HPChain

A Decentralized High-Performance Computing Power Network

Advancing AI and Cloud Gaming with our cost-efficient, high-performance GPU network.

Table of contents

Abstract

As we stand on the threshold of a fresh wave of technological revolution, 5G and Artificial Intelligence (AI) have primed our world for the next Internet evolution. NVIDIA, surpassing a trillion-dollar market cap, has risen to be the premier global chip manufacturer, signaling the emergence of the high-performance computing point in history. From the AI revolution to the cloud gaming industry, we are witnessing an unprecedented surge in computational power.

New technology is empowering today's significant evolution in computational power which previously had always been the bottleneck to the internet's expansion. The biggest paradigm shifts in the demand of high-performance computing are taking place in areas such as AI and cloud gaming. Current product supply chains and delivery mechanisms struggle to keep pace with these growing demands. Our objective is to apply our novel "Cloud & Chain" business model built on the blockchain using a distributed cloud platform as the point of entry to create the computational power supply and demand ecosystem.

Mission and Vision

Mission: At HPChain we aim to revolutionize computing power resources through an accessible, cost-effective, and decentralized high-performance computing network, reshaping how computing power is both accessed and utilized.

By using blockchain technology, we are able to create a worldwide network of superior computational power. By using the HPChain blockchain network, we create an open-sourced and distributed cloud computing platform within a trusted environment. HPChain is committed to delivering accessible, cost-effective computational power to developers and users worldwide engaged in AI and cloud gaming.

We are building a global computational power network to create a massively scalable high-performance computing (HPC) network. This network allows for a broader scale and scope of elastic scheduling creating globalized sharing of high-performance computing power. Algorithm workers are computational nodes managed by users creating the community communication hub helping to reduce operational costs and improve efficiency. AI users will be able to choose cost-effective computing power from anywhere in the world , while edge cloud computing users can opt for nearby computing resources to optimize their experience. For example, in cloud gaming where ultra-low latency is essential, HPChain will enable computing nodes from any location to participate freely.This creates a globally dispersed, high-density, and extensive coverage computing power network that can offer nearby resources to users.

Vision: HPChain envisions empowering hundreds of millions of individuals to access AI and enjoy cloud gaming. We believe in democratizing technology and making high-performance computing accessible and beneficial to everyone.

Part 1: Market Background

1.1 Emerging Surge in the High-Performance Computing Industry

Computational power equals productivity, a global consensus that makes computational capacity a critical driver for economic growth worldwide.

Rapid advancements in artificial intelligence (AI) and cloud gaming have led to an explosive increase in the demand for high-performance computing (HPC) power. AI requires a vast amount of high-performance computing resources to train and optimize models, while the cloud gaming industry requires a significant amount of real-time cloud rendering computation. Both domains rely heavily on robust high-performance computing power.

1.1.1 The Domain of Artificial Intelligence

AI Guided by Computing (AIGC) has sparked a fresh wave of revolution within the AI domain. In late November 2022, ChatGPT launched and quickly gained popularity on social media platforms. In just five days, it had over a million registered users. Two months after its launch, ChatGPT's monthly active users had reached 100 million, marking it as the fastest-growing consumer application in history. This trend continues unabated.

Companies like Stable Diffusion and Midjourney followed suit, achieving notable success. Stable Diffusion achieved a valuation of $1 billion within two years of inception, and Midjourney, a company made up of 11 employees specializing in AI image generation, generated over $100

million in revenue.

These recent breakthroughs in AI have been nothing short of remarkable. ChatGPT, for example, is a language model that can generate text, translate languages, write different kinds of creative content, and answer your questions in an informative way. It is so powerful that it has been called "near-omnipotent." High performance computing (HPC) is what powers these pioneering AI systems. HPC systems, such as GPUs, can perform millions of calculations per second, making them ideal for training and running AI models.

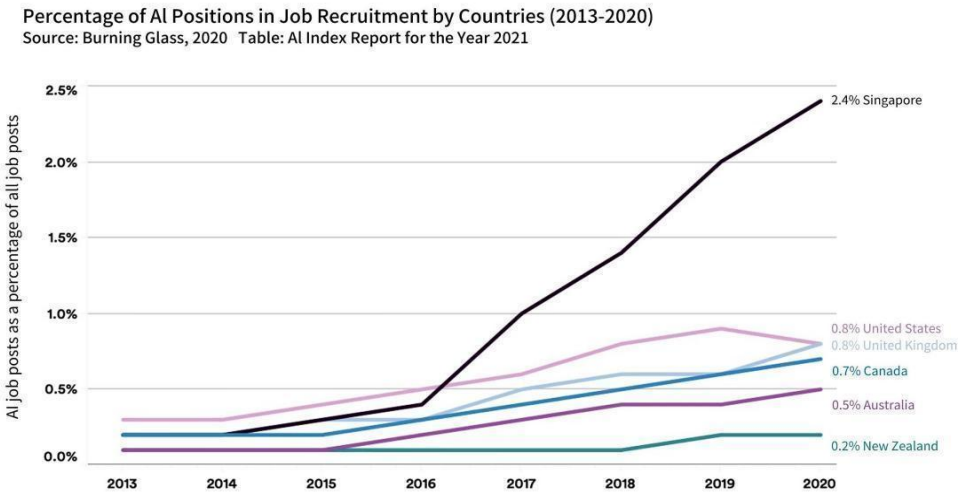
Microsoft-backed OPEN AI's investment of over $1 billion in GPU HPC power helped to make ChatGPT possible. This investment allowed OPEN AI to train ChatGPT on a massive dataset of text and code, which gave it the ability to generate such realistic and creative text. ChatGPT's generative AI requires enormous amounts of computational power for large-scale training. For instance, GPT-3's language model includes 175 billion parameters, almost 117 times more than its predecessor, GPT-2. GPT-4 is expected to have a staggering 1.6 trillion parameters, inevitably leading to a massive demand for parallel computational power.

