OctaSpace White Paper

OctaSpace Team

March 1, 2023

Contents

Abstract

OctaSpace is a network that connects the owners of spare compute to the users of compute-heavy applications.

The platform aims to be as easy to use as centralized services and largely scalable to millions of computers and users.

The project's objective is to provide a functional and transparent way to perform computations in a distributed environment to store, process, and serve any amounts of data.

The nature of the computations may vary from machine learning, simulations, rendering, or even gaming, any task that requires massive computational power.

1 Introduction

The backbone of the internet infrastructure is controlled by cloud compute providers that process and store essentially the entire world's data.

While these centralized compute providers work well enough for most tasks, they suffer from the inherent weakness of a large trusted corporate conglomerate.

The main issue being cost and centralization, making a global decentralization network not truly possible with these providers. With a handful of cloud providers gaining more share each year, it will lead to much higher prices for compute, which, in turn, decreases the services able to operate.

With centralization, the need for redundancies also spreads, which increases costs and resources. These weaknesses can be avoided by utilizing spare unused compute across the globe.

OctaSpace takes a distinct approach to building a scalable distributed compute system, which presents itself to compute users as a unified, coherent entity. This innovative method enables users to start renting Octa Nodes seamlessly, in a matter of minutes.

It's essential to have the system with unified interfaces and coherent behavior, providing a single entry point for the end-users' interactions similar to how cloud providers operate today.

OctaSpace is a distributed cloud platform to implement services built on top of hardware provided by decentralized users.

With OctaSpace, you can harness the power of CPU, GPU, storage, and traffic resources from Octa nodes across the world to tackle compute-heavy tasks.

By utilizing your spare hardware, it's possible to generate income by acting as a host and renting out your CPU/GPU or traffic.

The hardware required for this can range from basic, low-performance devices or inexpensive VPS to advanced mining rigs and carrier-grade servers located within data centers.

OctaSpace's initial goals:

- Utilize a powerful of Hybrid 51% attack proof blockchain to increase transparency.
- Provide easy to use interface for describing nodes and executing tasks and marketplace for renting computer focusing on GPU instances
- Provide infrastructure to deploy applications and databases mostly close to end users
- Implement a VPN marketplace using popular VPN technologies
- Implement distributed data storage and use it as a backbone for CDN service
- Increase OctaSpace accessibility globally

The sort of tasks the project focused on:

- Machine Learning
- CGI rendering
- Digital image processing
- Scientific modeling
- Data storing and distributing
- VPN technologies
- Other tasks/fields which require massive computation power to get be solved

