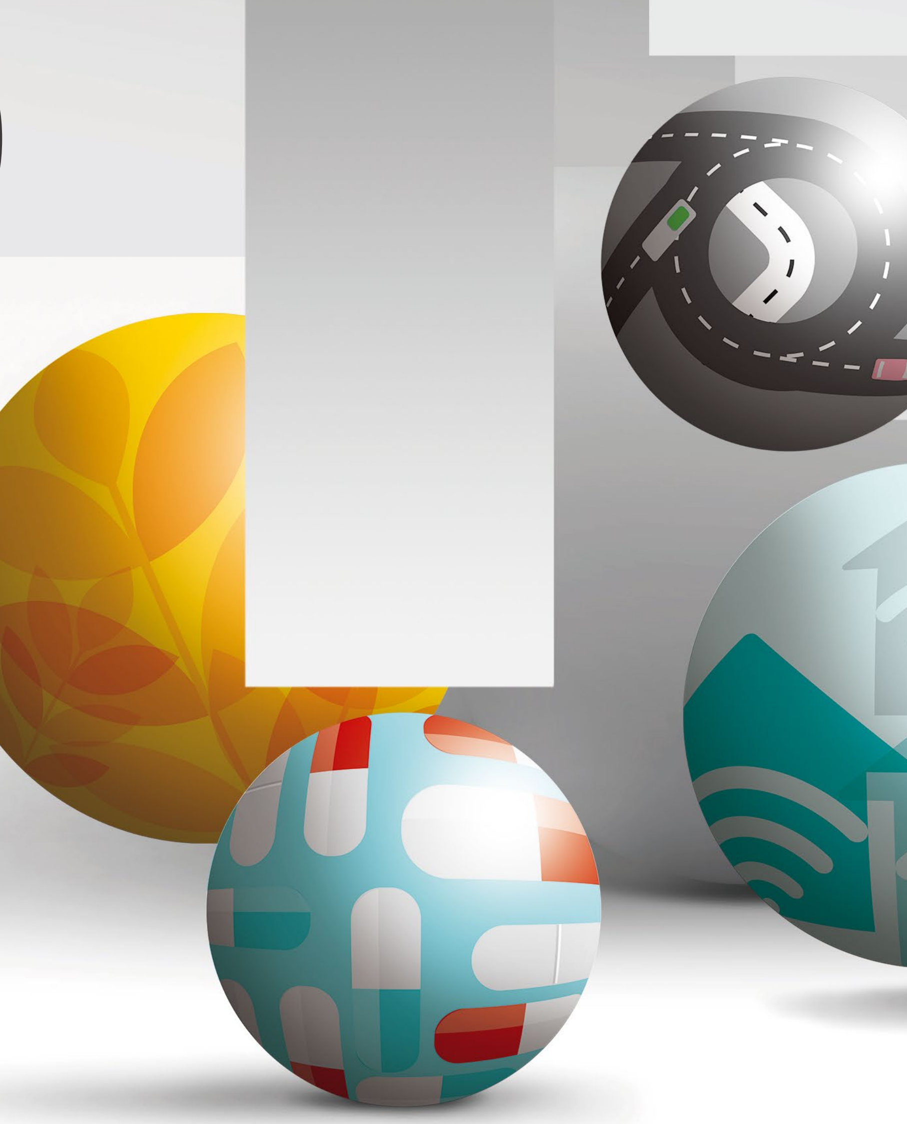
HUAWEI

Intelligent World 2030



Building a Fully Connected, Intelligent World

Foreword

George Gilder

Imagine an intelligent real-time world where waits and delays, queues and paperwork, hassles and drudgery dissolve in almost instantaneous provision of goods and services when and where you need them most.

Imagine a digital domain measurable in Yottabyte, ten to the 21th characters, where intelligence is as ubiquitous and broadband as air. Imagine nearly all homes and vehicles and people linked at a rate of at least 10 gigabits or billions of bits a second.

Imagine a world of “computable health services,” where no one has to wait in lines for care, where cancer and pandemic alike are cancelled...a world turning data into food to solve world hunger in farms that become skyscraping factories unaffected by climate or scarcity of land...a world where people commute in self-driving cars that are colleges, opening up a new “third space” of mobile learning and entertainment...a world where the self-driving revolution engulfs the skies, with electric taxis of the air available at a whim...

Imagine a world where cities are all-smart, all-optical, all-human, all connected, and all green...a world where 85 percent of all companies use unbreakable blockchain technology to protect personal identity and security and where more than half of all computing is “privacy enhanced”...a world where energy is clean, smart, and cloud- controlled...a world where companies are as flexible and resilient as human minds in new webs of digital trust.

In other words, an intelligent world.

As one of the world’s leading technology pioneers, spearheading the future of an ever more digital planet, Huawei has always subsisted on a luminous vision of coming innovations. Now, in Huawei Intelligent World 2030, the company cascades a new skein of oracular ideas and prophecies for the global economy.

Huawei’s “Intelligent World” research team met with some 1000 scholars without succumbing to their narrow disciplines. It tapped the ideas of scores of its customers without being captured by their short-term concerns. It consulted its partners without stinting on the contributions of its global rivals. It analyzed reams of data from international organizations, scientific journals and consultancies without losing its way in the mazes of jargon or special expertise. It held over 2000 workshops without falling into the cliches of conventional thought.

Soaring above its sources, this book integrates all narrow disciplines and specializations into a transcendent holistic unity guided by the theory of information.

In the information theory of economics, wealth is most essentially knowledge (the caveman after all had all the material resources we command today). Economic growth is learning, manifested in “learning curves” of collapsing costs throughout all competitive companies.Constraining the processes of learning is time—what remains scarce when all else becomes abundant.

This is the vision of Intelligent World. Here Huawei’s wealth of practical knowledge feeds on continuous learning. Here the limiting cadence of time restricts waste and assures ever growing efficiency and ever diminishing costs for all of us.

In this world, the test of worth is time-prices that measure value by the number of hours and minutes a worker needs to labor to gain the means to buy any particular good or service. Combining the rise in incomes with the decline in costs, time-prices reduce the metric of progress to one universal signature of time.

By this measure, we live in an age of ever-expanding abundance. From an era when people had to spend every waking hour chasing and collecting food just to live, a typical human today earns his food in a matter of minutes. People now can devote their hours to creating new things and new services, like the bonanzas of innovation envisaged in this book.

Gauged by time prices, economic growth has been at least twice as fast as is usually estimated. While the world’s population has risen 75 percent over the last 40 years, the time prices of the commodities of life have dropped 75 percent. Moreover, for the last 40 years, China at a rate of nearly 12 percent a year has massively led all other nations in the speed of its advance.

A key to this enriching cascade has been the prophetic visions of companies such as Huawei, as epitomized by this book.

George Gilder

Futurist, author, and venture capitalist

Wu Hequan

Today, we can clearly see the huge role industrial development has played in driving economic and social progress. However, we are also facing increasingly severe challenges. Population ageing and urbanization have been accompanied by a surge in healthcare demand, labor shortages, increased CO_2 emissions, urban traffic congestion, and environmental pollution. The outbreak and resurgence of COVID-19 over the past two years, and frequent disasters caused by extreme weather events, all highlight the vulnerabilities of today's world.

To address these challenges, the UN has set 17 sustainable development goals (SDGs) for 2030. The achievement of these goals will require scientific and technological innovation as well as supportive policies and regulations. We will need to embrace an intelligent world through digital and intelligent transformation.

Network infrastructure is the foundation of this digital and intelligent transformation process, and also the foundation of the future world that everyone will live in. So what requirements does the future world have for networks?

Internet penetration has reached 52% worldwide, and continues to increase by around three percentage points on average every year. Over the next decade, the number of Internet users will continue growing in developing countries, and by 2030, 70% of people will have access to the Internet. The number of network connections will increase even faster. The number of connected things already exceeds the number of connected people. The number of connections to the Internet of Things (IoT) worldwide has grown at a rate of 10% every year for the past decade. By 2030, the number of IoT connections is expected to be growing more than 10 times faster than the number of Internet users.

According to IDC, the total volume of global data increased by 30% year-on-year from 2016 to 2020. Assuming this rate continues, the world's total data will increase 14-fold over the next decade. Data drives the rollout of broadband. In the past decade, the average bandwidth required by consumer applications has increased more than tenfold. Over the next decade, 10 gigabit access will be available to massive numbers of applications. At the same time, the focus of Internet applications will shift from consumer use to industrial applications. Applications for industrial settings demand low latency, deterministic communications, and high-precision spatial positioning; they must also require high levels of security, reliability, and data protection.

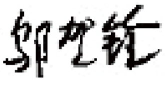
To ensure networks can meet the requirements of the future world, efforts have been made to plan future network technologies, targeting 2030. We are seeing the first fruits of these efforts today. The move to IPv6 will enable sensing applications and in-band signaling, based on which new routing protocols can be developed to support low-latency, high-reliability services. Flexible Internet protocols (IP) can be used to make networks adaptive to the services they carry.

Some new network architecture solutions have been proposed to provide highly secure, green, intelligent, and ubiquitous networks that combine deterministic services, endogenous security, and border security, and can meet the requirements of different applications. Moving forward, next-generation information technologies, such as 5G, cloud computing, IoT, big data, artificial intelligence (AI), and blockchain, will be added to the technology stack. They will enable intent-based services and applications and more efficient use of network resources, and ensure secure and trustworthy cloud-network synergy.

Huawei's Intelligent World 2030 report provides a glimpse of what the world will look like in 2030 in eight domains: healthcare, food, living spaces, transportation, cities, enterprises, energy, and digital trust. Powered by next-generation information technologies, health will be computable; people will enjoy higher quality of life; food production can be driven up by data; cities will become more livable; buildings will become greener; travel will become more convenient than ever; industries will become more intelligent; and trust will become more widespread in society.

The report lists the benefits delivered by some early showcase applications, and includes forecasts from global leaders in the eight domains, providing a clear picture of the world 10 years from now. The report also includes four separate industry reports on communications networks, computing, digital power, and intelligent automotive solutions.

As its title suggests, the report is a just a forecast; many unknowns loom ahead. The projections provided in this report are intended only as a reference for scientists and technology professionals, to help them envision the upcoming intelligent world. This report aims to stimulate innovative thinking and explorations into disruptive technologies to meet the challenges that lie ahead.



Member of Chinese Academy of Engineering (CAE)
Director of the Advisory Committee of Internet Society of China (ISC)
Director of China Standardization Expert Committee (CSEC)

Gao Wen

The past decade has witnessed huge advances in artificial intelligence (AI). Various AI applications have been woven into the fabric of our lives, including clothing, food, living spaces, and transportation. In particular, the major role that AI has played in tackling COVID-19 helped us see its value to all.

China has major advantages that can drive the development of AI: huge data volume, many industrial use cases, and a young and well-educated workforce. But improvements still need to be made in original algorithms, core components, and open-source and open platforms. We need to carefully consider and plan how to make rapid advances in these areas so that we can sustain healthy AI development in China.

The thriving of AI depends on three elements: data, algorithms, and computing power. The progress in AI to date has created a demand for levels of super computing power that far outstrip what Moore's Law can provide. For example, since the emergence of deep learning in 2011, the demand for computing power has been growing exponentially, doubling every 3.4 months.

In 2020, training Generative Pre-trained Transformer 3 (GPT-3), the natural language processing model with 175 billion parameters, consumed 3,640 petaFLOPS-days of computing power. In 2021, Pengcheng Pangu, the industry's first fully open-source Chinese pre-training language model with 200 billion parameters, demanded even more computing power: 25,000 petaFLOPS-days. It took Pengcheng Cloud Brain II, a supercomputer that has computing power measured by exaFLOPS, 50 days to complete the training of Pengcheng Pangu. By 2023, the demand for computing power from hyper-large models will reach millions of petaFLOPS-days, which will pose a severe challenge to existing computers and their computing power.

The development of AI relies on sharing of data, computing power, and ecosystem resources. Areas for exploration in the coming years will be how computing centers can work together for joint model training, while still ensuring data security, and the balance between privacy protection and data mining. Technologies such as federated learning and Trustworthy platform-based AI model generator have offered a solution to these problems. We believe that as these technologies mature, we will be able to use data resources to the fullest extent without undermining data security. Trusted computing platforms will also support further advances in computing networks.

This report gives a comprehensive account of Huawei's research to date. Huawei's insights into the computing industry can offer invaluable experience to those who are interested in the development of this industry.

Looking ahead to the next 5 to 10 years, the use of AI will be spread to more areas, such as AI-led 6G intelligent sensing networks, and Neuromorphic systems that will go beyond existing machine vision imaging systems. Let's work together to embrace a brighter future for the computing industry.

Academician, Chinese Academy of Engineering
Director, Peng Cheng Laboratory
Boya Chair Professor, Peking University

Chen Qingquan

When cars first appeared, they changed the world, making travel more convenient and flexible. However, cars have also been the cause of many woes in urban environments, like air pollution, traffic jams, and unsustainable energy use. This must change. In the next 10 to 20 years, we will see a new automotive revolution.

The automotive industry has been around for more than 100 years. Its "body" remains strong, but its "brain" is weakening. In contrast, the information and communications (ICT) industry has a powerful "brain", but its "body" lacks strength. If we want to ride the tide of the automotive revolution, these two industries must work together to develop a new generation of vehicles. This new revolution will be focused on the user. In essence, it is about moving the automotive industry towards connected, autonomous, shared, and electric (CASE) mobility. Its ultimate purpose is to support safe, comfortable, green, efficient, and at-will transport by making vehicles, roads, and cities more intelligent.

China, for example, has seen its economy shift from high-speed growth to high-quality development, in line with trends associated with the fourth industrial revolution. To cope with this revolution, we need to replace traditional linear thinking with strategic thinking. China's "4N4F" integration model is a prime example of how disruptive circular thinking can replace a linear approach. 4N4F refers to "four networks" and "four flows", where "four networks" refer to the nation's energy network, information network, transportation network, and human network; and "four flows" refer to energy flows, information flows, material flows, and value flows.

The 4N4F integration model will help bring human creativity to the energy revolution, information revolution, and transportation (mobility) revolution. This model is essentially about integrating information technologies, cloud, 5G, AI, and other leading technologies so that data can become useful information, and digital information can be analyzed to output knowledge and intelligence. This model also takes a step further by integrating intelligence into energy to create the intelligent energy sector. This new sector will be able to recycle waste energy while helping us achieve carbon neutrality and delivering increased economic and environmental benefits.

This marks a novel approach to integrating humans, cyber systems, and physical systems (HCPS). These composite systems will produce new kinds of productive arrangements, allowing us to tap into the huge potential of data monetization in the fourth industrial revolution. This revolution will drive exponential growth in productivity, building upon the advancements made in the previous three industrial revolutions. We will soon see a shift from electric vehicles to intelligent vehicles, to intelligent transportation, to smart cities, and ultimately to an intelligent world. During this process, we will be able to create greater added value and more opportunities for value creation.

We are already in the middle of an energy revolution, and have embarked a new journey to a new era. China is coming under huge environmental stress, so it is more desperate than any other country to transition to vehicles that use new energy sources. This is why developing new energy vehicles is a path that China must take. Now is the best moment to integrate the energy and Internet sectors.

Intelligent vehicles powered by new energy represent mobile units of energy production, energy storage, and information. We need to link the automotive revolution, energy revolution, and information revolution, use the 4N4F integration model, and turn energy flows into material flows. This will help us reduce our energy and carbon intensity and hit carbon neutrality targets. The 4N4F model also combines energy technologies with ICT, cloud, edge computing, AI, and big data through a smart energy operating system to deliver greater value. It will drive the energy transition from the fourth to the fifth industrial revolution.

A smart energy sector will need to provide more than just higher security and intelligence. It will need to turn previously useless energy into useful energy. Thermodynamically, this means changing entropy increase into entropy decrease, or turning entropy into exergy. At its core, this process is about turning disorder into order. The automotive industry should also be committed to carbon neutrality if it wants to grow steadily and sustainably.

We will soon enter an era of intelligent vehicles, where green, interconnected, and intelligent will be the way forward for technology in the automotive industry. Innovation will drive further development and transformation in vehicles. The automotive industry will see a transition to electric vehicles and facilitate a mobility revolution for all. This revolution will require more than just stronger research & development and manufacturing. It will also need intelligent networks, where connectivity technologies are key to enabling smart and agile transportation and building the smart cities of the future.

Vehicles are the ultimate products of deep integration between the Internet, big data, AI, and the real economy. Vehicles will be redefined by ICT, and become more connected and intelligent. Intelligent connected vehicles will be a key pillar of intelligent transportation. They will also present challenges in terms of technology, policy, laws, and regulations. The standardization of intelligent vehicles is still underway, and we will continue innovating as we build intelligent connected vehicles. We hope that our future transportation system will be safer and greener and more comfortable and convenient. Let's work together to make it happen.

陳清泉

Founding President of the World Electric Vehicle Association
Academician of Chinese Academy of Engineering
Fellow of Royal Academy of Engineering

David Wang

Exploring the Intelligent World

Technology is developing faster than we could have ever imagined. Digital technologies such as 5G, cloud, and AI are constantly being pushed to their limits, advancing in leaps and bounds. Innovation is no longer confined to a single discipline, but is intertwined with multiple disciplines. Innovation is no longer about a single technology, but about different technology classes. Innovation no longer happens in one industry, but in many industries. The digital, intelligent world is fast approaching.

The Internet currently connects 4.6 billion people and 20 billion devices, and has transformed how people live and work. But that same Internet is also evolving from a consumer-oriented to an industrial Internet. Huawei predicts that by 2030, there will be more than 200 billion global connections, and 10 gigabit connectivity will be available to every enterprise, household, and individual.

The smartphone in your pocket already carries more computing power than the Apollo shuttle that took us to the moon in 1969. By 2030, general computing power (FP32) will reach 3.3 ZFLOPS and AI computing power (FP16) will reach 105 ZFLOPS, a 500-fold increase over 2020. This will mean over 1 yottabyte of data will be generated around the world annually.

As we kick off this new decade, we can already see the outlines of the intelligent world of 2030. By 2030, the combination of ICT technologies and biological data will make health computable, improving healthcare and our quality of life. With big data, AI, and agronomic expertise, we can create scientific food systems that are resilient, sustainable, and "green".

Holographic communications and whole-house intelligent controls will be the key to adaptive, cozy, and friendly homes that understand our needs. The rapid development of new energy and autonomous driving technologies will make vehicles a mobile "third space" outside our home and workplace.

Digital technologies ranging from digital infrastructure, cloud computing, and blockchain will make cities more livable and city governance more efficient. Enterprises from countless industries will use more productive machines, such as collaborative robots, autonomous mobile robots, and digital employees. They will reshape our industrial and commercial models, and make companies more resilient.

An "energy Internet" will emerge, with digital technologies connecting power generation, grids, load management, and storage, systematically reducing carbon emissions, and digitalizing the entire clean energy process. Blockchain, digital watermarking, and privacy-enhancing technologies will lay a solid foundation for a sustainable digital culture.