The development of AI is rapidly accelerating, and the demand for HPC power is only going to grow. As AI systems become more powerful, they will require more and more computational resources to train and run. This will put a strain on current existing HPC infrastructure and developing new HPC systems has become essential to meet the demands of AI.

According to Verified Market Research the global GPU market size was estimated at $25.41 billion in 2020. Given the ever-increasing demand, it is predicted to reach $246.51 billion by 2028, a compound annual growth

rate (CAGR) of 32.82%. Since 2012, OpenAI's estimates suggest the computational power demand for training top global AI models doubles every 3-4 months, with the annual computational power requirement for top training models increasing tenfold. In contrast, Moore's Law suggests that chip computational performance doubles every 18-24 months. This disparity will inevitably accelerate the demand for computational infrastructure.

The future of AI is bright and will depend on the development of HPC systems like HPChain creates that can keep up with its demands. As HPC systems continue to evolve, we can expect to see even more amazing breakthroughs in AI.



1.1.2 Real-Time Cloud Rendering in the Field of Cloud Gaming

Real-time cloud rendering is a technology that allows users to experience high-quality, interactive graphics without the need for powerful hardware on their local devices. It works by offloading the rendering process to

remote servers in the cloud, which can then stream the rendered images back to the user in real time.

The three key components of real-time cloud rendering are:

- Real-time: The rendering process must be completed quickly enough to provide a smooth and immersive experience for the user.
- Cloud: The rendering process is performed on remote servers in the cloud. This allows users to access high-powered rendering hardware without having to purchase it themselves.
- Rendering: The rendering process involves converting 3D textures into images that can be displayed on a screen.

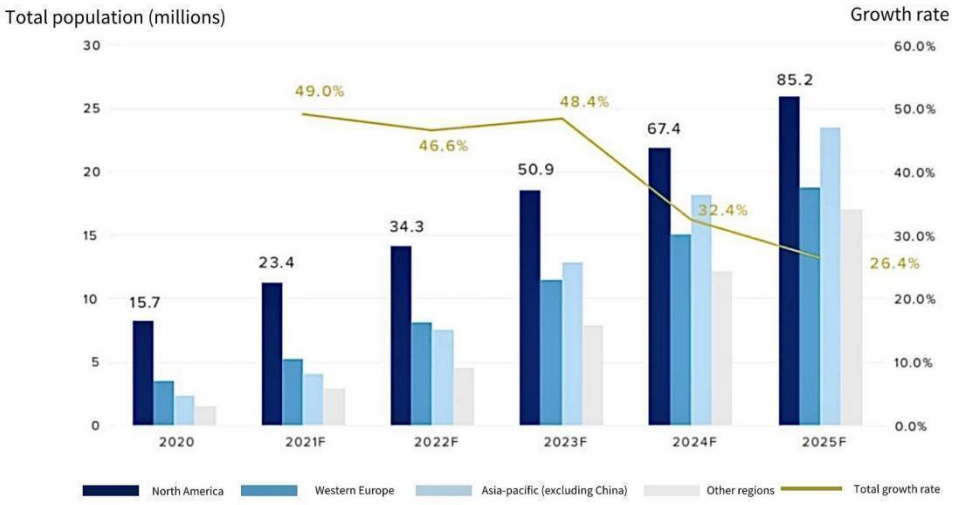
Rendered in the cloud and in real-time necessitates low latency to ensure a seamless user experience. The process can be computationally intensive, but can be performed much faster on powerful servers in the cloud. Cloud gaming is a type of real-time cloud rendering that allows users to play video games without having to download or install them on their local devices. The game is rendered in the cloud and streamed back to the user in real time, allowing them to play without any lag or latency.

High-performance computing power is also essential for real-time cloud rendering. The rendering process can be very computationally intensive and requires low latency with ultra high definition models relying heavily on large scale real-time rendering technology. Real-time cloud rendering is essential for cloud gaming because it allows users to experience high-quality graphics and gameplay without having to purchase expensive hardware. This makes cloud gaming a more accessible and affordable option for gamers of all levels. Real-time cloud rendering

technology used for cloud gaming creates even further demand for high-performance computing power.

