Bulletin of Chinese Academy of Sciences

Volume 37 | Issue 1 Article 2

1-20-2022

Thoughts on New IT Technique System

Ninghui SUN
Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing 100049, China, snh@ict.ac.cn

Recommended Citation

SUN, Ninghui (2022) "Thoughts on New IT Technique System," *Bulletin of Chinese Academy of Sciences (Chinese Version)*: Vol. 37 : Iss. 1 , Article 2.
DOI: https://doi.org/10.16418/j.issn.1000-3045.20211117002
Available at: https://bulletinofcas.researchcommons.org/journal/vol37/iss1/2

Thoughts on New IT Technique System

Abstract

The world is entering a new period of information society dominated by digital economy. To adapt to the requirements of the new development stage of China, and to achieve the key technologies development, it is necessary to accelerate the construction of a new selfreliance and self-improvement IT technique system. The article gives some thoughts on the demand, principle, and key method on above issues, and the new paradigm of research at Institute of Computing Technology, Chinese Academy of Sciences.

Keywords

IT system, self-reliance and self-improvement of technique, opening up and sharing, new paradigm of research

Citation: SUN Ninghui. Thoughts on New IT System [J]. Bulletin of Chinese Academy of Sciences, 2022 (1).

Thoughts on New IT System

SUN Ninghui[1,2]

1. Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China;
2. School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing 100049, China

Abstract: The world is entering a new period of information society dominated by digital economy. To adapt to the new development stage and make major breakthroughs in key technologies, China needs to accelerate the construction of a new IT system to achieve self-reliance and self-improvement. The article provides some thoughts on the demand, principle, and key method of establishing a new IT system and the new paradigm of research in the Institute of Computing Technology, Chinese Academy of Sciences. **DOI:** 10.16418/j.issn.1000-3045.20211117002-en

Keywords: IT system; self-reliance and self-improvement of technique; opening up and collaboration; new paradigm of research

The world is entering a new period of information society dominated by digital economy. The 19th National Congress of the Communist Party of China (CPC) set the strategic goal of becoming a global leader in innovation by 2035. Furthermore, the fifth plenary session of the 19th CPC Central Committee proposed that innovation should remain at the core of the national modernization, and self-reliance and self-improvement in science and technology be upheld as the strategy for national development. To adapt to the new development stage and make major breakthroughs in key technologies, China needs to establish a new information technology (IT) system to achieve self-reliance and self-improvement of technique.

To establish a new IT system, we need to comply with the general trend of IT development and think about China's international status in the new round of globalization. In addition, we should not only think about how to establish a new self-reliant and self-improved computing technology system but also consider how to build a new research paradigm suitable for the new system.

1 Necessity of establishing a new IT system for self-reliance and self-improvement

In the era of intelligence, will China continue to follow the technology system of the United States or develop its own technology system? Can China's technology system go global in the new round of globalization in the future? These are important issues worthy of in-depth consideration.

(1) From the current ecological status and development trend of processors. As the backplane of IT, processor has two dimensions: ecology and application, which can be expressed by the nine-square grid (Figure 1). In terms of ecological openness, processors can be classified as closed source instruction, authorized instruction, and extensible instruction. In terms of the application era, the development of processors experienced IT 1.0 (server/desktop), IT 2.0 (mobile application), and IT 3.0 (big data/AIoT) ①. The red diagonal in Figure 1 denotes the ecological development of AIoT which follows the Bell's law ②. At the same time, it is also the frontal battlefield of establishing a new IT system and accounts for the largest proportion of informatization. The Wintel alliance (Microsoft + Intel) was developed first, and then the dual A architecture (ARM + Android). The next one may be RISC-V/Ubiquitous OS (the fifth-generation reduced instruction set computer + ubiquitous operating system). The other two ecosystems in Figure 1 expand along the horizontal and vertical axes, respectively. The horizontal axis refers to the ecological development of PC processors, and the vertical axis represents the ecological development of server processors.

