Translation

The following white paper, issued by a PRC state-run think tank, calculates China's total computing power as 135 exaflops. The white paper also compares China's compute with other countries, and compares the computing power of different Chinese provinces and regions. The white paper omits Hong Kong, Macau, and Taiwan from its calculations. It also estimates general purpose compute numbers based on servers only, excluding PCs, mobile phones, and other terminal devices.

Title
White Paper on China's Computing Power Development Index
中国算力发展指数白皮书

Author
China Academy of Information and Communications Technology (CAICT; 中国信息通信研究院; 中国信通院). CAICT is a think tank under the PRC Ministry of Industry and Information Technology (MIIT; 工业和信息化部; 工信部).

Source
CAICT website. The white paper is dated September 2021 and was uploaded to the website on September 18, 2021.

The Chinese source text is available online at:
http://www.caict.ac.cn/kxyj/qwfb/bps/202109/P020210918521091309950.pdf
An archived version of the Chinese source text is available online at: https://perma.cc/YS34-L84C
U.S. $1 ≈ 6.4 Chinese Yuan Renminbi (RMB), as of November 10, 2021.

Translation Date	Translator	Editor
November 10, 2021	Etcetera Language Group, Inc.	Ben Murphy, CSET Translation Lead

White Paper on China's Computing Power Development Index

China Academy of Information and Communications Technology (CAICT)
September 2021

Copyright Statement

Preface

A new round of science and technology (S&T) revolution and industrial transformation is now reshaping the structure of the global economy. As a new productive force (生产力) in the digital economy era, computing power is the solid foundation supporting the digital economy's development. It plays an important role in driving S&T progress and promoting digital transformation, as well as in supporting digital society development. Computing power has become a new focus of global strategic competition. It is an important engine of national economic development, and the computing power of countries around the world shows a significant positive correlation with their economic development level.

The development of China's computing power accelerated in 2020 despite headwinds, and displayed the following characteristics:

The scale of computing power continued to expand, and the structure of computing power continued to evolve. China's computing power still maintained a vigorous development trend in 2020, increasing by 48 exaflops (EFlops) year-on-year to reach a total scale of 135 EFlops. In the context of the global epidemic, China's computing power nonetheless maintained a high growth rate of 55%, about 16 percentage points higher than the global growth rate, effectively supporting epidemic prevention and control efforts and economic and social development. In terms of computing power structure, basic computing power remained the main force, but intelligent computing power (智能算力) has increased rapidly. With its share of the total now exceeding 40%, it has become a driver of rapid growth in computing power.

The environment for computing power is being increasingly optimized, and application demand is rising constantly. First, constant optimization of the network environment is providing solid support for computing power development. On the one hand, network bandwidth continues to grow: In 2020, inter-provincial internet export bandwidth nationwide increased by 11%, boosting cross-regional dispatching of computing power. On the other hand, 5G coverage continues to increase, which is accelerating cloud-edge-terminal computing power collaboration. **Second,** constantly increasing investment in computing power is providing momentum for computing power development, with China's information technology (IT) spending reaching 2 trillion [Chinese] yuan Renminbi (RMB), an increase of 7.3%. **Third,** the internet remains the largest industry in terms of computing power demand, accounting for nearly 50% of total demand for computing power, and the telecommunications and finance fields are at industry-leading levels when it comes to the application of computing power.

As computing power boosts economic growth, the pace of development is

accelerating everywhere. The scale of China's computing power industry, represented by computers, reached RMB 2 trillion in 2020, directly driving RMB 1.7 trillion in economic output, and indirectly driving RMB 6.3 trillion in economic output. That is, each 1 RMB invested in computing power drives 3-4 RMB of economic output. The pace of computing power development has accelerated nationwide. The overall computing power development of Beijing-Tianjin-Hebei, the Yangtze River Delta, and the Guangdong-Hong Kong-Macau Greater Bay Area are at leading levels. The scale of computing power in Beijing, Shanghai, Guangzhou and their surrounding provinces and regions is especially prominent: Beijing, Guangdong, and Shanghai are the top three in rank, each with more than 15 EFlops. In terms of computing power environment, the four major urban agglomerations have good network support, computing power investment, and other environmental conditions. And in terms of computing power application, the eastern coastal provinces have greater consumer and industrial application demand, which has a significant pulling effect on computing power growth.

Accelerating the development of computing power is an important guarantee for China to create new advantages in the digital economy, to build the "dual circulation" (“双循环”) new development pattern (新发展格局), and to enhance the country's overall competitiveness. China attaches great importance to the development of computing power. It has specifically proposed laying out a nationwide computing power network of national hubs and nodes, and launched implementation of the "eastern data, western compute"[1] project, in order to build a national computing power network system. Driven by both demand and policy, all parts of the country are vigorously promoting the development of the computing power technology industry, infrastructure construction, and computing power application. This white paper systematically studies the development of computing power in China, and establishes a computing power development index, taking into consideration the characteristics and key influencing factors of computing power development. It comprehensively and objectively evaluates the level of computing power development in each province of China, in order to promote the development of the computing power technology industry, infrastructure construction, and computing power application, to strengthen the coordination and joint action of regions in China on computing power, and to provide effective support for the formulation of computing power development policies for the whole country and each province.

Of course, this white paper still has many shortcomings, and criticism and

[1] Translator's note: "Eastern data, western compute" (“东数西算”) refers to building data centers along China's densely populated and developed east coast, and building computing power in western China, where land and electricity are cheaper.

corrections from the public are sincerely requested.

Contents

List of Figures

List of Tables

I. Computing Power Connotations and Development Framework

1. Connotations of computing power

Viewed in its narrow sense, computing power is the computational ability of equipment to achieve the output of specific results by processing data. In a paper entitled *The Progress of Computing*, 2018 Nobel Prize in Economics winner William D. Nordhaus noted: "Computing power is the amount of information that equipment can process per second based on changes in internal states." Computing chips of various kinds—central processing unit (CPU), graphics processing unit (GPU), field-programmable gate array (FPGA), application-specific integrated circuit (ASIC), etc.—are the core for computing power realization, and massive data processing and various digital applications, hosted on computers, servers, high-performance computing clusters, and various types of intelligent terminals, depend critically on computing power-based processing and calculation. Higher computing power values represent stronger overall computational ability, and a commonly used unit of measurement is the number of floating-point operations per second (FLOPS, 1 EFlops=10^18 FLOPS). According to estimates, 1 EFlops is approximately 5 Tianhe-2A supercomputers, or 250,000 mainstream dual processor servers, or two million mainstream notebook computers.

Viewed in its broader sense, computing power is a new productive force of the digital economy era. It is the solid foundation supporting development of the digital economy. The key resources in the digital economy era are data, computing power, and algorithms, of which data are the new means of production, computing power is the new productive force, and algorithms are the new relations of production. These constitute the most fundamental cornerstones of production in the digital economy era. The rapid development of 5G, cloud computing, big data, the internet of things (IoT), artificial intelligence (AI), and other technologies in the current stage have driven the explosive growth of data and the increasing complexity of algorithms, and brought about a rapid increase in demand for computing power scale, computing power capabilities, etc. The progress of computing power has in turn supported innovation in applications, thereby achieving technology upgrading, innovative development of applications, constant enlargement of the scale of industry, and continuous economic and social progress. With the accelerating pace of commercial use of 5G, the ever-deepening connections between things, the expanding application of computing power in fields such as autonomous driving, smart security, and smart cities, and the growing demand for edge computing and fog computing, the scope and boundaries of computing power are still expanding continuously.

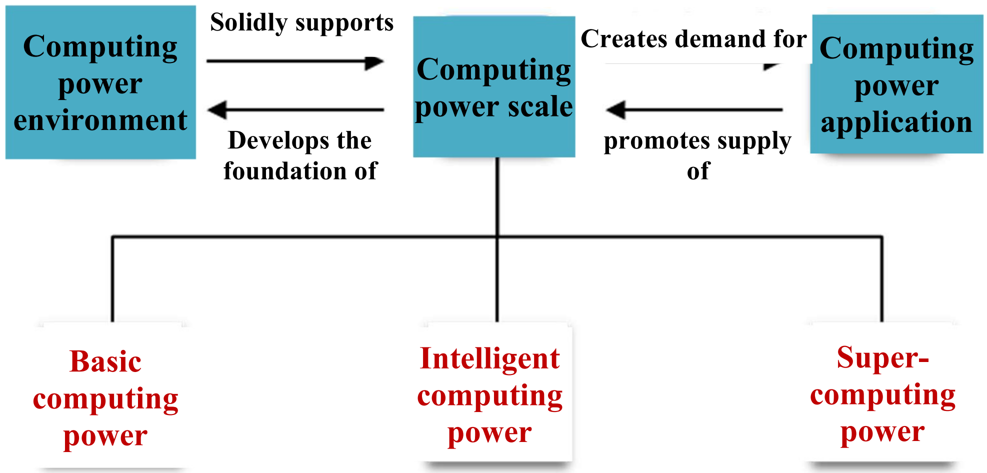
2. Computing Power Development Framework

Taking into consideration the connotations and characteristics of computing power, this white paper establishes an overall computing power development framework from three perspectives: computing power scale, computing power environment, and computing power application. In the digital economy era, the scale of computing power is an important indicator of the level of digitalized productive force development of countries and regions, the computing power environment is an important factor in China's development of new productive forces, and the application of computing power reflects the demand conditions for this new productive force in China. The computing power environment provides solid support for computing power scale development, the application of computing power encourages growth in computing power scale, and the three promote each other and develop synergistically.

The focus of computing power scale at the current stage includes three parts—basic computing power, intelligent computing power and supercomputing power[2]—which provide basic general purpose computing, AI computing, and science and engineering computing, respectively. Basic general purpose computing power is mainly the computing power provided by CPU-based servers; AI computing power is mainly the computing power for AI training and inference provided by accelerated computing platforms based on GPUs, FPGAs, ASICs, and other chips; supercomputing power is mainly the computing power provided by supercomputers and other high performance computing clusters. The computing power environment mainly includes factors such as the network environment and investment in computing power. Continuous optimization of the network environment provides solid support for computing power development, while large-scale investment in computing power generates direct and indirect driving effects on the growth of computing power. Computing power application mainly includes consumer application and industrial application. Consumer and industrial application have brought about rapid increases in the demand for computing power scale, computing power capabilities, etc., and conversely, advances in computing power have driven application development.

The framework for calculating the computing power index is as shown in Figure 1.

[2] Computing power mainly considers more intensive and larger-scale forms of computing power. Personal and desktop, terminal, edge, and other forms of computing power are not included in these statistics.



Source: CAICT

Figure 1 Overall framework of computing power development

II. Global Computing Power Development Enters a New Stage

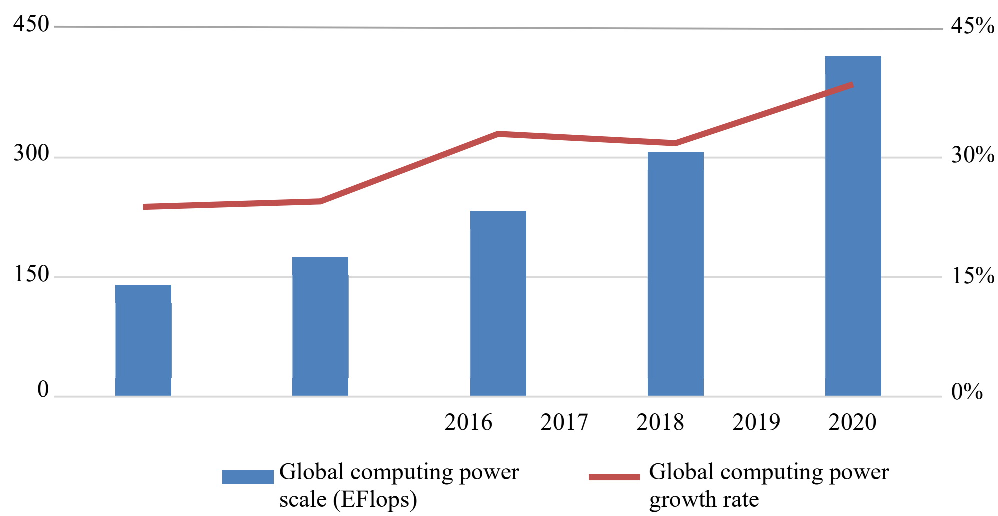
In the digital economy era, computing power is becoming a new productive force. At present, innovation and breakthroughs are accelerating in next-generation information and communication technologies such as 5G, IoT, cloud computing, big data, AI, and blockchain. The explosive growth of data and the increasing complexity of algorithms, as well as the ever-growing diversity of application scenarios, are driving the continuous upgrading of computing power demands and requirements. With the diversification of global computing power becoming increasingly pronounced and the pace of innovation accelerating further, computing power has become a new engine of the digital economy and a new focus of strategic competition.

