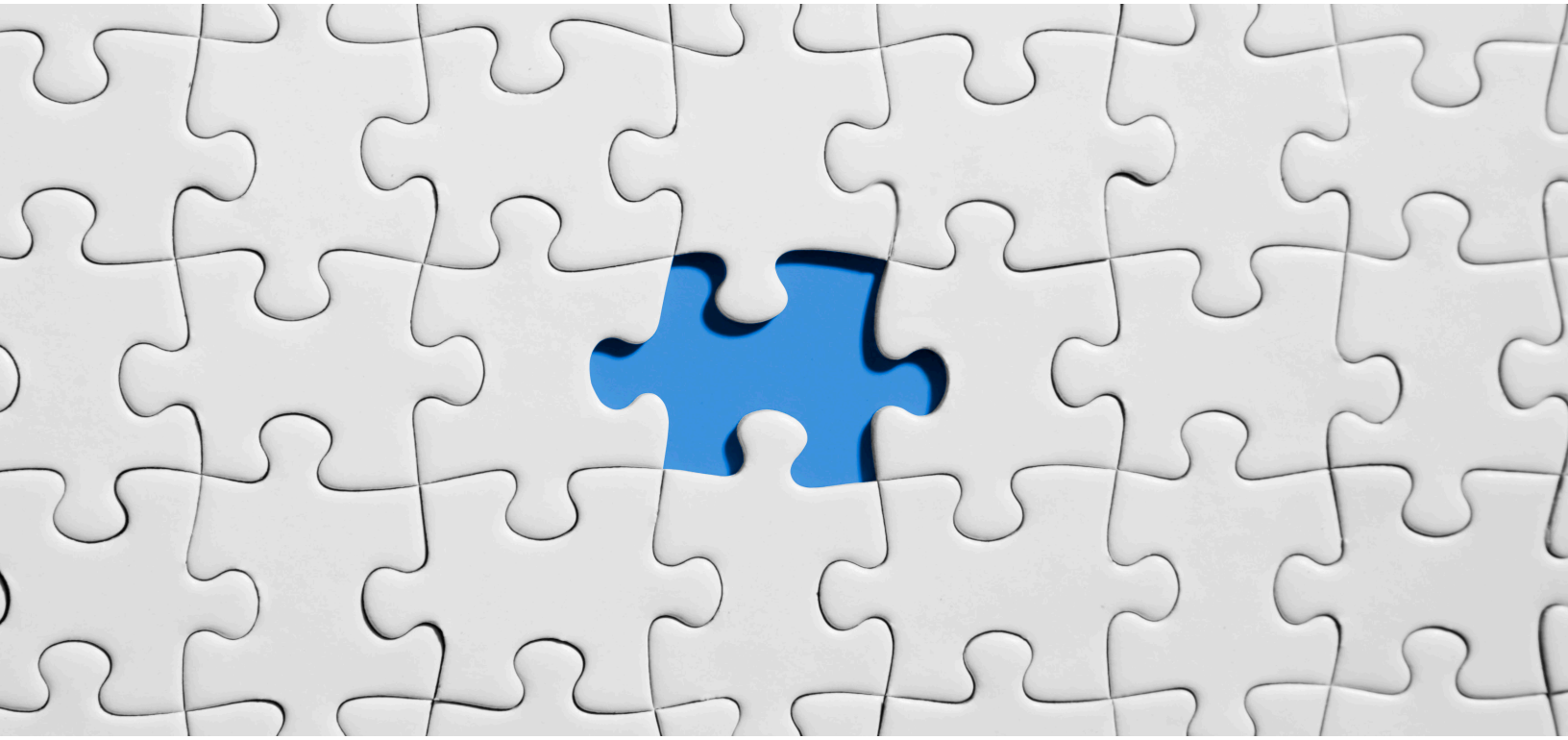
DELL EMC

DISTRIBUTED CLOUD STORAGE WITH BLOCKCHAIN TECHNOLOGY



Kanishk Kumar

Sales Engineer Analyst

Dell EMC

Kanishk.kumar@dell.com

DELL EMC PROVEN PROFESSIONAL

Knowledge Sharing Article

Table of Contents

Introduction

Everywhere we turn these days “the cloud” is being talked about. While “the Cloud” is just a metaphor, Cloud computing is what people are really talking about. Now is the right time to say the future application stays in cloud. Many of the enterprise applications and customer’s demands are mostly serviced from cloud computing technologies. Cloud service providers like Dell EMC and Amazon are attracted to cloud-based services considered superior to traditional data centers in terms of cost and technical dimensions.