(2) From the overall pattern of processors. The overall pattern of processors can be represented by three ecological control systems which are referred to as the A, B, and C systems in this paper. ① A system. It is an ecosystem controlled by the United States and originated in the 1990s. It is an information highway technology system developed under the leadership of the United States [1]. For China, the A system is indispensable. Therefore, solving the strangled problems should be put in priority. At the same time, binding with the United States can be taken as a way to avoid hard decoupling. ② B system. It is an independent and controllable IT system that China must establish for national security and

① AIoT (Artificial Intelligence Internet of Things) = AI (Artificial Intelligence) + IoT (Internet of Things).
② A new type of computer appears every 10 years with a tenfold increase in number.

Received: 2021-12-24
Supported by: Special Project of Strategic Research and Decision Support System Construction of Chinese Academy of Sciences (GHJ-ZLZX-2021-06)

an ecological system controlled by China. The B system mainly imitates and tracks the A system in key technologies, and innovates according to the domestic demand. For China, the B system is flexible. Although we have held the bottom line through the IT innovation project for in-situ replacement in the last 20 years, the application of the B system is still limited. ③ C system. In the next 30 years, we need to focus on independent innovation of core technologies, make original, basic, and core achievements in the cutting-edge and core technologies, and establish a C system to achieve self-reliance and self-improvement of technique in the era of intelligence. The C system is a technological system jointly established by China and the world, and it is an ecosystem controlled by the world. The C system can solve the problems both in China and the world. For China, it is a balanced system established by joint effort.

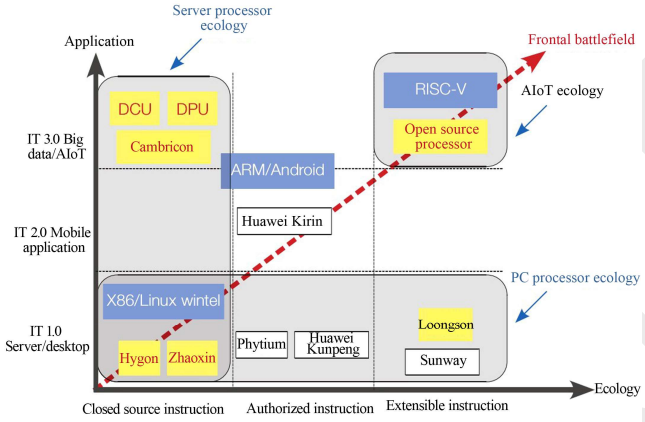

Figure 1 China's processor ecosystem diagram

The A, B, and C systems are all indispensable. Several representative companies related to the technology transfer of the Institute of Computing Technology, Chinese Academy of Sciences (the ICT of CAS) are the backbones of the three systems. The representative company of the A system is Hygon Information Technology Co., Ltd. which aims to meet the needs of informatization to the greatest extent and squeeze the profit space of main competitors by relying on China's cost, human resources, and policy advantages. The representative company of the B system is Loongson Technology Co., Ltd. which aims to prepare for the worst scenario and ensure the basic needs of China's core departments and key industries. The representative company of the C system is Cambricon Technologies Co., Ltd. which aims to compete with major rivals for excess profits in the future mainstream and international markets, especially the dominant position in global ecology by relying on the pioneering advantages of original technology.

2 Basic principles of establishing the new IT system

To achieve self-reliance and self-improvement in the IT field, we first need to clarify the basic principles of system establishment. Facing the new demands, new challenges, and new markets in the era of intelligence, the C system expands in two dimensions. One refers to the workload, design, and manufacturing technology of the chip. The other refers to the end, edge, network, and cloud throughout the information infrastructure. There are always philosophical ideas behind everything, and the design of a new IT system is no exception. The basic principles of establishing a new IT system are not sufficient conditions for success but are necessary conditions, and they can bring great changes. The establishment of a new IT system for self-reliance and self-improvement of technique and the breakthrough of a new generation of IT require us to give full play to the innovative elements with Chinese characteristics and the advantages such as strong collective consciousness and governance ability, rich innovative scenarios, numerous engineering and technical talents, and large markets. In general, to establish a new IT system surpassing the original technology system, we need to follow five basic principles (Figure 2).