Huawei has been tireless in our exploration of the future. Over the past three years, Huawei has hosted more than 2,000 seminars, bringing together over 1,000 academics, customers, and partners. Based on these discussions, our experts from around the world have compiled our thoughts on the next decade into this report: Intelligent World 2030.

We lay out eight directions for exploration, spanning multiple disciplines and domains, to answer questions like: How can ICT help people overcome the problems and challenges that we face? Over the next 10 years, what opportunities do organizations and individuals need to be ready for? In terms of our own industry, this report also systematically addresses future technologies and lines of progress in four domains: communications networks, computing, digital power, and intelligent automotive solutions.

More than 30 years ago, we set out to connect every family using phones and enrich life through communications. 10 years ago, we decided to connect our networks to every corner of the globe and build a fully connected world. Now, we are committed to bringing digital to every person, home and organization for a fully connected, intelligent world. We believe that a wonderful intelligent world is fast approaching.

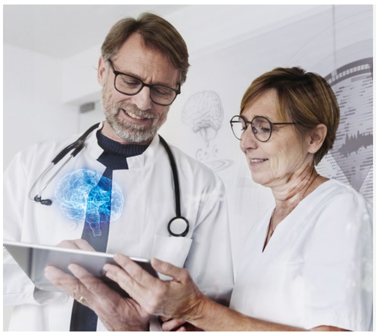
Our imagination is the only limit on how far we can go, but it is the actions we take now that will determine how quickly we can get there. The best way to embrace the future is to build it ourselves. There may be ups and downs on our journey towards the intelligent world of 2030, but we will overcome them by working together and innovating continuously. Together, let's create a better, intelligent world.

David Wang

Executive Director
President of ICT Products & Solutions, Huawei

Contents

Executive Summary 16



Outlook for Healthcare

Making Health Computable, Bettering Quality of Life 22

Outlook for Food

Data-driven Food Production for More Bountiful, Inclusive, and "Green" Diets 36

Outlook for Living Spaces

Personalized Spaces with Novel Interactive Experiences 48

Outlook for Transportation

Smart, Low-carbon Transport Opens up the Mobile Third Space 60

Outlook for Cities

New Digital Infrastructure Makes Cities More Human and Livable

80

102

Outlook for Enterprises

New Productivity, New Production Models, New Resilience

Outlook for Energy

Intelligent, Green Energy for a Better Planet

118

Outlook for Digital Trust

Technologies and Regulations Shape a Trusted Digital Future

140

157 Industry Trends

161 Future Network Use Cases

175 Vision for Future Networks and Their Defining Features

192 Recommendations

193 Appendix A

197 Macro trends

199 Future computing scenarios

211 Vision and key features of Computing 2030

239 Call to action

240 Appendixes

Digital Power
2030

245 The next energy transformation: Low-carbon and sustainable energy

250 Power electronics and digital technologies are a key driver of energy transformation

254 The era of digital power: Digital and intelligent transformation for integrated information and energy flows, and green and low-carbon operations

265 Conclusion

267 Foreword: ICT enables an intelligent automotive industry and helps carmakers build better vehicles

269 Macro trends: Cross-sector collaboration for shared success

276 New scenarios: Bringing digital to every vehicle

281 Technology projection: New components will drive sustained innovation in intelligent vehicles

Intelligent Automotive Solution
2030

313 Data Prediction Methodology

314 Definitions of the Metrics

Executive Summary

We are making strides towards an intelligent world. When looking ahead to 2030, we hope that the future will bring improved quality of life, sustainable and green diets, and more comfortable living spaces. We also look forward to the end of traffic congestion and pollution in cities, fully green energy, and a wide range of new digital services. We dream of robots that can do repetitive and dangerous work for us so that we can devote more time and energy to more valuable, creative work, and to our personal interests. These are the goals that drive exploration in every industry.

Huawei is committed to bringing digital to every person, home and organization for a fully connected, intelligent world. In this report, we examine the prospects for the intelligent world over the next decade by analyzing macro trends in healthcare, food, living spaces, transportation, cities, enterprises, energy, and digital trust.
We believe in the infinite possibilities of the intelligent world, but constant collaboration and exploration among many different industries will be required to build a better future.

Outlook for Healthcare:
Making Health Computable, Bettering Quality of Life

By 2030, sensitive biosensors will be in widespread use, and massive amounts of health data will be stored on the cloud, making health computable. People will be able to proactively manage their health, shifting focus from treatment to prevention. Driven by technologies such as IoT and AI, personalized treatments will become a reality. Portable medical devices will enable people to access coordinated telemedicine services from the comfort of their homes.

Huawei predicts that by 2030

Global general computing power (FP32) will reach 3.3 ZFLOPS, a 10-fold increase over 2020.

AI computing power (FP16) will reach 105 ZFLOPS, a 500-fold increase over 2020.

Outlook for Food:
Data-driven Food Production for More Bountiful, Inclusive, and "Green" Diets

By 2030, we will be producing visualized data graphs, which will make precision farming possible. Collecting data will enable us to control factors affecting crop growth, such as temperature and humidity, so that we can build vertical farms unaffected by the uncertainties of climate and weather. 3D printing technologies are also introducing the possibility of artificial meat designed according to taste and dietary requirements. By 2030, we will be building more resilient and sustainable food systems and relying on firm data rather than the vagaries of the heavens.

Huawei predicts that by 2030

There will be 200 billion connections worldwide.

1YB of data will be generated annually worldwide, a 23-fold increase over 2020.

Outlook for Living Spaces: Personalized Spaces with Novel Interactive Experiences

By 2030, we will no longer have to live with clutter. We will manage our possessions with a digital catalog powered by a 10 gigabit network, holograms, and other technologies. Automatic delivery systems will bring household items from shared warehouses to our doors whenever we need them. Intelligent management systems that control our physical surroundings for automatic interactions will mean that the buildings where we live and work may produce net zero carbon.

Next-generation IoT operating systems will enable people to live and work in adaptive environments that understand their needs.

Huawei predicts that by 2030

- There will be 1.6 billion fiber broadband subscribers.
- 23% of homes will have access to 10 gigabit fiber broadband.

Outlook for Transportation: Smart, Low-carbon Transport Opens up the Mobile Third Space

In 2030, the transport system will see innovations across many different dimensions. Vehicles using green energy and controlled by autonomous driving technology will provide us with a mobile third space. Electric vertical take-off and landing (eVTOL) aircraft will make emergency rescue faster, reduce the costs of delivering emergency medical supplies, and may even change how people commute. Mobility solutions will be efficient, customized, and shared, meaning that vehicles will be used much more consistently and travel will become greener.

All of these will require secure and stable autonomous driving algorithms; cost-effective, reliable sensors; high-speed, stable space-air-ground integrated networks; and a central brain with massive computing power for traffic management. These technologies will be indispensable for developing connected, autonomous, shared, and electric vehicles that deliver a low-carbon transport experience.

Huawei predicts that by 2030

- 50% of new vehicles sold will be electric vehicles.
- Whole-vehicle computing power will exceed 5,000 TOPS.

Outlook for Cities:
New Digital Infrastructure Makes Cities More Human and Livable

The spread of new digital infrastructure will make for better management of the urban environment, with more efficient use of resources and more effective city governance. Centralized digital platforms for government processes and services will make government services user-friendly and easier to access. This will help create more comfortable and livable cities.

Huawei predicts that by 2030

40% of companies will have access to 10 gigabit Wi-Fi networks.

Outlook for Enterprises:
New Productivity, New Production Models, New Resilience

By 2030, digital transformation will have brought a new wave of modernization to enterprises. They will use more productive machines, such as collaborative robots and autonomous mobile robots. New business models will be more people-centric, with increased flexibility in manufacturing, logistics, and other activities. Digitalization will help companies interweave and graphically monitor their supply chains for better resilience in the face of dynamic market environments.

Huawei predicts that by 2030

Every 10,000 workers will work with 390 robots.

5G

One million companies are expected to build their own 5G private networks (including virtual private networks).

Cloud services are forecast to account for 87% of enterprises' application expenditures.

AI

AI computing will account for 7% of a company's total IT investment.

Outlook for Energy: Intelligent, Green Energy for a Better Planet

Energy will be greener and more intelligent in 2030. Power plants will be generating electricity from renewable energy sources in lakes and near-shore marine areas. An "energy Internet" will emerge, with digital technologies connecting generation-grid-load-storage, including virtual power plants and an energy cloud. Zero-carbon data centers and zero-carbon telecom towers could possibly become a reality.

Huawei predicts that by 2030

- Renewables will account for 50% of all electricity generation globally.

Outlook for Digital Trust: Technologies and Regulations Shape a Trusted Digital Future

In 2030, digital trust will be a basic requirement for our social infrastructure. We will need a combination of technical and organizational measures: blockchain, AI fraud detection, and privacy-enhancing computation. This will need to be combined with privacy and security regulations such as the General Data Protection Regulation (GDPR). Used together, these measures will deliver an intelligent world with digital trust.

Huawei predicts that by 2030

- Privacy-enhanced computing technologies will be used in more than 50% of computing scenarios.
- 85% of enterprises will adopt blockchain technology.

Healthcare

Making Health Computable,
Bettering Quality of Life

Over the past decade, the health of humanity as a whole has improved markedly. According to the World Health Organization's (WHO) World Health Statistics 2021, global life expectancy at birth has increased from 66.8 years in 2000 to 73.3 years in 2019[1]. This means that most people can now enjoy long lives, making a better quality of life the new priority.

The pace of population aging is accelerating worldwide. Projections indicate that 16.5% of the global population will be 60 years old or over by 2030[2]. This is expected to drive a surge in demand for healthcare services. According to the WHO's 2019 findings, spending on health is growing faster than the rest of the global economy, accounting for 10% of global gross domestic product (GDP). The WHO also predicts that by 2030, there will be a global shortfall of 5.7 million nurses[3] and 18 million[4] health workers in total.

In addition, chronic diseases and suboptimal health status are increasingly affecting people's quality of life. According to World Health Statistics 2021, seven of the 10 leading causes of deaths in 2019 were non-communicable diseases (NCDs) and global premature NCD mortality reached 17.8%[5].

At the same time, we are seeing wide disparities in the global distribution of medical resources – disparities that become especially clear when viewed in terms of population growth. According to the UN's World Population Prospects 2019, Africa's population is expected to reach 2.49 billion in 2050 while Europe's will be 710 million[6]. However, when we look at the distribution of medical resources, Germany alone has 10.5 times more physicians per 1,000 people than Nigeria, which is one of the more economically vibrant countries in Africa[7].

Looking to the future, new methods of reducing healthcare costs, diversifying healthcare resources and services, and creating new prevention and treatment methods are desperately needed to increase quality of life and make medical treatment more accessible and affordable for all. Many innovative solutions are emerging that may find application within the next ten years.

Direction for exploration: Using computing technologies to identify potential health problems, shifting the focus from treatment to prevention

According to the WHO, 60% of illnesses are caused by lifestyle factors[8], so maintaining good habits is a crucial part of staying healthy. Real-time health monitoring and health data modeling can help us cultivate healthy lifestyles and weave disease prevention into the fabric of our daily lives. This shifts the paradigm governing our healthcare system from treatment to prevention.

Snapshot from the future: Building a knowledge graph to achieve real-time and efficient health management

The growing popularity of wearables and portable monitoring devices combined with advances in technologies such as the Internet, IoT, and AI will make personal health data modeling a realistic prospect in the near future.

Technologies such as big data and IoT will enable doctors to build knowledge graphs based on users' data, including health indicators, medical diagnoses, and treatment results. Doctors can compare and analyze the knowledge graphs to formulate personalized health solutions for users.

In addition, we can take intervention measures which include guidance on nutrition, exercise, and sleep, as well as mental health support to incrementally improve our lifestyles. One digital health company built a knowledge graph to examine relationships between diet and disease.

The company used the knowledge graph to help individuals improve their sleep quality and manage their weight. The health management survey conducted by the company showed that the participants recorded an average 35 minutes more sleep daily, and a total body weight roughly 1.5 kg lighter across the year[9], which translated into lower probability of disease.

We can also combine health knowledge graphs with medical knowledge graphs to predict users' disease risks and future health status, and help them obtain more accurate information about their medical conditions, including information about symptoms, medicine, risk factors, and doctors' diagnoses. Doctors can use this information to achieve more rapid and accurate diagnosis.

Snapshot from the future: Making infectious disease prediction more accurate

New digital technologies, such as natural language processing, can also broaden the amount of data that can be used for epidemiological management. These technologies allow public health institutions to collect and analyze news articles, reports, and search engine indexes to track major public health events around the world. These institutions can extract valid information from the collected data, build new scientific models, and conduct intelligent analyses of the data so that they can respond to incidents faster and more effectively.

ICT technology can also be used to monitor and predict the spread of infectious diseases. For example, a technology company has used natural language processing and machine learning to gather data from hundreds of thousands of public sources, including statements from official public health organizations, digital media outlets, global airline ticketing agencies, as well as livestock health reports and population demographics, to analyze the spread of disease 24 hours a day[10].

Direction for exploration: Precision medicine and treatment plans

A well-designed medical treatment plan is an indispensable part of administering effective, safe, and convenient treatment to patients. When designing a medical treatment plan, doctors have to track and evaluate the patient's symptoms and assess the efficacy of different treatments based on the patient's specific condition. If a patient's clinical symptoms are complex, finding the optimal solution among multiple potential treatment plans can be extremely challenging, but new computing technologies may be the best tool for this challenge.

Snapshot from the future: More accurate drug trials shifting treatment from "one-size-fits-all" to "bespoke"

Some of the factors that must be taken into consideration when designing a medical treatment plan include the physical condition of the patient, appropriate drug categories, timing, dosage, and treatment durations, and possible drug interactions and side effects. These plans, once formulated, also need to be regularly updated based on the treatment progress. All of this puts significant strain on physicians, whose resources are already stretched extremely thin, forcing them to rely on general expertise and experience rather than the patient's specific condition to quickly formulate a general treatment plan. However, AI can help doctors develop personalized treatment plans by analyzing thousands of pathology reports and treatment plans, and determining which would be most appropriate for each patient.

One research institute in Singapore has even created an AI-powered pharmaceutical platform that optimizes medication dosages. The platform can quickly analyze a patient's clinical data, provide the patient with a recommended drug dose or combination regimen based on their specific condition, and revise tumor sizes or biomarkers levels based on available data. In addition, doctors can use the data to determine new courses of treatment for patients[11].

Snapshot from the future: Achieving safe & precise identification of cancer cells with AI

Precision medicine can help the fight against cancer. Estimates showed that there were 19.29 million new cancer cases and 9.96 million cancer deaths[12] worldwide in 2020. Global cancer cases are projected to rise 75%[13] by 2030. Technology can improve cancer diagnosis and treatment, deliver better medical services, and help patients recover more quickly.

During traditional radiation therapy, radiation is directed at the general location of the cancerous tissue to kill cancer cells. Since the targeted area is quite broad, the radiation typically also kills a large number of healthy cells, putting significant strain on the patient's body and resulting in serious side effects.

With the help of AI technology, adaptive radiation therapy (ART) systems can automatically identify changes in lesion positioning and more accurately outline the target areas for radiation treatment. This helps focus the radiation on just the cancer cells and reduces damage to healthy tissue.

AI is already enabling accurate identification and automatic contouring of target areas for various types of medical imaging, including CT, ultrasound, and MRI. AI-based image registration can automatically identify organs and target areas based on clinical needs, making image registration faster and more accurate. With the help of AI, a contouring workload that would once have taken hours can now be completed in less than a minute, and the damage caused by radiation therapy to healthy tissue can be reduced by 30%[14].

Direction for exploration: Home-spital: Using cloud-edge-device synergies to bring healthcare services into patients' homes

The changing role of technology in disease prevention and treatment has increased demand for medical products and personalized medical services that can manage health in real time. The healthcare industry is actively exploring ways to build coordinated telemedicine systems through synergy between cloud computing, edge computing, and on-device computing. Such systems would allow people around the world to access high-quality medical resources anywhere, whether in hospitals, community healthcare centers, or homes.

Snapshot from the future: Cloud-based diagnosis services connecting patients with top experts online

Traditionally, patients have to go to hospitals or specialized treatment facilities for medical examinations and in-person diagnoses. The uneven distribution of medical resources means that top experts are often concentrated in big cities. This often leaves patients in small towns and rural areas without access to accurate diagnostics, resulting in ineffective or delayed treatment.

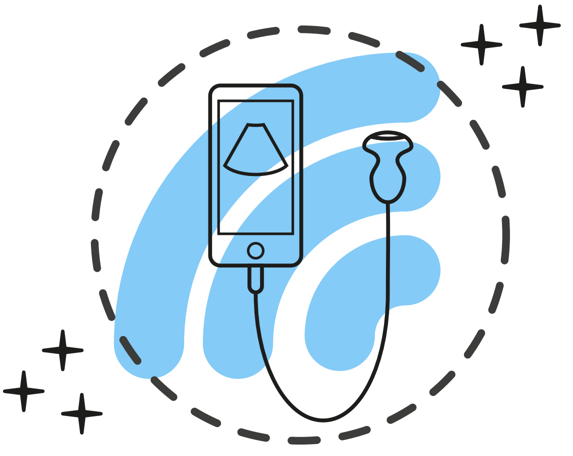
In the near future, the complex processing workloads of imaging devices will be handled in the cloud. Doctors will soon be able to access medical imaging results and AI-assisted diagnostic tools through the cloud. Medical images, examination results, and medical records will be synchronized in real time. This will help doctors provide patients with remote diagnoses and treatment services, wherever they are.

With the application of a "device data collection + 5G + cloud computing" model, medical images can be more easily shared between community hospitals and medical centers. After images are captured on medical devices in community hospitals, they can be automatically or manually uploaded to the cloud. Experts in medical centers can then access the images on the cloud and issue diagnostic reports.

Snapshot from the future: Portable devices that lower the threshold for professional medical examinations

As components get smaller and chip technology advances, large medical devices that previously could only be used in hospitals are now becoming portable, making mobile medical examinations a reality.

Handheld ultrasound scanners are a prime example of this trend. Compact and portable scanners are equipped with the same functions as bulky conventional scanners. This is possible because the functions of conventional ultrasound probes are integrated into a single chip used in portable scanners. In addition, ultrasound information can be collected through smartphone applications. Powered by cloud computing and deep learning, handheld ultrasound scanners can enable powerful functions, such as real-time composite imaging and automatic scanning, that people can use anytime, anywhere. A handheld ultrasound scanner costs only a few thousand US dollars, but has the same functionality of a bulky conventional scanner, on which hospitals often spend more than US$100,000[15].



Conclusion:

Making health computable and bettering quality of life

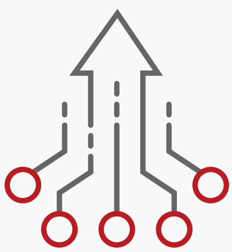
By 2030, we will be able to track our own physical indicators in real time with sensitive biosensor technologies and intelligent devices. We will also be able to build health knowledge graphs to manage our health independently, reducing the reliance on doctors.