1. The diversification of computing power is becoming increasingly pronounced

Global computing power continues to grow. In 2020, total global computing power grew at a 39% to reach 429 EFlops, of which basic computing power (in FP32[3])

[3] FP32 is single-precision floating-point format, in which numbers are expressed using 32-bit binary codes. It is commonly used in multimedia and graphics processing calculations. FP16 is half-precision floating-point format, in which numbers are expressed using 16-bit binary codes. It is suitable for deep learning applications. FP64 is double-precision floating point format, in which numbers are expressed using 64-bit binary codes. It is commonly used in scientific calculations that deal with a large range of numbers and require precise calculations.

was 313 EFlops[4], intelligent computing power (converted to FP32) was 107 EFlops[5], and supercomputing power (converted to FP32) was 9 EFlops[6]. With the dawning of the "everything sensing, everything connected, and everything smart" era, according to International Data Corporation (IDC) forecast data, the number of IoT devices in 2025 will exceed 40 billion, the quantity of data generated will approach 80 zettabytes (ZB), and terminal or edge computing ability will be relied on to process more than half of that data. It is estimated that over the next five years, the global scale of computing power will grow at a rate greater than 50% and reach an overall size of 3,300 EFlops.



Sources: CAICT, IDC, Gartner, TOP500

Figure 2 Trends in the scale of global computing power

Demand diversification is accelerating computing power diversification and upgrading. Basic computing power: Cloud computing has maintained high growth and rapid innovation trends. According to IDC data, the size of the global cloud computing market reached U.S. $312 billion in 2020, growing 24.1% year-on-year.

---

[4] The scale of basic computing power is estimated based on the total computing power of the global stock of servers. According to asset allocation standards, the service life of servers is generally six years. The total stock of servers is close to the scale of server shipments globally over the most recent six years. Global basic computing power = ∑last 6 years (annual scale of server shipments*average server computing power in the current year).

[5] The scale of intelligent computing power is estimated based on the total computing power of the global stock of AI accelerator chips. The total stock of accelerator chips is estimated using the scale of global accelerator chip shipments over the most recent six years. Global intelligent computing power = ∑ last 6 years (annual scale of accelerator chip shipments*average computing power of accelerator chips in the current year).

[6] The scale of supercomputing power is estimated mainly based on TOP500 global supercomputer data, and by referring to relevant data from supercomputer manufacturers.

Driven by the digital transformation of industries, cloud-native technologies continue to be implemented. This is driving the all-round improvement of the technology architecture, increasing application efficiency and the benefits of moving to the cloud, deepening the development of cloud-digital integration, cloud-intelligence integration, and high-performance computing, and driving the development of high-quality cloud-digital-intelligence integration. **Intelligent computing power:** Along with the evolution from active learning to transfer learning and then to reinforcement learning, the capabilities needed have also evolved constantly, from "human-computer interaction and digital integration capabilities" to "large-scale pre-training and data processing capabilities" and then to "information collection and probability calculation capabilities." There are also significant differences in the computing power requirements for model training and for inference deployments in AI. The computational accuracy in the training phase directly affects the accuracy of final models, which puts greater emphasis on 32-bit or 16-bit floating-point computing capabilities. The inference stage combines the deployment requirements of various types of devices from cloud to terminal, and can use low-precision 16-bit or 8-bit fixed-point operations to improve computing efficiency. **Supercomputing power:** The law of thousandfold increases every decade continues to hold. For the world's TOP500 supercomputers, petaflop (P)-level supercomputing has become the current threshold for entry. In June 2021, Japan's Fugaku supercomputer topped the list with a peak speed of 537 petaflops, opening a prelude to exascale (E) computing (one quintillion operations per second) entering into applications.

2. The pace of computing power innovation is accelerating further

At present, Moore's Law is showing a gradual slowing trend, and the classical model of computing system evolution dominated by advanced process upgrading is being challenged by "power consumption wall" and "memory wall" bottlenecks. In order to respond to thousandfold increases in demand for computing power from massive application innovations and major technological innovations in the age in which all things are intelligent, the important initiatives now for supporting further computing power upgrading are to deeply tap technological growth potential at the levels of computing devices, computing chips, computing systems, and computing theory, and explore collaborative innovation involving more dimensions and more elements.

From a computing chip perspective, Moore's Law still holds. Relying on fin field-effect transistor (FinFET) miniaturization and design-technology co-optimization, as well as support from new structures, new equipment, and new materials such as gate-all-around (GAA) nanosheets, high numerical aperture extreme ultraviolet photolithography machines, and ruthenium interconnects, the future path for upgrading the integrated circuit manufacturing process is gradually becoming clear. Taiwan Semiconductor Manufacturing Company Limited (TSMC)'s 3 nanometer (nm) process is expected to achieve mass production applications in the second half of 2022; exploration in other directions such as forksheet field-effect transistors (FETs) and complementary field effect transistors (CFETs) is expected to further extend advanced processes to the 2nm process mark. Another important direction for the overall performance of computing chips and systems to achieve continued Moore's Law upgrading is to realize heterogeneous integration between different processes and different types of chips with the help of chiplet, 2.5D, and 3D-based advanced packaging technologies.

From a computing architecture standpoint, multi-level heterogeneous computing is increasingly popular. Heterogeneous computing improves computing parallelism and efficiency through a mixed collaboration mode with a variety of computing units. Its presence in mobile internet, AI, cloud computing, and other types of typical applications has increased significantly, and it achieves the best balance between performance, power consumption, and cost, mainly through on-chip heterogeneity and node heterogeneity modes. System on a chip (SoC) chips are typical representatives of on-chip heterogeneity. Apple's M1 chip, for example, achieves performance superior to general purpose CPUs by integrating CPUs, GPUs, neural processing unit (NPU) and other cores. The domestically self-developed Sunway (申威) supercomputing CPU chips also introduce a master-slave multi-core architecture to improve computing flexibility. Node heterogeneity mostly uses collaboration between the CPU and various kinds of accelerators to improve overall capability. For example, heterogeneous CPU+GPU chips are mostly selected for use in AI, while cloud computing uses the CPU+GPU+data processing unit (DPU) mode more. In addition, Intel, Nvidia, Huawei, and other companies are promoting the layout of heterogeneous software ecosystems such as heterogeneous acceleration libraries, compilers, and toolchains for diverse chips. For example, the OneAPI heterogeneous framework released by Intel achieves full-stack heterogeneous capability building across xPUs, and seizes heterogeneous development opportunities.

From a computing systems perspective, construction of computing power infrastructure continues to speed up. Supercomputing: Global supercomputing is set to enter the exascale computing era, and cloud-based service models are being explored to provide flexible and elastic computing support. In the future, cloud supercomputing will serve as an important supplement to supercomputing to meet the dramatically growing demand for high-performance computing. **Intelligent computing centers:** Swiftly rising demand for AI computing has further spawned rapid development of AI-oriented intelligent computing centers. The principal nations and leading enterprises around the world are engaged in widespread construction of intelligent computing platforms to provide the computing services, data services, and algorithm services required for AI applications, and they efficiently support open data sharing, smart ecosystem construction, and industrial innovation clustering. **Data centers:** As pan-terminal computing platforms such as smartphones, intelligent vehicle, and home gateways become increasingly popular, the deepening accumulation of cloud-native technologies, and coordinated supplementation by edge computing power, as well as rapid growth in the computing power and ecosystem interconnection of different levels and among different systems, will accelerate the deep integration and coordinated development of cloud-edge-terminal computing power. This will broadly support an abundance of application scenarios such as cloud mobile phones, virtual reality (VR) media, government and enterprise mobile offices, online education, and so on.

From a computing theory perspective, cutting-edge disruptive systems have become an important direction for future exploration. First, storage-calculation integrated architecture realizes calculation in the storage unit. It thereby overcomes the limitations of separate “storage” and “calculation” in the "von Neumann" system, and is suitable for analysis and processing tasks involving vast quantities of data, such as big data and AI tasks. There are already many efforts underway within the industry to explore industrialization. **Second,** in solving special problems like Gaussian boson sampling, quantum computing achieves computing power far beyond typical computers—close to ten billion times as much computing power as in supercomputing[7]. Current research is attempting to apply it in scenarios like combinatorial optimization, chemical simulation, and drug development. **Third,** photonic computing uses the refraction and interference properties of optical devices to perform calculations, which is more advantageous for specific complex tasks such as signal processing and AI. **Fourth,** brain-inspired computing chips simulate the way neurons in the human brain work, based on spiking neural network (SNN) algorithms, and are suitable for application in low power consumption scenarios such as IoT.

3. Computing power has become an engine of the digital economy

Computing power directly drives digital industry development. The development of core digital industries, such as electronic information manufacturing, telecommunications, software and digital technology services, and the internet industry, is closely related to computing power. The internet is the industry with the

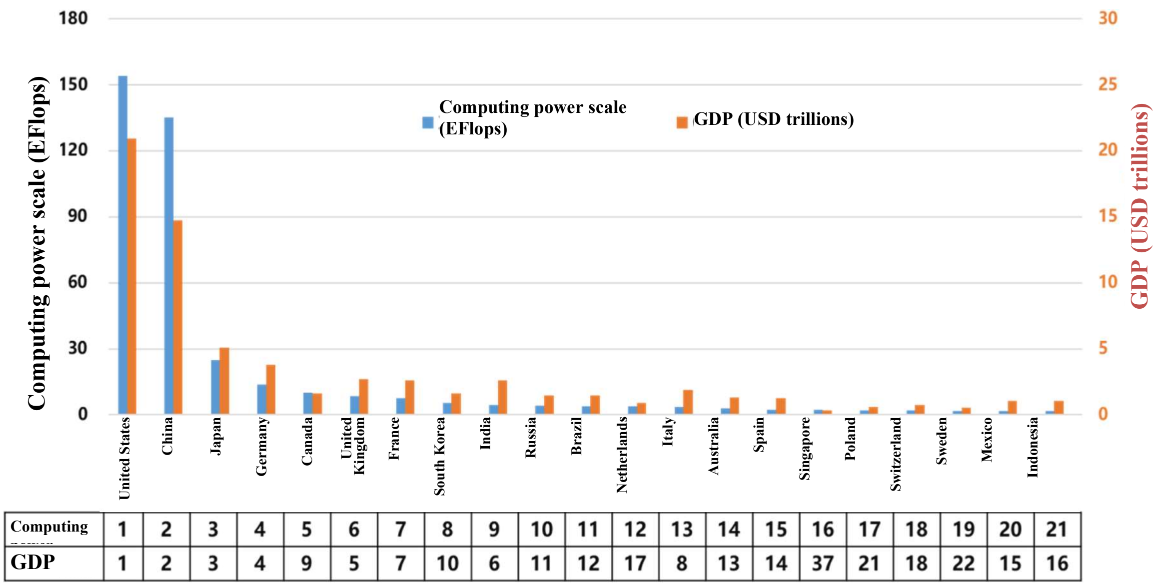
---

[7] China's "Jiuzhang" (“九章”) 76-photon quantum computing prototype took 200 seconds to solve a Gaussian boson sampling algorithm that would have taken the world's current fastest supercomputer, Japan's Fugaku machine, 600 million years to solve, having nearly ten billion times as much computing power as that supercomputer.

largest investment in computing power. Amazon, Microsoft, and Google typically invest a total of over U.S. $25 billion in capital expenditures each quarter, most of which is spent deploying hyperscale data centers, making the internet industry representative of advanced productive forces: It not only supports the accelerated penetration of internet technology into payment, e-commerce, services and content, but also gives the industry the potential to transform other production chains and value chains, further promoting its evolution from consumer internet to industrial internet.

Computing power boosts the digital transformation and upgrading of industries. The digital intelligence technology brought about by computing power investment not only produces growth in industrial output value for many industries such as manufacturing, transportation, and retail, but it also brings extended benefits, including production efficiency improvement, business model innovation, and user experience optimization. Its pulling effect on economic growth is becoming more and more pronounced. Take the manufacturing sector for example: Investment in computing power, as represented by cloud computing, edge computing, and intelligent computing, and their large-scale application, can significantly improve production efficiency, and empower and reconfigure the manufacturing industry, from demand perception, research and development (R&D), procurement, production, and marketing to after-sales and other links in the production chain, creating a highly collaborative intelligent manufacturing ecosystem.

The development of computing power strongly drives the digital economy forward and powers national economic development. In the industrial age, electric power was an important indicator for assessing gross domestic product (GDP) growth. In the digital economy era, computing power is a key factor in information and communications technology (ICT) industry development. It plays an important role in driving S&T progress, promoting the digital transformation of industries, and supporting economic and social development, and it has become a new core indicator. Computing power has a significant role in leading the development of the digital economy and GDP. Between 2016 and 2020, the scale of global computing power grew 30% per year on average, while the digital economy and GDP grew by 5% and 2.4%, respectively. The scale of computing power in countries around the world is closely related to their level of economic development: The higher the level of economic development, the larger the scale of computing power. Of the top 20 countries ranked by scale of computing power, 17 are ranked among the top 20 economies, and the top four are the same in both rankings.