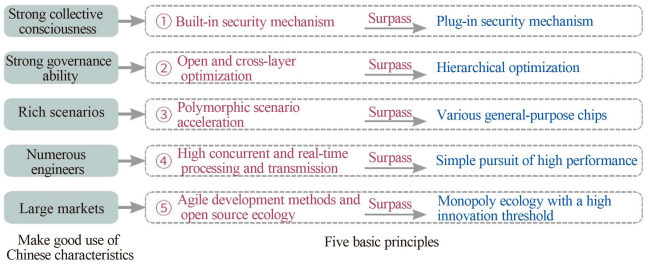

Figure 2 Five principles of the C system of IT in China

(1) The built-in security mechanism surpasses the plug-in security mechanism. The A system is a security mechanism designed according to the plug-in security mode. The design of the plug-in security mechanism is like that of a house. At the beginning of the design, we did not take the quakeproof or theftproof properties into account. After it is done, we are reminded of strengthening the beam and installing security doors and windows, and cameras. These supplementary measures are plug-in security measures. The same is true in the design of the A system. Security is ignored in terms of transmission, network, and computing. In the transmission layer, the transmission technology is designed based on the Shannon's theorem, which focuses on capacity improvement. In the network layer, the technology is based on TCP/IP protocol ③, DNS (domain name resolution system), and PKI (public key infrastructure), and the security mechanism is

③ TCP/IP protocol (Transmission Control Protocol/Internet Protocol) refers to a protocol suite that can achieve information transmission between different networks.

patched, with the security of the network architecture failing to be fully considered. In the computing layer, von Neumann computing architecture was proposed in the 1950s, which focused on improving computing performance without considering security. In the future, the design of the computer system should take security into account.

(2) Open and cross-level optimization surpasses hierarchical optimization. The goal is to exert China's institutional advantages and fully mobilize all possible factors by formulating interface standards based on the traditional hierarchical development of IT. In this way, the cross-layer and vertical optimization of the industry can be achieved, and the system performance can be improved [2]. The rocket technology system and the IT system (Figure 3) employ different structures. The rocket technology system is characterized by a big base and narrows as it goes up, with heavy-lift launch vehicle at the top. The overall innovation system is not easy to be strangled. However, the IT system widens as it goes up. The top refers to application and software, and the bottom represents equipment, chip, and technology. The system narrows as it goes down, which indicates that the market is small, with limited suppliers and fragile supply chains, so it can be easily strangled. The A system provides hierarchical optimization, with Intel, IBM, and Google in different layers and optimization in different layers. In such a system, each company can be great when it gets each layer done well, while it is difficult to achieve cross-layer and vertical optimization. IBM, an international giant, once adopted a full-stack and connected layout involving technology, materials, chips, equipment, software, and applications and made huge profits in high-value fields such as banking and insurance. However, in the internet era, it was surpassed by the enterprises (e.g., Google, Intel, and TSMC) adopting hierarchical optimization with higher cost performance. Therefore, we need to develop a more cost-effective vertical technology system to ensure both hierarchical optimization and vertical optimization, so as to break through the market barriers established by international monopoly enterprises.