2 Architecture

The following picture shows the core components and interfaces

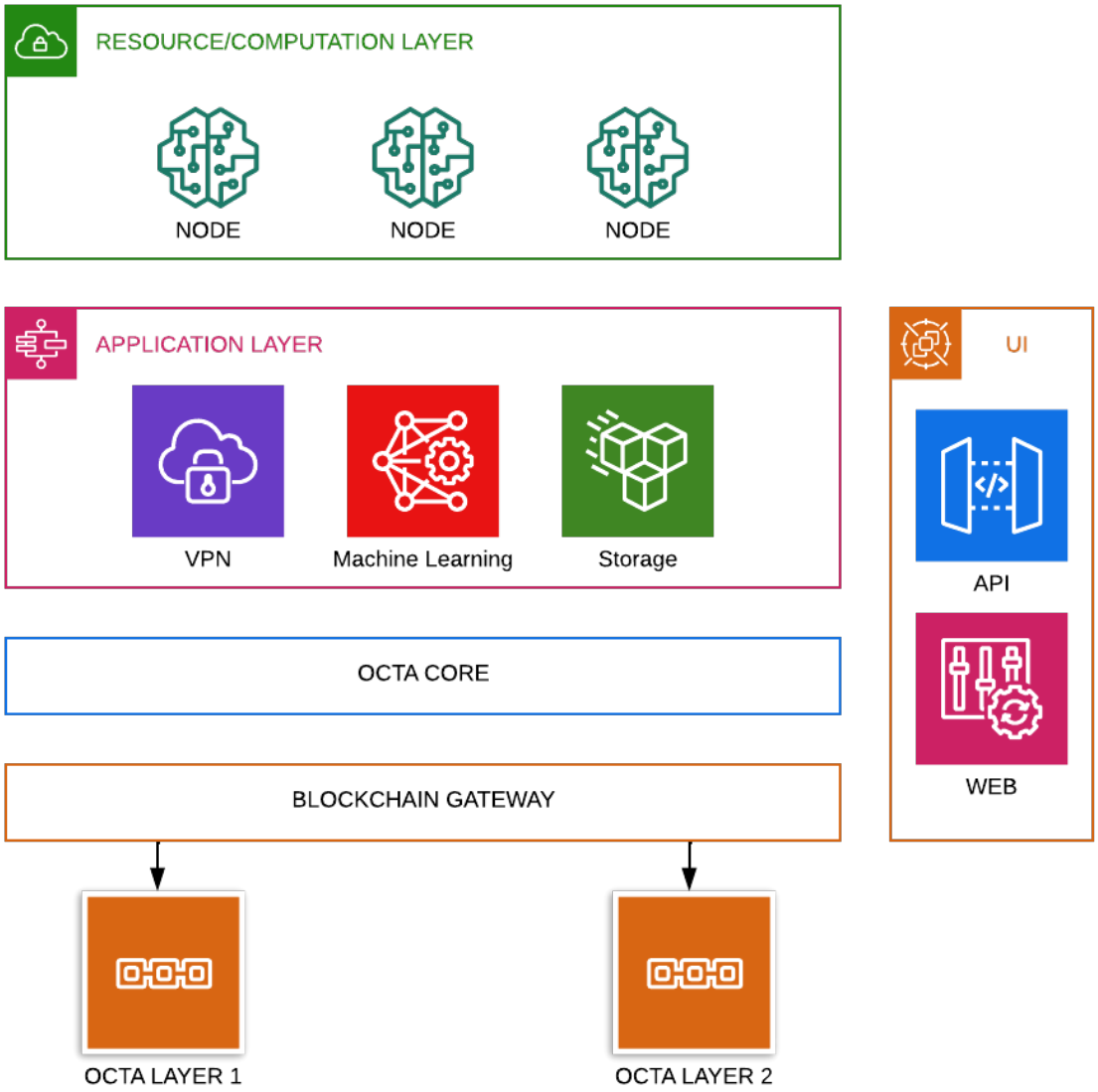

Figure 1: High Level Architecture

2.1 Blockchain

OctaSpace utilizes a Layer 1 PoW blockchain that is secured using the pirl 51% guard technique. This security protocol provides protection against attacks on the network, making it a secure platform for users to transact on.

In addition to the Layer 1 PoW blockchain, OctaSpace also employs a Layer 2 PoA blockchain that is used to speed up transactions for billing operations. This blockchain is based on validators and is not used to secure the Layer 1 PoW blockchain.

By utilizing this Layer 2 PoA blockchain, OctaSpace is able to process a large number of transactions quickly and efficiently for billing operations such as charging users for the services they have used.

Overall, OctaSpace's use of the pirl 51% guard and Layer 2 PoA blockchain demonstrate its commitment to security and efficiency. By employing these techniques, OctaSpace is able to provide a secure and stable platform for its users while also maintaining fast and efficient transactions for billing operations.

2.2 Layer 1 network

OCTA Layer 1 is PoW[9] blockchain network is used for the frontend user financial operations using the native coin OCTA.

Network based on go-ethereum[4] codebase with the following specification:

- Block time is 15 seconds
- Total supply is unlimited[1]
- Block reward and halving implemented according to Monenary policy
- PirlGuard is used as protection mechanism from 51% attack
- Transaction fee is 21 Gwei

OctaSpace network is designed to be fair and transparent, without any premine or presale. To ensure an equitable distribution of rewards, the genesis difficulty was set at 100Gh, preventing the instant allocation of rewards.

2.3 Layer 2 network

This is a side chain for Layer 1 network, implemented as PoA[8] network with a set of validators.

Used for handling internal transactions for the node services its lightning-fast performance and its seamlessly handling of high-frequency, high-usage operations, resulting in reduced operational costs, speedy transactions and a massive boost in charging operations throughput.

2.4 Blockchain gateways

These gateways provides unified API for **OCTA CORE** layer to give it ability work with both blockchain networks.

This API is private and not accessible outside.

2.5 OCTA CORE and Application Layer

The engine of our system, the compute rental manager, seamlessly handles all requests for compute rentals and communicates with the nodes and user applications to make it effortless for users to access and use the resources of connected nodes.