Computing power ranking	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
GDP ranking	1	2	3	4	9	5	7	10	6	11	12	17	8	13	14	37	21	18	22	15	16

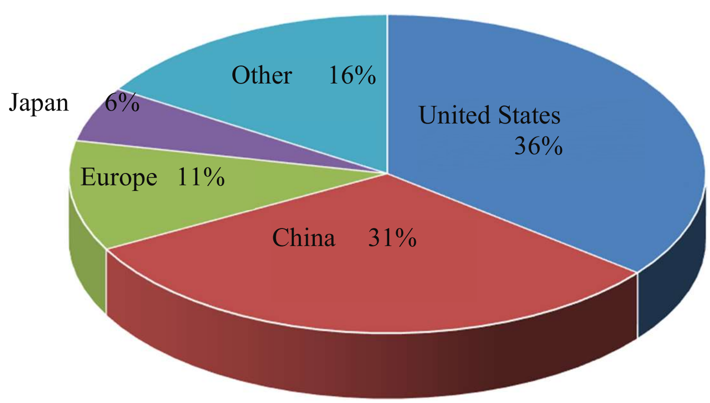
Sources: CAICT, IDC, Gartner, World Bank

Figure 3 Relationship between the scale of global computing power and GDP

4. Computing power has become a new focus of strategic competition

Major countries and regions around the world have accelerated deployment of computing power. Computing power has now become a reflection of countries' core competitiveness, and countries around the world are stepping up their strategic deployment efforts. In November 2020, the United States released *Pioneering the Future Advanced Computing Ecosystem: A Strategic Plan*, which would make advanced computing ecosystems national strategic assets to ensure the United States' leadership in science and engineering, its economic competitiveness, and national security. In 2019, Japan launched plans for a new generation of domestically produced supercomputers, with an investment of around 130 billion yen [U.S. $1.15 billion] to build the world's fastest supercomputer. The EU is deploying supercomputing and quantum computing with a focus on digital sovereignty. In 2018, it proposed the "European High-Performance Computing Joint Undertaking," and in September 2020 it proposed investing eight billion euros [$9.3 billion] to support the research and innovation of a new generation of supercomputing technologies and systems, mainly exascale computing and quantum computing, and to maintain and enhance Europe's leading levels in the supercomputing and quantum computing fields.

Global computing power competition is intensifying. In terms of computing power levels, the shares of the United States, China, Europe, and Japan in global computing power are 36%, 31%, 11% and 6%, respectively. As for basic computing power, the United States and China are the first echelon in global competition, and the gap between China and the United States is shrinking continuously. The United States ranks first in share of global basic computing power, with 43%, while China ranks second with a 26% share; in intelligent computing power, China and the United States are in the lead, with 52% and 19%, shares of global computing power, respectively; in supercomputing, the United States, Japan, and China have obvious advantages when it comes to indicators of overall performance, with 31%, 23% and 20% shares of total computing power, respectively, and China is in the lead in terms of absolute quantity. **With respect to globalization**, the United States, Japan, and Europe got off to an early start on informatization (信息化), the development histories of the relevant enterprises are long, and they were also earlier than China at developing international markets. Thus, their shares of the international market are significantly higher than China's, and the competitiveness of Chinese computing power in international markets is significantly weaker than it is in the domestic market. For example, although Alibaba Cloud's share of the global cloud market is ranked fourth, the vast majority of its business is concentrated in China's domestic market, and there is still a large gap with Amazon, Microsoft, Google, and other U.S. companies. At the same time, as a result of the intensification of the game being played between the great powers (大国博弈), the globalized production chain and supply chain pattern is facing reshaping and reconstruction, which will bring new challenges for computing power technology innovation and industrial ecosystems. China's computing power-related equipment manufacturing ability continues to improve, but still faces a serious "chokepoint" (“卡脖子”) problem if it is to completely catch up with the United States and other leading countries.



Sources: CAICT, IDC, Gartner, TOP500

Figure 4 Distribution of global computing power

III. Overall Trends in China's Computing Power

In 2020, the global spread of the COVID-19 epidemic took a large toll on China's economy, exacerbating cyclical economic fluctuations and bringing heavy downward pressure on economic growth. China's computing power has maintained strong growth in this complex and harsh development environment. The structure of computing power has continued to evolve, the innovation level, development environment, and application demand have continued to improve, and computing power has become a solid foundation supporting the high-quality development of China's economy.

1. China's computing power has expanded continuously

China's computing power has continued to expand in scale. During the 13th Five-Year Plan period [2016-2020], the scale of China's computing power overall maintained a steep growth trend, with rapid growth maintained in the numbers of data center racks, general purpose servers, AI servers, and supercomputers. Under accelerating technological innovation and iteration, the performance of individual computing power devices also rose consistently. China's computing power overall reached a scale of 135 EFlops in 2020, accounting for 31% of the global total, and maintained a high growth rate of 55%, some 16 percentage points higher than the global growth rate. China's basic computing power (in FP32) reached 77 EFlops[8], which was 26% of the global total, while intelligent computing power (converted to FP32) reached 56 EFlops[9], 52% of the global total, and supercomputing power (converted to FP32) was 2 EFlops[10], representing 20% of the global total.

The structure of computing power continues to evolve. Much like the global development situation, China's computing power structure has evolved constantly in recent years, along with changes in application requirements. In particular, the basic computing power proportion of overall computing power decreased from 95% in 2016 to 57% in 2020, while the intelligent computing power proportion increased from 3% to 41%. As demand for internet and cloud computing has risen, AI computing

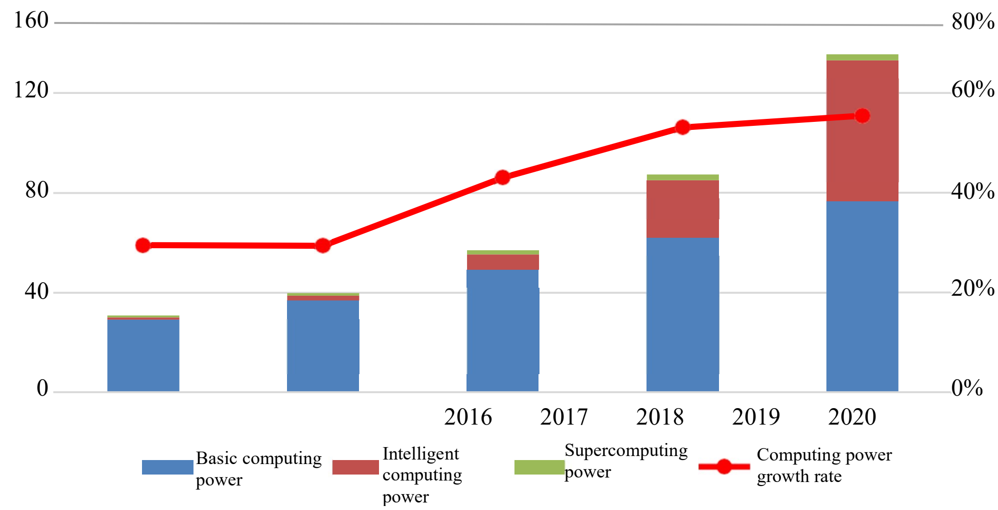
---

[8] The scale of basic computing power is estimated based on the total computing power of China's stock of servers. According to asset allocation standards, the service life of servers is generally six years. The total stock of servers is close to the scale of server shipments in China over the most recent six years. China's basic computing power = ∑ last 6 years (annual scale of server shipments*average server computing power in the current year).

[9] The scale of intelligent computing power is estimated based on the total computing power of China's stock of AI accelerator chips. The total stock of accelerator chips is estimated using the scale of China's accelerator chip shipments over the most recent six years. China's intelligent computing power = ∑ last 6 years (annual scale of accelerator chip shipments*average computing power of accelerator chips in the current year).

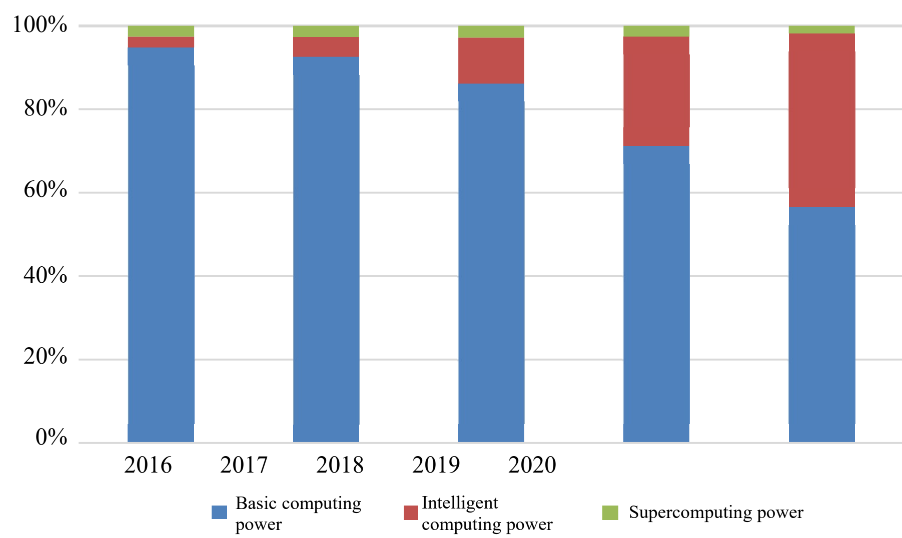
[10] The scale of supercomputing power is estimated mainly based on TOP500 global supercomputer and China HPC Top100 data, and by referring to relevant data from supercomputer manufacturers.

infrastructure represented by intelligent computing centers has developed rapidly, the pace of intelligent upgrading of data centers has accelerated, and intelligentized (智能化) upgrading of computing has become a general trend. The share of supercomputing power in China's overall computing power has been relatively stable, at about 2%, mainly for fields such as scientific computing and engineering computing.



Sources: CAICT, IDC, Gartner, TOP500, High-Performance Computing (HPC) TOP100

Figure 5 Scale and growth rate of China's computing power



Source: CAICT

Figure 6 Internal structure of China's computing power

2. The level of computing power innovation has risen greatly

1. Basic general purpose computing power has developed steadily

The support capability of basic computing power has strengthened greatly. With the informatization of the economy and society, computing power infrastructure, as represented by data centers, has seen tremendous development. Together with the communication network infrastructure, it constitutes a new generation of information infrastructure, contributing an important impetus to the building of China into a cyber powerhouse[11] and to building a Digital China, and providing important support for many fields such as industrial production and social livelihood. China's basic computing power has grown rapidly in the last five years, growing from 23 EFlops in 2015 to 77 EFlops in 2020, for an average annual growth rate of over 27%. In 2020, China's basic computing power played an important supporting role in epidemic prevention and control. It ensured the steady development of epidemic tracing, virus research, disease diagnosis and treatment, vaccine R&D, drug screening, etc., through the development of digital platforms and applications, launching of basic computing power facilities and algorithms, and other ways. Many new applications such as big data trip cards were derived as well, helping China's full recovery of all social production (社会生产) work. As of August 2021, service had been provided on over 12 billion "Telecom Big Data Itinerary Card"[12] queries, with a steady daily volume of more than 25 million queries and a basic computing power demand of more than 50 PFlops, becoming an important means of epidemic prevention and control in different places.

The ecosystem for basic computing power is improving. In terms of whole machines, against the backdrop of a steadily growing global market for servers, domestically produced servers have swiftly become the main force in China's basic computing power. Domestic server makers Inspur, Huawei, H3C, and Lenovo all occupy top-five spots in China's server market, and their overall market share has reached 74%. With regard to CPUs, X86 architecture CPUs dominated by Intel and Advanced Micro Devices (AMD) continue to dominate the server market, having a share of over 96%; CPU chips from China's Kunpeng, Phytium, and Hygon are now achieving large-scale applications, their overall performance level has reached the mid-to high-end, and the industry ecosystem is gradually improving. With respect to operating systems, Linux, characterized by being open source, highly stable, secure, and free (自由), dominates the server market, with a 70% market share. China's Unity and Kylin operating systems, developed based on Linux, have used the Party, government, and military sectors as points of entry, and are gradually penetrating the

[11] Translator's note: Alternate English translations for the Chinese term wǎngluò qiángguó (网络强国)—here translated as "cyber powerhouse"—include "cyber superpower," "network powerhouse," "network superpower," and so on. For a more thorough discussion in English of the meaning of the term wǎngluò qiángguó, see: https://www.newamerica.org/cybersecurity-initiative/digichina/blog/lexicon-wangluo-qiangguo/

[12] Translator's note: The "Telecom Big Data Itinerary Card" or "Itinerary Card" (通信大数据行程卡; 通信行程卡; 行程卡) is a PRC smartphone app that digitally verifies users' travel histories for COVID-19 monitoring purposes.

financial, telecommunications, healthcare, and other industries.