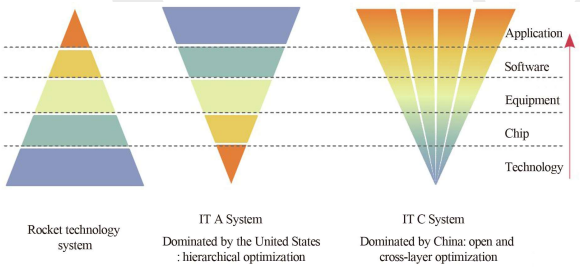

Figure 3 Structure of rocket technology and IT systems in China

(3) Polymorphic scenario acceleration surpasses various general-purpose chips. Comparatively speaking, the United States has few people and engineers, large high-value markets, and advanced semiconductor manufacturing industries. For the United States, the most cost-effective and feasible way is to use various CPU (central processing unit) and GPU (graphics processing unit) general-purpose chips and rely on advanced semiconductor technology to cover a wide range of scenarios. China has numerous people and engineers, rich scenarios, and large but fragmented markets. The semiconductor manufacturing industry in China has lagged behind for a long time and is still strangled by the extreme ultraviolet (EUV) lithography machine. Therefore, can China gain competitive advantages by using 100 kinds of domain-oriented special chips and relying on the second principle to optimize the connection of software and hardware? Recently, the "Siyuan 270" NPU (embedded neural network processing unit) developed by Cambricon Technologies Co., Ltd. is 10 times better than the GPU Tesla V100 developed by NVIDIA in the same period, in terms of comprehensive performance/power consumption and price ratio for performing artificial intelligence (AI) tasks, and this is exactly an example of polymorphic scenario acceleration.

(4) High concurrent and real-time processing and transmission surpasses the simple pursuit of high performance. China has a large population but relatively limited resources, so it faces serious high concurrency. The United States has a small population but rich resources, so it can pursue high performance. Just like the traffic jam in the morning rush hour, if the entrance and exit designs are unreasonable, and the test is not available, then the effectiveness of infrastructure investment will drop at an inflection point. This is why we propose that China should pursue high throughput. The construction of a national high-throughput information infrastructure aims to increase the data volume and the bandwidth by more than 1 000 times and 100 times, respectively, support the real-time connection of 100 billion things, reduce the processing delay to less than one fifth with the end-to-end delay at the millisecond level, which will help to achieve the collaborative optimization of data transmission and processing, and effectively support the applications sensitive to real-time delay.

(5) Agile development methods and open source ecology surpass monopoly ecology with a high innovation threshold. Agile design is the most effective way to deal with fragmented application scenarios, and open source can effectively break the technological monopoly [4]. Gaining cost advantage through extreme optimization, occupying the middle and low-end markets, pushing competitors to the high end, and winning the time window are effective strategies for the weak to defeat the strong, and this is how China's manufacturing industry develops. China is confident during the trade war because it has an extreme cost advantage in the manufacturing industry, occupies the middle and low-end markets in an irreplaceable way, and then gradually develops towards the middle and high-end markets. At the 20th Academician Conference of the Chinese Academy of Sciences, the 15th Academician Conference of the Chinese Academy of Engineering, and the 10th National Congress of the China Association for Science and Technology, General Secretary Xi Jinping pointed out that human beings are entering an era of intelligent interconnection of all things with

the human-cyber-physical system integration. Here, a review of the historical role of open source in various stages of human-cyber-physical systems is needed. In the era of cyber, the open source operating system made the dominant position of IBM's minicomputers occupied by X86 servers, and Intel, benefiting from the open source Linux, accounted for 90% of the data center market. In the era of human-cyber system integration, the open source Android system made the dominant position of X86 occupied by ARM, and the smartphone industry benefited from the Android. Now in the era of human-cyber-physical system integration, will the open source ubiquitous operating system and open source instruction set lead to the replacement of ARM by RISC-V?

Here, the establishment of the chip technology system of the C system is illustrated. The chip technology system of the C system contains (1) lithography machine, photoresist and large silicon wafer with domestic 28 nm technology; (2) 28 nm chip manufacturing technology with ultra-high cost performance; (3) open source electronic design automation (EDA) toolchain, open source CPU core, open source key IP, and open source chiplet technology; (4) cloud-based agile design tools and platforms of the chip; (5) chip innovation platform with a domain-specific architecture (DSA) created for massive scenarios.