The development and maturation of 5G communication technology which makes it possible to stream high-quality video games without lag or latency, has led to a surge of interest and investment into the cloud gaming field ushering in a period of rapid expansion. Tech giants like Amazon, Microsoft, and Google have launched cloud gaming platforms to engage users. The growing number of successful cloud gaming cases has led to an increase in market users and premium paying users. The exponentially increased demand for high-quality graphics and gameplay can only be delivered via cloud gaming.

As of 2022, the global gaming population exceeded 3 billion, with over 200 million dedicated cloud gaming users. Projections suggest that by 2025, cloud gaming users will surpass 500 million, marking the onset of a period of explosive growth with significant potential.



1.2 Evolution in the Demand for Computational Power Industry

The reshaping of "new businesses" by "new computing" continually drives the transformation of "new computing" itself, a cyclical process that is destined to occur. New computing technologies have always reshaped new businesses. For example, the development of the internet led to the creation of new businesses such as Amazon, eBay, and Google. These businesses could not have existed without the internet, which provided them with a platform to reach a global audience.

In turn, the success of these new businesses has helped to drive the development of new computing technologies. For example, Amazon's need to process and store massive amounts of data has led to the development of new cloud computing technologies.

This cyclical process continues to reshape new businesses. HPChain is a new computing technology creating new opportunities for businesses. And as businesses succeed, they will help to drive the development of even newer computing technologies.

1.2.1 Transition in Customer Demographics

The inherent characteristics of customers are evolving, shifting from a B2B-centric approach to an emphasis on B2C. During the PC and mobile internet era, B2B customers were the primary patrons of cloud computing. However, as we transition into an age dictated by artificial intelligence and virtual internet, the B2C sector will attain a comparable significance to the B2B in terms of computational power, hence, mainstreaming the consumer-end users. In the future, the cost of computing power will become as commonplace to the majority of internet

users akin to current bills we pay for the internet and electricity.

This shift to B2C will have a number of implications for the cloud computing industry. The HPChain platform enables new services and pricing models to meet the needs of B2C customers. The platform also allows users access to new technologies to improve the performance and scalability of its services, creating further innovation and growth.

Cloud gaming is a prime example of how the progression of real-time cloud rendering technology is addressing issues related to high-quality user experience and low-entry barriers. With real-time cloud rendering, users can experience high-end games without having to purchase expensive hardware. Their device only needs to have basic video decoding capabilities and network connectivity. This reduces the device threshold and allows PC and console games to be smoothly transitioned to mobile platforms. As a result, more consumer-end users are exploring gaming through the cloud.

AI on Edge Computing (AIGC) is another example of how high-performance computing power is making AI products more accessible to individual users. AIGC allows AI models to be run on edge devices, such as smartphones and laptops. This means that users do not need to have powerful computers to access AI-powered applications.

1.2.2 Transformation in Demand Patterns

The demand for computational power is shifting from traditional cloud computing to edge computing. Consumer-end users are inherently global and immensely dispersed, meaning that they need access to computational power from anywhere in the world. Traditional cloud computing is not well-suited to meet this need as it requires users to

connect to a centralized data center. T Edge computing is a more decentralized approach to computing that places computing resources closer to the user, improving latency and performance. Jensen Huang, NVIDIA's founder, stated that “next-generation computing will not be limited to supercomputing centers. It will be distributed across a multitude of nodes, extended to the edge, and connected to remote sensors”

Real-time cloud rendering demands ultra-low latency. This is because the user needs to see the rendered image as soon as possible after they interact with the application. If there is too much latency, the user will experience a choppy or unresponsive experience. Many of AIGC's future inference scenarios are generated in real-time, prioritizing immediacy, which requires computational power to be as proximate as possible. Hence, to deliver a seamless interactive experience, computational services must be provided by edge cloud nodes at the infrastructure level. This means that the computational resources for real-time cloud rendering and AI inference should be located as close to the user as possible. This will ensure a high-performance, low-latency, real-time interactive content experience. Traditional cloud computing power will face challenges in terms of network platform reliability and edge politics when globalizing, and may struggle to meet these widely dispersed edge demands. HPChain’s edge cloud computing infrastructure is a more decentralized approach to computing that can address these challenges by creating edge cloud nodes located closer to users.

1.2.3 Transformation in Product Utilization

Traditionally, the utilization of cloud computing power is continuous. However, the future utilization of computational power, whether it's B2B

or B2C, will become elastic. HPChain addresses the different scales of computing power required by different business needs through its proprietary dynamic elastic utilization scheduling platform.

There are three main reasons for this shift to elastic utilization:

1. Varying business needs require different scales of computing power. For example , in the realm of cloud gaming, the peak and off-peak periods of online gamers entail significantly different computing power demands. Elastic use of computing power can address this demand effectively, by dynamically scaling up or down the number of servers and computing resources available to cloud gaming providers, based on the actual demand.