2. Intelligent computing has emerged rapidly

Intelligent computing centers are an important direction for AI computing development. In the past decade, computing resources for AI training models have proliferated, the computational complexity of AI training has surged tenfold each year, and AI computing has become a mainstream form of computing. In 2020, intelligent computing power accounted for 41% of the total computing power in China, and intelligent computing power's share is expected to increase to 70% by 2023. The computing requirements of complex AI models and scenarios generate further demand for AI-oriented computing power infrastructure, of which intelligent computing centers are the current focus of development. "Intelligent computing centers" mainly refers to new computing power infrastructure that is based on the latest AI theories, uses advanced AI chips, and provides general purpose computing power, calculation, and algorithm services for diverse AI scenarios involving massive data processing, intelligent algorithm acceleration, high-speed training and inference, and other computing needs. Such centers are receiving wide attention from the government, enterprises, and research institutes. In April 2020, the National Development and Reform Commission (NDRC) included intelligent computing centers under the computing power infrastructure category for the first time and proposed to lay out about 10 regional-level data center clusters and intelligent computing centers nationwide.

Intelligent computing shows a multi-path development trend. First, GPUs occupy a dominant position in intelligent computing centers. In 2020, Nvidia GPU chips accounted for about a 95% share of the domestic AI server market. A large number of domestic startups have also emerged—Iluvatar CoreX, MetaX, Biren Technology, Moore Threads, etc. For example, in March 2021, Iluvatar CoreX released the BI cloud training general purpose GPU chip. Manufactured with a 7nm process, it delivers 147 trillion calculations per second (FP16), able to complete the AI processing of hundreds of channels from video cameras. **Second,** ASICs and other non-GPU accelerator chips will be adopted more and more in a variety of industries and fields to satisfy the performance, power consumption, and cost requirements of differentiated scenarios. Enflame Technology launched the second generation Yunsui T20 AI training chip in July 2021, with a single-precision tensor computing power of 160 TFlops; and Cambricon has formed a complete product line layout with the SiYuan 290 cloud training chip, SiYuan 270 inference chip, and SiYuan 220 edge computing chip.

Many locations in China have accelerated their investment in intelligent computing center layouts. By joining forces with Huawei, SenseTime, Inspur, and other leading AI enterprises, industry clusters such as Shanghai, Guangzhou, Jinan, Wuhan, and Shenzhen have adopted a government-enterprise collaborative construction and operation approach to promote implementation. Huawei has cooperated with Shenzhen, Wuhan, and other local governments to promote implementation of intelligent computing centers. The Peng Cheng Cloud Brain II project involved cooperation with Shenzhen Peng Cheng Laboratory and launched successfully in October 2020. It can provide one-EFlops-level AI computing power and is currently ranked first in the world's top 500 AI computing power rankings; and the Wuhan Artificial Intelligence Computing Center (武汉人工智能计算中心) project, which has a planned computing power at the 100-P level, is expected to have the conditions in place to go live in 2021. SenseTime's new generation AI computing and empowerment platform project is also expected to be up and running in 2021. The platform can meet the needs of four megacities simultaneously, and provides access capability to 8.5 million video channels.

3. High-performance computing development has accelerated

China has made great strides in the supercomputer field. In terms of computing power services, the National Supercomputing Service Grid (国家超级计算服务网格) has been connected to eight national supercomputing centers in Tianjin, Guangzhou, Shenzhen, Changsha, Jinan, Wuxi, Zhengzhou, and Kunshan. A host of localities, industries, and universities are also promoting the construction of high-level supercomputing centers, which currently serve more than 20,000 users and support the research of over 2,000 national scientific computing projects and important engineering projects of various kinds. In terms of computing power capabilities, China ranked first in the world in number of supercomputers in the global TOP500 supercomputing rankings as of June 2021, accounting for nearly 40%. It was ranked third in supercomputing scale after the United States and Japan, and it held two of the top ten spots on the list: "Sunway TaihuLight" and "Tianhe-2A" were ranked fourth and seventh, respectively. In terms of market supply, domestic supercomputer manufacturers are far ahead by market share. Lenovo, Inspur, and Sugon ranked first, second, and fourth in the world on the global TOP500 supercomputer list, having delivered 184, 57, and 39 units, respectively, and had a cumulative market share of 56%.

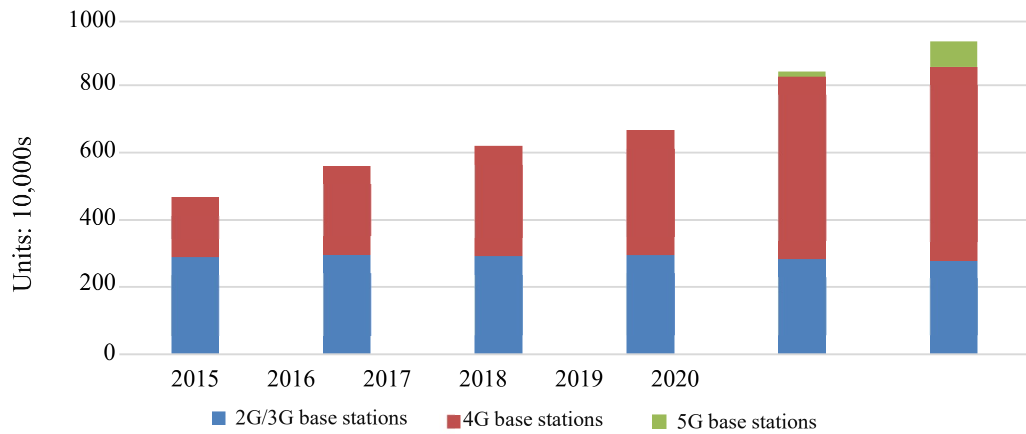
Domestic E-level supercomputing system development work has made steady progress. At present, the development of three exascale supercomputing systems—Sunway (神威), Tianhe, and Sugon, led respectively by Jiangnan Institute of Computing Technology (江南计算技术研究所), National University of Defense Technology, and Sugon—is advancing side-by-side. In terms of the configurations of relevant exascale supercomputer prototypes, the Sunway system uses domestic Sunway multi-core processors, the Mellanox unlimited bandwidth interconnect architecture and self-developed network chipsets, and the heat dissipation system is liquid cooling mode. The Tianhe system integrates the Phytium ARM architecture FT-2000 chip and a 128-core digital signal processor (DSP) chip, and uses a 400GB/s photoelectric integrated high-speed interconnect architecture and hybrid liquid-air-cooled targeted heat dissipation technology. The Sugon system adopts a heterogeneous computing architecture with the Hygon X86 processor and self-developed DCU[13] accelerator. It is equipped with the 6D-Torus high-dimensional hierarchical network system with a bandwidth of 200GB per second, and uses two-phase immersion cooling (全浸式相变冷却) technology to achieve extremely low power usage effectiveness (PUE).

3. The environment for computing power development is improving

A continuously improving network environment provides solid support for computing power development. The level of computing power depends on the computing power of servers and terminals, but it is also affected by network transmission capabilities. At the same time, the ubiquitous distribution of data is pushing the gradual spread of computing from the cloud to IoT and edge computing, with synergies between different fields of computing. As the telecommunications network increasingly moves to the cloud, computing-network fusion (算网融合) will become an important trend in computing power development. With 5G bringing a pipeline with low latency, high bandwidth, and high connection density, the boundaries separating data centers from the edge and terminals are breaking down, and synergies between computing and networks can promote improvement in computing power levels. China's network infrastructure capabilities continue to be upgraded, nationwide promotion of the internet backbone network and the urban area network structure continues, and with stepped up efforts to expand inter-provincial export bandwidth, the average inter-provincial internet export bandwidth of provinces reached 22 Tbps in 2020, an increase of 11%. With regard to mobile communications, the total number of mobile base stations nationwide reached 9.31 million in 2020. That included a total of 5.75 million 4G base stations, with urban areas achieving deep coverage. In accordance with the principle of taking the lead at a reasonable pace (适度超前), 5G network construction is steadily advancing. With over 770,000 5G base stations, prefecture-level cities and key counties and cities nationwide are covered, and the 5G coverage rate is 13%, an increase of 10 percentage points. China's IoT development has obtained positive results. The number of narrowband internet of things (NB-IoT) base stations has surpassed 700,000, cities at the prefecture level and above are covered nationwide, the user base is the world's largest, and the number of mobile IoT

[13] Translator's note: "DCU" is a term PRC CPU manufacturer Hygon uses to refer to one of its accelerator microchips. It is unclear what "DCU" stands for.

connections has exceeded 1.08 billion.

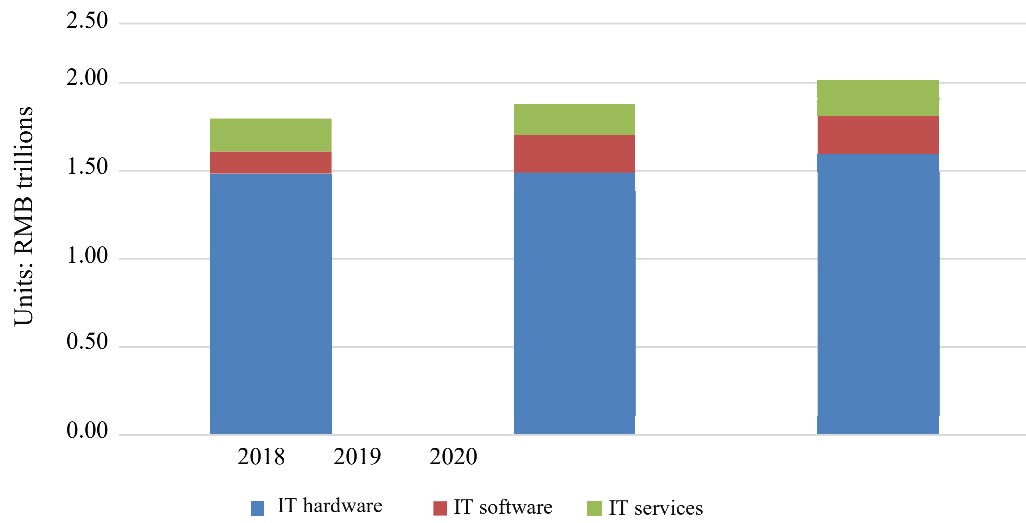


Source: Ministry of Industry and Information Technology (MIIT)

Figure 7 Development of mobile communication base stations in China

Constantly growing investment in computing power has an important pulling effect on computing power development. China's constantly growing spending in the fields of computing hardware, software, and services has further boosted the computing power promotion effect of emerging applications such as AI, autonomous driving, and smart city development. IDC data show that China's IT spending rose 7.3% to reach RMB 2 trillion in 2020, driving the strong recovery of China's economy. China's IT spending is expected to reach RMB 2.21 trillion in 2021, which would be an increase of 10.0% compared to 2020. At the same time, there has been a gradual diversification of entities investing in China's computing power. The development of the digital economy has caused a shift in investment entities, from single-entity investment by governments or enterprises to multi-party development involving cooperation between governments and social capital.[14] There are more and more such multi-party investment models featuring government-social capital cooperation. The public-private partnership (PPP) model in particular is widely used, and more flexible investment models further promote the sustained growth of China's computing power investment.

[14] Translator's note: The Chinese term 社会资本, translated literally as "social capital," and its synonyms "social funding" (社会资金), "social investment" (社会投资), and "social financing" (社会融资), refer to any source of funding outside of government budget outlays. These terms encompass investment by private individuals and private institutions. However, investment from state-funded entities such as state-owned enterprises (SOEs), including state-run banks, also falls under the umbrella of "social capital."