Driven by ICT technologies, personalized treatment solutions will become a reality in the future. For example, strong computing power and highly intelligent deep learning systems will be widely used in areas such as precision medicine, adaptive radiation therapy, and rehabilitation robots.

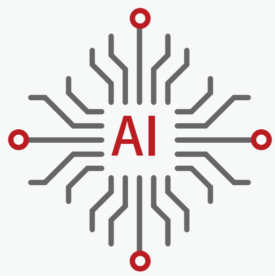
Portable medical devices powered by advanced software and hardware, cloud-edge-device computing, and stable networks will be available in grassroots-level hospitals, communities, and households. These devices will collect medical data in real time and upload the data to the cloud for processing, allowing users to access coordinated telemedicine services and keep track of their health from the comfort of their homes.

By 2030, ICT technology will be enabling access to various applications that will help us stay healthier than ever before. We need a lot of computing power to achieve this. Huawei predicts that by 2030, the global general computing power (FP32) will reach 3.3 ZFLOPS, a 10-fold increase over 2020. AI computing power (FP16) will reach 105 ZFLOPS, a 500-fold increase over 2020.

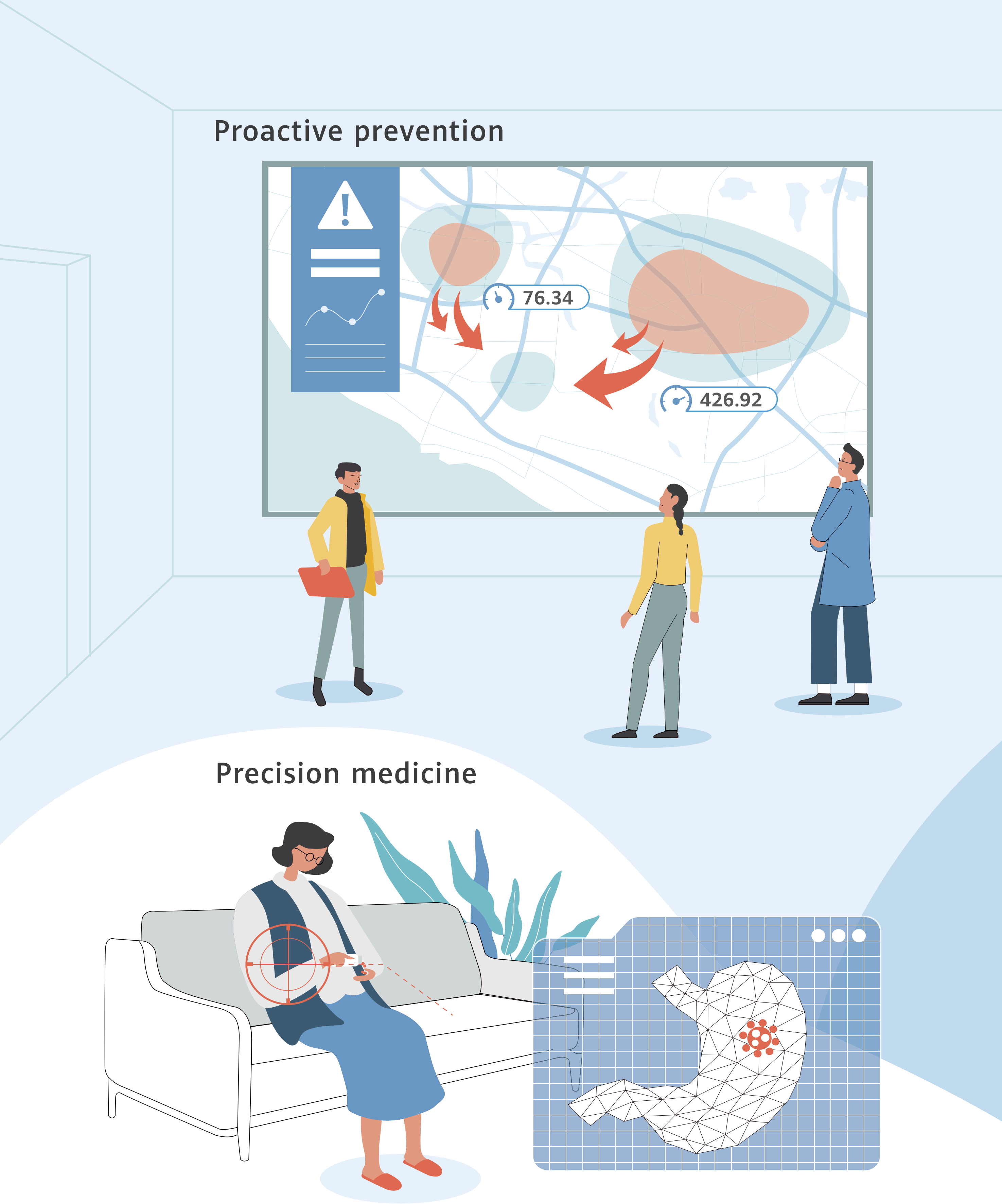
Huawei predicts that by 2030,

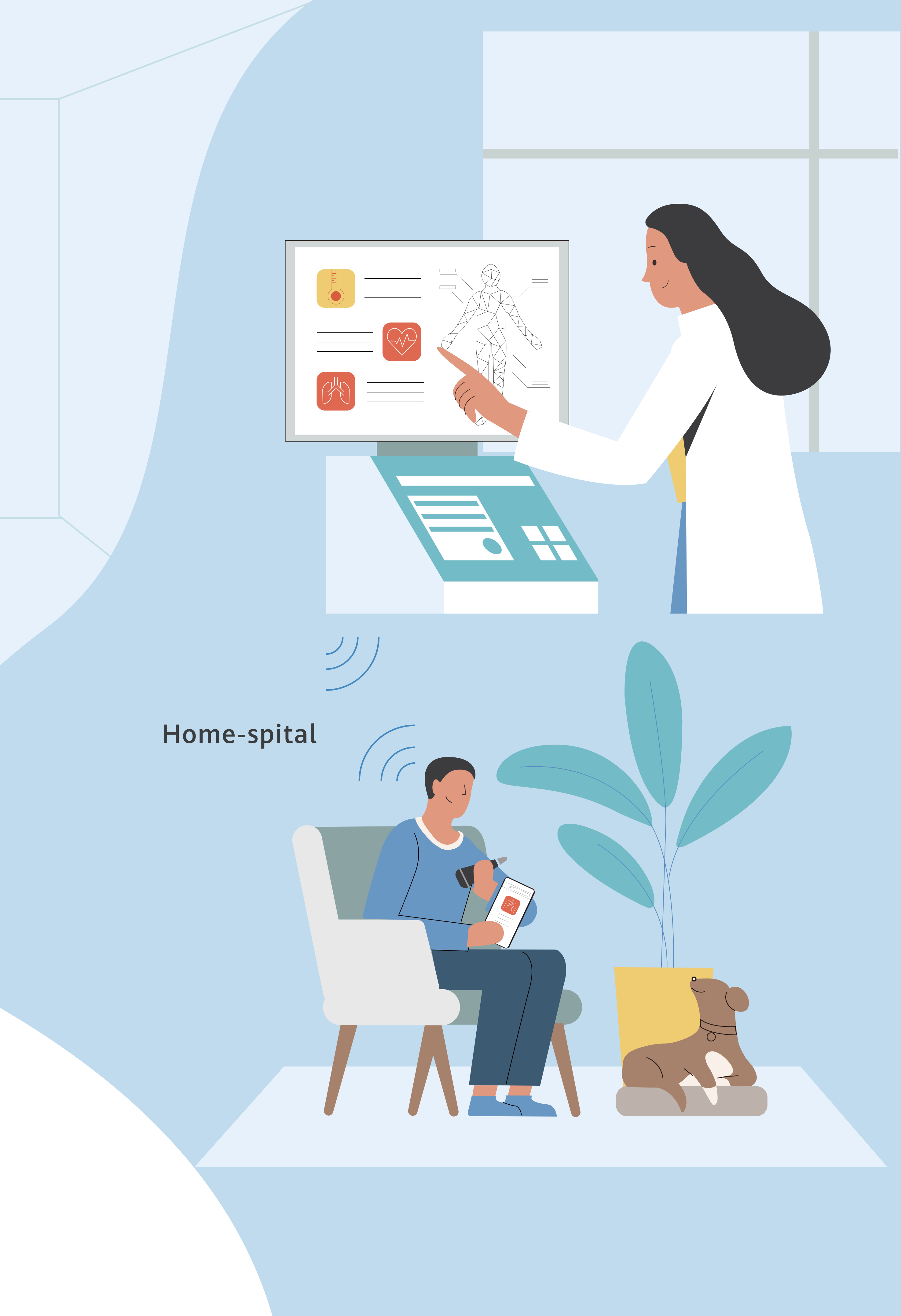


Global general computing power (FP32) will reach **3.3** ZFLOPS, a **10**-fold increase over 2020.



AI computing power (FP16) will reach **105** ZFLOPS, a **500**-fold increase over 2020.


Proactive prevention
76.34
426.92
Precision medicine


Home-spital

Food

Data-driven Food Production for More Bountiful, Inclusive, and "Green" Diets

Food is a necessity for all, so the UN has made "Zero Hunger" one of its Sustainable Development Goals (SDGs) for 2030. Current estimates show that nearly 690 million people are hungry, and if recent trends continue, the number of people affected by hunger would surpass 840 million by 2030[1].

In the past, food shortages were addressed by focusing on resource insufficiencies, such as poor climate and soil conditions, or labor shortages. However, environmental change and accelerating urbanization mean traditional agricultural technologies and abundant natural resources are no longer enough if we want to achieve the UN's 2030 SDGs[2].

The agriculture workforce is shrinking: According to the International Labor Organization, the proportion of the world population working in agriculture has dropped from 43.699% in 1991 to 26.757% in 2019[3].

Arable land per capita is decreasing: According to World Bank data, arable land per capita has fallen from 0.323 hectares to 0.184 hectares from 1968 to 2018 – a drop of 43%[4].

Overuse of pesticides is causing severe soil pollution: According to statistics, 64% of global agricultural land (approximately 24.5 million square kilometers) is at risk of pesticide pollution, and 31% is at high risk[5].

Simultaneously, the focus of people's diets around the world is shifting from "Does this taste good?" to "Is this good for me?" This has resulted in more nutrition and food safety standards. For example, 13,316 food products in China received some kind of green certification in 2018. This number increased to 14,699 in 2019, up 10.4% YoY[6].

This higher demand for green-certified products results in higher requirements on agricultural conditions and technologies.

As we move towards 2030, our food supply faces new challenges and demands. We can see that technology is key to empowering agriculture, helping it overcome traditional growth constraints, increasing food production across the board, and bringing "green" food to every table around the world.

Direction for exploration: Using accurate data, not experience, to guide cultivation

As the saying goes, there is "a time to plant and a time to uproot". Farmers normally rely on calendars to determine the best times to sow. In addition, they also rely heavily on personal experience to determine when to sow, fertilize, and use pesticides. However, this leaves a lot of uncertainty, and whether any given year yields a good harvest is still ultimately up to fate.

Snapshot from the future: Precision farming based on visualized data graphs

Even within a single crop field, the moisture content, available nutrients, and crop conditions can vary. But with modern tools like sensors and mobile devices, farmers can remotely and accurately monitor soil moisture, ambient temperature, and crop conditions in real-time. This makes it possible to flexibly adjust agronomic measures, like sowing, irrigation, fertilizer, and seed adjustment, based on diverse data sets, to better align crops with the soil available. Take maize, for example. Data-powered adaptive sowing can increase crop yield by 300 to 600 kilograms per hectare of land[7].

Precision agriculture also relies on in-depth analysis of collected data and cloud-based visualized data graphs. Data from these graphs help farmers make informed decisions regarding soil fertility, water, and nutrient delivery across key crop growth stages. These graphs can also help farmers better understand information such as local topographical characteristics, climate conditions, and crop diseases or pests so that they can better estimate crop yields, implement agricultural measures, and adjust budgets accordingly.

Visualized data graphs can also be used to monitor and manage agricultural production in real time, helping farmers make proactive, quick, and precise responses to changes in their environment. For example, in the event of extreme weathers, farmers can use this data to rapidly locate affected areas, develop solutions, and mitigate negative impacts on their yields.

Direction for exploration: Taking a "factory-like" approach to protect agricultural production from environmental conditions

While precision agriculture is an effective tool for increasing agricultural yields, it is not sufficient for meeting the surging food demands caused by population booms, shrinking arable land per capita, pesticide pollution, and worsening climate change.

Under precision agriculture, data is used for analysis and calculation so that the best cultivation solution can be found. However, the environment is constantly changing, so data can only be used at the moment it is collected. Therefore, the results of agricultural data analysis cannot be used iteratively.

In addition to precision agriculture, we can also take a more "factory-like" approach to agriculture, creating "vertical farms" in enclosed environments. Not only can vertical farms be used to collect more data, they can also allow farmers to directly adjust parameters to allow crops to grow in an optimal environment. Both countries with little arable land, like Japan, South Korea, and Singapore, and countries with abundant land resources, like the US, are proactively developing vertical farm technologies.

Snapshot from the future: A new form of agriculture in intelligent vertical farms

One typical example of industrialized agriculture is indoor, vertical farms that use data to build standardized growth environments without needing to consider geographical constraints. In vertical farms, every step of the cultivation process, from sowing, to fertilizing, to harvesting, is closely monitored, with farmers precisely controlling light, temperature, water, and nutrient delivery based on the needs of each crop. By controlling every stage of crop growth and adjusting environmental parameters as needed, farmers are able to artificially create the ideal environment for their plants.

Vertical farms have three main advantages:

◎ **They don't need pesticides or soil, and reduce agricultural water waste:** Hydroponics and aeroponics, which are common in vertical farms, utilize solutions that are more efficient at delivering nutrients to plants, with any remaining nutrients being recaptured together with water. These methods use less than one tenth of the water used in traditional agriculture, making the entire process more eco-friendly and reducing pollution.

◎ **They are not affected by climate, providing consistent and ideal conditions for fresh produce:** As vertical farms create closed environments, automatic control systems can be used to ensure reliable, large-scale cultivation. This makes it possible to grow vegetables in a wider variety of locations and climates. Vertical farms can be built on rooftops, in office buildings, abandoned workshops, and basements, or in deserts, rivers, or seas.

◎ **They provide smart agricultural models that are globally replicable:** The ICT control system and data model used in one vertical farm can be used anywhere in the world to achieve almost the same results. The vertical farm model allows anyone to emulate the environment that grew the best wine-making grapes in history, and even arid regions that see little daylight can be used to grow warmth-loving cherries.

Recent pilot programs for vertical farms have found that, if harvested every 16 days, a 7,000 square meters area can yield a staggering 900,000 kilograms of vegetables every year[8].

Direction for exploration: Low-carbon, 3D-printed solutions for meat

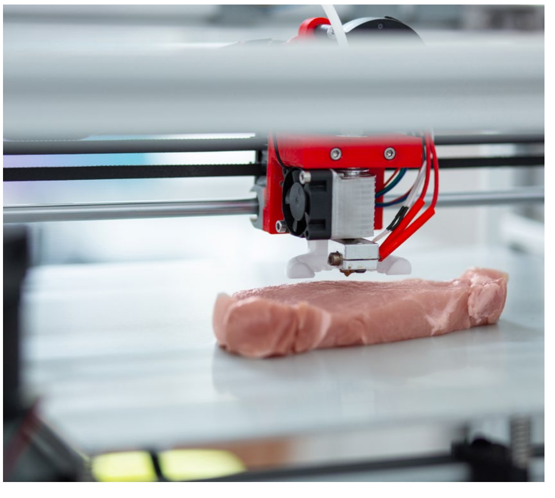
In addition to grains and vegetables, meat is an integral part of most dietary traditions. Traditional animal husbandry though is not only inefficient and unable to meet demand, it also has negative environmental impacts. It is estimated that animal husbandry produces about 7.1 billion tons of CO_2e every year, accounting for 14.5%[9] of total anthropogenic greenhouse gas emissions. Methane, a volatile and powerful greenhouse gas, is the second-largest contributor to climate change. Moreover, the majority of methane emissions caused by human activity come from livestock. There are roughly 1 billion cattle worldwide[10].

According to the UN's Food and Agriculture Organization (FAO), the global demand for meat is projected to increase 70%[11] by 2050 as the global population grows. This makes meat production an area in critical need of proactive solutions. The food industry as a whole has already shown significant interest in artificial meat.

Snapshot from the future: Healthy and sustainable meat supply with 3D printing

Recent applications of 3D printing technologies have successfully improved the quality of artificial meat, making it better tasting and better looking to consumers. It has proven capable of turning both plant proteins and animal cells into artificial meat.

3D printing can use plant-based photoproteins to build fibrous skeletons that closely mimic the texture of real meat. Alternatively, 3D printing can also be used to stack nutrient elements made of real animal cells to create the musculature and fat layers normally seen in animal tissue. Currently, 3D printing can be used to create many types of artificial meat, including pork, chicken, and beef, with the price of artificial beef quickly approaching the market price of real beef.



Conclusion:

Turning data into food to solve global hunger

In the future, we can use the Internet of Things (IoT) to monitor and analyze soil conditions and crop growth, and to increase yields based on collected data. We can also use historical data to predict changes in the natural environment so as to take proactive intervention measures that reduce the risk of yield reduction. Science-based precision agriculture systems powered by big data, AI, and agricultural know-how would make it possible for farmers to precisely water and fertilize crops and use drones to apply pesticides more accurately.

Intelligent farming models, such as data-based vertical farms will free agricultural production from the constraints of climate dynamics. These models can be replicated worldwide for more inclusive and green diets.

3D printing technologies are also making it possible for people to create artificial meat using their data models based on taste and dietary requirements. By 2030, ICT technology will enable us to connect key agricultural production factors, such as farmland, farm tools, and crops, and to collect and utilize data on climate, soil, crops, etc., to increase yield. Huawei predicts that by 2030, the data generated worldwide will reach 1 YB every year, a 23-fold increase over 2020. There will be 200 billion connections worldwide, and IPv6 adoption will reach 90%. With the wider application of data in agriculture, we will build a more resilient and green food system.

Huawei predicts that by 2030,

There will be **200 billion** connections worldwide.