However, unlike any other storage technology it has its own drawbacks, security being the most common of them all. Moving business data to the cloud means that the responsibility of data security becomes shared with the cloud provider. The overlapping of trust boundaries and increased exposure of data can provide malicious cloud consumers with greater opportunities to attack IT resources and steal or damage business data.

The answer to this challenge in cloud security can be tackled by the use of a “Distributed cloud model” using blockchain technology.

Blockchain is a hot topic and many are looking for new, secure, cost-efficient methods to store their ever growing data libraries. Originally devised for the digital currency, Bitcoin, blockchain eliminates the need for trusted third parties. Additionally, blockchain database isn’t stored in any single location. This creates a transparent, traceable peer to peer system that’s almost impossible to hack.

Traditional Cloud Storage Model

A traditional cloud storage model consists of a front end platform which could be a client or a mobile device, a back end platform which could be a server or storage and a network, possibly internet or an intranet.

Google Drive is a typical example of traditional cloud storage. When you upload data to the cloud, Google stores it in one of their datacenters. When you want to access the data from a mobile device/laptop, a request is sent to the data center and you can access your data.

The problem(s)

Running vast data centers are expensive. The tech in these data centers need to be refreshed on a regular basis. Moreover, there are operational costs because of cooling, maintenance and updates.

Safety is another aspect to consider. All cloud service providers have strict safety processes in place but there is always room to penetrate and gain access to confidential data. The recent iCloud hack of celebrities was one such occurrence.

And it isn’t just human error that puts your privacy in danger. Large companies have the ability to search non-encrypted files. Their privacy terms outline a lot of different scenarios where they can legally access and share your data.

Blockchain

Blockchain technology is not new; rather, it is a combination of proven technologies applied in a new way. It is the combination of the internet, private key cryptography and distributed open ledger protocol.

Let use an example to explain Blockchain in simple terms.

Suppose there are four entities/nodes A, B, C and D and they want to transfer a certain amount to each other.

Let us assume A has $10 and A transfers $5 to B and B transfers $3 to C and C transfers $1 to D.

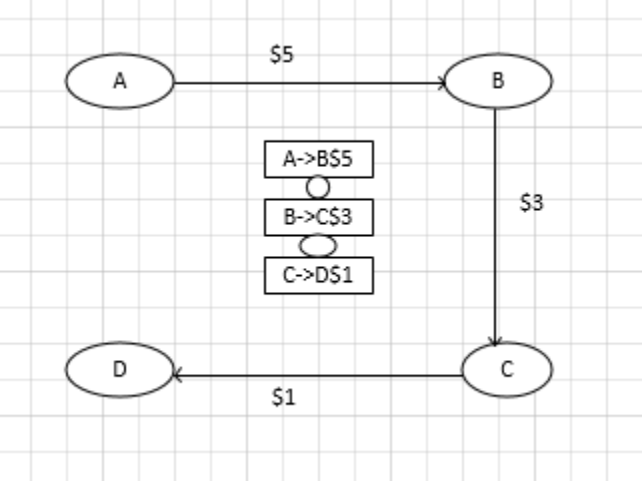


Fig1: Open Ledger Protocol

All these transactions are recorded centrally and linked to each other and all the entities are aware of these transactions. The record where all the transactions are stored is called a “ledger” and the protocol described above is an open ledger protocol.

All the entities are aware of each transaction and it is verifiable. In case D wants $15 from A, A cannot transfer this amount since it is only left with $5 and all the other entities are aware of the same.

Blockchain uses a same approach with the only difference being that the data, i.e. transactions, are not centrally recorded but each entity will have a copy of the transactions.