3 Portfolio of research directions for establishing the new IT system

The IT ecology based on the C system is characterized by openness, controllability, security, and "federalism." (1) Openness. The dual cycle of economic development emphasizes that the new round of globalization based on the domestic cycle must be more open, which requires the open ecology of the new system. (2) Controllability. Any link cannot be strangled. (3) Security. In the era of intelligence, the physical world and the virtual world are deeply integrated, which puts forward higher requirements for security. (4) "Federalism". The economic model corresponding to the C system is like "federalism." IT is closely related to the economy, so the development of IT must consider the elements of the economy. The international IT industry is full of monopolies, and the economic model is like an imperial system. We can employ the "federalism" to build a new ecology of the information industry and a global community with a shared future for the digital space according to the idea of building a community with a shared future for mankind proposed by General Secretary Xi Jinping, which may be a new idea for Chinese IT enterprises to go global.

According to the structure of C system, the research in network computing (Figure 4) can be divided into chip, system, internet, distributed system, and application layers. (1) In the chip layer, polymorphic accelerators such as open source CPU, NPU, and DPU (data processing unit) are required. (2) In the system layer, new computing systems such as high-throughput computers, intelligent computers, and edge computers are required. (3) In the network layer, new network equipment including the virtual router and integrated space-ground network is required. (4) In the distributed system layer, a computing network is required, with computing resource considered as the infrastructure. The proposed comprehensive test site of the information superbahn information infrastructure refers to a test computing network. The past information superhighway is a data network, and the information superbahn is a new generation of the wide-area distributed computing system. (5) The top layer represents application, and the informatization of sectors such as public opinion, agriculture, and discipline inspection is like special trains of the information superbahn. Some layers depend on each other and others support each other.

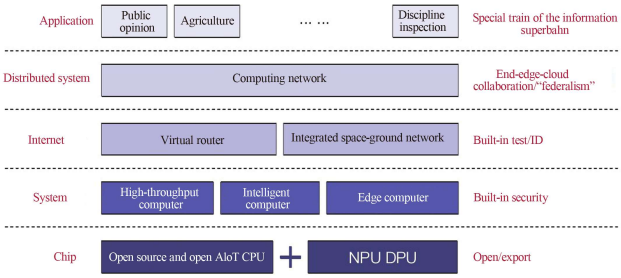

Figure 4 Portfolio of research directions of C system in IT of China

4 New research paradigm for the new IT system

The establishment of a new system requires a new research paradigm. The following part will take the ICT of CAS as an example to explain how the research layout and paradigm should be adjusted for establishing the new IT system.

(1) New layout of the ICT of CAS. In a narrow sense, the C system is equivalent to the strangled chip and information superbahn in the ICT of CAS. In a broad sense, the intelligent computing system, data science, and computational intelligence should also be added to form a new layout of the ICT of CAS, as shown in Figure 5. The vertical part represents the intelligent computing system, which can connect the chip, computing system, and data science. The chip being strangled is the base, and information superbahn is the information infrastructure. Data science and computational intelligence are the momentum for the C system in the new era, which used to be computing science and network science in the past era. IT 1.0 and IT 2.0 are mainly driven by computing science and network science, respectively. IT 3.0 is driven by data science and computational intelligence. For the ICT of CAS, the research on intelligent algorithms and applications should play a leading role and be tightly coupled with the special train of information superbahn, special machine of intelligent computer, DSA chip, so as to drive the whole C system forward.

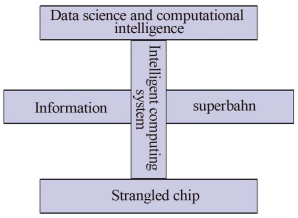

Figure 5 Research during the 14th Five-Year Plan period in the Institute of Computing Technology, Chinese Academy of Sciences