This layer is designed to be user-friendly, so even those without technical expertise can take advantage of the available resources with ease.

Its the core engine of our system, it's responsible for the following operations:

- Communicates, monitoring and low level interaction with nodes
- Handle requests for computing resources
- Provide interface for creating services on a top of resources provided by nodes
- Services usage charging and billing operations
- Provide API for automation or integration with third party systems
- Fraud control
- Statistic and telemetry of system usage

2.6 Resource Layer

While the Octa Chain is very capable at processing large amounts of transactions the real aim of the project is to provide real world applications and to bridge the gap of computational owners and computational users. Octa nodes were built to accomplish this.

This layer consist of hardware(nodes) connected to the OctaSpace cloud.

[1] Total supply will be reviewed after Mahasim fork

Nodes are the foundation of our compute and services marketplace, providing the necessary computational power to meet the demands of tasks.

These computers are equipped with a blend of CPUs, GPUs, memory, and disk space that allows them to handle distributed workloads with ease.

The nodes are connected to the OctaSpace, which enables them to seamlessly work together to deliver optimal performance and efficiency.

The tasks and services may vary, well equipped machines with powerful GPU may performs AI/ML tasks, common machines may acts as VPN gateway, provide disk storage for services like file sharing or host applications deployed by users.

2.7 API and UI

To work with system the following interfaces available:

- Web applicaton with user-friendly interface: `https://cube.octa.space`
- RESTful API
- **octactl** command-line utility that provide user friendly interface to RESTful API

3 Nodes

In the nutshell node is a Linux machine with special software installed.

This software is called **ORC**, using it **OCTA CORE** is able to establish secure communication channel to the node.

Communication between **OCTA CORE** and **ORC** is doing in RPC[10] like manner.

Secure channel is implemented using HTTPS[5] protocol with validating each request using security token.

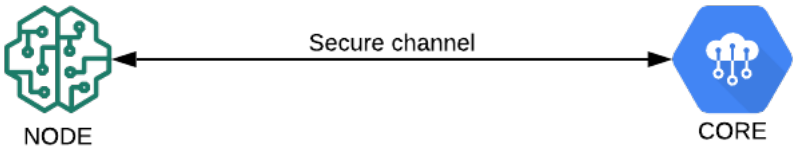

Figure 2: Secure communication channel

ORC is responsible for the following activities:

- Detect installed hardware: CPU, GPUs, RAM, volume of disk storage
- Collect metrics about hardware usage, like free/used disk/ram space, CPU and GPU load, temperature and fans speed
- Manage Docker[2] containers and Firecracker[3] microVMs

There are two types of nodes: **blockchain** and **service**

The first one is responsible for support of OCTA Layer 1 blockchain by running network node software. As a result such nodes making network more stable, distributed, more latency fair and speed up the synchronization.

Service node provides resources we used to implement services for the end users.

3.1 Hardware and software requirements

To cover a wide range of supported hardware, node software can be installed on any machine with x86_64[12] or ARM[1] architecture.

In order to run the node, the hardware must meet the minimum system requirements, which are as follows: 1 CPU, 1 Gb of RAM, and 10Gb of free disk space.

However, the requirements for hardware may vary depending on the intended purpose of the node. For instance, a node that provides only VPN services may only need to meet the minimum requirements. Conversely, nodes that are designed to perform AI/ML tasks will require powerful GPUs connected with a high bandwidth PCIe interface, as well as ample disk storage.

It's worth noting that both NVIDIA and AMD GPUs are supported by the system.

To determine the performance of a node, the following measurements are taken:

- Network upload/download speed
- Disk write speed
- GPU performance using AI benchmark (only for NVIDIA)

These performance metrics help users to choose the hardware they need for their tasks.

The node software can be installed on any Linux distribution; however, we primarily focus on Ubuntu LTS or Debian as the recommended operating systems.

Support for Windows in development.

3.2 Security

To eliminate possible security risks for users running the software **ORC** on their machines, we follow a set of rules and guidelines, which include:

- **ORC** is open sourced, this may possible to make audit by other people that software does not have malicious code
- Keep code base of **ORC** is small as possible for ease auditing
- Node software run under non privilaged user and don't have any permissions which not need for operation
- Regular software updates to ensure that the latest security patches are applied. Along with comprehensive testing of software to detect and address any potential security issues.