Sources: CAICT, IDC

Figure 8 IT hardware, software, and services spending in China

4. Computing power application demand continues to grow

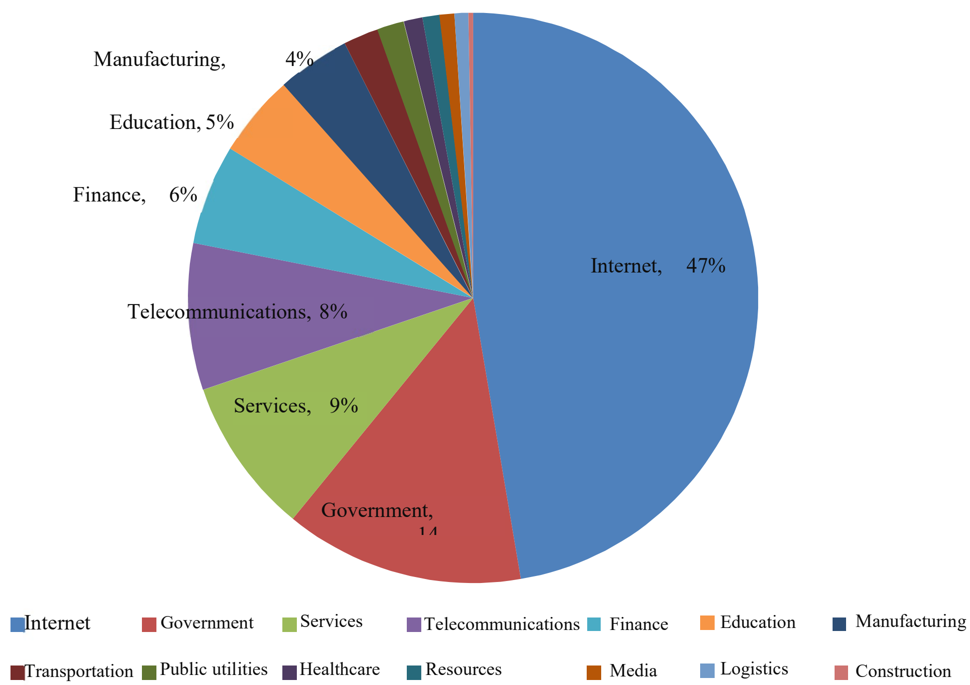
At the same time that China's computing power has seen innovation and development, the potential demand for computing power has continued to grow. With the rapid development of new technologies and applications such as 5G, the industrial internet, AI, cloud computing, and big data, computing power serves as an indispensable digital foundation for economic and social operations, effectively supporting the vigorous development of China's digital economy and promoting the digital transformation of different industries and fields.

1. Computing power enables digital transformation in a myriad of industries

Computing power development is driving the deepening digital transformation of China's industries. Supported by computing power infrastructure, a succession of new digitalization models, such as e-commerce, the platform economy, and the sharing economy, have emerged in China, while the industrial internet, smart manufacturing, and so on have accelerated across the board, generating strong momentum for the sustainable and healthy development of China's industrial digitalization. The scale of industrial digitalization in China reached RMB 31.7 trillion in 2020, accounting for 31.2% of GDP, a nominal increase of 10.3% year-on-year, and its proportion of the digital economy increased from 74.3% in 2015 to 80.9% in 2020. The digital economy portions of services, industry, and agriculture accounted for 40.7%, 21.0% and 8.9%, respectively, of the value added of those industries.

Computing power provides strong support for the digital transformation of industries. With regard to industries, the internet industry is still the largest in terms of computing power demand, accounting for nearly 50% of computing power overall.

Internet giants represented by Alibaba, Tencent, Baidu, and ByteDance have more urgent demand for computing power, while the centralized deployment of computing power also makes the internet industry representative of advanced productive forces. The government, services, telecommunications, finance, education, manufacturing, and transportation industries are ranked from second to eighth, respectively. Informatization and digitalization started earlier in the telecommunications and finance industries, which are traditionally the industries in China with relatively large-scale application of computing power, and their application of computing power is at leading levels in their industries. The digital transformation of the manufacturing industry is still in the early stages, and needs more support in the form of large-scale, general purpose public computing power infrastructure.

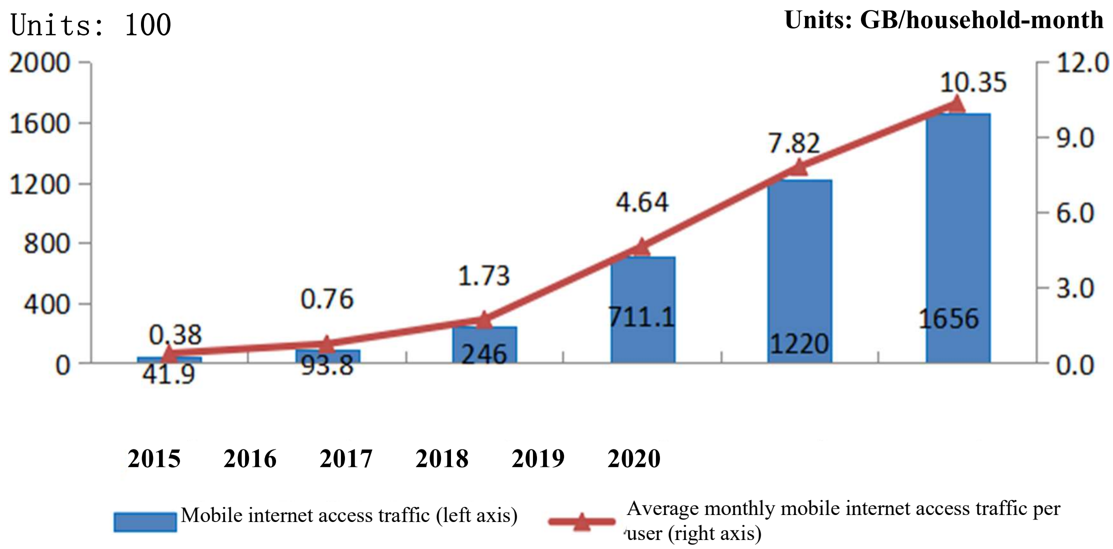


Sources: CAICT, IDC

Figure 9 Distribution of computing power application among industries in China

2. Information consumption continues to generate computing power demand

Expansion of intelligent terminal consumption and mobile data traffic consumption has continued. Constantly rising computing power has driven the spread of emerging applications, including high-definition (HD) content, video streaming, and augmented reality (AR) navigation, which in turn has further promoted smart terminal consumption growth. The mobile phone penetration rate in China has increased steadily, with the 4G user penetration rate exceeding 80% and the number and proportion of 4G users continuing to expand. By the end of 2020, the number of 5G network users in China exceeded 160 million, accounting for approximately 89% of total 5G users worldwide. The number and proportion of 5G cell phone shipments in China continue to rise, with domestic 5G cell phone shipments totaling 163 million units in 2020, representing 52.9% of cell phone shipments in the same period. Meanwhile, the scale of mobile data traffic consumption has continued to expand. In 2020, the average monthly mobile internet access traffic per user (DOU) crossed into the 10GB range, with access traffic consumption reaching 165.6 billion GB, up 35.7% over the previous year. Cell phone internet traffic reached 156.8 billion GB, up 29.6% over the previous year, and accounted for 94.7% of the total traffic.



Source: MIIT

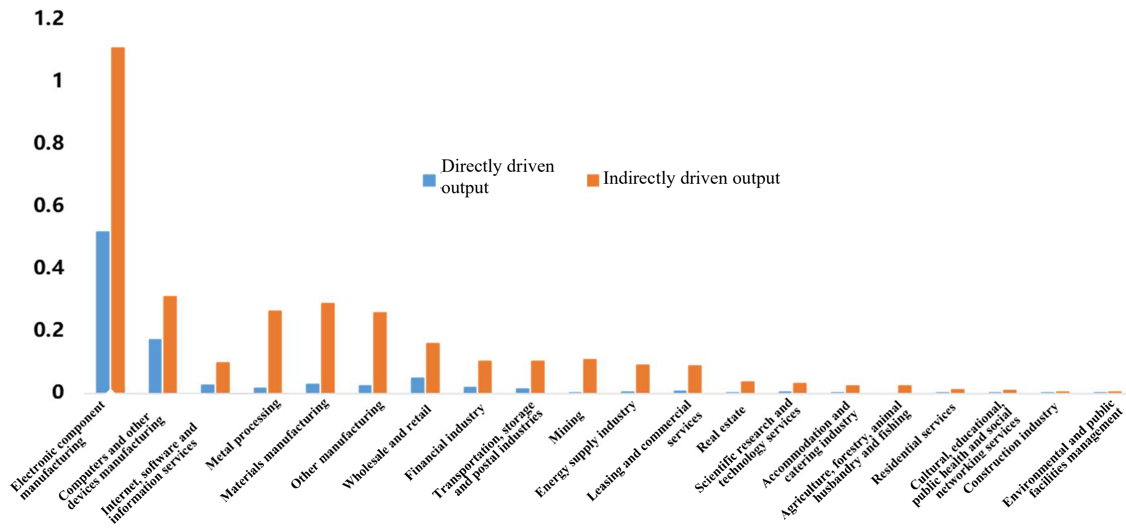
Figure 10 Growth in mobile internet traffic and monthly DOU, 2015-2020

The growing ubiquity of computing power has promoted a diversification trend among intelligent terminals. With the large-scale construction of the 5G network and edge computing, emerging applications will accelerate and drive the diffusion of data processing from the cloud to the edge and terminals. Computing power at the edge continues to grow, and a trend toward ubiquitous computing power has appeared, driving huge demand for various computing devices. In the future, as edge-terminal devices grow more abundant, personal PCs and even home gateways may come to serve as computing power nodes; and the popularity of cell phones, intelligent vehicles, and other intelligent terminals, meanwhile, has given rise to near-data processing and ubiquitous computing processing scenarios, which will also promote the collaboration and linkage of computing power across different distances and scales in the information space around the user, resulting in a three-level "cloud-edge-terminal" computing architecture. At present, "pan-terminal 5G" (泛终端 5G) has reached 20 terminal types, covering VR/AR headsets, customer premises equipment (CPE), industrial-grade routers/gateways, drones, robots, vehicle on-board units (OBUs), and many other types. Non-cost-sensitive fields such as industry and healthcare will take the lead in the popularization and iterative evolution of 5G, which will produce disruptive change in culture, education, leisure and entertainment, etc.

5. Computing power development boosts economic growth

Computing power has become a strong foundation supporting China's high-quality economic development. With the continued progress of the new round of S&T revolution and industrial transformation, coupled with the impact of COVID-related factors, computing power has become a key driver promoting sustained and stable growth of the digital economy and the national economy, and it has played an important role in seizing the double victory in epidemic prevention and control and in economic and social development. Computing power has become the most dynamic, innovative, and widely distributed information infrastructure in China at present, and it has become a key indicator for gauging the vitality of the digital economy. During the 2016-2020 period, the scale of China's computing power grew an average of 42% per year, while the digital economy grew by 16% and GDP grew by 8%. Compared to the rest of the world, the pulling effect of computing power on the growth of the digital economy and GDP in China is pronounced.

The computing power industry's role in driving the energy levels of China's economy, society, and industries is growing stronger by the day. Calculated using a national input-output model, the computing power industry, as represented by computers, reached RMB 2 trillion in 2020, directly driving RMB 1.7 trillion in economic output, and indirectly driving RMB 6.3 trillion in economic output. That is, each 1 RMB invested in computing power drives 3-4 RMB in economic output. For industries like electronic components, computers, materials, software, and IT services in particular, the computing power industry has a relatively large direct pulling effect, directly driving up to RMB 1.5 trillion in economic output; among different industries, computing power investment has relatively significant direct driving effects on economic output in fields such as manufacturing, the internet, and finance.



Sources: National Bureau of Statistics, CAICT

Figure 11 Driving effect of investment in computing power on overall economic output in China

IV. China Computing Power Development Index Assessment

China attaches great importance to computing power development. In May 2021, NDRC, the Office of the Central Cyberspace Affairs Commission,[15] MIIT, and the National Energy Commission jointly issued the *National Integrated Big Data Center Collaborative Innovation System Computing Power Hub Implementation Plan*, which specifically proposed laying out a nationwide computing power network of national hubs and nodes, and launched implementation of the "eastern data, western compute" project, in order to build a national computing power network system. In July 2021, MIIT issued the *Three-Year Action Plan for the Development of New Data Centers (2021-2023)*, which specifies that in three years, a new data center development pattern will have basically taken shape, one that is rationally laid out, technologically advanced, green and low-carbon, and in which the scale of computing power is commensurate with the growth of digital economy.

Driven by both demand and policies, all parts of the country are vigorously promoting the development of the computing power technology industry, infrastructure construction and computing power application. In order to comprehensively sort through and objectively evaluate the computing power development situation in China, and arrive at a more scientific and concrete understanding of China's computing power, CAICT has established the China Computing Power Development Index, taking into consideration the characteristics and

[15] Translator's note: The Office of the Central Cyberspace Affairs Commission (中央网络安全和信息化委员会办公室; 中央网信办) is effectively the same organization as the Cyberspace Administration of China (CAC; 国家互联网信息办公室; 国家网信办), as they share the same personnel and the same offices.

key influencing factors of computing power development. It comprehensively and objectively evaluates the level of computing power development in each province of China, thereby providing effective support for the formulation of computing power development policies for the whole country and each province.

1. Basis for index establishment

Based on the above-mentioned analysis of the definition of computing power and its development characteristics, and a comprehensive referencing of research on computing power measurement and related index systems by IDC,[16] Roland Berger,[17] Huawei, Inspur, and other domestic and foreign institutions and enterprises, and on the basis of fully seeking expert opinions, this white paper selects relevant indicators from three dimensions—computing power scale, computing power environment, and computing power application—to establish the China Computing Power Development Index, in order to comprehensively and objectively evaluate the development of computing power in China, and analyze the current level of computing power development in each province. The indicators for the China Computing Power Development Index were selected according to the principles of scientific soundness, representativeness, and independence, taking into account the characteristics and key influencing factors of computing power development, and incorporating consideration of the accessibility and comparability of data at the provincial level.