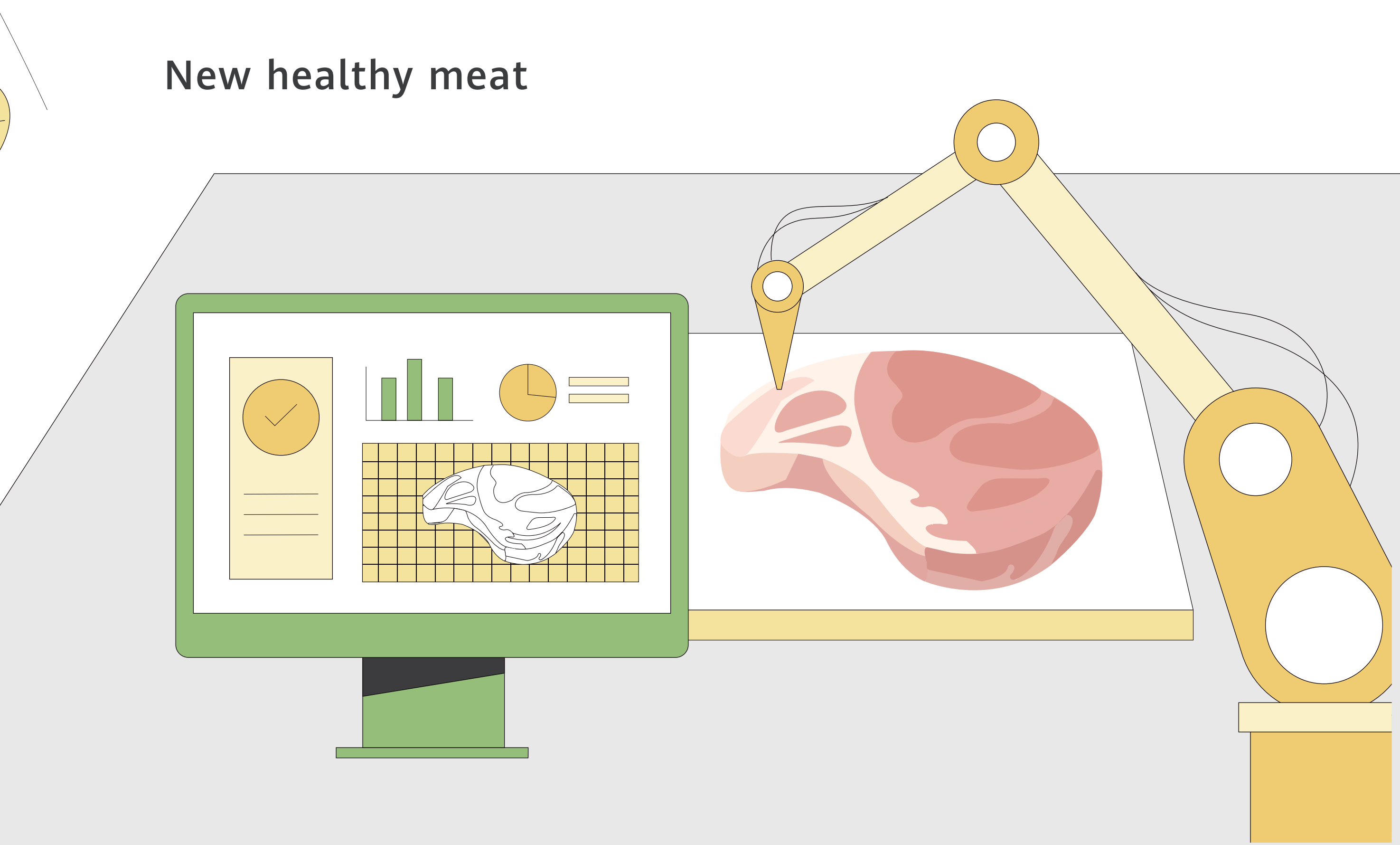
IPv6 adoption will reach **90%**.

1YB of data will be generated annually worldwide, a **23**–fold increase over 2020.

Precision farming

"Factory-like" agricultural production

New healthy meat



Living Spaces

Personalized Spaces with Novel Interactive Experiences

The spaces people live in have dramatically changed throughout history. Over time, ancient caves have been replaced by modern buildings, and people have moved from rural communities to dense urban districts. The role of living spaces has also transformed. What used to simply be a shelter from the elements has become the vault for our most precious possessions.

Industrial advancements have led to an abundance of material wealth, and while material possessions bring us joy, they have also filled our living spaces to the brim. The average American home has more than 300,000 items[1], and 1 out of every 10 Americans rents offsite storage[2]. Similarly, a British study found that the average 10-year-old child owns 238 toys but plays with just 12 on a daily basis[3]. While many find buying new objects satisfying, the downsides of hyper-materialism are becoming increasingly apparent to society as a whole. This dilemma opens up new possibilities for future home design.

It has been reported that in 2019, CO_2 emissions from the operation of buildings worldwide reached 10 $GtCO_2$, or 28% of total global energy-related CO_2 emissions[4]. Moreover, the number of new buildings all over the world is set to explode. The International Energy Agency estimates that global building stock will rise by 5.5 billion square meters per year on average until 2050[5]. The UN Environment Programme's Global Status Report 2017 predicts that the world will add 230 billion square meters (2.5 trillion square feet) in new construction by 2060, which is almost equal to current global building stock. This means an amount of construction equal the size of New York City will be added to the planet every 34 days over the next 40 years[6]. In response, the World Green Building Council has stated that, if we want to meet the goals of the Paris Agreement, all new buildings should operate at net zero carbon by 2030 and all buildings should achieve this by 2050[7].

As demand for personalized home experiences continues to rise, ICT-enabled smart home technology is gaining popularity. A survey found that about 80% of millennials and 69.2% of baby boomers are interested in smart home technologies[8]. In the UK, 80% of consumers are now aware of smart home technology and it is second only to mobile payments in consumer awareness of a basket of tech trends.

Interoperability has risen as one of the most important buying considerations[9]. Interest in smart living spaces that offer enhanced convenience and safety is also on the rise.

Direction for exploration: New infrastructure provides comprehensive services for communities

Smart doors, smart smoke detectors, falling object alerts, delivery notifications, and many other smart services are becoming increasingly widespread. This means that residents are much more closely connected with their communities and local authorities. In the future, new communities will deliver comprehensive services to residents, powered by the Internet of Things (IoT), Gigabit fiber networks, and other new advanced infrastructure. Services such as virtual community events and smart pet management will bring residents and their communities more closely together. Groundbreaking new design concepts will also start changing the way our homes look at the household level.

Snapshot from the future: Digital cataloguing and automated delivery for offsite storage

One potential solution to the overwhelming amount of possessions that now fill households is offsite storage. Some proposed solutions include digitalization and cataloguing of all household items, with technologies like 3D scanning, and then storage in local shared warehouses. This would mean when you decide to go to a party, you can flick through a 3D hologram menu to pick out the dress and accessories that you want, and, at the touch of a button, have those items delivered to your door, either by robot[10] or through the building's internal delivery system[11].

A similar system could power a shared "library" of useful household items. For example, in a typical household, electric drills are not needed often, so instead of buying your own, you could search for one in the shared library's online resource catalogue and borrow it for a few days.

Automated delivery systems will bring the drill to you and take it back when you are finished.

Direction for exploration: Net-zero-carbon buildings with IoT and intelligent management systems

According to the World Green Building Council, a net zero carbon building is "a highly energy efficient building that is fully powered from on-site and/or off-site renewable energy sources and offsets." Net zero carbon is achieved when the amount of carbon dioxide emissions released on an annual basis is zero or negative. Minimizing building energy consumption through new designs and eco-friendly materials is the first step to achieving net zero carbon.[12] The next step requires not only clean energy sources, but also information and communications technologies (ICT).

One day, net-zero-carbon buildings will be able to automatically interact with their environment through sensors.

- Sensors monitor and generate data about the building in real time, including its environment and condition.
- The Internet of Things connects sensors, cloud-based control systems, and core systems such as lighting, electricity meters, water meters/pumps, heaters, fire alarm systems, and water chillers.
- Intelligent, cloud-based systems utilize sophisticated algorithms and real-time data to automatically decide how the building can minimize energy use. For example, a complete automated system could use IoT devices to check the number of people in a building in real time, and then decide when to switch air conditioners and lights on or off in different parts of the building. Such a system would also be able to manage elevators, hallways, and shutters, depending on actual human activity.

In addition to the environmental benefits, net-zero-carbon buildings will also make people's lives more comfortable. Automated systems can keep indoor temperatures at agreeable levels, while soundproofing materials can keep outside noise down to a minimum. There will also be health benefits. Automated systems can decide how much sunlight should pass through a window, to help limit UV exposure, encourage natural sources of vitamin D, support regular sleeping patterns, and combat seasonal affective disorder (SAD).

Snapshot from the future: Automated building management systems for museums

Some key museums are already upgrading their energy systems with automated controls. For example, one museum in Australia installed a building management system that constantly monitors 3,000 different indoor environment data points, and automatically adjusts utilities to provide the right conditions for visitors and the objects on display. Heating, ventilation, air conditioning, lighting, and water efficiency have all been upgraded, reducing the museum's GHG emissions by 35%, and electricity costs by 32%[13].

Direction for exploration: Adaptive home environments that understand your needs

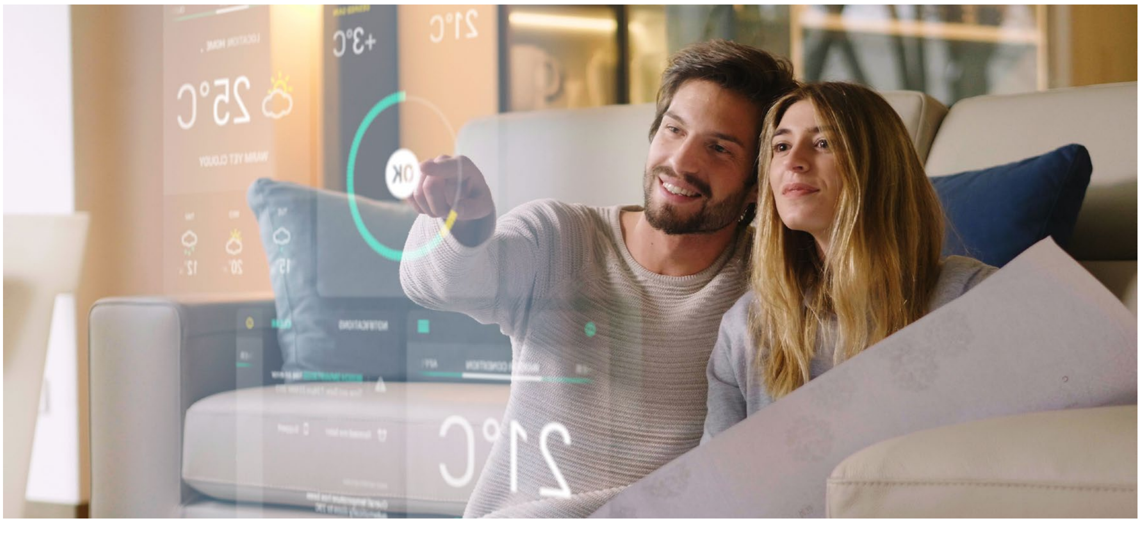
Today, we expect more from our homes than ever before. Homes should be more than just places to live: They should also offer superb experiences. The "homes of the future" will intuitively understand all of our needs. The moment we arrive home after a long exhausting day, the lights, sound systems, air filters, and television will switch on automatically. When we walk into the kitchen, the refrigerator will push healthy meal suggestions adapted to our personalized dietary needs. In the bedroom, air conditioners will check air quality and automatically adjust temperature and humidity based on what we are doing.

There are perceptible and imperceptible factors that determine how comfortable our homes are. Perceptible factors are those we can instinctively feel, such as temperature, humidity, lighting, and ease of access to household items. Imperceptible factors usually include indoor air quality and safety. Both types of factors can be controlled in real time by intelligent automated systems.

Snapshot from the future: Whole-house intelligence that understands usage and creates intuitive experiences

Smart home systems collect data from a wide range of smart home appliances and sensors, over highly-reliable, high-speed networks that reach every corner of your home. They use AI engines to determine what is happening and run appropriate applications. The AI engines, in turn, need distributed processing and computing to understand your behavior, indoor environment, and hardware systems, and then make smart decisions to configure your home appliances. These steps could be taken independently or in collaboration with other systems, to meet your needs. When implemented properly, smart home systems deliver immersive, personalized, and intelligent experiences that evolve as your usage needs change. In the future, the way we interact with home appliances will also change through touch panels, apps, voice commands, and gestures. Sometimes interactions will be so subtle that we won't even be aware of them.

The variety of smart home appliances we will see in the coming years is expected to explode. They will work together to intelligently anticipate and meet your needs in different situations. Everything, from smart beds and pillows to lights and audio devices, will be able to collaborate. A sleep support solution could easily be created for the bedroom by designing a system that automatically adjusts the softness of your mattress and pillow to suit your body and sleeping habits, and changes your bedroom lighting to stimulate the production of melatonin – the hormone that helps you fall asleep. Bedroom speakers could play music to relax you, and air conditioners could keep track of temperature, humidity, and oxygen levels. Such a system could even identify snoring and curb it by rapidly adjusting the softness of your mattress and pillow. Temperature and humidity regulation could also be achieved to stop you from tossing and turning in bed[14].



Conclusion:

Personalized spaces with novel interactive experiences

In the future, new intelligent infrastructure will play an integral role in smart communities. With the help of ICT technologies, community management systems will aggregate massive amounts of data generated by smart equipment, and use that data to holistically manage how communities operate in real time, delivering superior services to residents.

Net-zero-carbon buildings will be made possible by eco-friendly designs and clean energy sources. Passive design for energy conservation cannot achieve zero carbon goals alone, and energy management systems can contribute to these efforts by effectively managing energy sources and accurately controlling indoor environments to minimize energy consumption.

5G and Artificial Intelligence of Things (AIoT) will help smart home systems autonomously adapt to user needs. These systems will rely on superfast network connections and sophisticated algorithms that enable them to promptly sense user needs and provide intuitive services.

In the future, ICT technologies, especially sensors, IoT, and AI, will reshape our living spaces, from communities to buildings to homes. The result will be personalized living spaces that offer all kinds of new interactive experiences: a space that knows you as well as you know it.

At that time, your home may be full of smart appliances that bring a new level of interactivity to your lifestyle and entertainment. The building you live in may be supported by a great variety of smart control systems, and smart functions may be more widely available in your local community. However, none of this will be possible without connections that deliver high bandwidth and extremely low latency. Huawei predicts that by 2030, there will be 1.6 billion fiber broadband subscribers and 23% of homes will have access to 10 gigabit fiber broadband.

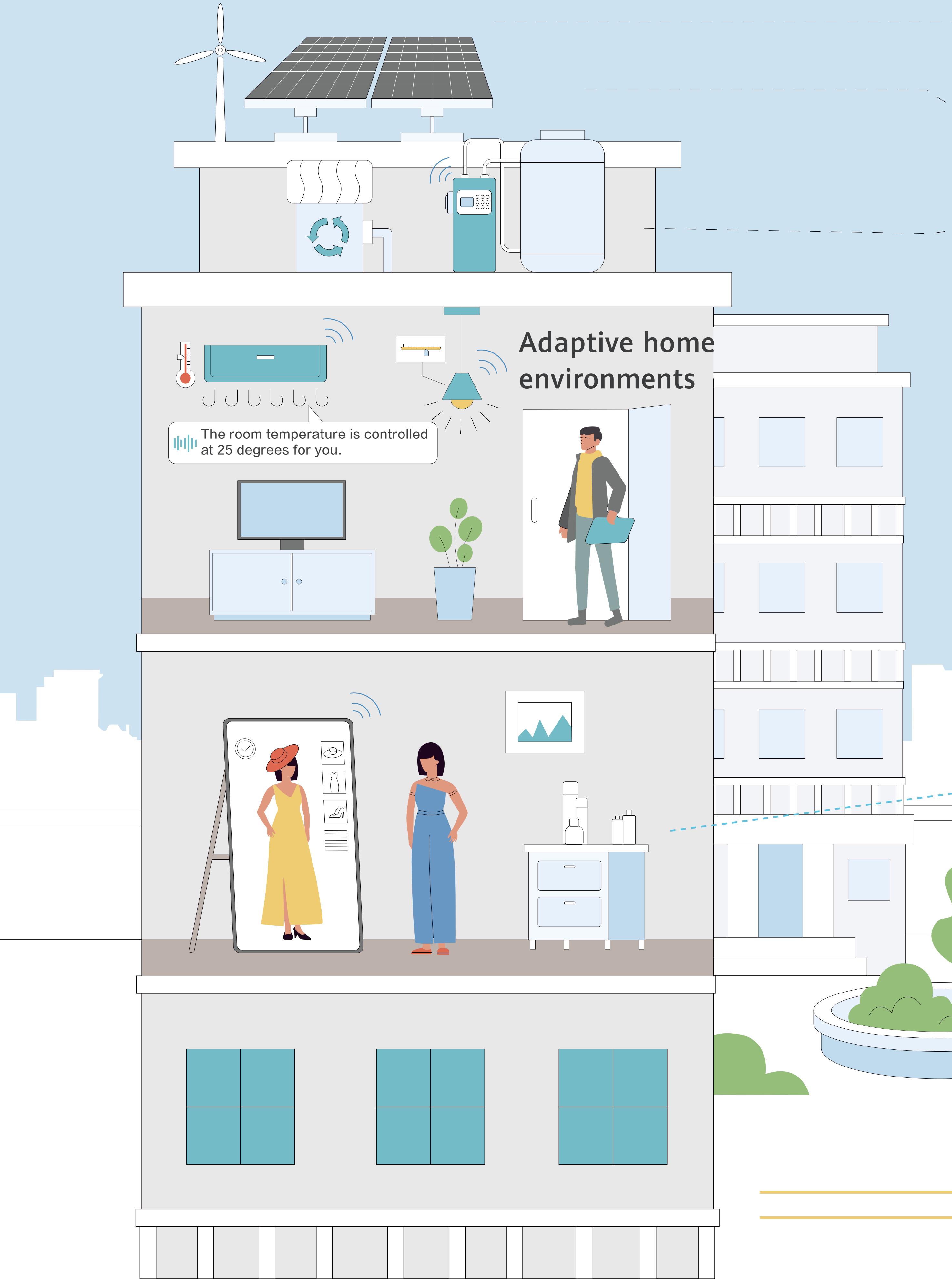
Huawei predicts that by 2030,

There will be

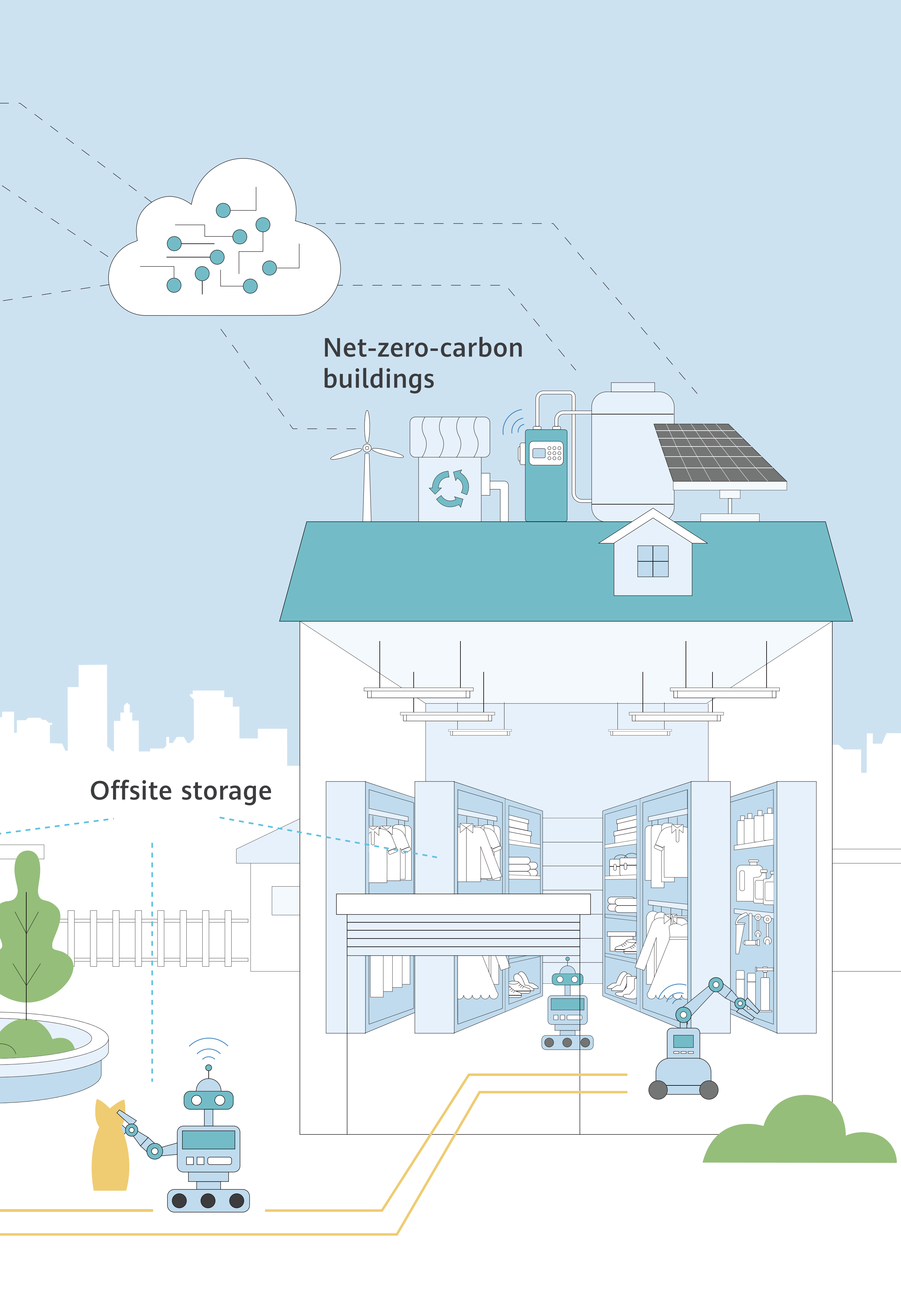
1.6 billion

fiber broadband subscribers.