3.3 Verification

Ensuring that node infrastructure operates smoothly and reliably is crucial for providing high-quality services to end-users.

Each new node that joins the cloud must be verified and confirmed to meet the necessary requirements and provide services.

Periodically, we conduct re-verification checks on verified nodes. Therefore, it's crucial to monitor the status of your nodes to avoid having them changed to an unverified status.

The checks performed to ensure that the node is properly configured may include but are not limited to:

- Meet minimum hardware requirements
- System clock is synchronized
- All necessary network ports are open
- GPU driver is correctly installed

The list of checks will be expanding in the future to ensure even greater accuracy and reliability.

The following restrictions are applied for the unverified nodes:

- Nodes can't provide services
- Nodes can't participate in staking

4 Services

OctaSpace is committed to providing services that address the needs of end-users, offering solutions to everyday problems. These services range from performing complex calculations to simple file sharing between friends, and more. Our team is continually striving to bring new and innovative services to our users, and we are excited to share with you some of the services that we have already implemented or plan to implement in the near future.

We believe that our services will be of significant interest to users and will have practical applications in their daily lives. Our goal is to continually improve and evolve our offerings to meet the changing needs of our users.

4.1 GPU marketplace

The unique GPU marketplace where users can rent or rent out GPU compute power, our GPU rental service provides users with the ability to leverage powerful GPUs for tasks in AI/ML, CGI rendering, and other fields that require high-performance computing. Our support for both NVIDIA and AMD GPUs expands the range of tasks that can be tackled.

Users can access their rented GPU instances via the secure SSH protocol, and also have the option of using Jupyter[6]and LiveBook[7] systems for interactive access. These rented GPU instances can be combined into a cluster, allowing for the development and execution of distributed programs. For instance, users can use TensorFlow or PyTorch for distributed training of ML models.

By providing powerful and flexible GPU rental options, we aim to empower users with the resources they need to solve complex problems and accelerate their projects.

4.2 VPN

OCTA VPN[11] offers a variety of key benefits to its users. One of the main advantages is its ease of setup, which is made possible by utilizing non-modified and open-source software that is compatible with a wide range of platforms.

Additionally, users have the flexibility to choose from a variety of VPN technologies to suit their needs. There are also no limitations on the number of devices that can be connected simultaneously.

To add to that, the billing model is pay-as-you-go, which means you are billed only for the amount of data you use. This gives you complete control over your usage and costs.

Service don't any logging of the user traffic or DNS requests.

VPN access is implemented using the following technologies:

- WireGuard
- ShadowSocks
- OpenVPN

In the future more VPN types will be added, including some to bypass China's golden shield (Great Firewall of China).

4.3 Remote Video Gaming

Many have tried and failed with Stedia recently falling. All these services essentially rely on massive centralized server farms which inherently introduces huge latency issues resulting in bad gaming experience if your not located with 100km of the facility.

But with OctaSpace and enough nodes it may finally be possible for a remote gaming service to be widly used utilizing nodes across the globe allowing gamers to use share unused hardware. By connecting to say your neighbours unused pc would solve the major problem of latency in remote gaming services.

4.4 HashCache

HashCache is a Password Recovery service for performing password cracking operations on a large scale. It utilizes multiple nodes working together in a coordinated manner to speed up the cracking process. This is

achieved by dividing the cracking workload among the nodes, allowing them to work in parallel on different parts of the task.

The system is designed to be highly configurable and can be optimized for different types of cracking operations, such as dictionary attacks, brute force attacks, and others.

4.5 Distributed Rendering

Another service example would be a distributed handbrake video encoding network. To allow content creators or studios to quickly render out massive edits economically.

4.6 Instant File Sharing

This service provide a easy way to upload a file, share a short link, and after it is downloaded, the file is completely deleted.

Key benefits:

- All files are encrypted
- It's possible to set expiration (Time To Live), after such period file will be deleted
- Simple RESTful API

5 Monetary Policy

OCTA is the main payment instrument and the currency used to pay for the services provided by the project, rewards to node owners and dividend payments for OCTA holders.