Dimension one: Computing power scale. Computing power scale is measured mainly based on the three aspects of basic computing power, intelligent computing power, and supercomputing power. Basic computing power is based mainly on the computing power level of CPUs in the stock of servers. Single-precision floating point format (FP32) is used to measure the computing power performance of each server[18]. Intelligent computing power is mainly the level of heterogeneous computing power supported by GPUs, FPGAs, ASICs, or other accelerators with highly parallel and high-density computing capability. The half-precision floating point (FP16) computing power currently used in deep learning has become the mainstream of intelligent

[16] The *2020 Global Computing Power Index Evaluation Report* jointly released by IDC and Inspur carries out evaluation mainly with four types of indicators focusing on computing power, computing efficiency, application level, and infrastructure support.

[17] *Ubiquitous Computing Power: The Cornerstone of an Intelligent Society*, jointly published by Roland Berger and Huawei, provides a system of indicators for global computing power measurement, and carries out estimation of the overall computing power of countries around the world based mainly on cloud-edge-terminal.

[18] The general purpose computing power of servers is evaluated using single-precision floating-point (FP32) computing power, where server computing power = number of processor chips*number of single-precision floating-point operations performed per clock cycle*processor clock rate*number of processor cores.

computing. Supercomputing power is estimated mainly based on global and Chinese supercomputing power data in the internationally well-known TOP500 list and China High-Performance Computing (HPC) Top100 list, and by referring to relevant data from supercomputer manufacturers. Supercomputing power is estimated mainly using the LINPACK benchmark test run on supercomputing systems. The test uses double-precision floating-point (FP64) computing power to measure the computing power performance of supercomputers. Computing power scale calculations are uniformly converted to computing power in single-precision floating-point (FP32) format for statistical purposes.

Dimension two: Computing power environment. Computing power environment measurement is based mainly on investment in the network environment and computing power. Continuous optimization of the network environment provides solid support for computing power development, while large-scale IT investment generates direct and indirect driving effects on the growth of computing power. The network environment focuses on inter-provincial internet export bandwidth and 5G mobile base station coverage. It mainly reflects the service capability of each region regarding data transmission to and from other regions in China for business such as data and internet, as well as the edge computing support capability of the region. The synergy between computing and networks can promote improvement in computing power levels, and the network transmission service capacity and edge computing support capacity of each region determine the overall level of computing power dispatching, which promotes computing power supply-demand balance. AI, autonomous driving, smart cities, and other emerging applications drive the development of computing power, spurring the growth of computing hardware, software, and service expenditures.

Dimension three: Computing power application. Measurement of computing power application is based mainly on the level of consumer application and industrial application. Computing power drives the development of consumer applications and industrial applications, while consumer and industrial applications drive the growth of computing power. The consumer application level mainly focuses on average monthly mobile internet traffic. There is a significant correlation between mobile internet access traffic and the scale of computer power[19]. Computing power is the carrier of massive mobile internet data, and the growth of data traffic is the core driver of the rapid growth in computing power scale. Mobile internet applications such as smartphones, telecommuting, online meetings, and mobile payments have driven the growth of back-end computing power infrastructure, greatly promoting the booming growth of computing power. The industry application level mainly focuses on industrial

[19] A data center industry research report indicates that mobile internet access traffic and data center market size are significantly correlated, with a correlation coefficient of 0.92.

digitalization. It reflects the application of computing power in internet, manufacturing, finance, and other fields. The industrial internet, fusion between informatization and industrialization (两化融合), intelligent manufacturing, internet of vehicles (IoV), the platform economy, and other convergent emerging new business formats (新业态) and new models provide broad space for computing power development.

2. Indicator system establishment

In the evaluation process, the determination of weights, assignment of values, and calculation of scoring stages were performed for each indicator in accordance with scientific research and analysis methods to obtain a composite index of China's computing power development. The process of forming the composite index can be divided into four stages:

1. Index system formation: On the above basis for establishing the index system, experts' opinions were sought, the current situation in terms of evaluating China's computing power development was sorted through, and the characteristics and key influencing factors of computing power development were factored in to build the computing development index system from the three dimensions of computing power scale, computing power environment, and computing power application. The index involves three primary indicators, including computing power scale and computing power environment, seven secondary indicators, such as basic computing power and intelligent computing power, and eight tertiary indicators, such as server computing power scale and AI accelerator computing power scale.

Table 1 China computing power development indicator system

Level 1 metrics	Level 2 metrics	Level 3 metrics	Organization
Computing power scale	Basic computing power	Server computing power scale	EFlops
	Intelligent computing power	AI accelerator computing power scale	EFlops
	Supercomputing power	Supercomputing power scale	EFlops
Computing power environment	Network environment	Inter-provincial internet export bandwidth	Tbps
		5G coverage rate	%

	Investment in computing power	IT expenditures	RMB 100 million
Computing power application	Consumer application level	Average monthly mobile internet traffic	EB
	Industry application level	Scale of industry digitalization	RMB 100 million

Source: CAICT

2. Determining indicator weights by the analytic hierarchy process (AHP) method: Based on expert scoring methods, the AHP was used to obtain the relative weights among each primary, secondary, and tertiary indicator in the evaluation index system.
3. Assigning values to the indicators according to the actual situations of regions: The actual values of each indicator were obtained according to the actual computing power development situations in 31 provinces[20], and the data were standardized to obtain the assigned values of each indicator. The total score of each indicator is 100, and the score of each province in the indicator is obtained by comparing the value of the current year with the target value after 5 years. Calculation of the target value after 5 years is based on data from authoritative organizations, institutions, and enterprises in various fields of industry, and is set by expert investigation and analysis.
4. Calculating the composite index: Lastly, based on the specific assigned values of indicators and corresponding weights, the composite index of each region is finally formed.

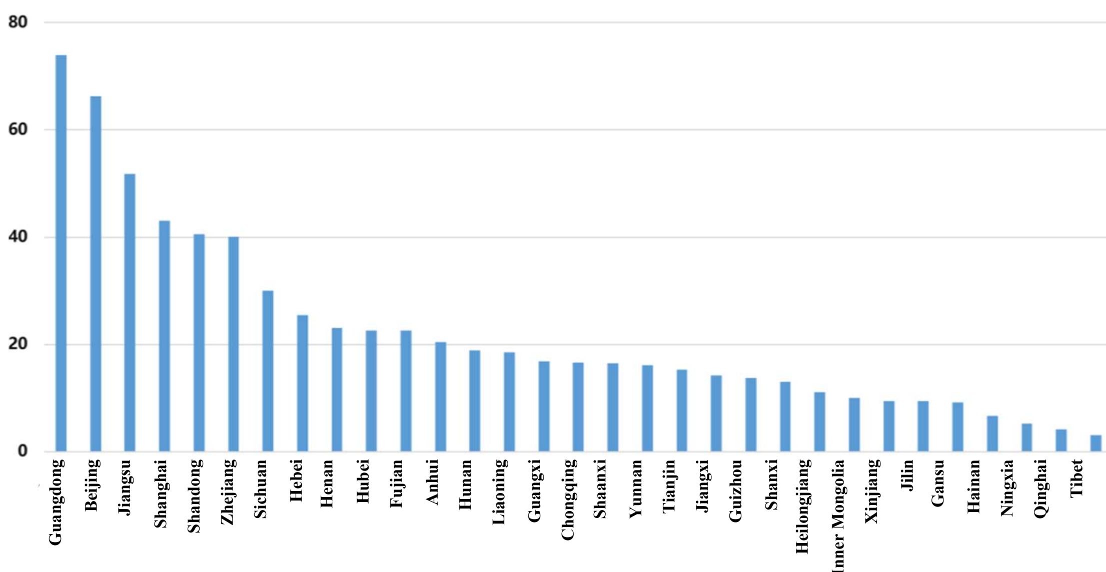
3. Evaluation of China's computing power

1. Computing Power Development Index

Beijing-Tianjin-Hebei, the Yangtze River Delta, and the Guangdong-Hong Kong-Macau Greater Bay Area are at leading levels. Overall, the computing power index is generally higher in Beijing, Shanghai, Guangzhou, and their surrounding provinces and regions. Ranking highest are Guangdong, Beijing, Jiangsu, Shanghai, Shandong, and Zhejiang, with computing power development index values over 40. Beijing, Shanghai, Guangzhou, and neighboring provinces with overall high development index values have grasped computer power development opportunities, accelerated promotion of computing power technology industry layout and computing power infrastructure construction, created good computing power development environments, and actively promoted the wide application of computing power.

[20] Due to data availability and continuity limitations, this report's calculations do not include Hong Kong, Macau, and Taiwan.

The computing power development of core provinces in the central and western regions is gathering steam. Computing power demand in Beijing, Shanghai, Guangzhou, and neighboring provinces is brisk. Land and energy being scarce, it is difficult to develop computing power on a large scale. Supply has fallen short of demand, and computing power development is subject to certain limitations. Provinces in the central and western regions have ample resources, but bottlenecks exist, such as narrow network bandwidth, high inter-provincial data transmission fees, and limited computing power demand. With the advancement of the "eastern data, western compute" project, the computing power development potential of core provinces in the central and western regions is greater. At present, Henan and Hubei, representing central provinces, as well as Sichuan, representing western provinces, are in the top ten, with computing power index values of over 20. Sichuan is an important platform for the western China development strategy (西部大开发) and provides strategic support for the Yangtze River Economic Belt; Hubei radiates to and drives the central and middle reaches of the Yangtze River, supporting the development of computing power in the Yangtze River Economic Belt; and Henan leads the Central Plains in the development of computing power and supports the rise of central China. In the future, the central and western provinces may achieve synergistic joint development with other regions including Beijing-Tianjin-Hebei, the Yangtze River Delta and the Guangdong-Hong Kong-Macau Greater Bay Area.