23% of homes will have access to 10 gigabit fiber broadband.



Adaptive home environments
The room temperature is controlled at 25 degrees for you.


Net-zero-carbon buildings
Offsite storage


71
24°C
Autonomous Mode
21 °C

Transportation

Smart, Low-carbon Transport Opens up the Mobile Third Space

Travel is an important part of modern life, taking many forms, but particularly private cars. The US has long been known as a "nation on wheels", and in 2020, the vehicle-miles traveled across the US totaled 2.83 trillion miles[1], which is more than 30,000 times the distance between the Earth and the sun. In Europe, vehicles travel more than 12,000 kilometers a year on average[2]. As urban areas keep expanding, more and more commuters face the challenge of a long daily commute. In China, more than 10 million people commute for over 60 minutes per day[3]. And the trend is continuing: Car-based mobility, in terms of passenger kilometers and predicted vehicle stock, is expected to increase by 70% globally by 2030[4]. Clearly, existing transportation systems will continue to face many challenges.

More congestion, lower efficiency: Traffic jams are becoming increasingly serious and frequent worldwide. Studies have found that the average American commuter wastes 54 extra hours a year in traffic delays[5]. In Colombia's capital Bogota, the world's most congested city, drivers wasted an average of 272 hours in traffic jams during 2018[6], equal to more than 11 days. Congestion also causes significant economic losses. For instance, congestion caused losses of US$88 billion in the US in 2019[7].

Huge demand for travel also poses more serious challenges to the environment: According to the International Energy Agency (IEA), the transportation industry contributed 26% of global carbon emissions in 2020, far exceeding the manufacturing and construction industries. The adoption of circular economy practices combined with accelerated electrification in the automotive industry has the potential to reduce carbon emissions by up to 75% and resource consumption by up to 80% per passenger kilometer by 2030[8], accelerating low-carbon development.

Future transport will not only be congestion-free and low-carbon, but also hassle-free. Travelers will no longer need to think about their route, only about what to do during the journey. They will enjoy a quiet and personal space, where they can watch TV, concentrate on work, or even get an in-seat massage to relax. Commuters will no longer feel exhausted at the end of the journey, instead feeling a sense of satisfaction for making the most of their time.

In order to realize these changes, transportation networks will need to be further upgraded. All of the key elements (vehicles, traffic lights, pedestrians, etc.) need to be connected using ICT technologies so that each phase of a journey can be automated.

Direction for exploration: Electric vehicles (EVs) for green transport

As transportation consumes increasing amounts of energy, many countries and regions around the world are making efforts to develop low-carbon travel. Energy saving and emission reductions in transportation have become the key to carbon neutrality. In July 2021, the European Commission officially launched the European Green Deal, specifying the goal of reducing greenhouse gas emissions by 55% by 2030 compared with 1990 levels, and achieving carbon neutrality by 2050. The Deal also proposes to reduce carbon emissions from transportation by 90% by 2050 compared with 2021 levels. Land transport is the focus of this policy as it contributes 20.4% of the EU's greenhouse gas emissions.

To save energy and cut emissions from land transport, countries are vigorously developing vehicles that use renewable energy, including pure EVs, plug-in hybrid EVs, and fuel cell vehicles. Many countries and regions have set timetables for reducing combustible fuel vehicles. For instance, the EU has set the goals of reducing emissions from passenger cars and vans by 55% and 50% respectively by 2030 (previous goals: 37.5% and 31%). It has also set a new goal of ensuring that by 2035, all new vehicles sold will be zero-emission vehicles, which is equivalent to banning the sale of combustion-engine vehicles from 2035[9]. Japan plans to increase the share of EVs in domestic sales from 50% to 70% by 2030[10]. China has suggested that traditional fuel vehicles will be phased out of its market from 2030[11].

Snapshot from the future: Renewable energy boosts green mobility

The adoption of new energy vehicles in urban public transportation, including buses and taxis, started early and is now well-advanced in many cities. For example, every one of Shenzhen's 16,000 buses were electric by 2017, making it the world's first city with a 100% electric bus fleet[12]. In Europe, EVs make up over 78% of Denmark's new buses, and about two-thirds of new buses are emission-free in Luxembourg and the Netherlands[13].

There are two reasons for the fast progress in public transportation. First, public transport vehicles are replaced at a relatively fast rate,

providing the opportunity to plan and quickly implement the adoption of EVs. Government subsidies plus efficient O&M solutions can also reduce the operating costs of electric fleets to levels close to or even lower than conventional vehicles. These factors reduce the obstacles to introducing new energy vehicles.

Second, these publicly-owned vehicles are centrally stored and maintained in specialized facilities that can easily be upgraded into multi-functional spaces with charging piles for EVs. Therefore, the charging problem is not a major obstacle for the electrification of public transportation.

Public transport vehicles also travel longer distances every day and generate more carbon emissions than private vehicles. Therefore, the wide adoption of electric public transport is a highly efficient way of reducing vehicle emissions. In Beijing, for example, 71,000 private pure EVs saved 89 million liters of gasoline and 199,000 tons of CO_2 emissions in 2018. By contrast, just 9,400 electric taxis in Beijing made a similar contribution, saving 65 million liters of gasoline and 145,000 tons of CO_2 emissions[14].

According to the IEA, although the global automotive market shrank by 16% in 2020, following the COVID-19 pandemic, the number of newly registered EVs hit a new high of 3 million, up 41% from the previous year. The number of EVs in use worldwide in 2020 exceeded 10 million, and that strong momentum has continued. The sales of EVs in the first quarter of 2021 were almost 1.5 times higher than Q1 2020. Moreover, consumer spending on EVs grew by 50% in 2020, reaching US$120 billion, while government subsidies were just US$14 billion. Government subsidies as a percentage of the total spending on EVs have declined for five years in a row. This indicates that although government subsidies can stimulate the market, sales of EVs are increasingly driven by consumers' own choices.

The number of electric cars, vans, heavy trucks, and buses on the road worldwide is expected to reach 145 million by 2030. If governments accelerate efforts to achieve global climate and energy goals, then the global EV fleet will reach 230 million vehicles by 2030[15]. The IEA also predicts that more than 300 million EVs will be put into service by 2040, reducing oil consumption by 3 million barrels per day[16].

Snapshot from the future: Clean energy aircraft trials

The aviation industry is actively developing clean energy aircraft, both small local planes and larger passenger aircraft, in an effort to protect the environment, reduce pollution, and cut O&M costs. The continuous development of urban air mobility (UAM) is also driving the aviation industry to increasingly use electric power, setting it on a greener path.

In terms of carbon emissions, the aviation industry contributed about 2% of global anthropogenic CO_2 emissions in 2019. If such emissions are not effectively curtailed, this percentage is expected to increase to 25% by the middle of this century[17].

In terms of O&M and fuel expenses, global maintenance, repair, and operations (MRO) costs were US$69 billion in 2018, accounting for 9% of airlines' total operating costs. Maintenance costs for engines accounted for 42% (US$29 billion) of total MRO costs. Fuel costs for the global aviation industry were US$188 billion in 2019, accounting for 23.7% of total operating expenses. It is expected that MRO costs will grow to US$103 billion in 2028[18].

At present, three main types of clean energy aircraft are being developed: hybrid-electric, pure electric, and hydrogen-powered. In addition to improvements in increasing energy efficiency and reducing pollution and noise, clean energy aircraft also represent an opportunity to try new designs, such as blended wing body aircraft. This design can significantly reduce aircraft's drag and energy consumption, and improve flight performance. In addition, this design can increase the amount of space in the aircraft cabin, which is very valuable as it increases the aircraft's carrying capacity.

In June 2020, France announced an investment of EUR15 billion that included EUR1.5 billion for the development of large clean energy passenger aircraft. France plans to complete the maiden flight of the aircraft by 2035, and more than 1,300 companies in the industry will participate in the plan. For this, France has drawn up a clear roadmap which starts with a revamp of Airbus's A320 product line to develop a hybrid electric "successor" model to the A320. The prototype of this new model will be unveiled between 2026 and 2028, and make its maiden flight by 2035.

Direction for exploration: Autonomy opens up the mobile third space

From horse-drawn carts to modern cars, humans have developed many generations of vehicles to help us move faster than we could ever run. However, the upcoming autonomous driving era means that humanity will soon equip vehicles with their own brains. As an important factor for the way we travel, autonomous driving technology will reshape our travel experience, and thoroughly transform business models in the transportation industry.

As artificial intelligence replaces the human brain to make decisions for vehicles, drivers' hands, feet, and eyes will be freed during travel. This will enable new mobile entertainment, social engagement, shopping, and work scenarios on the road, turning vehicles into our mobile third space.

Autonomous driving technology is classified into six levels, L0–L5, by the standards of the US Society of Automotive Engineers and the US National Highway Traffic Safety Administration. Specifically, L0 refers to traditional human driving with no automation; L1–L3 include AI-assisted driving with low or moderate automation; and L4–L5 represent that a vehicle can be completely controlled by the AI system, with no human operation needed.

Autonomous driving technology involves multiple industries, such as ICT, manufacturing, and transportation. Its development requires cross-industry collaboration, which can in turn stimulate economic growth. After being deployed at scale, autonomous driving will significantly improve road safety and transport efficiency, and create positive social and economic benefits in terms of energy saving and emission reductions. Autonomous driving is expected to create about US$3.2–6.3 trillion in economic benefits for the US by 2050, leading to nearly US$800 billion in annual social and consumer benefits[19].

In the report titled On the road to automated mobility: An EU strategy for mobility of the future, the European Commission sets out the goal of making self-driving vehicles commonplace across the EU by 2030. The strategy states that the deployment of driverless mobility, when fully integrated in the whole transport system, is expected to contribute significantly to achieving the Vision Zero, i.e. no road fatalities on European roads by 2050[20]. In China, 11 ministries including the National Development and Reform Commission jointly issued the Smart Car Innovation and Development Strategy in February 2020. The strategy proposed to have conditional autonomous vehicles in large scale production and high-level autonomous vehicles commercialized for specific environments by 2025, and to build a mature, standardized intelligent vehicle system by 2050[21].

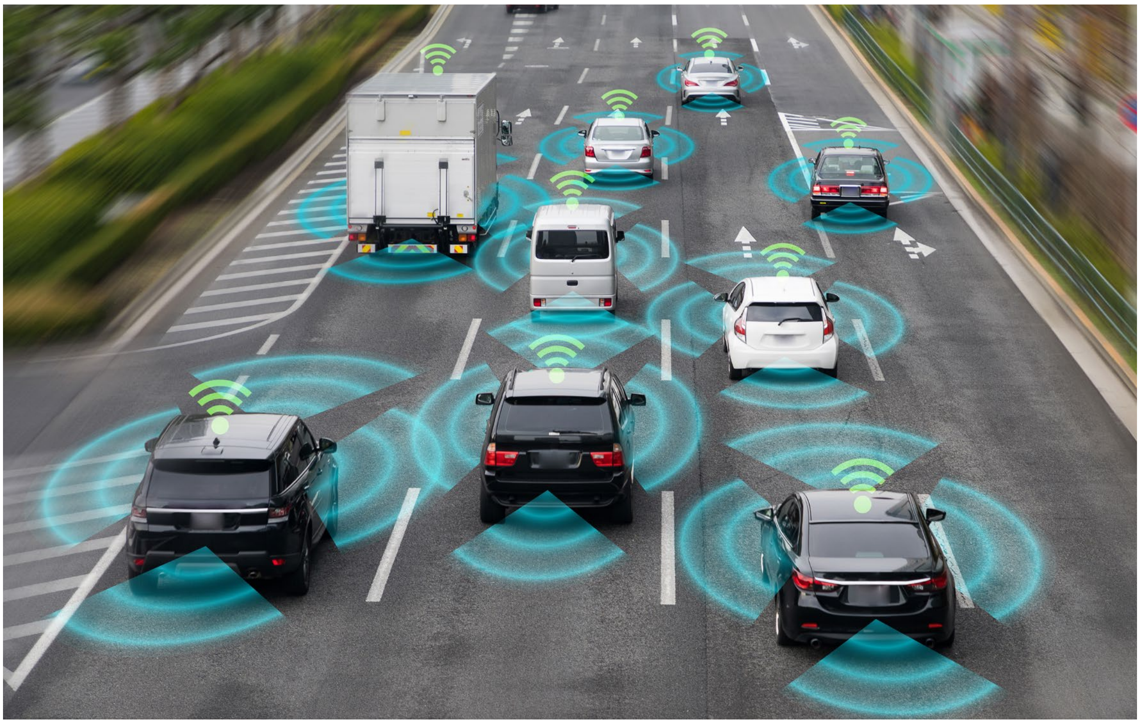
Snapshot from the future: Self-driving vehicles in the fast lane

Self-driving vehicles are achieving higher levels of automation, from L2 and L3 to L4 and L5. Buses, taxis, low-speed logistics, and industrial transport (logistics and mining) will be the first commercial applications of autonomous driving.

Low-speed public roads: Self-driving vehicles have delivered positive results in fields such as logistics and distribution, cleaning and disinfection, and patrolling. Unmanned vehicles for logistics and distribution can successfully drive at low speeds on roads with less complicated conditions. This means they can provide safe unmanned delivery services on public roads. Low-speed unmanned vehicles have provided valuable support during the fight against COVID-19, especially in the transportation and distribution of medical supplies, cleaning and disinfection, patrolling, and checking temperatures. These vehicles have proved their practical value, laying a foundation for their adoption in other markets.

High-speed semi-closed roads: Heavy trucks are expensive, so the price of sensors is not a limiting factor. Sensors such as lidar can be installed in these trucks for better sensing of their environment. Heavy trucks are mainly used in high-speed cargo transportation, ports, and logistics parks, which means the driving environment is less complex and routes are generally fixed. Heavy trucks are rarely seen on complex urban roads. This reduces the complexity of the driving environment that autonomous driving systems have to handle. Truck drivers are expensive, and they frequently breach rules by overloading their vehicles and working overtime. So autonomous driving of heavy trucks would quickly help industries cut costs and work more efficiently, making this a compelling business case. According to a Deloitte report on smart logistics in China, technologies like unmanned trucks and artificial intelligence will mature in a decade or so, and will be widely used in warehousing, transportation, distribution, and last mile delivery[22].

Special non-public roads: Autonomous driving is playing an increasingly important role in environments like mines and ports. Some companies are working with ports to test self-driving container trucks. We have already seen unmanned trucks working in multiple fleets and even during night shifts at mines. At the Yangshan Port in Shanghai, 5G-powered L4 smart-driving heavy trucks can drive at speeds of up to 80 km/h, and the distance between vehicles can be shortened to 15 meters. Thanks to the centimeter-level precision of the Beidou GPS system, the vehicles can come to a stop within 15 seconds of a command, with an error of only 3 centimeters. The use of autonomous vehicles has brought a 10% improvement in vessel loading/unloading times[23].



With its high level of safety and efficiency, autonomous driving will first demonstrate its commercial value in mining. While working autonomously, many mechanical vehicles, such as mining trucks, excavators, and bulldozers can work together. In the past, one driver took care of each mining truck, but now a commander will take care of an entire group of trucks. In the event of a fault or danger, the commander can remotely pilot the vehicle to a safe area from the control center, and send warnings to nearby vehicles.

Public roads: Robotaxis are an obvious business model for self-driving companies, and are one of the best ways to get returns from their initial investment. According to one study, robotaxis could replace 63% of carshare and taxis and 27% of public transport[24].

Autonomous driving technologies will lead to more innovative changes in the designs of car bodies. All possible configurations will be explored and exploited. Cars can become the mobile third space, catering to many different scenarios. This will disrupt the business models of existing industries like catering. Self-driving food trucks may become the standard of the future, and dinner with friends and family may take on a whole new form: After you book a lunch, a self-driving food truck will pick you up and carry you along whatever scenic route you choose. You can enjoy the views while dining and chatting, all within a private space. This model would eliminate the need to visit restaurants and ensure privacy during the meal. For a restaurant, the size of the premises would no longer be a limiting factor for the size of its business, and location would no longer restrict the clientele it could attract. Business results could be disconnected from footfall.

Snapshot from the future: Urban air mobility

In the future, airspace will become an important resource for urban transportation. An efficient air-based urban transportation network will greatly free up roads, reduce travel times, and improve the efficiency of logistics and emergency services.

As defined by NASA in the Urban Air Mobility Airspace Integration Concepts and Considerations, safe and efficient air traffic operations in a metropolitan area for manned aircraft and unmanned aircraft systems consist of aircraft, a command and dispatch platform, a navigation and positioning system, a charging system, and a tarmac.