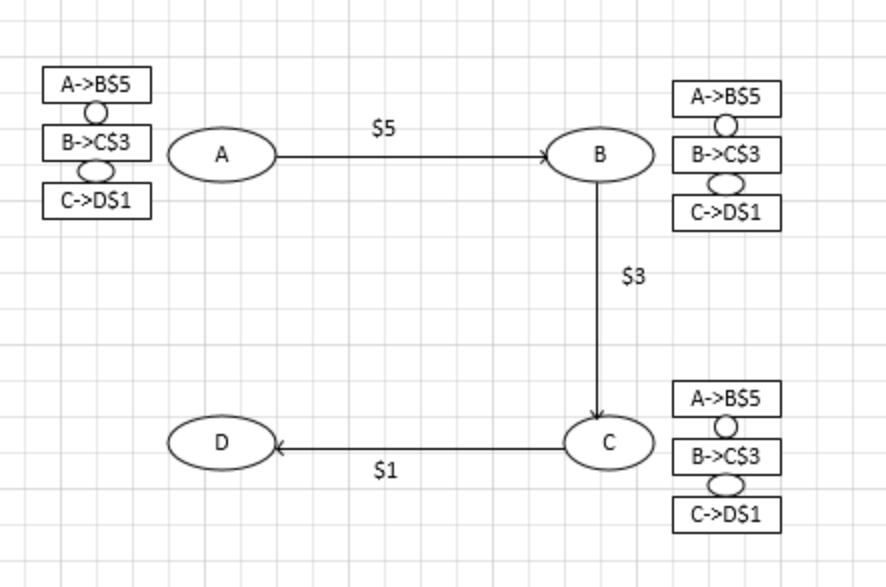


Fig2: Distributed Open Ledger Protocol

By storing data across its network, the blockchain eliminates the risks that come with data being held centrally.

Blockchain security methods include use of public-key cryptography. Value tokens sent across the network are recorded as belonging to that address. A private key is like a password that gives its owner access to the data.

This is where blockchain has its advantage. Decentralized data is transparent to everyone involved. Every node in a decentralized system has a copy of the blockchain. Transactions are broadcast to the network using software. Mining nodes validate transactions, add them to the block they are building, and then broadcast the completed block to other nodes. This also maintains sync between the different copies of data that various nodes are holding onto.

Let us take the above example and see how this works.

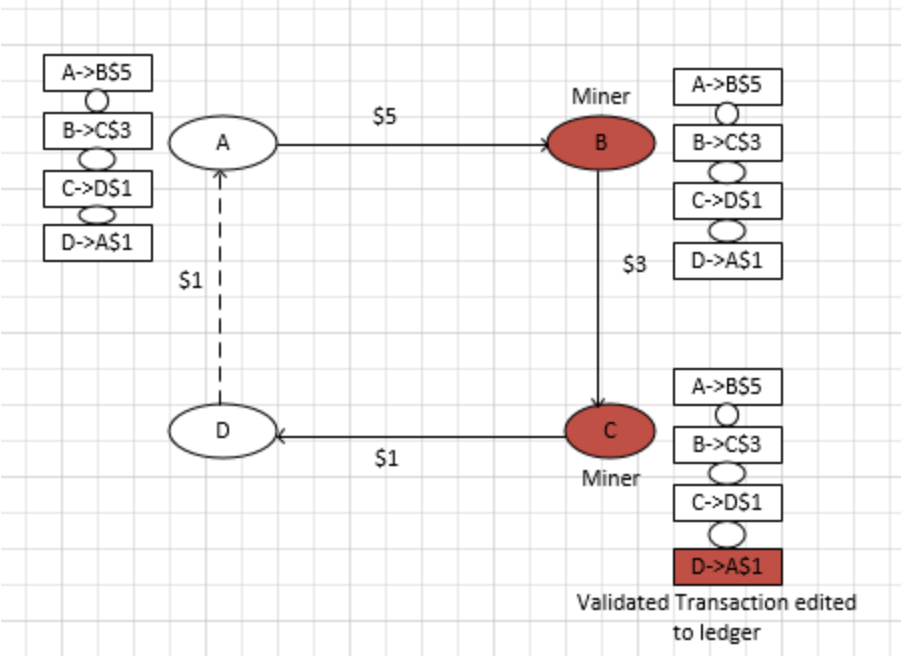


Fig3: Data transaction in a distributed open ledger

Suppose D wants to transfer $1 to A. D encrypts the data with a public key and sends it to A which has the private key to decrypt and access the data. A broadcast message regarding this transaction will be sent to the network.

In a blockchain network there will be certain special nodes called “miners” which can validate a transaction. Assuming C and B are the miners they will try to validate the transaction from D to A. In this case validation can be done using two points.

If D has the funds to do this transaction

Decrypting the public key

Since all the nodes in this network have a copy of the previous transactions, both C and B are aware that D has the funds to carry this transaction. Decrypting the public key however is not simple; a miner repeatedly generates keys and tries to guess the correct key until it finds one. If C finds the key that matches the key for this transaction it validates this transaction and edits it to its own ledger. All the remaining nodes will do the same since this transaction is already validated.