2. As the demand for computing power escalates, customers or enterprises need to procure a vast number of servers and computing resources to meet daily requirements. This necessitates high investment and could potentially lead to resource wastage. Elastic utilization of computing power allocates computing resources dynamically thereby saving money t and optimizing consumption. .

3. Elastic utilization allows businesses to save money on computing costs and reduce the need for new data centers. When businesses can dynamically scale their computing resources up or down, they don't need to build new data centers to accommodate peak demand. This helps to reduce the environmental impact of data centers and makes elastic utilization of computing power a preferred method of utilization

1.3 Challenges in the Computing Power Industry

Factors such as customer transitions, alterations in demand, and shifts in usage patterns have resulted in a supply-demand imbalance in the computing power sector. The industry is currently grappling with a "triple conundrum" encompassing cost, efficiency, and user experience.

Customer transitions: Customers are increasingly moving away from on-premises data centers and towards cloud-based solutions. This has put a strain on the supply of computing power, as cloud providers need to invest in new infrastructure to meet the growing demand.

Changes in demand: The demand for computing power is also changing in terms of the types of applications that are being used. For example, the rise of artificial intelligence and machine learning is driving demand for more powerful computing resources. This is putting a strain on the supply of computing power, as providers need to upgrade their infrastructure to meet the growing demand for these applications.

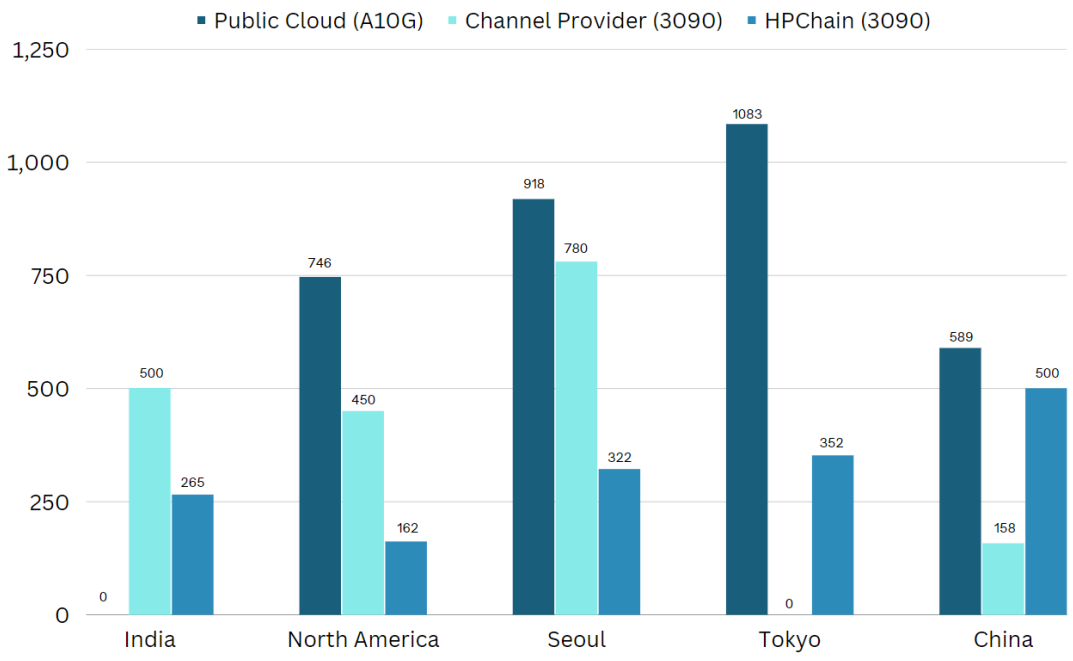
Shifts in usage patterns: The way that people use computing power is also changing. For example, the rise of mobile devices is leading to more decentralized computing. This means that people are using computing resources that are located closer to them, rather than relying on centralized data centers. This is putting a strain on the supply of computing power, as providers need to build more data centers in different locations to meet the growing demand.

The "triple conundrum" that the computing power industry is facing is a complex challenge. The HPChain cloud platform provides a number solutions to help address the imbalances by increasing the supply of computing power, improving the efficiency of computing power,

improving the user experience and making computing power accessible.