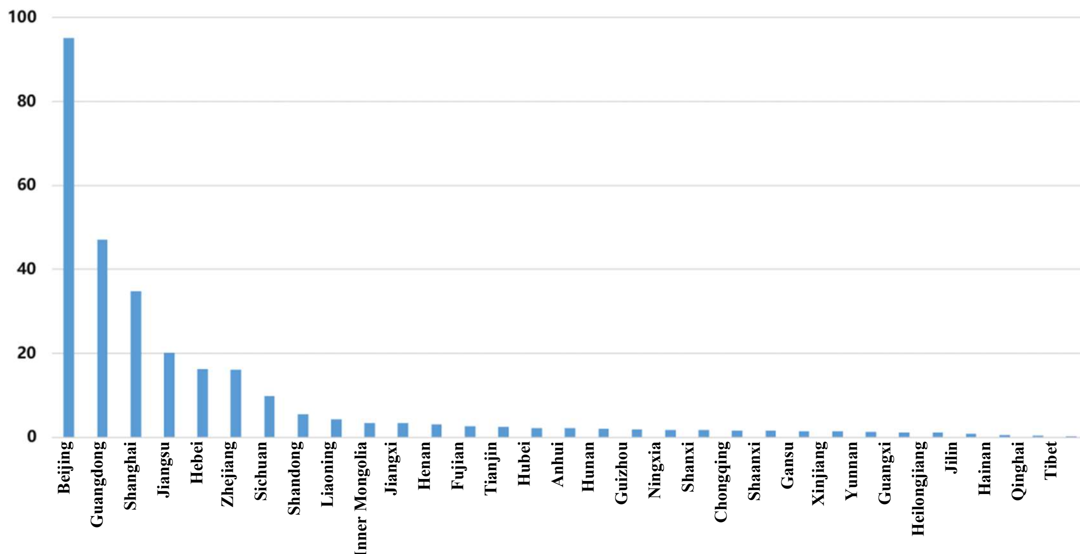


Source: CAICT

Figure 12 Computing power development index of some Chinese provinces in 2020

2. Computing power scale sub-index

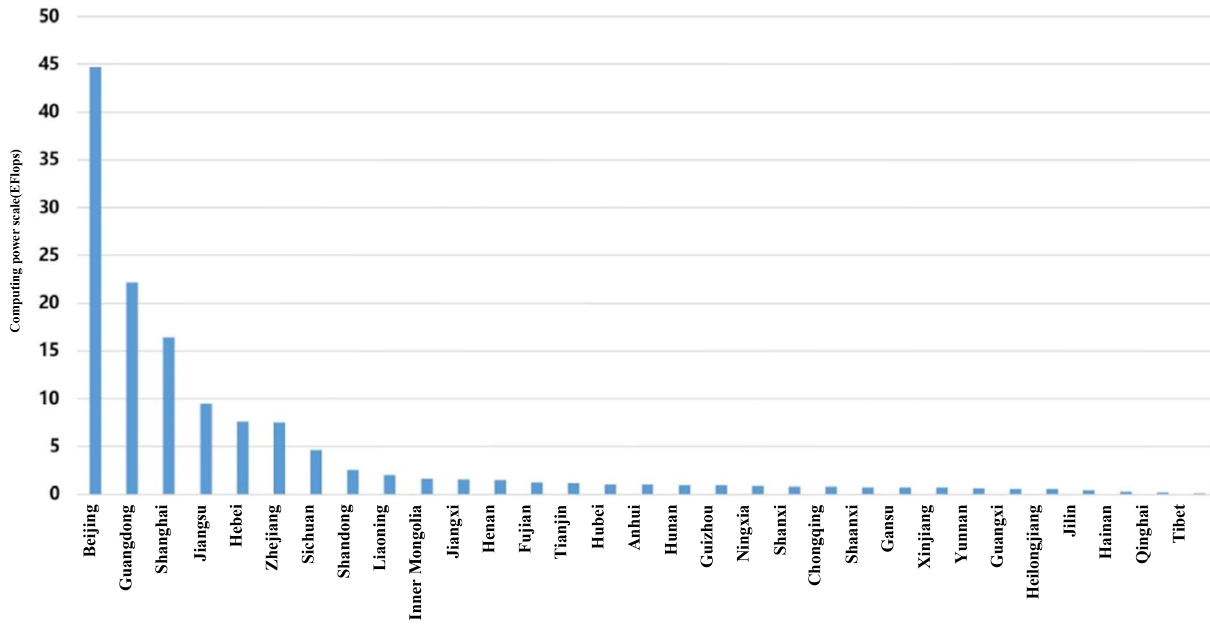
The scale of computing power in Beijing, Shanghai, Guangzhou, and their surrounding provinces and regions is particularly high. With the rapid development of mobile internet and the IoT, as well as the digital transformation and upgrading of traditional industries, demand for computing power in hotspots such as Beijing, Shanghai, and Guangzhou has climbed rapidly, driving the continuous growth of local computing power. With computing power scale index levels over 30, Beijing, Guangzhou, and Shanghai make up the top three. With hotspot regions such as Beijing, Shanghai, and Guangzhou limited by electricity and land resource constraints and locked in by policies, the neighboring provinces of Hebei, Jiangsu, and Zhejiang have gradually taken up the spillover demand from those regions, driving the expansion of computing power in Hebei, Jiangsu, and Zhejiang. Also in the front ranks in computing power scale are regions with booming big data industry development—Sichuan, Inner Mongolia, Jiangxi, and Henan in central and western China, as well as Liaoning in the northeast. They have made the most of local big data resources and accelerated construction of large-scale data centers and other computing power infrastructure.



Source: CAICT

Figure 13 Computing power development scale index of some Chinese provinces in 2020

Computing power scale development is steadily rising everywhere. In terms of total amounts, Beijing, Guangdong, and Shanghai are the top three in computing power, each with over 15 EFlops; Jiangsu, Hebei, Zhejiang, Sichuan, Shandong, Liaoning, and Inner Mongolia complete the top ten. **With respect to growth rates,** Zhejiang and Guangdong are ranked first and second, with growth rates over 60%. Beijing, Jiangsu, Jiangxi, Shandong, Chongqing, Shanghai, and other provinces have computing power growth rates of more than 50%, while the remaining provinces have growth rates between 40% and 50%. **As for the different dimensions of computing power,** Beijing, Guangdong, and Shanghai are China's top three in basic and intelligent computing power. They lead the nation in computing power scale, with a total national share of 62%. In terms of supercomputing power, Jiangsu, Beijing, and Sichuan are China's top three in the scale of supercomputing power. Jiangsu Province in particular has strongly supported construction of the Wuxi and Kunshan national supercomputing centers. Beijing has taken the lead in realizing supercomputing cloud services, and Chengdu Supercomputing Center in Sichuan Province is the first national supercomputing center built and put into operation in the west.

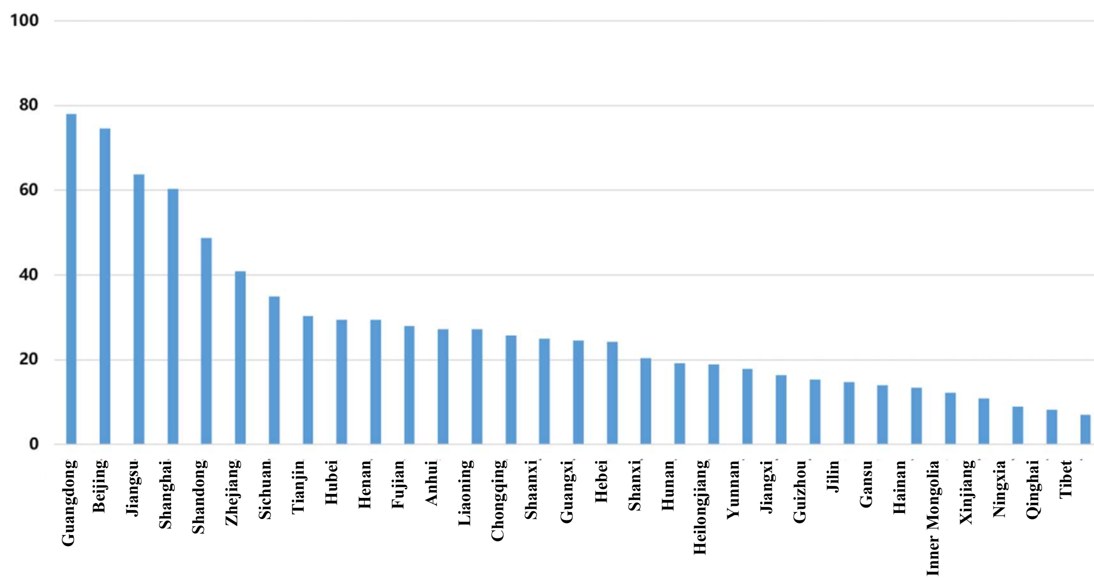


Source: CAICT

Figure 14 Computing power scale of some Chinese provinces in 2020

3. Computing power environment sub-index

The four major urban agglomerations—Beijing-Tianjin-Hebei, the Yangtze River Delta, the Guangdong-Hong Kong-Macau Greater Bay Area, and the Chengdu-Chongqing Economic Circle (成渝双城经济圈)—have higher computing power environment index levels. On the whole, the environment for computing power development in the provinces continues to improve, the computing network environment continues to be refined, and investment in computing power continues to increase. Guangdong, Beijing, Jiangsu, and Shanghai are the top four, with computing power environment index levels of over 60, while Shandong, Zhejiang, Sichuan, Tianjin, Hubei, and Henan fill out the top ten. **In terms of computing power network environment,** the top five are Jiangsu, Guangdong, Shanghai, Shandong, and Beijing, with relevant index levels of over 50. Jiangsu, Guangdong, and Shandong lead the country in inter-provincial internet export bandwidth, and Shanghai and Beijing rank at the top in the country in 5G coverage rate, at over 28%. **In terms of the strength of computing power investment,** Beijing, Guangdong, Shanghai, Jiangsu, and Zhejiang are the top five, with relevant index levels of over 50. Beijing and Guangdong in particular have IT hardware, software, and service expenditures of over RMB 230 billion, putting them ahead of other provinces and regions in computing power investment strength. In addition, Guizhou, Inner Mongolia, Gansu, Ningxia, and other nodes are rich in renewable energy and have a suitable climate, and they have great potential for green development of computing power.



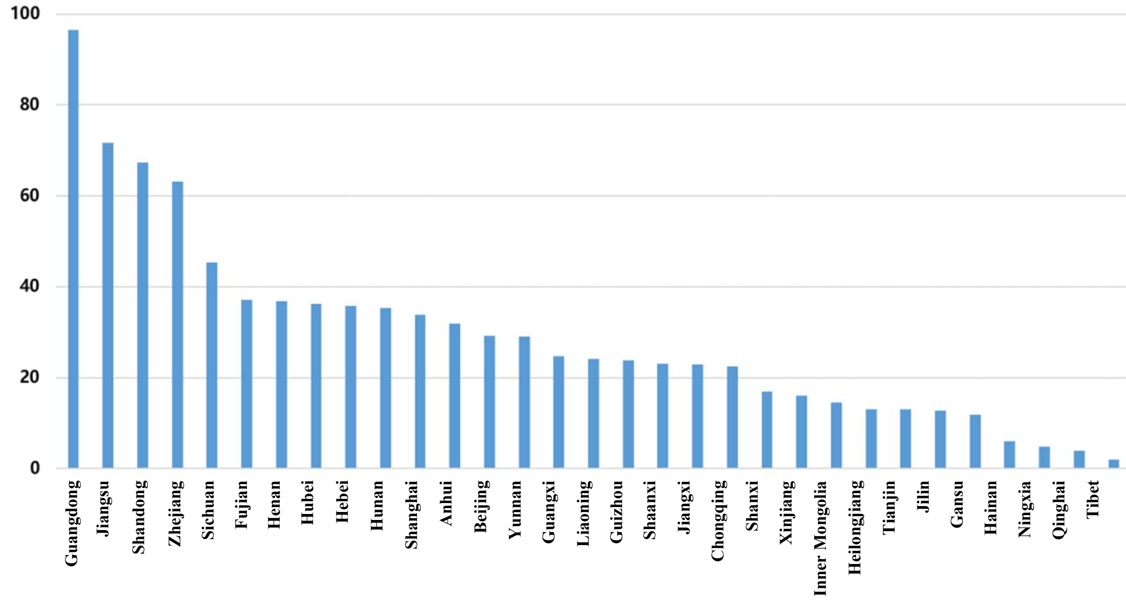
Source: CAICT

Figure 15 Computing power environment sub-index levels of some Chinese provinces in 2020

4. Computing power application sub-index

China's eastern coastal provinces generally have higher computing power application index levels. On the whole, the eastern provinces of Guangdong, Jiangsu, Shandong, Zhejiang, Fujian, and Hebei have relatively high computing power application index levels. Guangdong, Jiangsu, Shandong, and Zhejiang are ranked as the top four, with computing power application index levels over 60. The central provinces of Henan, Hubei, and Hunan and the western province of Sichuan are also among the top ten, with computing power application index levels over 35. **In terms of consumer application levels,** Guangdong, Jiangsu, Shandong, Zhejiang, and Sichuan are ranked as the top five, with consumer application index levels exceeding 50. They lead the country in smart terminal consumption and mobile data traffic consumption, with average monthly mobile internet traffic in excess of one exabyte (EB), and have greater computing power demand for mobile internet applications. **In terms of industrial application levels,** computing power produces a strong impetus for the sustained and healthy development of industrial digitalization in provinces, and the pulling effect on the digital transformation of industries is especially obvious.

Guangdong, Jiangsu, Shandong, Zhejiang, and Shanghai are the top five by industry application index, with index levels over 50. Guangdong Province in particular is far ahead in industrial digitalization development, with the scale of industrial digitalization exceeding RMB 3.5 trillion. Industrial digitalization also exceeds RMB 2 trillion in the Jiangsu, Shandong, and Zhejiang regions.



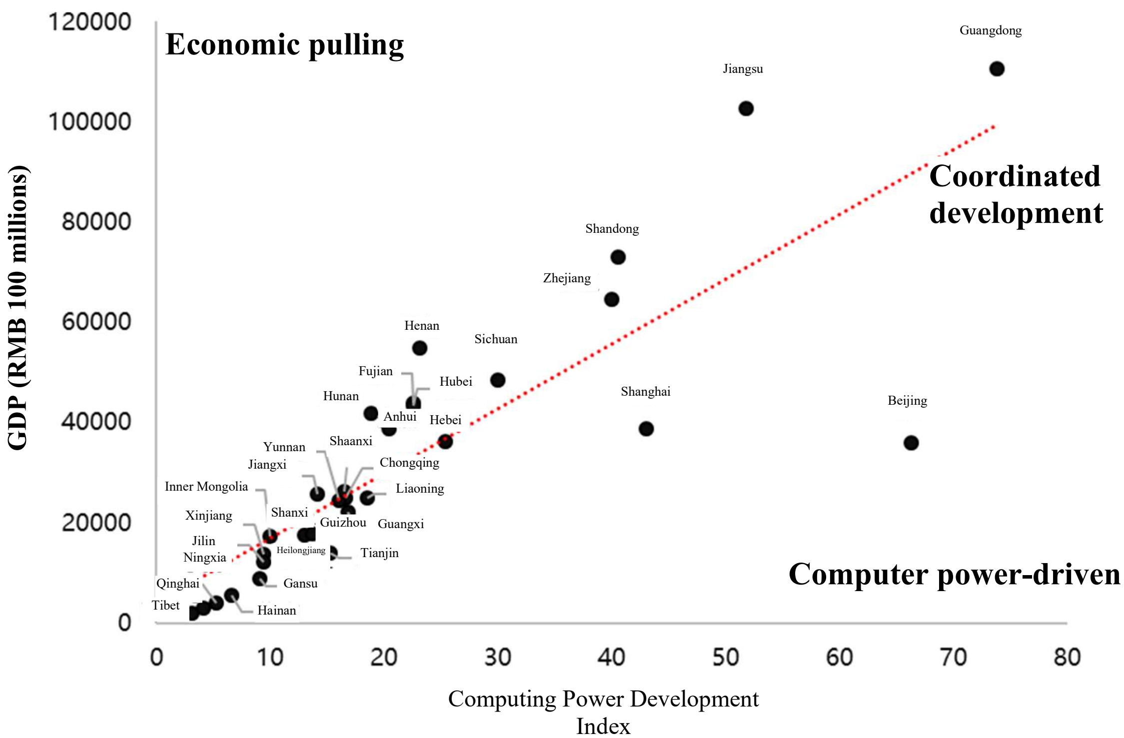
Source: CAICT

Figure 16 Computing power application sub-index levels in some Chinese provinces in 2020

4. Relationship between the computing power index and the economy

A significant positive correlation is evident between the computing power development index levels of provinces and the size of their GDP. Computing power has a strong driving effect on the economic development of provinces, and the provinces with higher levels of digital economy and GDP development in 2020 also had higher levels of computing power development. For every 1-point rise in the computing power development index, GDP grows by approximately RMB 129.3 billion (representing about 0.13% of national GDP). Increases in the computing power development index are determined by increases in the development levels of computing power scale, computing power environment, and computing power application, as coordinated development of supply and demand is achieved.