The development of electric vertical take-off and landing (eVTOL) aircraft has attracted investment from innovative companies around the world, and their performance has seen solid progress. Currently, four-seat aircraft made by several companies can deliver a cruising range of about 100 kilometers[25]. Some companies are working on eVTOL aircraft with seven seats or more[26]. Some are exploring hydrogen-fueled aerial vehicles[27] for longer ranges (more than 600 kilometers). These new aircraft may be used in various scenarios, including emergency medical services, urban air mobility (UAM), regional air mobility (RAM), freight transport, and personal aircraft.

Air emergency rescue systems: Over the past decade (2010–2020), skyscrapers have sprung up in major cities around the world. The number of skyscrapers will continue to grow over the next decade as global urbanization continues. The rapid rise of skyscrapers may make for impressive skylines, but they also create safety risks. Firefighting and emergency medical services in skyscrapers will be a new challenge for cities. Air emergency rescue offers a new solution to these challenges, allowing firefighters and medical personnel to better protect lives and property by quickly reaching higher floors to put out fires or assist people.

Air metro/air taxis: Convenient and efficient transportation is a core need of urban residents. Batteries with higher energy density are enabling electric aircraft to work longer and have a larger capacity. eVTOL will prove to be an effective tool to improve the urban transport experience. Pilot projects have begun for air passenger transport services. In 2019, a Chinese company launched the world's first urban air mobility service in Zhejiang, cutting road trips that normally took 40 minutes to a five-minute air hop[28]. According to NASA's projections, air metro will support 740 million passenger trips by 2030.

Of course, these scenarios require a fast and stable space-air-ground integrated network and positioning system, cost-effective and reliable visual sensors and lidar, secure and stable automatic flight algorithms, and an efficient, real-time command and dispatch platform.

Direction for exploration: Sharing vehicles for faster, low-carbon transport

When a centralized transport management system schedules the vehicles and resources, passengers can be provided with tailored mobility solutions based on their individual needs; at the same time, the vehicles will be used much more efficiently. The shared vehicle model avoids wasteful, carbon-intensive transport where a single passenger takes a vehicle to a single destination.

Snapshot from the future: Mobility as a Service (MaaS) available on demand

According to the International Road Transport Union, MaaS is to put the user at the core of transport services, offering them tailor-made mobility solutions based on their individual needs. MaaS is the integration of various forms of transport modes into a single mobility service accessible on demand. It combines all possible transport modes, enabling users to access services through a single application and single purchase[29].

The key objective of MaaS is to provide integrated and convenient public transport services and develop green transport. MaaS systems aim to integrate local transport (e.g. buses, rail, shared cars, and shared bikes) and intercity transport (e.g. planes, high-speed rail, and long-distance coaches) and provide useful local information like dining, accommodation, shopping, and local tourist attractions. These systems will build on the intelligent scheduling functions of public transport systems, and identify passenger travel models while prioritizing green transport. With online payment functions integrated, MaaS systems can offer travel booking, one-tap itinerary planning, seamless connections between different transport modes, and one-tap payments. MaaS will improve satisfaction with transport services while providing green transport options.

Many EU cities are building MaaS showcase projects. Different cities have different levels of integration in terms of facilities, fares, payments, information, communications, management systems, and transport services. Gothenburg, Hanover, Vienna, and Helsinki were the first cities to explore MaaS. These cities have made full use of digital technologies to optimize their transport systems, including buses, shared cars, bicycles, and urban deliveries. This will help them incubate emerging transport service providers and drive urban decarbonization[30].

MaaS can bring tangible benefits: Individuals can cut their transport costs while enjoying better safety and a better experience. Governments can optimize their investment in transport infrastructure for more sustainable urban management and higher citizen satisfaction. In addition, MaaS will create more opportunities for transport service providers, as they can cut service costs and expand their services. After MaaS is widely deployed, we will see integrated scheduling of transport resources, better shared resources, a user-centric experience, and low-carbon transport.

Direction for exploration: Connected vehicles for safer, faster,and larger-scale autonomous driving

Intelligent and self-driving vehicles that are unconnected face challenges in areas like target detection, trajectory prediction, and driving intention prediction, especially when visibility is limited or weather conditions are harsh. However, when vehicles are connected, there will be well-coordinated information exchange, sensing, and decision control. This will give each individual vehicle a much greater ability to sense what's around it. The introduction of new intelligent resources like high-dimensional data will make collective intelligence a reality. This means that the technical challenges facing individual vehicles in autonomous driving will no longer exist, and these vehicles will become safer, better drivers. As a result, the operation design domain (ODD) of autonomous driving will expand. ODD means the operating conditions under which a given driving automation system is specifically designed to function, including, but not limited to, environmental, geographical, and time-of-day restrictions, and/or the requisite presence or absence of certain traffic or roadway characteristics.

Affordability is a major challenge in the large-scale commercialization of autonomous driving. High-level autonomous driving requires a large number of on-board sensors. Generally, an L4 self-driving vehicle requires 6–12 cameras, 3–12 millimeter wave radars, 5 laser radars, 1–2 global navigation satellite systems (GNSSs) or inertial measurement units (IMUs), and 1–2 computing platforms. This means that costs are high for each vehicle. If the vehicles are connected and supported by roadside sensing and scheduling, the cost per vehicle will go down. This is how intelligent vehicles will evolve, and the only way to deploy automated driving systems on a large scale[31].

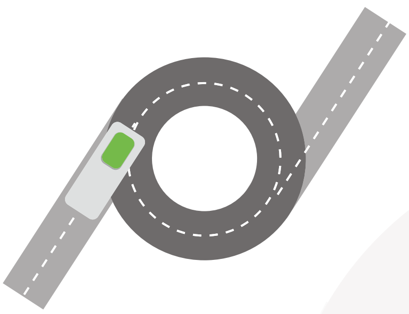
Connecting vehicles requires continuous network coverage. Currently, global mobile communications services cover only about 20% of the land area (land only covers 29% of the earth's surface), and less than 6% of the earth's surface. For example, more than 95% of the sea area of China is not covered by terrestrial mobile communications networks[32].So a space-air-ground integrated network is needed to provide continuous coverage. As in-vehicle and in-flight entertainment on large screens and holographic conferences are becoming more popular, terrestrial networks alone will not give users the consistent experience they demand in their entertainment and work. A space-air-ground integrated network will be needed to provide large bandwidth and high availability.

Snapshot from the future: Safer, more efficient dispatch services

In the past decade, pioneers have begun exploring the use of elevated rails to transport containers in busy ports. Containers are sent to rails similar to cable railways. The railway system dispatches the containers based on their destination and sends them to railway stations, truck warehouses, or even waterless ports in inland cities. This makes container transportation much faster at a very low cost. In the future, space-air-ground integrated networks will support the dispatching of unmanned vehicles and drones, providing safer and more efficient dispatching services for each vehicle and aircraft, and making large-scale autonomous driving a reality.

Snapshot from the future: Broadband in the air, just as at home

Moving forward, broadband coverage will extend beyond the ground into the air and beyond. Broadband connections will be available to devices at various heights, such as drones less than 1 kilometer above the ground, aerial vehicles 10 kilometers above the ground, and low-orbit spacecraft hundreds of kilometers above the ground. The integrated network will consist of small cells covering hotspots with a radius of 100 meters, macro cells with a radius of 1 to 10 kilometers, and low-orbit satellites with coverage over a radius of 300 to 400 kilometers, providing users with unbroken access to broadband of up to 10 Gbit/s, 1 Gbit/s, and 100 Mbit/s, respectively[33].



Conclusion：

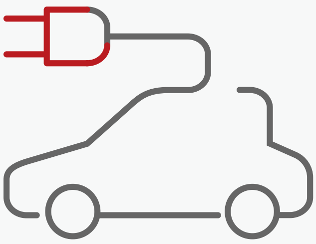
Smart, low-carbon transport opens up the mobile third space

Future transport will be a multi-dimensional innovative system. The shift to electric, autonomous, shared, and connected vehicles will create an intelligent, convenient, low-carbon transport experience. To make this happen, we need innovative applications of new energy technologies; secure and stable autonomous driving algorithms; cost-effective, reliable sensors; a high-speed, stable space-air-ground integrated network; and a traffic management brain based on great computing power.

The mobile third space will reshape the transport experience, incubate innovative mobility services, and drive the emergence of new business models. An intelligent urban transport management system can optimize resource allocation, enable more efficient sharing of transport resources, alleviate traffic congestion, and reduce the environmental pollution caused by traffic. This is how we will resolve the conflict between the surging demand for transport and the urgent need to decarbonize.

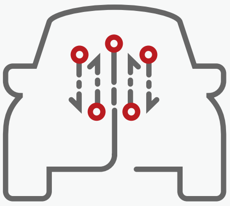
Huawei predicts that by 2030, 50% of new vehicles sold globally will be EVs, and 20% of new vehicles sold in China will be autonomous vehicles. In addition, by 2030, the whole-vehicle computing power will exceed 5,000 TOPS, and 60% of new vehicles sold will support C-V2X.

Huawei predicts that by 2030,



50% of new vehicles sold will be electric vehicles.

20% of new vehicles sold in China will be autonomous vehicles.



Whole-vehicle computing power will exceed **5,000** TOPS.

60% of new vehicles sold will support C–V2X.

23：SAIC Iveco Hongyan Commercial Vehicle Co., Ltd., https://www.hongyantruck.com/html/news/20200630/1219.shtml

24：China’s First Survey Report on Robotaxi Passengers, Institute of transportation Engineering of Tsinghua University and WeRide, https://www.weride.ai/static/中国首份自动驾驶出租车（Robotaxi）乘客调研报告 -65333c4d8c753aa16d8e00a1b4f5c1bd.pdf

25：The Roadmap to scalable urban air mobility, Volocopter, https://volocopter-statics.azureedge.net/content/uploads/20210324_Volocopter_WhitePaper_Roadmap_to_scalable_UAM_m.pdf

26：Lilium，https://lilium.com/

27：Skai，https://www.skai.co/

28：EHang, https://www.ehang.com/cn/

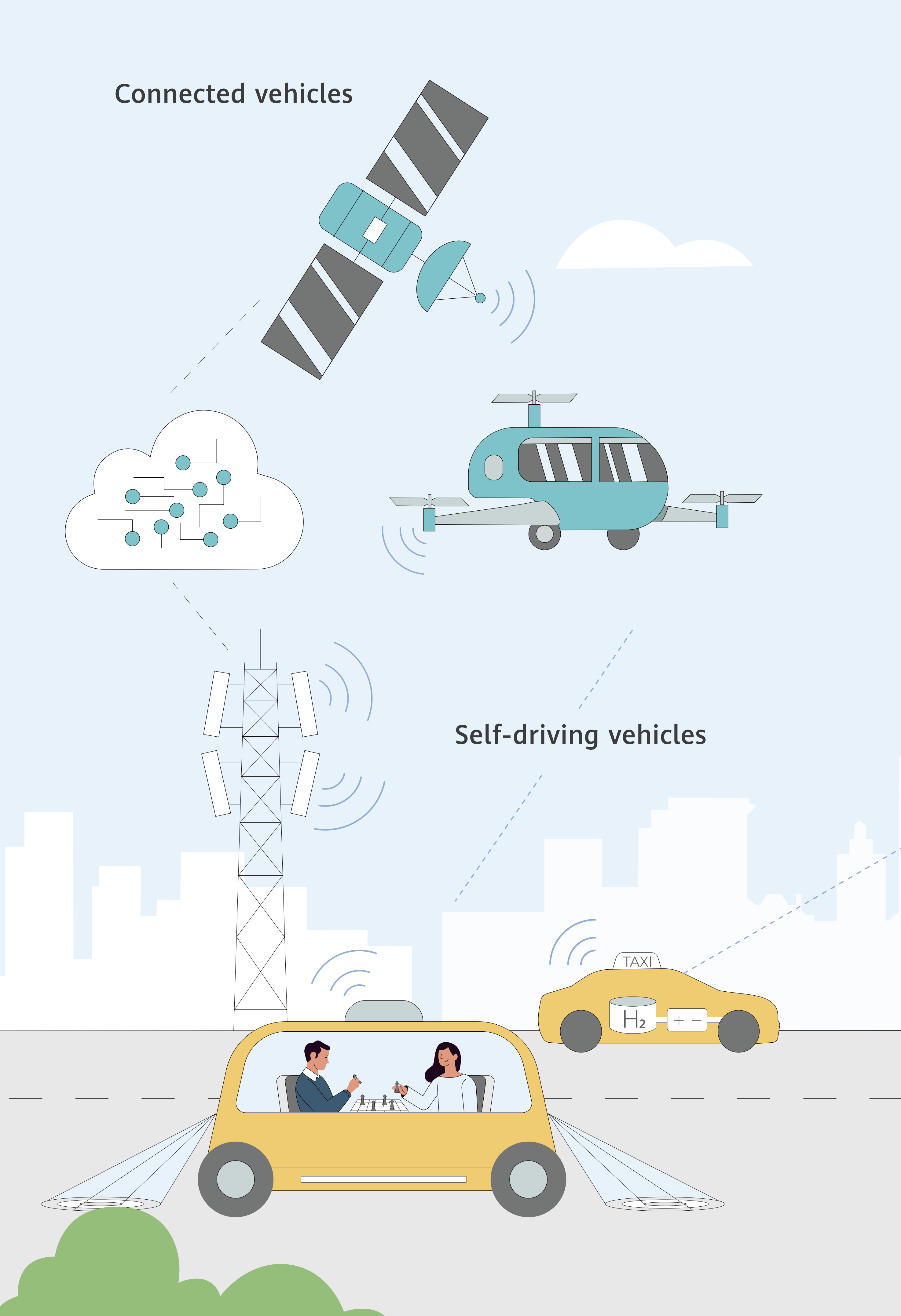
29：IRU, https://www.iru.org/who-we-are/where-we-work/europe/maas-mobility-service

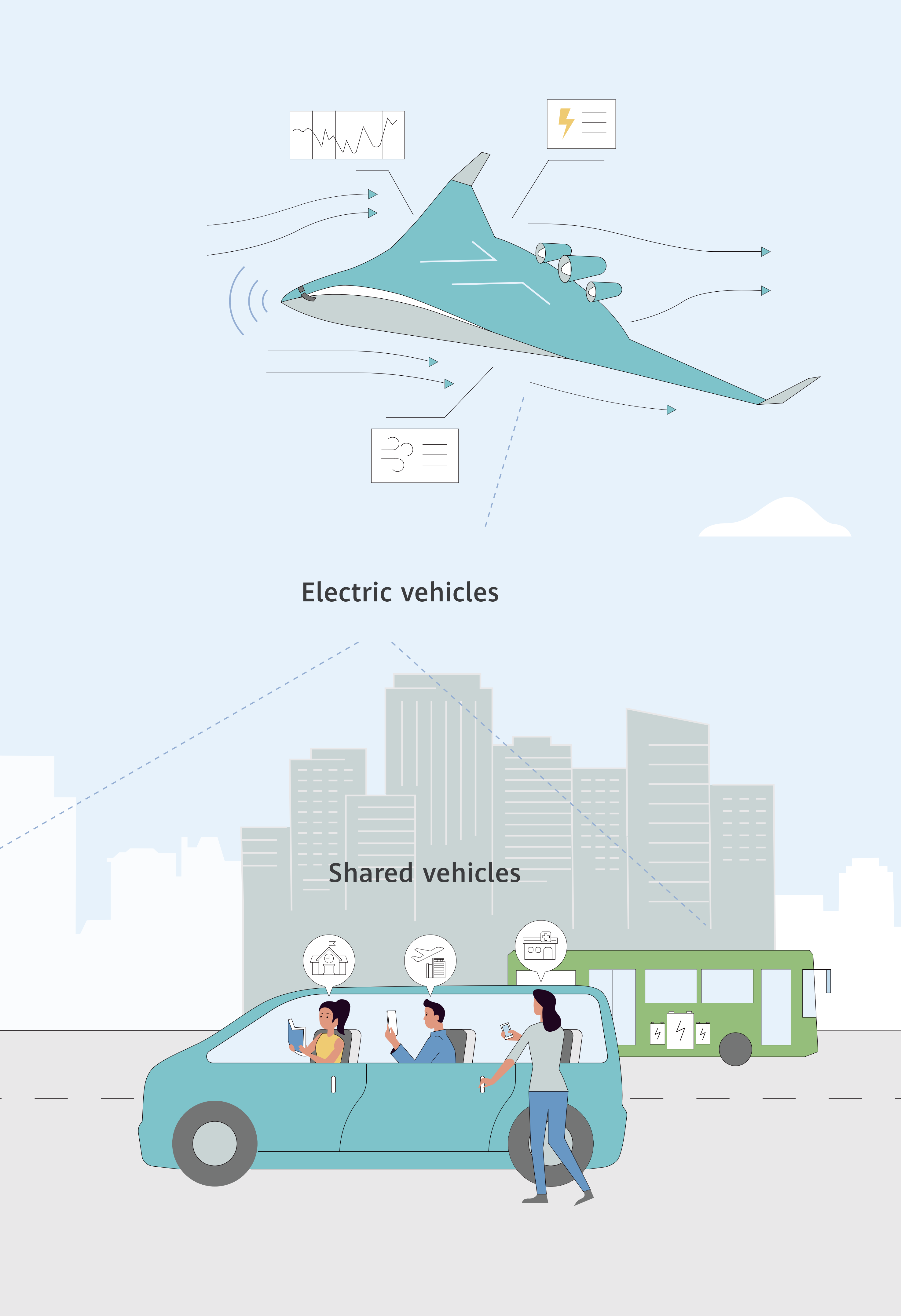
30：ITS Deployment Evaluation, https://www.itskrs.its.dot.gov/its/benecost.nsf/ID/d1c266bdeefee830852581bb005d27be

31：Key Technologies and Outlook for Vehicle-Road Coordination for Autonomous Driving, Institute for AI Industry Research of Tsinghua University, and Baidu Apollo, https://apollo.auto/docment/ 面向自动驾驶的车路协同关键技术与展望 .pdf

32：Vision, Requirements, and Technology Trend of 6G: How to Tackle the Challenges of System Coverage, Capacity, User Data-Rate and Movement Speed, Shanzhi Chen et al., https://ieeexplore.ieee.org/document/9003618

33：Network 2030 : A Blueprint of Technology, Applications and Market Drivers Towards the Year 2030 and Beyond, FG-NET-2030, https://www.itu.int/en/ITU-T/focusgroups/net2030/Documents/White_Paper.pdf


Connected vehicles
Self-driving vehicles
TAXI
H_2
+ –


Electric vehicles
Shared vehicles

Cities

New Digital Infrastructure
Makes Cities
More Human and Livable

Urbanization is a global mega trend of this century. More than half of the world's population now lives in urban areas, and this figure is expected to rise to 60% by 2030[1]. Take China as an example: Morgan Stanley projects that the percentage of China's population in urban areas could increase from 60% currently to 75% by 2030, translating into 220 million new urban dwellers[2]. China will form five super-city clusters, such as the Guangdong–Hong Kong–Macau Greater Bay Area and the Yangtze River Delta[3]. In 2030, there will be 43 megacities with populations of 10 million or more[4].