Distributed Cloud Storage

Distributed cloud storage is envisioned where all aspects of cloud storage such as transport, processing, or storage of data are entered into the blockchain. Later, what happens to the data can be verified by anyone who has the access to the blockchain. Such a system provides complete traceability, accountability, and transparency to the cloud.

Distributed Cloud model enables users to store data in a secure and decentralized manner. This is done by using blockchain features such as ledgers, public/private key encryption, and so forth which we discussed earlier in this paper. These features are putting the user back in control over their data and devices. The decentralized aspect ensures there are no central servers to be compromised.

Distributed Cloud Design Considerations

To design a high performance distributed cloud architecture that meets both current and future challenges, the following principles should be taken into consideration:

Resilience: Even if some nodes fail, computation continues on other nodes.

Efficiency: The users receive excellent performance even if the nodes involved are heterogeneous.

Ease of deployment: The nodes can be deployed in any configuration without disrupting the other nodes.

Adaptability: The architecture of the network should be able to adapt to the changing environment and broaden its use to meet the increasing needs and demands of customers.

Performance: Linear performance is always needed in a distributed network.

Security: Data protection, confidentiality and information security must be adequately addressed.

Distributed Blockchain Cloud Architecture

The proposed model consists of the following four steps:

1. **Selection of Resource**: The cloud user must select the resource provider from the service provider pool in the blockchain base distributed cloud.
2. **Provision services**: The selected service provider will provide required services, such as data management, task execution and the provision of servers to that user.
3. **Registration of services**: After providing the requested services, the service provider registers the transaction in the form of blockchain and shares it with all distributed peer service providers.
4. **Payment**: The user will pay and reward the service provider and all the peers will copy this transaction/movement in their own blockchain.

This is a trusted peer-to-peer network maintaining a distributed ledger that consists of validating nodes/miners that update the ledger and respond to requests. Requests can be invoked through client SDKs or REST API calls.

Multiple peers endorse/sign the results, which are then verified and sent to the ordering service. After consensus is reached on the order, results are grouped into cryptographically secured, tamper-proof data blocks and sent to peer nodes to be validated and appended to the ledger.

Members can be added rapidly to the Blockchain network whether they are next door or across the world. Once their instance is provisioned they can join the blockchain by exchanging digital certificates and securely conduct data transactions with their peers.

Not all business data exchanged between members is suitable for sharing with all participants. In such an environment we can isolate peers into subnets and create private ledgers. Blockchain members can then conduct private and confidential transactions while coexisting with members on the same blockchain. Peers can only join the chain when approved by other organization on that chain. Client requests are routed to a specified channel and once endorsed, the results are updated in that channel's ledger, which is only accessible to its member peer nodes.

Types of Blockchain networks

1. **Consortium blockchains:** In a consortium blockchain, the consensus process is controlled by a pre-selected group such as a group of corporation. The right to read the blockchain and submit transactions to it may be public or restricted to participants. Consortium blockchains are considered to be "permissioned blockchains" and are best suited for use in business.
2. **Semi-private blockchains:** Semi private blockchains are run by a single company that grants access to any user who satisfies pre-established criteria. Although not truly decentralized, this type of permissioned blockchain is an appealing option for government applications.
3. **Private blockchains:** Private blockchains are controlled by a single organization that determines who can read it and participate in the consensus process. Since they are 100 percent centralized they are only useful as sandbox environments but not for actual production.
4. **Public blockchains:** Anyone can read a public blockchain and participate in the consensus process. They are considered to be "permission-less". Every transaction is public and users can remain anonymous. Bitcoin and Ethereum are prominent examples of public blockchains.

Benefits

1. **Full Decentralization and true redundancy**: Using blockchain, we can build a cloud storage where data is stored in dozens of discrete nodes intelligently disbursed across the globe.
2. **Transparency**: Information in blockchains is viewable by all participants and cannot be altered. This reduces risk and fraud and creates trust.
3. **Security**: The distributed nature of blockchain means it is almost impossible to hack.
4. **Fewer Intermediaries**: Blockchain is true peer-to-peer network that reduces reliance on third parties like banks, brokers, gateway, etc.
5. **Automation**: Blockchain is also programmable – which make it possible to automatically trigger actions or events once the conditions are met.
6. **Faster Processes**: Blockchain can speed up process execution in multi-party scenarios – and allow faster transactions that aren't limited by office hours.