1.3.1 Cost Concerns

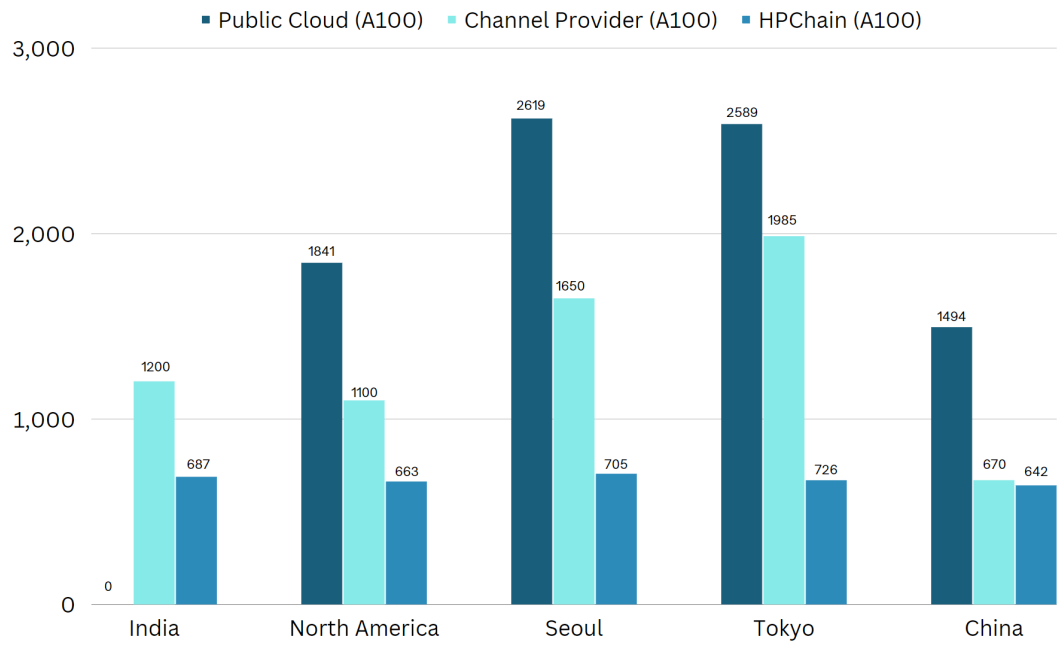
The cost of computational power is a combination of infrastructural costs (which include land, energy, and network expenses), labor costs, and the expenses associated with purchasing and operating computational servers. Developed countries and regions typically have higher infrastructure costs due to the high cost of land, energy, and labor. Developed countries with superior artificial intelligence capabilities are at a disadvantage because they have to pay more for computational power. One of the foremost challenges in the AI industry today is reducing the cost of computational power usage.



There is a substantial amount of idle computational power globally. This is due to the fact that traditional computational power manufacturers are not interconnected. This lack of connectivity creates "computational power islands," which prevents the global sharing and logical allocation

of computational power to achieve maximum utilization. This lack of optimization significantly contributes to the high cost of computational power.

The dominance of chip manufacturers like NVIDIA has also led to limitations in using consumer-grade GPUs for enterprise applications, with the cost of utilizing enterprise-grade GPUs being particularly high. This is because enterprise-grade GPUs are designed for specific tasks and require specialized drivers and software.



1.3.2 Efficiency Challenges

The traditional internet era had a large customer market that allowed traditional cloud computing providers to meet customer needs efficiently through centralized services. However, the advent of extensive AI models has resulted in substantial GPU computational power requirements. These needs are so vast that even top cloud computing providers face limitations

in their delivery capacity, leading to delayed responses for customers. This issue is further amplified in the small-to-medium enterprises (SMEs) and individual users' market.

According to data from June 2023, high-end enterprise graphics card cloud computing market customers have experienced a queue time exceeding three months without price adjustments. Such delays are simply untenable for SMEs and individual developers. Hence, the improvement of input and operational efficiency in AI computational power is an urgent problem that needs addressing In the cloud gaming market, the need for real-time cloud rendering requires low latency which calls for edge computing. The edge computing paradigm implies a large scale of investment and highly dispersed operations. Addressing these challenges is paramount in the evolving field of computational power, especially concerning cloud gaming.

1.3.3 User Experience Complications

Establishing global trust in computational power usage is challenging the framework of centralized cloud computing platforms. There are several reasons for this.

First, these centralized platforms amass substantial computational resources and data, creating a concentration of risk. In the event of an attack or failure, users could face significant losses, consequently reducing their trust in cloud computing platforms.

Secondly, the lack of transparency in the way these platforms utilize computational resources and data provided by users is problematic. Users often are unable to verify whether their resources and data are being used as agreed, resulting in mistrust and an increased sense of risk.

Thirdly, data privacy concerns arise as these platforms handle and process user data. Users may be uncertain about whether their data is adequately protected and used in compliance with legal standards, further eroding trust in the platform.

Finally, geopolitical issues present significant challenges for cloud computing manufacturers and users alike. Traditional cloud computing manufacturers have reaped substantial benefits during their market development process, whereas the benefits for the people that created the ecosystem have been lacking. This imbalance does not align with the evolving needs of the new era.

It's become critical to develop a new set of rules that are globally applicable, trustworthy, and mutually beneficial to address these user experience issues effectively.