As urbanization accelerates, cities worldwide will be forced to negotiate the contradiction between increasing size and limited resources. Typical urban problems, such as high energy consumption, pollution, traffic jams, and unequal access to digital infrastructure, will become even more pressing.

According to UN-Habitat, cities consume about 75% of global primary energy and emit between 50% and 60% of the world's total greenhouse gases[5]. By 2030, the world is expected to generate 2.59 billion tonnes of waste annually[6]. Up to 53 million tonnes of plastic could end up in rivers, lakes, and oceans every year by 2030[7]. Another worrying statistic is that air pollution kills an estimated seven million people worldwide every year[8].

The conflict between growth and limited resources will be the biggest headache for cities. They will need to make the most efficient use of their resources. Rapid advances in new technologies, such as 5G, cloud, AI, blockchain, and intelligent sensing are opening up more possibilities for cities, which represent the best places to create and incubate new applications for these technologies.

Over the past decade or so, countries across the globe have been making their cities more digital and exploring ways to leverage technologies to support sustainable development. In 2020, nearly 1,000 exploratory smart city projects were underway worldwide. China was home to about 500 of these, with 90 in Europe and 40 in the US[9]. Global spending on smart city initiatives is increasing year by year. In 2020, this spending totaled nearly US$124 billion, an increase of 18.9% over 2019[10]. Advancing digital transformation has become one of the key pathways to sustainable development for the world's leading cities.

Direction for exploration: New digital infrastructure is the engine of digital cities

The unstoppable expansion of cities is putting great pressure on the environment and on our limited resources. The biggest challenge for the future of urban development is how to use technology to significantly improve urban governance. Cities need advanced, refined systems that deliver sustainability while minimizing resource use.

In the past, cities provided physical public infrastructure, including water, electricity, gas, and road networks, to support rapid industrialization. Moving forward, a major direction for exploration will be new digital infrastructure that supports the development of digital and intelligent urban systems.

We believe that new digital infrastructure will consist of four layers. The bottom layer is an intelligent sensing system that can accurately sense dynamic situations and the heartbeat of the city in real time. The second layer is intelligent connectivity. High-speed wired and wireless connections will connect the city, creating an organic whole. The third layer is an intelligent hub, which will serve as the city's "brain" and decision-making system. This layer aggregates massive data and enables city-wide data sharing so that AI systems can deliver the maximum possible value, making granular, data-driven, highly-automated city governance a reality. The fourth and top layer is smart applications. A comprehensive ecosystem of smart applications will be built on top of the digital infrastructure, covering the "last mile" of service delivery and creating infinite possibilities for smart urban growth. These four layers will interconnect and support each other, ultimately producing a smart city that embodies the new goal of ubiquitous intelligence.

Snapshot from the future: Nanosensors track the pulse of the city

Digital cities depend on data, which comes from a wide array of sensors scattered throughout the city. Just as people perceive their surroundings using their senses of sight, hearing, smell, taste, and touch, a city needs its own sense organs, deployed around the city, to sense what is changing. These sensors will provide the data that underpins the growth of a digital city.

The MIT Technology Review listed Sensing City as one of the 10 Breakthrough Technologies 2018. We believe that in the cities of the future, restricted sensing systems will merge into a comprehensive sensing network. Sensors across a city will be connected using a range of different transmission systems. Analytics using these massive flows of comprehensive data will generate a more accurate picture of the latest developments within a city. Breakthroughs and advances in sensing technology will drive leapfrog advances in sensing cities.

One particularly cost-effective and disruptive technology – nanosensors – is expected to drive the next revolution. Nanosensors have huge potential and can be deployed in massive numbers to form a wireless nanosensing network. This will greatly enhance a city's ability to sense, leading to major advances in climate monitoring, health monitoring, environmental protection, and other domains.

Nanosensors are very small and precise, and will vastly improve sensing performance. Working at the atomic scale, they are expanding our understanding of what sensors can be, driving new advances in sensor manufacturing, and opening up new fields of application. Early applications of nanosensors cover many different fields, including biology, chemistry, mechanics, and aerospace.

Graphene gas nanosensors are ultra-sensitive to odors. A nano-coating on the sensor's surface where it makes contact with the gas improves sensitivity and performance. The sensor collects odor molecules with a metal-organic film, and then amplifies the chemical signals using plasma nanocrystals. The most common application is for detecting carbon dioxide, but the sensor can also quickly detect hazardous and toxic gases. An American university has created a novel type of nano-coating using graphene, and when the coating is applied as a nanofilm on gas sensors, it delivers a 100-fold increase in molecular response compared to the best available sensors that use carbon-based materials[11].

In the near future, these sensors will be able to accurately identify hazardous, toxic, or explosive gases in the air, thereby greatly improving safety in scenarios like factories and at customs inspections.

Nanocrack-based acoustic sensors are able to recognize specific frequencies of sound. They are more sensitive than other acoustic sensors, as the spacing between cracks in nanocrack-based sensors can be as little as just a few nanometers. Researchers build the sensor frame by adding

a 20-nanometer-thick platinum layer to the surface of a viscoelastic polymer. When the platinum layer deforms and stretches, it pulls apart from the underlying polymer. Researchers then measure the conductivity of the sensor surface. In an environment with 92 decibels of noise, the nanocrack sensor performed far better than conventional microphones at separating out sounds in a given frequency range. For example, when these nanosensors are placed on the surface of a violin, they can accurately record every note of a tune and "translate" it so that a connected device can accurately recreate an electronic version of the tune. When this kind of sensor is worn on the wrist, it can accurately monitor a person's heartbeat. Breakthroughs in this technology will greatly enhance acoustic monitoring in urban infrastructure[12].

Snapshot from the future: All-optical, 10 gigabit cities

The digital transformation of cities will require massive flows of information. The newest generation of connectivity technologies, such as 5G, F5G, and gigabit Wi-Fi, is enabling high-speed networks with universal coverage in urban spaces. High speed information flows require optical networks. With optical networks as the foundation, cities will be able to merge their operational infrastructure into their communications infrastructure. As a result, new types of people-centered government services will be accessible and affordable to every person, home, and organization.

Several major cities have already conducted preliminary research into this area and found tentative signs that all-optical cities will unleash tremendous value and growth potential.

In April 2021, Shanghai became the world's first all-optical smart city. Owing to its F5G optical network, the city is able to deliver stable connections with latency below 1 millisecond anywhere in the urban area. The deployment of this high-speed optical network has laid a solid foundation for Shanghai's future digital transformation[13].

In Adelaide, Australia, 1,000 buildings are now connected to 10 gigabit networks. Companies in these buildings can access cloud services at a speed of 10 Gbit/s, creating huge opportunities for industries such as education, video, IT, and software engineering[14].

We believe that all-optical infrastructure will drive leapfrog improvements in the communications networks of future cities, in terms of capacity, bandwidth, and user experience: Uplink and downlink rates will reach 10 Gbit/s; latency will be reduced to microseconds; and the number of connections will see a 100-fold increase.

By 2030, many leading cities will have 10 gigabit connectivity, with 10 gigabit wireless services available to organizations, homes, and individuals.

The future architecture of an all-optical city will consist of four parts:

All-optical network access: All network connections will be optical, including homes, commercial buildings, enterprises, and 5G base stations. The all-optical transmission network will be extended into edge environments like large enterprises, commercial buildings, and 5G

base stations.
This will enable the digital transformation of many different industries, and support the development of F5G+X and 5GtoB industrial applications.

All-optical anchors: Connections originating in home broadband, enterprise broadband, 5G networks, or data centers will be routed and transmitted through all-optical networks; optical networks will support multiple different fixed access technologies and provide one-hop connections to the cloud.

All-optical switching: One-hop access to services through urban optical networks. All-optical cross-connect technologies are used to build multi-layer optical networks that support one-hop access to services; high-speed inter-cloud transmission; and high synergy between cloud and optical networks.

Fully automated O&M: Real-time sensing of network status with proactive, preventive O&M. This supports elastic network resources, and automated service provisioning, resource allocation, and O&M.

Snapshot from the future: Intelligent hubs cut out the human factor from urban management

During the digital transformation of cities, data barriers will be broken down, and all data will ultimately form a single data lake. AI will play an increasingly important role, and many public service and regulatory issues will be decided by expert algorithms instead of being left up to human judgment. Narrow AI applications will expand until they can deliver data-driven wisdom for all city governance scenarios.

Technological advances are driving important changes in the approach to city governance. For example, there is a shift from reactive services to proactive services, from broad-brush to granular regulation, and from post-hoc response to real-time response and even incident prediction and prevention.

These changes will pose new challenges: For example, AI will lead to the creation of new regulatory authorities. Supported by massive data and the superior performance of algorithms, AI will itself become a part of city governance, and will then use its position to drive further changes. AI will be integrated into more and more sectors of government, connected to all the various sensors deployed across the city, so that it can support the accurate and efficient provisioning of urban resources. AI ethics and the principles of humanism and fairness must be constantly deployed to correct any deviations as the technology evolves.

Huawei believes that as the goals of city governance evolve, cities will need powerful, intelligent hubs. These hubs will help us overcome the challenges of technology, and will aggregate data and support applications. They will also have the capacity to iterate and improve themselves. These hubs will aggregate massive amounts of data from every corner of a city, and mine that data for insight to support better city governance. This will benefit every industry and greatly improve the efficiency of city governance and the experience of users of government services.

Early-stage explorations by Toyota: In Toyota's plan for the city of the future, each house, building, and vehicle will be equipped with sensors. Data from these sensors will then be aggregated into a city's data operating system, through which people, buildings, and vehicles are all connected. When the data is aggregated to the intelligent hub, AI will be used to analyze people's surroundings and then guarantee the safety of pedestrians and drivers by keeping them separated. In addition to new technologies like indoor robots, citizens can use AI to check their health at home. Wearable medical sensors for home use will transfer data to the data operating system, which will provide instructions for healthcare[15].

Snapshot from the future: Smart ecosystems spread intelligent services across all use cases

When cities have ubiquitous, high-speed connectivity, intelligent hubs, and massive, real-time data from city sensors, smart applications will emerge in every aspect of urban life. Starting with government services, AI will expand to industry support and smart lifestyles. The key to this process will be an ecosystem for innovation in smart applications that bridge the last mile in citizen services. These services will be vital to unleashing the value of new digital infrastructure.

Some major cities are already early adopters of this idea. In China, Huawei has partnered with Guangming District in Shenzhen to build a green, all-optical, smart district as a showcase. Once kicked into full gear, this project will accelerate the innovation in services to key urban industries, such as smart manufacturing, life sciences, and optical networks.

Huawei and Guangming will build China's first innovation center for life sciences and intelligent manufacturing. It will include two service platforms – EI Health and Fusion Plant. By bringing together both upstream and downstream players to serve their respective industries, Huawei and Guangming will accelerate their modernization and use of AI.

The industry platforms will provide services to support business innovation: a powerful public cloud computing platform and massive storage. For the life sciences, they will provide trained algorithms for image analytics, gene analysis, and drug R&D data analytics. For manufacturers, they will offer industrial Internet services. This will accelerate digital transformation for biomedical and industrial enterprises, and support the emergence of smart applications and application providers.

Thanks to new digital infrastructure, cities will be able to create their own smart innovation ecosystems. These ecosystems, in turn, will take advantage of infrastructure to bridge the gaps in digital innovation. We expect to see the ecosystems serve industries and the industries give back to and nurture the ecosystems. This will unleash the tremendous value of digital infrastructure across cities, benefiting countless industries. Smart ecosystems are key to spreading intelligent services across all use cases.

Direction for exploration: Smart government services make cities more human

In China, accessing government services used to involve many visits to different agencies and administrative departments. Today, most Chinese cities have government service halls and most government services can be delivered within a single hall. Since the outbreak of COVID-19, more government services have been delivered through smartphones, and most people no longer even need to visit service halls.

Government service designed for people is an approach being progressively implemented in China and certain leading cities worldwide. At the same time, emerging technologies such as cloud, AI, and blockchain are all evolving. What will these developments bring to cities? They can guarantee swift and easy access to government services and make our cities friendlier places to live. This is one of the major directions of development for our cities.

Snapshot from the future: Proactive, precise provision of government services

Machine recognition technology makes contactless services possible. Today, in most of China's developed provinces, citizens do not need to go to government offices to access government services. They can now access them directly through their smartphones. Over the next decade, the digitalization of government services will be taken to the next level.

1. Digital identity authentication will be widely adopted. The ID cards, drivers' licenses, social security cards, and bank cards that people carry at all times will be digitalized. It is estimated that the total addressable market for global electronic identity authentication services will be worth US$18 billion by 2027[16].

2. Digital credit will underpin and restructure many public service processes and the customer experience. It will be one of the founding technologies for digital government. Most citizens are already familiar with forms like electronic library cards, social security cards, and car rental services that require a credit rating. As these services continue to improve, they will deliver a better experience to millions of citizens.

3. Universal access to one-stop, e-government services will soon be realized. In the future, all government services will be remotely accessible and government service halls may cease to exist.

Technological advances will give rise to new digital government practices and government services. Today's centralized digital government

model is a great example. In many Chinese cities, the local government has created a single network. This model uses big data and IoT, and enables complex operations involving multiple different departments, regions, and levels of government. It offers an architecture that addresses the needs of the municipal government and of local residents. In the future, as governments aggregate more massive data and their AI technologies mature, they will be able to deliver government services in a more precise and proactive way; manage their municipalities more efficiently; and improve their service experience.

Let's look at smart care for the elderly as an example. Communities in Shanghai have installed smart water meters for elderly people who live alone and agree to the installation. If the total water used within a 12-hour period falls below 0.01 cubic meters, the meter will send an alarm to the central network and community workers. These workers will then visit the elderly person in question to check whether everything is normal. Such attentive care demonstrates a real human touch and concern for the elderly[17].

Snapshot from the future: Blockchain-based data sharing

Data is the most important element in the digital transformation of cities. However, data sharing services have been confined to a single center, which is not conducive to broad sharing or openness. There are also many other challenges, including data security risks during data exchanges, flaws in standards and specifications, information silos, and inconsistent solutions for data access control, encryption, and access audits. Together, these challenges make it difficult to share data effectively between government departments or between governments and private companies.

Blockchain technology can be fully integrated with emerging information technologies such as cloud computing, big data, and AI. It offers a solution to the problems of trust in data transmission, sharing, and use that arise in digital cities. Blockchains are maintained by multiple parties, and use multiple cryptographic technologies to ensure secure transmission and access. They meet the needs of many different government data applications. Blockchain technology uses hash pointers to prevent data tampering, and is a secure and trustworthy medium for data sharing. In a complex service environment such as digital government services, blockchain could be used to establish trust among government departments and dramatically increase the efficiency of data sharing.

The application of blockchain as a base technology in digital government will support the provisioning of government services. We can examine how the technology could be used in different scenarios. Clear use cases include information infrastructure, smart transportation, and energy sectors such as electricity.

Dubai is one of the world's fastest adopters of digitalization. This city is applying some of the most innovative ideas to real-world scenarios, aiming to become a smart metropolis powered by blockchain. Within its smart city program, one pilot initiative is garnering much public attention. The initiative in question plans to use blockchain technology to track, transport, and deliver import and export goods. By integrating this initiative into the city's administration of foreign trade. Dubai plans to create a secure and transparent platform. This blockchain-powered system will give a huge boost to document processing efficiency and is predicted to save US$1.5 billion and 25.1 million working hours. Once this system is launched, citizens will no longer need to wait in long queues to access government services[18].

Direction for exploration: Intelligent environmental protection for livable cities

The growth of a city always puts pressure on the environment. Cities cause air pollution, carbon dioxide emissions, solid waste, and water pollution. Infrastructure for protecting the environment usually lags far behind the economic development and population growth of a city. Therefore, a key aspect of constructing future cities relates to addressing the conflict between city growth and the environment. This leads us to examine how digital and intelligent technologies can be used to protect the environment more efficiently and make cities more livable for every resident.

Snapshot from the future: Automatic waste disposal makes zero waste cities

Cities produce huge amounts of solid waste everyday, and effective waste disposal has always been a daunting challenge for city administrations. The zero waste city is a concept that emerged from building innovative, green cities, and sharing the benefits of city development. Under the zero waste city initiative, cities will grow and live in a green way, reduce solid waste, use resources more efficiently, and minimize the impact of solid waste on the environment.

"Zero waste city" initiatives have been launched around the world. For example, the European Union (EU) has kicked off the European Green Deal. Key elements of the waste proposal in this deal include:

- A common EU target for recycling 65% of municipal waste by 2030;
- A common EU target for recycling 75% of packaging waste by 2030;
- A binding landfill target to reduce landfill to a maximum of 10% of all waste by 2030;

- A ban on landfilling of separately collected waste;

- Concrete measures to promote re-use and stimulate industrial symbiosis – turning one industry's by-product into another industry's raw material[19].

The C40 Cities issued an Advancing Towards Zero Waste Declaration, pledging to achieve zero waste cities by:

- Reducing municipal solid waste per capita by at least 15% by 2030 compared to 2015;

- Reducing the amount of municipal solid waste disposed through landfills and incineration by at least 50% by 2030 compared to 2015; and

- Increasing diversion away from landfill and incineration to at least 70% by 2030[20].

In 2019, China started piloting its zero waste city initiative in 11 + 5 cities[21], exploring ways to construct a solid waste classification and recycling system.

With countries setting goals to construct zero waste cities, solid waste processing technologies and innovations are set to develop rapidly, generating many new methods and best practices in this sector.

The Songdo Smart City Hub in South Korea has introduced an automated waste disposal system, which uses negative pressure to suck domestic waste to a waste processing center through underground pipes. In Malaysia, a company has developed a waste disposal system[22] that transports municipal or domestic solid waste through underground pipes at high speeds, taking the waste from waste chutes and outdoor load stations into sealed containers located up to 2.5 km away. Once the containers are full, they are collected at specific times by a flatbed armroll truck. This system dramatically speeds up the entire waste collection process whilst reducing manpower.

A company in Europe has developed an automated waste sorting robot. Powered by AI, this robot can automatically identify different types of waste on a conveyor belt, classify the waste as required by customers, and then process and recycle the waste. The robot can sort waste several times faster than a normal worker and can work 24/7, significantly speeding up the entire waste sorting process. In the future, with the help of such intelligent waste sorting robots, waste sorting in cities is set to be much faster, as it becomes fully automated and unmanned.

With the help of AI, the entire waste management process in a future city, from collection and transportation to sorting and processing, will be automated and intelligent. Intelligent waste recycling bins, driverless garbage trucks, automated waste sorting robots, automated garbage recycling devices, and other innovative applications will emerge one after another. Hopefully, this will help to make more and more zero waste cities possible.

Snapshot from the future: Optical detection makes water sources safer

The uneven distribution of water resources and water pollution have long been problems for cities. Many face water shortages. At present, slightly less than one half of the global population, amounting to about 3.6 billion people, live in areas that suffer from water scarcity at least one month each year[23]. Industrial wastewater and agrochemical water pollution are far more serious today than in the past.