(2) New collaborative paradigm. Establishing a new system is much more complex than developing new equipment because it requires joint efforts of multiple teams from research institutes, universities, and enterprises. If an analogy is drawn between this and military operations, a research institute is like a military sub-district while the research center is a basic combat unit and an independent regiment. The regiment fights independently, which is the best organizational way of a research institute. Therefore, better research teams will establish an enterprise to form a collaborative paradigm between the research center and the enterprise. The establishment of a new system depends on a new organizational paradigm and requires a regimental commander system that can command regiments to fight with clear objectives and mixed coordination. Therefore, it is essential to establish a new collaborative paradigm with the regimental commander system in the ICT of CAS and other related research institutions in the future. Here are two examples of collaborative research. ① For the information superbahn, the ICT of CAS needs the participation of N research teams, M enterprises, and Nanjing Institute of Information Superbahn. ② For the intelligent computer, the ICT of CAS requires the joint efforts of N research teams, one research center of Zhejiang Lab, and one innovation center of a related institution. In the establishment of a new collaborative paradigm, we should learn from the successful experience of China's audio video coding standard (AVS) mode: AVS teams are distributed all over the country and can do a great thing through cooperation. At the same time, we also need to think about how to use the educational resources of the four universities affiliated to the CAS (F4 universities) to promote the integration of science and education and the integration of industry and education in the future. Furthermore, we need to think about and explore the orientation and division of industrial resources of enterprises affiliated to the ICT of CAS and how to combine high-quality innovation forces.

(3) New relationship featured by four circles. More than 10 years ago, when making the strategic planning, the ICT of CAS summarized the internal and external relationship into four circles. Under the new situation of establishing a new system, the positional relationship of the four circles has changed. As shown in Figure 6, the four circles are no longer spreading out from the center, and the ICT of CAS moves from the center to the left. The circles constantly enlarge and become coupled from the left to the right. The left-most ICT of CAS is the source (the first circle), and the second circle represents cooperative institutions, enterprises affiliated with the ICT of CAS, national laboratories, and F4 universities. The third circle refers to the joint laboratory enterprises and cooperative universities, and the fourth circle is the Turing enterprise (enterprises invested by the Turing fund of the ICT of CAS), the entrepreneur club of the ICT of CAS, and the alumni association. The four circles are coupled with each other, just like a board with chess pieces. How to work together and connect them is something we need to think about when establishing a new IT system.

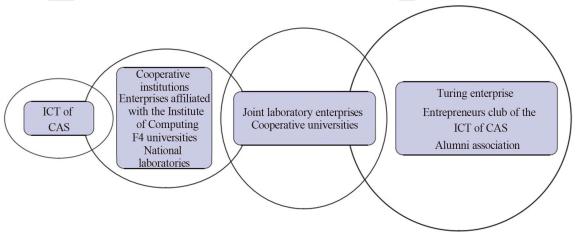

Figure 6 Relationship among four innovation circles of Institute of Computing Technology, Chinese Academy of Sciences

References

1 Broad W. Clinton to promote high technology, with Gore in charge. New York Times, 1992-11-10 (01).
2 Leiserson C E, Thompson N C, Emer J S, et al. There's plenty of room at the Top: What will drive computer performance after Moore's law? Science, 2020, doi: 10.1126/science. aam9744.
3 Wang X H, Wang S, Tang H W, et al. Building substrate for national strategy of new infrastructure construction—Practice and thought of Information Superbahn testbed. Bulletin of Chinese Academy of Sciences, 2021, 36 (9): 1066–1073. (in Chinese)
4 Sun N H. Spirit of open source. Communications of the CCF, 2021, 17 (5): 1. (in Chinese)

(Translated by WANG YX)

SUN Ninghui, Academician of Chinese Academy of Engineering. He is the Scientific Director of Institute of Computing Technology (ICT), Chinese Academy of Sciences (CAS). He is also the Director of State Key Laboratory of Computer Architecture. He serves as Editor-in-Chief of Chinese Journal of Computers, and Editor of Bulletin of Chinese Academy of Sciences. He is the leader of the expert group on strategic studies of information technology development roadmap in CAS. His main research interests include high performance computing and computer architecture. E-mail: snh@ict.ac.cn