Distributed Cloud Use Cases

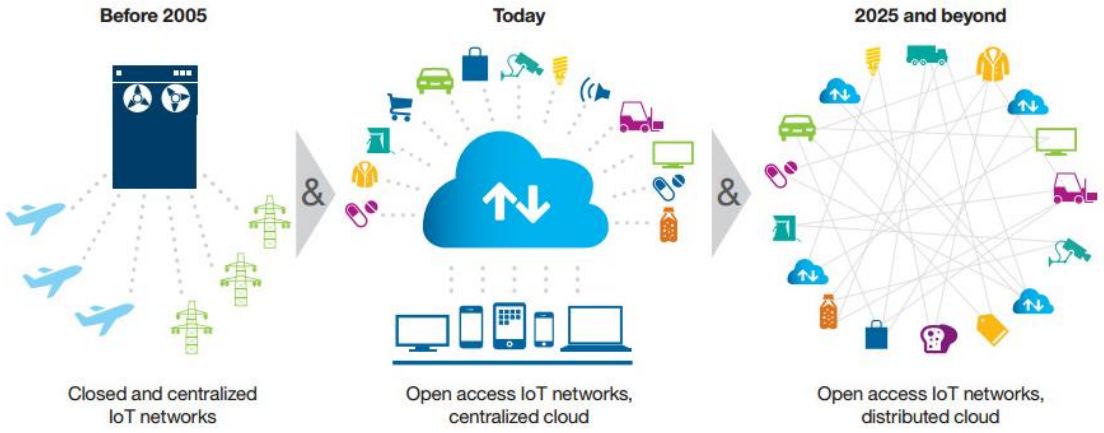


Fig 4: Migration to Distributed Cloud Infrastructure

1. **Network applications/NFV:** The NFV evolution has made it possible to distribute virtual network functions (VNFs) in a more flexible way. The infrastructure for NFV is an important starting point for the distributed cloud evolution. Distributed Cloud infrastructure enables intelligent placement of VNFs, mobile cores and RAN functions.
2. **Content delivery Networks:** Content delivery solutions have been run as applications on generic computing and storage platforms. This means these platforms must support distribution across regional and hub sites as well as across multiple service providers. A decentralized architecture provides better response time as well as efficiency in transport and peering costs.
3. **Data storage with regulatory compliance:** Enterprises are increasingly using cloud service providers for scalable storage of various data sets where security and regulatory constraints are major concerns. A decentralized architecture enables compliance with regulations and ensures control of cost and policy with regards to the cloud service providers.
4. **Hybrid enterprise cloud:** Enterprises that want to use cloud service providers for elasticity and scalability reasons, but also want to control where applications are executed. A cloud platform can be deployed across on-premises and cloud resources such that applications and data can be placed according to policy and performance constraints and intents.
5. **IoT and data stream processing:** Applications that collect and process data are often composed of several parts that include components for data collection, data throttling, data pruning, anomaly detection and storage. We can improve the improve scalability and performance by placing these components at an optimal location in the network which in turn will lead to better response times and efficient transport and peering costs.

Conclusion

In this paper, we have described the Traditional cloud model and its short comings. We discussed the Blockchain Technology and how we can design distributed blockchain cloud architecture. We went through its benefits and some of the use cases it can serve.

By 2020, Cisco predicts that 50 billion devices will be connected to the internet – which means more and more online content will be generated and the requirement for a secure storage will be more than ever. The traditional cloud model that we have today is not going to cut it in future. A well-designed and publicly accessible distributed cloud can replace many of the functions that we currently rely on cloud intermediaries for providing a trustworthy trading environment, guarding against fraud and handling, ensuring contract compliance and financial transactions.

Distributed cloud storage on the blockchain seeks to provide a new solution for a problem that is only growing in size. We are on the edge. Some say the blockchain and its potential is much where the internet was in the early 90s.

References

"Blockchain and Distributed Ledger Technology"
https://www.sap.com/india/products/leonardo/blockchain.html#

Martin Korling "Future Digital Blog" http://cloudblog.ericsson.com/digital-services/distributed-cloud-infrastructure-nfv June 12, 2017

"Integrate Your Business Network with the Blockchain Platform"
https://cloud.oracle.com/opc/paas/ebooks/Oracle_Blockchain_Cloud_Service.pdf

"How Blockchain Tech Is Changing Cloud Storage" https://www.belugacdn.com/blog/162663814938-how-blockchain-tech-is-changing-cloud-storage