Part 2: Business Model and Strategy

2.1 Design philosophy of HPChain

Our founding team is passionate about addressing the challenges inherent in the application of computational power within next-generation high-performance computing. We leverage years of collective expertise in artificial intelligence, cloud gaming, and high-performance cloud computing to develop innovative solutions that are both technically sound and commercially viable. We are committed to working with our partners to accelerate the adoption of next-generation high-performance computing. We believe that this technology has the potential to revolutionize a wide range of industries, and are excited to play a role in making it a reality.

Our business model is established on comprehensive industry knowledge and commercialization expertise:

1. Promoting the global dispatchment of AI computational power from low-cost regions to high-cost regions.
2. Reducing computational power costs. For instance, exporting computational power from the United States to Japan, South Korea, or Singapore can lead to a decrease in costs of at least 30%.
3. Building a global computational power-sharing platform, akin to the business model of Uber, creating a peer-to-peer marketplace for computational power..Individuals and businesses will be able to rent out their unused computational resources to others. This will allow businesses to significantly reduce their computational power costs,

while also helping to reduce the environmental impact of AI.

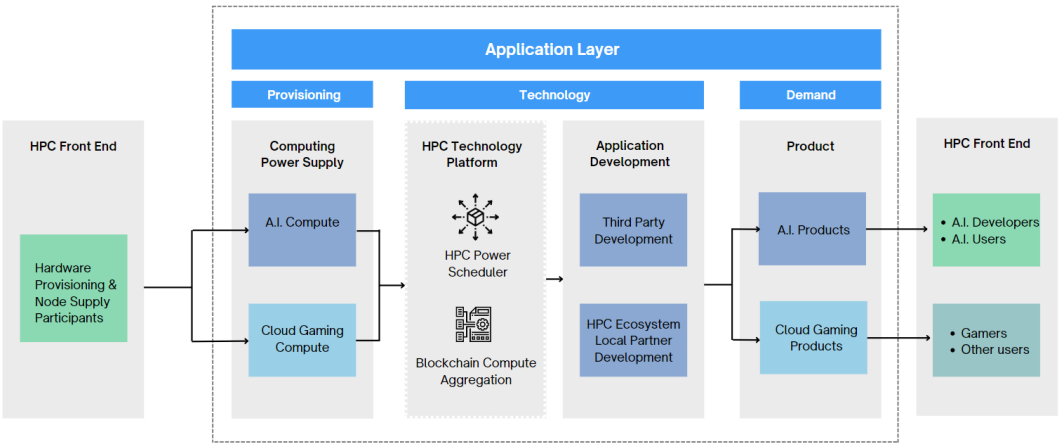
4. Establishing a C2C platform that enables the use of consumer-grade GPUs for business applications, leading to a further reduction in computational power costs. The C2C platform will also help to counteract the monopoly of chip manufacturers allowing individuals and businesses to rent out their unused consumer-grade GPUs to others.

5. We are committed to offering localized services for businesses, which allows us to effectively circumvent geopolitical risks.

6. To facilitate cooperative decision-making and mutual benefit, we are developing a set of decentralized rules using blockchain technology. The decentralized rules will empower ecosystem participants to make decisions, establish rules, develop the platform, and share the benefits making sure the platform is sustainable in the long term.

2.2 HPChain solutions

HPChain relies on a two-layer network mechanism, combining blockchain and distributed cloud technology. The blockchain layer pools global high-performance computing power, while the cloud layer is designed to cater to industry-specific applications, allowing us to provide a comprehensive and flexible ecosystem for businesses of all sizes.



2.2.1 HPChain's Great initiative

A prominent achievement of HPChain is the democratization of the cloud computing market. Previously, the cloud computing market was dominated by a handful of corporate giants. However, HPChain has transformed this market into an inclusive marketplace, inviting participation from everyone. This has ensured that all participants share in the development dividends of the cloud computing

2.2.2 Commercial design of HPChain

On the supply side, HPChain harnesses the potential of blockchain technology to amass computational power. Enterprises and individuals alike can contribute computational power that meets market demand , enabling a global computing power sharing network. Computational power nodes can take on various forms, including GPUs, FPGAs, and more. They can range from large-scale clusters to small and medium-sized nodes, or even individual idle computational power nodes. The primary revenue source for these nodes derives from rental fees and token rewards paid by users. The rental fee is essentially the payment made by high-performance computing power users to purchase

computational power resources through HPChain, with all transactions conducted using HPChain. The reward tokens are earned in accordance with the reward rules of HPChain's entire network computing power.

On the application side, any team or individual with a need for cloud HPC or holding resources demanded by computing power users can establish their self-operated brand cloud computing platform within the HPChain network. Our target markets are primarily the AI and cloud gaming industries. HPChain not only empowers utilizing computing power at a more affordable cost with greater convenience but also enables users to earn HPChain token rewards. Operators of distributed cloud HPChain who create their own brands based on the HPChain network platform can also earn rewards. HPChain relies on smart contracts for transactions and rule building, and all transactions are conducted through HPChain .