Due to unlimited emission to prevent too much inflation the following monetary policy is implemented to decrease inflation:

Era	Start block	Total	Miner	Staking	Dev
Octa	1	8	6.5	0	1.5
Arcturus	650_000	8	5	1.5	1.5
Oldenburg	1_000_000	7.5	4	2	1.5
Zagami	1_500_000	7	3.5	2.5	1
Springwater	2_000_000	6.5	3	3	0.5
Polaris	2_500_000	6	2.8	2.8	0.4
Mahasim	3_000_000	5	2.3	2.3	0.4

Table 1: Reward distribution according to era

This policy should decrease inflation by changing amount of block reward dependents of era.

Era	Start date	End date	Total coins	Duration
Octa	Jun 19 2022	Oct 09 2022	5_200_000	≈112 days
Arcturus	Oct 09 2022	Dec 08 2022	2_800_000	≈60 days
Oldenburg	Dec 08 2022	Mar 04 2023	3_750_000	≈86 days
Zagami	Mar 04 2023	May 09 2023	3_500_000	≈86 days
Springwater	May 09 2023	Aug 23 2023	3_250_000	≈86 days
Polaris	Aug 23 2023	Nov 17 2023	3_000_000,	≈86 days
Mahasim	Nov 17 2023	∞	2_500_000	∞

Table 2: Approximate calculation of era timeline and reward distribution

Total amount of coins mined before Mahasim era is ≈18_350_000

6 Staking

To incentivize interest in holding, a staking mechanism is introduced. People who lock some amount of coins and run a node are rewarded. To activate staking, the following requirements must be met:

- Collateral - 100_000 OCTA
- Node reliability for the last 30 days $\geq 75\%$
- Node must be verified

To begin staking, you need an address wallet with sufficient funds and link it to an existing node. Rewards will be sent to the provided wallet.

If the system detects a balance lower than the collateral, staking will be disabled for that wallet-node pair for several rounds.

6.1 Reward distribution scheme

The billing period is 1 week (every 7 days).

Assuming that block time is always 15 seconds, the amount of blocks mined would be:

- 1 minute - 4
- 1 hour - 240
- 1 day - 5760
- 7 days - 40320
- 30 days - 172800

We will use the reward amount for the Arcturus and Oldenburg eras to calculate; 1.5 OCTA from each block goes to the staking fund.

In total, 60480 OCTA will be mined in 1 week.

Let's say 10 nodes are in operation; the 60480 OCTA will be shared between them in the following proportion:

- 60% (base reward) - 36240 OCTA, 3624 for each node
- 20% (gpu reward) - 12080 OCTA, 1208 for each node which provide renting service with GPU
- 10% (vpn reward) - 6040 OCTA, 604 for each node which provide VPN service
- 10% (rent reward) - 6040 OCTA, 604 for each node which provide renting service

If there are no nodes to which 40% of coins (except for the base reward) can be distributed, the coins will remain in the staking fund for the next round.

The most profitable variant is to have a node with a GPU that provides both services. Such a node will receive 6040 OCTA per week, as well as additional payments for service usage.

Using the current scheme, we can calculate a minimal monthly return on investment (ROI) as follows:

$$ROI = \frac{MonthReward/Collateral * 100}{N} = \frac{M}{100} * 60$$

Where **MonthReward** is 172800 blocks multiply by staking reward according to current era.

N is amount of nodes participated in staking, for example 10.

Therefore the minimal ROI for Arcturus and Oldenburg eras will be:

$$ROI = \frac{172800/100000 * 100}{10} = \frac{17.28}{100} * 60 = 10.36\%$$

References

[1] *ARM architecture*. https://en.wikipedia.org/wiki/ARM_architecture_family.

[2] *Docker*. https://www.docker.com.

[3] *Firecracker*. https://firecracker-microvm.github.io.

[4] *Go implementation of the Ethereum protocol*. https://github.com/ethereum/go-ethereum.

[5] *HTTP Over TLS*. https://tools.ietf.org/html/rfc2818.

[6] *JupyterLab*. https://jupyter.org.

[7] *LiveBook*. https://livebook.dev.

[8] *Proof of Authority*. https://en.wikipedia.org/wiki/Proof_of_authority.

[9] *Proof of Work*. https://en.wikipedia.org/wiki/Proof_of_work.

[10] *Remote Procedure Call*. https://en.wikipedia.org/wiki/Remote_procedure_call.

[11] *Virtual Private Network*. https://en.wikipedia.org/wiki/Virtual_private_network.

[12] *x86-64 architecture*. https://en.wikipedia.org/wiki/X86-64.