Provinces are mainly divided into three types according to computing power development: computing power-driven, economic pulling type, and coordinated development type. First is the computing power-driven type represented by Beijing and Shanghai. Having made sustained investments in large-scale computing power infrastructure construction for many years, these places have larger-scale computing power and better computing power environments. This not only serves the computing power application of these provinces, but also provides computing power support for the consumer and industrial applications of other provinces. **Second** is the economic pulling type represented by Guangdong, Jiangsu, Shandong, and Zhejiang, where local computing power demand is strong and the levels of consumer and industrial application of computing power are high. For these provinces, computing power cultivates new momentum and opens up new space for sustainable and healthy digital economy development and economic growth. **Third** is the coordinated development type represented by Hebei and Chongqing, where the digital economy and computing power have developed synergistically. Coordinated development layouts have gradually formed in the Beijing-Tianjin-Hebei, Yangtze River Delta, Guangdong-Hong Kong-Macau Greater Bay Area, and Chengdu-Chongqing economic circle regions.



Source: CAICT

Figure 17 Relationship between the computing power development index and GDP

V. Accelerating Computing Power-Based High-Quality Development to Support Construction of the New Development Pattern

The 14th Five-Year Plan period [2021-2025] is a major period of strategic opportunity for China to achieve leapfrog development in computing power. Taking supply-side structural reform as the main line and implementing expansion of domestic demand as the strategic starting point, we should solidify the foundation, optimize industries, apply traction, and open up cooperation, so as to achieve a higher-level dynamic balance where supply creates demand, and demand elicits supply, thereby accelerating the cultivation of new dynamism (新动能), and comprehensively supporting the "dual circulation" new development pattern.

First is to accelerate construction of computing power infrastructure: Focusing on strengthening support for digital transformation, intelligent upgrading, and integrated innovation, accelerate the laying out of data centers, intelligent computing centers, supercomputing centers, and other computing power infrastructure construction, so as to enhance data operations (数据运算) capabilities. Accelerate the construction of a national integrated big data center system, strengthen the overall smart dispatching of computing power, build several national hubs and nodes and big data center clusters, and build exascale and 10 EFlops-level supercomputing centers. Continuously guide the linking of computing power supply and demand in each region, so as to improve application levels.

Second is to build a good environment for computing power development. Enhance the ability to support new computing power networks (新型算力网络) and optimize regional computing power interconnection capability. Continuously promote the structural optimization and expansion of key links of the internet backbone network and urban area networks, and steadily promote 5G network construction. Channel social capital to participate in the construction of computing power infrastructure, and encourage financial institutions and others to increase their support for computing power infrastructure. Encourage qualified financial institutions and enterprises to issue green bonds and support qualified enterprises to seek initial public offering (IPO) financing.

Third is to raise the competitiveness of the production chain and supply chain. Focus on creating supply chain advantages for the computing industry's production chain, continuously strengthen the capacity of the whole computing system, and accelerate software-hardware integration and heterogeneous collaborative innovation, in order to meet the computing needs of massive data and diverse scenarios. Strive to make up supply chain shortcomings in the production chain, promote industry-academia-research institute collaboration, and intensify research efforts in important products and in key and core technologies, so as to accelerate improvement of the computational efficiency of computing power and promote diversification of the computing power supply chain. Enhance basic research and multi-path exploration, step up promotion of cutting-edge computing system innovations such as in-memory processing (存算一体), quantum computing, and brain-inspired computing, and strengthen original innovation capacity.

Fourth is to strengthen the demand pull of computing power application. Strengthen demand pull and supply-demand linkages, tap digital consumption demand for computing power, raise the level of consumer application of computing power, and accelerate development of new forms of integrated online-offline consumption of education and training, healthcare and elder care, transportation and travel, etc., which will promote the structural optimization and upgrading of consumption. Further promote the digital transformation of traditional industries, strengthen the application of computing power in various industries, encourage enterprises to use the cloud in depth, deepen digital applications in all aspects of industry, cultivate new business formats, build industry chains in the "cloud," incubate data-driven enterprises, and enhance the digital transformation capability of industry. Promote the ubiquitous application of public computing power to meet the needs of government services and the people's livelihoods, improve the supply of public computing power resources, optimize the computing power service system, reduce the cost of computing power use, and raise the levels of computing power service dispatching and application empowerment support.

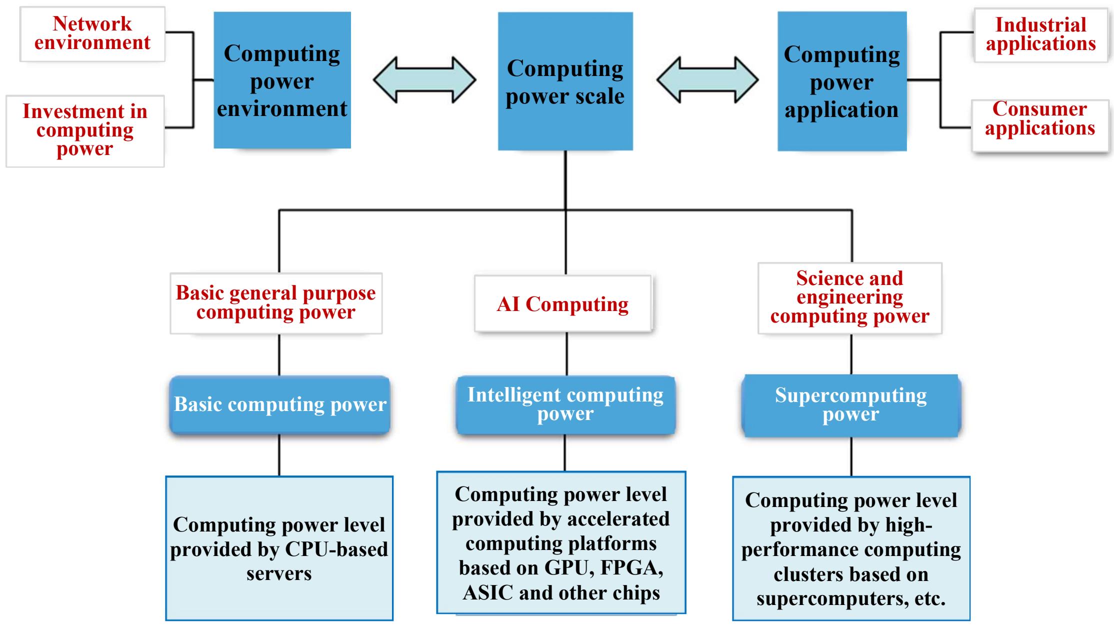
Fifth is to continuously expand international cooperation. Seize opportunities for integrating the digital economy and the real economy, strengthen “Belt and Road”[21] cooperation, enhance cooperation in the fields of computing power infrastructure, computing power technology industry, and digital transformation, create partnerships based on mutual trust, mutual benefit, inclusiveness, innovation, and win-win benefits, expand the vast space for digital trade development, and build a community of common destiny in cyberspace (网络空间命运共同体) among countries along the route.

Appendix 1: Framework for calculating the computing power index

Taking into account the definition of computing power, the computing power index includes three parts: computing power scale, computing power environment, and computing power application. The key elements of computing power scale include basic computing power, intelligent computing power, and supercomputing power, which provide basic general purpose computing, AI computing, and science and engineering computing, respectively. The computing power environment mainly includes the network environment and investment in computing power. Continuous optimization of the network environment provides solid support for computing power development, while large-scale investment in computing power generates direct and indirect driving effects on the growth of computing power. Computing power application mainly includes consumer applications and industrial applications. Consumer and industrial applications have brought about rapid increases in the demand for computing power scale, computing power capabilities, etc., and

[21] Translator's note: The "Belt and Road" (“一带一路”) refers to the Silk Road Economic Belt (丝绸之路经济带) and the 21st Century Maritime Silk Road (21 世纪海上丝绸之路).

conversely, advances in computing power have driven the development of applications. The framework for calculating the computing power index is as shown in the following figure.



Source: CAICT

Appendix Figure 1: Framework for Calculating the Computing Power Index

1. Computing power scale sub-index calculation method

The computing power scale sub-index is calculated by weighting the values of secondary indicators—basic computing power, intelligent computing power, and supercomputing power. Standardized processing is performed separately on basic computing power, intelligent computing power, and supercomputing power to obtain the value assignments for each indicator.

1. Basic computing power. Reflects the computing power of each region provided by CPU chip-based servers, and is measured mainly using an indicator of the scale of server computing power.

Scale of server computing power = $\sum_{\text{last 6 years}}$ (annual scale of server shipments*average server computing power in the current year).

2. Intelligent computing power. Reflects computing power of each region for AI training and inference provided by accelerated computing platforms based on GPUs, FPGAs, ASICs, and other chips, and is measured mainly using an indicator of the scale of AI accelerator computing power.

Scale of AI accelerator computing power = $\sum_{\text{last 6 years}}$ (annual scale of accelerator chip shipments shipments*average accelerator chip computing power in the current year).

3. Supercomputing power. Reflects the computing power of each region provided by high-performance computing clusters based on supercomputers, etc., and is measured mainly using an indicator of the scale of supercomputing power.

Supercomputing power = scale of supercomputing power = ∑ supercomputer computing power

2. Computing power environment sub-index calculation method

The computing power environment sub-index is calculated by weighting the values of secondary indicators—network environment and computing power investment. Standardized processing is performed separately on network environment and computing power investment to obtain the value assignments for each indicator.

1. Network environment. Reflects the service capability of each region in data transmission with other regions in China for business such as data and internet, as well as the edge computing support capability of the region. It is measured mainly using tertiary indicators such as inter-provincial internet export bandwidth and 5G coverage rates, and is calculated by standardized processing and weighting of values.

Inter-provincial internet export bandwidth = the sum of the urban area export bandwidths of all operators

5G coverage ratio = number of 5G base stations/number of 4G base stations

2. Investment in computing power. Reflects each region's investment in computing power, and is measured mainly using an indicator of the scale of IT expenditures.

Scale of IT expenditures = sum of each region's investment in IT hardware, software, and services

3. Computing power application sub-index calculation method

The computing power application sub-index is calculated by weighting the values of secondary indicators—consumer application level and industrial application level. Standardized processing is performed separately on the consumer application level and industrial application level to obtain the value assignments for each indicator.

1. Consumer application level. Reflects each region's level of computing power application in consumer fields such as mobile internet, and is measured mainly using an indicator of average monthly mobile internet traffic.

Average monthly mobile internet traffic = number of mobile phone users in each province*average monthly mobile internet traffic per user

2. Industrial application level. Reflects each region's level of computing power application in traditional industrial fields such as agriculture, industry, and services, and is measured mainly using an indicator of the scale of industry digitalization.

Scale of industry digitalization = each region's increased output and efficiency gains (value added) from the convergence and penetration of ICT products and services in other fields

Appendix 2: Data Sources

1. Basic data, including population data, economic value added, industrial value added, and national input-output tables come from the relevant data of the National Bureau of Statistics and provincial statistics departments.

2. Data center rack numbers, 5G base station numbers, mobile phone user numbers, and average monthly mobile internet traffic per user for each of China's 31 provinces comes from MIIT statistical data.
3. Server and AI accelerator shipment quantities globally and in China come from IDC and Gartner statistical data, and are used to calculate and assess basic computing power and intelligent computing power globally and in China.
4. Data on the scale of supercomputing power globally and in China come from the internationally well-known TOP500 list and data provided by relevant firms.
5. Expenditures on computing power hardware, software, and services by each of China's provinces comes from the National Bureau of Statistics and relevant IDC statistical data, and is used to assess investment in computing power by each of China's provinces.