We envision a scenario where a college student in Singapore can build their own artificial intelligence cloud computing company , connected to mainstream deep learning frameworks such as TensorFlow, Caffe, CNTK, and more. This would ensure that users requiring computing power can tap into the high-performance computing power of the HPChain network from anywhere. We plan to collaborate with globally renowned game manufacturers and AIGC players to launch classic application scenarios and set benchmarks for high-performance computing power applications.

At its core, the role of HPChain network design is to provide the essential infrastructure for high-performance computing operations and a platform that is cost-effective, efficient, reliable, global, and friendly to both developers and users. This forms a synergistic ecosystem that fosters mutual growth and advancement.

2.2.3 The Core Values of HPChain

- Cost

In terms of cost, global computing power sharing disrupts the isolation or 'island effect' of computing power. This helps to reduce resource idle rates and operational costs of computing power, which can lead to lowered overall expenses.. Channeling AI computing power from low-cost regions to high-cost areas helps to further diminish user costs. Our groundbreaking approach enables the utilization of consumer-grade graphics cards by users. This not only significantly brings down the cost of using computational power but also challenges the chip market monopolies.

- Efficiency

Efficiency is another cornerstone of HPChain's value. We believe in mobilizing social capital and resources to optimize both the investment in and operation of computational power to help enhance efficiency.

- Experience

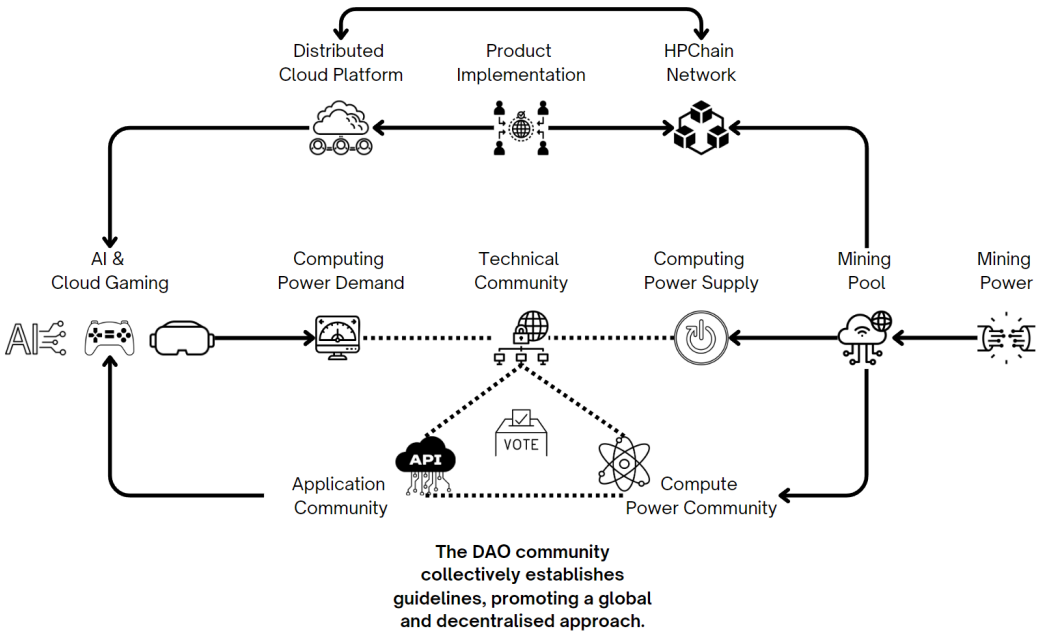
Stemming from our commitment to offering an exceptional user experience, we aim to actualize the global scheduling of computational power while providing localized services. This will allow us to bypass geopolitical risks and provide users with a seamless experience, regardless of their location.

2.2.4 Highlights of HPChain's Business Model

Cloud and Blockchain technologies, forming a closed business loop.

HPChain is a decentralized ecosystem that is built and shared by the global community. This participatory model encourages everyone to engage in the HPChain ecosystem, contribute computational power, develop users, and obtain rewards from the HPChain network. Blockchain technology safeguards all rules, making them immune to alteration or tampering. Management and decision-making are executed through the Decentralized Autonomous Organization (DAO), or the HPChain community. This ensures that the rights and interests of all participants are protected, and that the ecosystem is robust, equitable, and progressive.

2.3 HPChain Ecosystem diagram



2.4 Core Technology

HPChain's core technology is an ensemble of cloud and blockchain. This multifaceted network is designed to meet the commercial orientation of

HPChain's entire business design.

2.4.1 Leading Global AI Cloud Platform Technology

1. Decentralized computing engine: This proprietary computing engine, characterized by unstructured and dynamically adjustable capabilities, allows for flexibility and adaptability.

2. The deliberate design of a loose P2P network: supporting tens of millions of nodes, enabling a more expansive and interconnected system through the collaborative work of distributed nodes.

3. Shared system resources: including CPU and GPU, a balanced computational load is maintained, ensuring efficient resource utilization and task execution.

