/the-blockchain-transformation-vector-for-the-future-of-the-automotive-industry/v

Disclaimer: The views, processes or methodologies published in this article are those of the authors. They do not necessarily reflect Dell Technologies' views, processes, or methodologies.

Dell Technologies believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED "AS IS." DELL TECHNOLOGIES MAKES NO RESPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying and distribution of any Dell Technologies software described in this publication requires an applicable software license.

© 2023 Dell Inc. or its subsidiaries. All Rights Reserved. Dell and other trademarks are trademarks of Dell Inc. or its subsidiaries. Other trademarks may be trademarks of their respective owners.