4. Efficiently compressed protocol encoding: The employment of efficient compressed protocol encoding technology fortifies data integrity and transmission efficiency.

2.4.2 Global Leading Cloud Gaming Technology

1. Virtualization infrastructure and PC Farm: Resource scheduling via virtualization technology allows dynamic online resource expansion without system interruption. By using a customized high-density server, one server can be configured with multiple independent graphics cards, which improves the price-performance ratio.

2. Graphics cards pass-through and GPU virtualization: By supporting GPU pass-through which enhances the clouding of various 3D applications without any further modification. Maximizing GPU resources and encompassing both consumer-level and professional graphics cards. The risk of GPU vendor lock-in will be avoided.

Through GPU virtualization, GPU slices are allocated to virtual machines for use, hence that multiple virtual machines can share one GPU, and the video memory of each virtual machine is independent of each other and not shared.

3. Low latency and low jitter network transmission: Providing 10 Gigabit bandwidth with low latency and zero jitter transmission for best real-time gaming experience.

4. Encoding technology: Following hard-coded technologies such as H264 and H265 and combined with Intel's dedicated hardware decoding which will provide efficient video encoding and decoding functions to reduce the bandwidth requirements of network transmission.

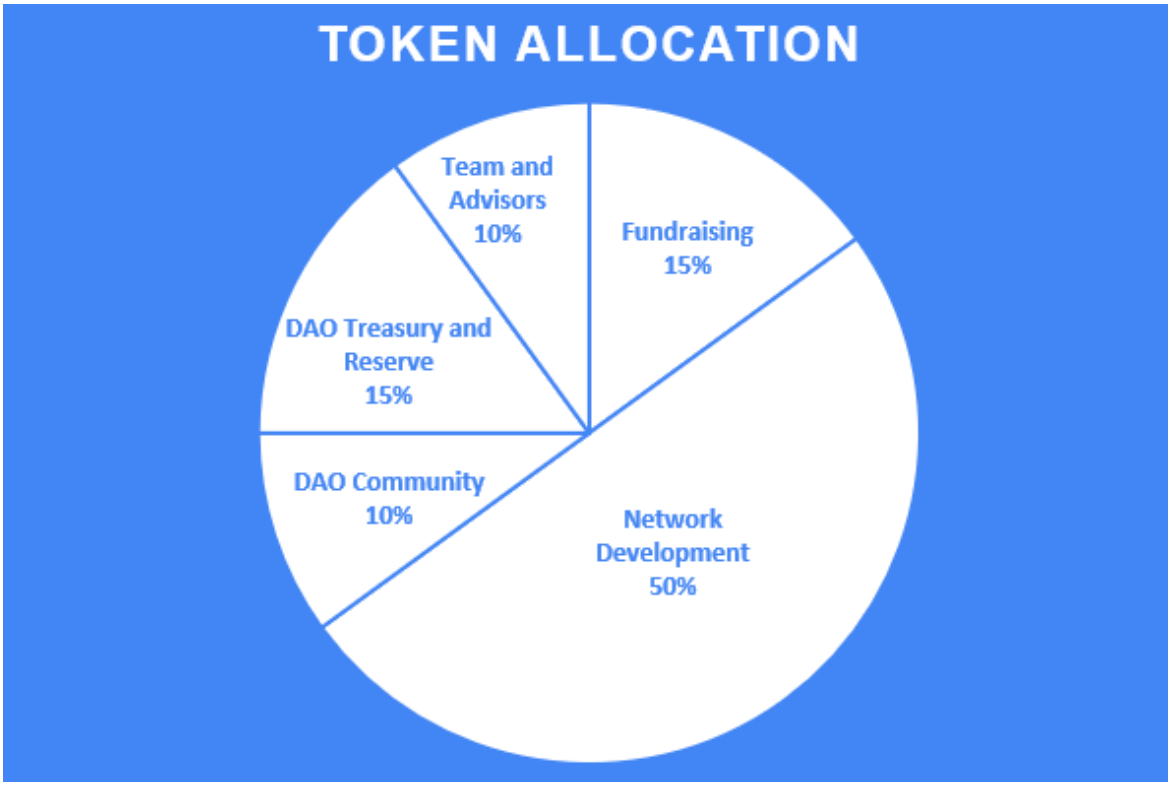
2.4.3 Application-Oriented Blockchain Technology

1. By adopting modular blockchain technology, the functions of each part of the blockchain are layered on the system architecture.

2. Increase the degree of decentralization of the blockchain and improve the throughput and capacity of the blockchain. At the same time improve the reusability, flexibility, convenience of maintenance and upgrades, security and scalability that enable the construction and deployment of the blockchain applications become easier and more efficient.

Part 3. Token distribution

Token Name: HPChain Token (HPC) **Token Supply:** 500 Million



The above rules are for reference only, and the details are based on the final mainnet launch, and other economic model details are to be announced.

Part 4. Roadmap

HPC Roadmap