The management and use of water resources in most cities worldwide is still split across several different industrial sectors, siloed, and lacks any kind of overview. In the future, cities will need to properly manage their full water cycle, from intake and use to discharge. A holistic, AI-managed system will be introduced to manage water resources, while cities rebuild their water facilities. An AI system will be able to maximize the use of water resources within a city by refining every stage of water intake, use, and discharge, using forecasts for the weather and water consumption. This process will also involve precise, scheduled use of water resources and a reduction in the total energy consumed.

Water quality monitoring, especially regarding the treatment of industrial wastewater, is another key concern for the conservation of city water resources. The latest technologies can be applied in this sector. Wastewater often goes through a chemical treatment process. However, this approach usually takes a long time and has many restrictions. In contrast, a new optical detection technology does not suffer from any of the drawbacks of chemical treatment. Each substance produces a unique spectral pattern. Leveraging this feature, the new technology could be used to monitor water quality in real time throughout the entire process, detecting wastewater whenever it is present.

A research team in the US has developed optical sensors to detect sewage contamination in the Great Lakes. Statistical relations between optical properties and genetic bacteria markers will be used to calibrate field sensors for the detection of sewage contamination, including the sources and timing of contamination[24].

Optical technologies can also be further integrated with analytics from the IoT, AI, and cloud computing. Sensors for water quality and deep data analytics will move us closer to 24/7, efficient, real-time, automated, intelligent water quality monitoring. This will provide quick warnings in cases of water contamination.

Optical technology can be used alongside AI to explore the hidden relationships between water quality parameters and treatment processes, so as to upgrade and transform urban sewage treatment processes more methodically.

We will thus form a closed loop of prevention, control, monitoring, treatment, and better prevention.

Snapshot from the future: Real-time AI air quality monitoring

Over recent years, air pollution has posed an increasing threat to people's health, and urban air quality is attracting more attention than ever before. According to the World Health Organization, nearly 90% of cities worldwide fail to meet its air quality standards. This problem is currently only getting worse. Industrial exhaust, coal burning, automobile exhaust, and other types of air pollution have all become major challenges to public health.

Most cities will opt to deploy cost-effective and reliable air quality sensors and build a monitoring network. This way, they can monitor air quality and weather across the entire city and take targeted measures to improve air quality and the urban environment. One company has developed a highly-integrated, real-time air monitoring system[25]. The integrated sensors and software can monitor the concentrations of environmental pollutants in urban environments, such as PM2.5, PM10, CO, NOx, SOx, and O_3, as well as other environmental parameters such as noise, temperature, humidity, air pressure, rainfall, and floods. Data is transmitted to the cloud platform in real time through wireless connections. The cloud platform provides a real-time visual dashboard for clear monitoring and management of the overall environment in key areas of the city.

In the future, we can consider integrating sensors with AI. Supported by machine learning, sensors will be trained to detect their surroundings and

make preliminary judgments regarding potential changes. This intelligent upgrade of sensors will substantially improve the ability of cities to automatically sense environmental changes in real time.

The use of AI-powered sensing technology in the fight against COVID-19 is a prime example of how the technology can be applied. Every time a person exhales, small droplets are released into the air. If someone is infected, droplets in the air can transmit the virus to others. The lower the humidity or temperature, the longer the infectious aerosol can stay in the air and the higher the probability infecting another person. An AI sensing system can monitor volatile organic compounds, humidity, and air temperature, and determine whether the environment is conducive to virus transmission. In addition, the system can automatically and centrally control ventilation and air conditioning systems to reduce the risk of infection.

Novel applications in this sector will boost the ability of cities to monitor and improve air quality and simplify the process of managing our atmospheric environment.

Conclusion:

New digital infrastructure makes cities more human and livable

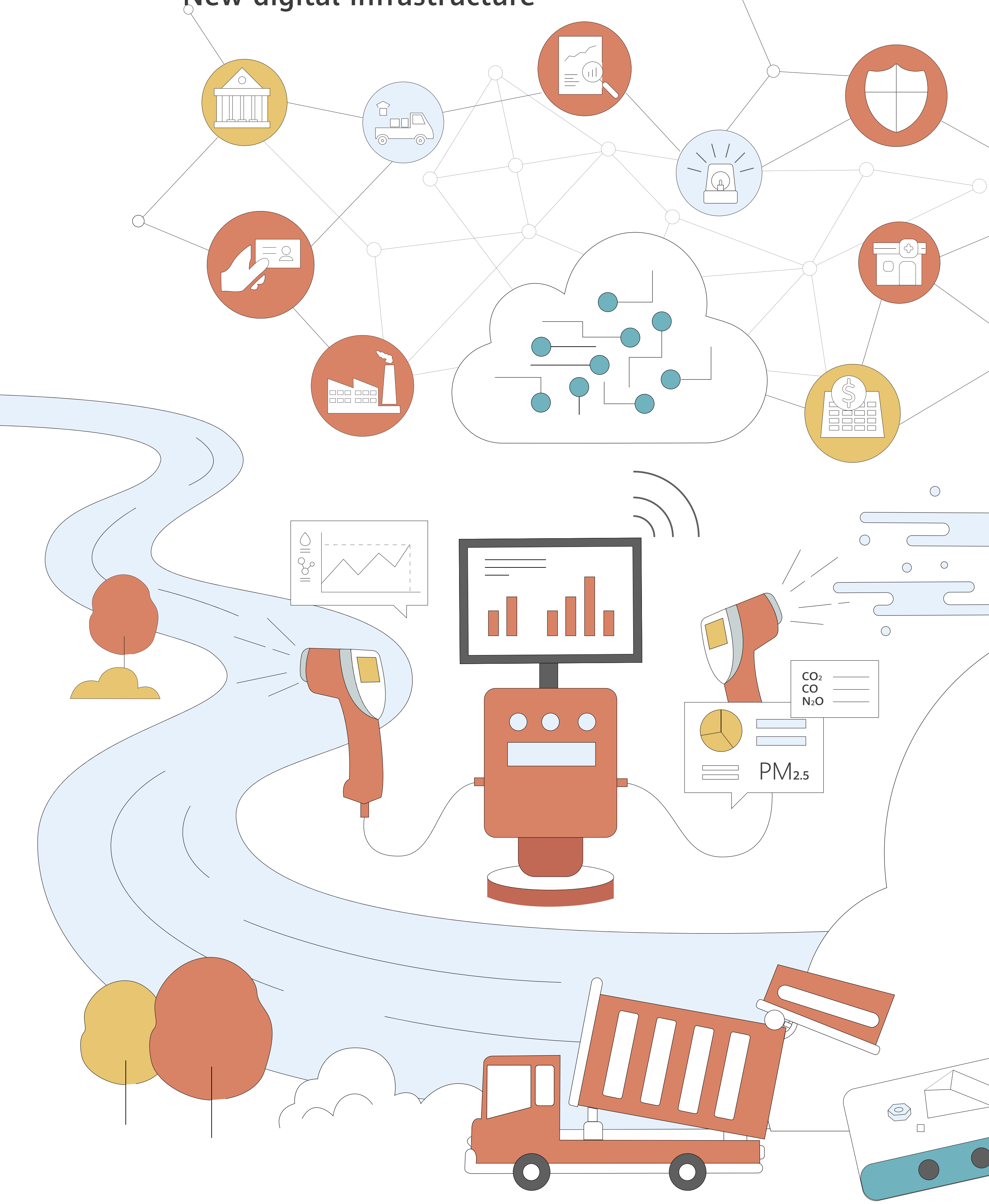
ICT technologies like 5G, optical networks, AI, cloud, and blockchain will all be rolled out rapidly over the next decade. Cities will soon welcome a period of 10 gigabit connectivity, with 10 gigabit wireless services becoming available to organizations, homes, and individuals. Huawei predicts that by 2030, 40% of companies and 23% of homes worldwide will have access to 10 gigabit Wi-Fi networks.

The application of these ICT technologies in cities will significantly boost our ability to make use of limited resources, to manage our cities efficiently, and to give citizens a positive experience. ICT technologies will help cities achieve their sustainable development goals, and make our cities more human and livable.

Huawei predicts that by 2030,

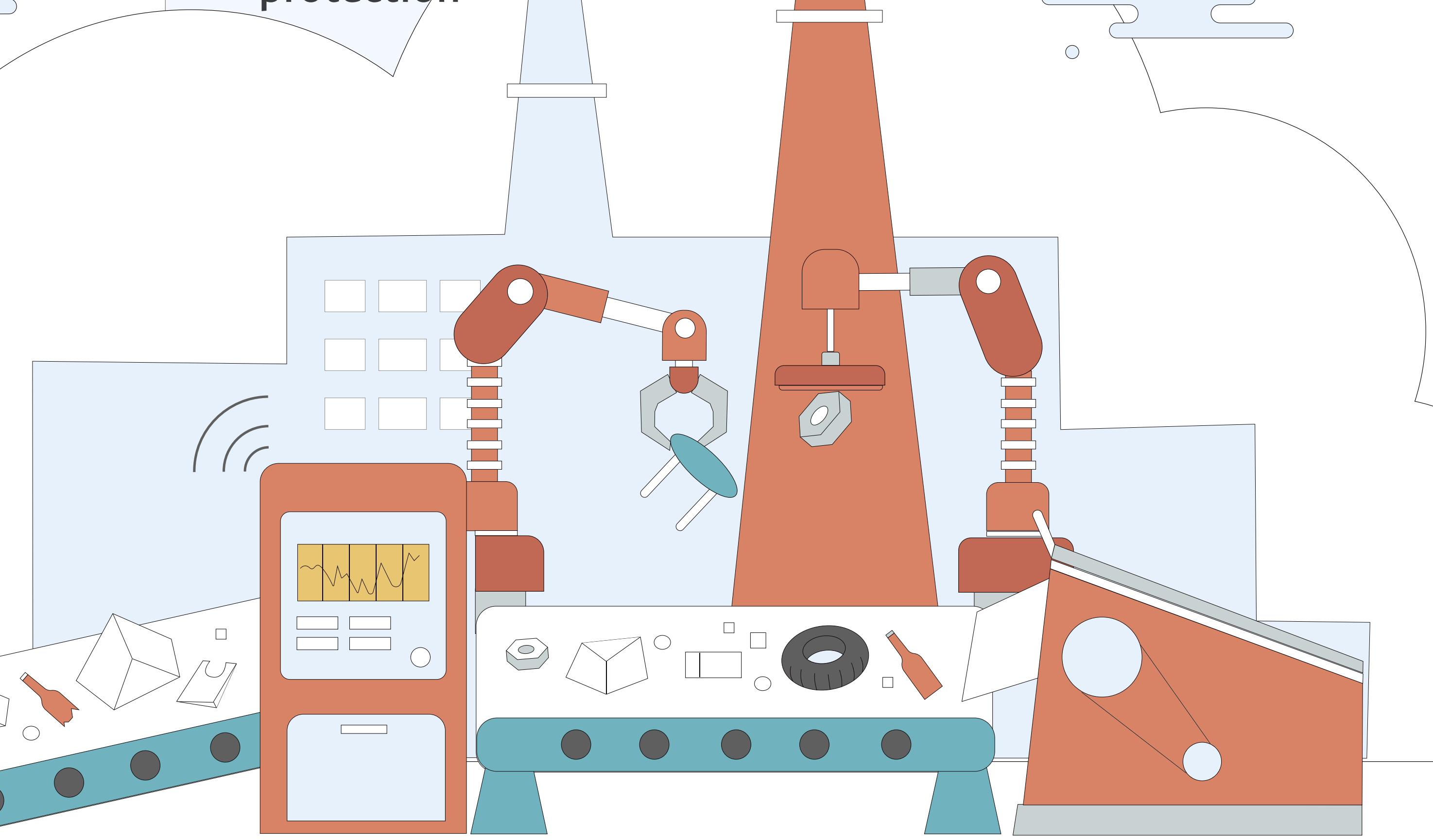
40% of companies will have access to 10 gigabit Wi-Fi networks.

New digital infrastructure



Smart government services

Intelligent environmental protection



Enterprises

New Productivity,
New Production Models,
New Resilience

Over the next decade, the world's population will age significantly and irreversibly. According to a report published by the UN[1], the global population aged 65 and over is projected to exceed 12% of the total population by 2030, while the global population aged under 25 will decrease from 41% in 2020 to 39% in 2030.

Population ageing will lead to a huge worldwide labor shortage. By 2030, we can expect a deficit of 85.2 million workers around the world – more than the current population of Germany[2]. The size of the workforce is an important factor in economic growth for every country. Take manufacturing for example. By 2030, this sector is estimated to face a global labor shortage of 7.9 million workers, leading to unrealized output of US$607.14 billion[3].

Consumer demand is set to become much more diverse, which will profoundly change production models, forcing businesses to innovate. Companies that aim to grow and expand their business will need to capture, stimulate, and nurture the increasingly diverse consumer demand. Companies of the future must rapidly respond to new consumer demands by launching products with innovative functions. For example, as the "singles economy" gains traction, companies can rapidly adjust their products by targeting solo dining, small home appliances, and mini karaoke booths. In addition, companies need to take the initiative and stimulate demand through emotional appeal, and rapidly produce combinatorial designs for product appearance, images, and implications. For example, they can customize limited-edition products or launch co-branded products within the shortest time possible.

Black swan events pose new challenges to business continuity. For example, the global spread of COVID-19 has had a negative impact on the economy, resulting in factory shutdowns, material shortages, and disruptions to global logistics and supply chains. Due to the pandemic, it is estimated that global GDP suffered US$3.94 trillion in lost economic output in 2020[4]. According to a survey by McKinsey, although economic recovery is gaining momentum in many parts of the world, supply chain disruptions are the most-common risk to company growth, which respondents now say pose a greater risk than when asked in prior surveys[5]. Therefore, determining how to enhance supply chain resilience is now a vital challenge that companies must take very seriously.

Direction for exploration: Bringing unmanned operations to manufacturing and services to make up for labor shortages

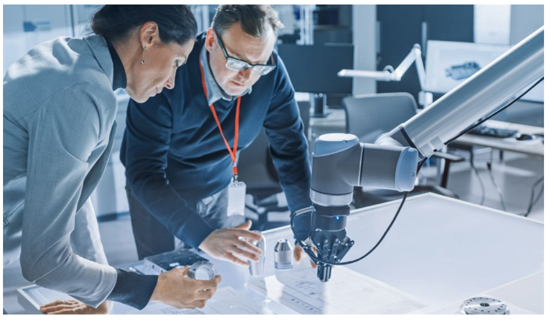
In order to expand, companies need to promptly seize business opportunities. When they receive a large order, they must quickly expand their production capacity. However, more and more companies are missing out on opportunities due to labor shortages. This is where new forms of productivity come in.

People are also trying to introduce new forms of productivity to help solve chronic issues in education, healthcare, and many industries, such as unequal distribution of resources and shortages of professional talent.

Snapshot from the future: Collaborative robots

Collaborative robots are a type of industrial robot. They were initially designed to meet the customized and flexible manufacturing requirements of small- and medium-sized enterprises, and perfectly align with the development trends of the manufacturing industry. Collaborative robots are suitable for jobs that people are unwilling to do, such as highly repetitive work like sorting and packaging. Collaborative robots have several unique advantages:

Safer: Collaborative robots are compact and intelligent, and their sophisticated sensors enable them to stop in an instant. Unlike traditional industrial robots, collaborative robots do not need to be isolated by physical fences. The scope of their movement is restricted by virtual digital fences. This means they can be placed at any



location along a production line on demand, and work closely together with human workers on the production line to get the work done.

Faster and more flexible deployment: Traditional industrial robots require professionals to plan and program their movement paths and actions, so they take a long time to deploy and are very costly. In contrast, collaborative robots feature user-friendly programming, such as programming by demonstration, natural language processing, and visual guidance. They can be placed in new positions at any time, and programing and commissioning can be completed rapidly, so they can start working very quickly.

Lower total cost of ownership (TCO) and shorter payback period: The price and annual maintenance cost of collaborative robots are significantly lower than those of traditional industrial robots. According to China's Forward Industry Research Institute, the average selling price of collaborative robots has halved over the past several years[6]. As collaborative robots are adopted more widely, their price will fall even further, meaning that investment in these robots will quickly turn cash flow positive.

Collaborative robots are currently most widely used in the manufacturing of computers, communications equipment, consumer electronics products, and automobiles. They are also starting to be used in the medical industry for analysis and testing, liberating medical professionals from repetitive and time-consuming procedures (e.g., urinalysis) and reducing the risk of infection among medical workers by taking care of tasks like throat swabs.

Snapshot from the future: Autonomous mobile robots

Autonomous mobile robots (AMRs) are a key enabler to help the manufacturing industry become flexible and intelligent. They will reshape the production, warehousing, and logistics processes.

AMRs generally need rich environmental awareness. They feature dynamic route planning, flexible obstacle avoidance, and global positioning. The AMRs used in industrial manufacturing and logistics are mainly powered by the simultaneous localization and mapping (SLAM) technology to enable autonomous navigation. Their environment does not need to be tagged to enable them to navigate[7].

On production lines, AMRs make automated and unmanned logistics possible. This includes unmanned execution; unmanned interaction between AMRs and other equipment for material collection, feeding, and unloading; and unmanned material handling.

In warehouses, AMRs implement goods-to-person picking, executing intelligent picking, movement, and stock-in and stock-out procedures. In this model, the control system receives an order and assigns an AMR, which then lifts the shelf containing the required goods, moves it to the operator console, and unloads the goods to complete the order. After picking is completed, the robot moves the shelf back to its original position.

The distribution and picking of materials are not confined to factory buildings; AMR systems can be expanded to an entire campus. For example, when goods are unloaded, robots can automatically move them into their designated warehouse. Goods will be automatically logged into and out of warehouses, and the movement of goods between factories or warehouses will be automatically registered. These functions require the robots to support outdoor autonomous navigation, using features such as laser navigation, visual navigation, and satellite positioning.

Snapshot from the future: AI-powered adaptive teaching

Conventional education uses the same model to deliver the same course content to different students. AI can transform this industry. It analyzes learning models and individual differences between students. This improves the quality of education and makes it possible to teach students in accordance with their aptitudes.

For example, as technologies such as big data, cloud computing, Internet of Things (IoT), virtual reality (VR), and augmented reality (AR) evolve, AI-assisted education will break down learning and teaching behavior in a more granular way and build more robust and precise education models. VR and AR technologies can be used to present materials in a more engaging way and deliver interactions that suit students' personal preferences, helping students better master their course content.

AI liberates teachers from the repetitive and tedious grading of exam papers and administration, allowing them to focus on the creative work of educational research and one-to-one communication with students. With the support of huge amounts of data generated through educational activities, AI will help teachers better understand the educational needs of their students, and provide key recommendations on the most effective teaching methods and the best way to organize course content.

In schools, AI can be deployed anywhere, and can simulate the best teachers of every subject, bringing the highest-quality education and content to the most remote schools. AI-based education offers multi-channel engagement with students, including video and audio, which can help make up for the scarcity of teaching resources in some areas (for example, in understaffed schools, a teacher may have to teach four or even five different subjects). In this way, AI promotes educational equity[8].

Direction for exploration: New production models geared towards personalized needs

The role of consumers in the production process is changing remarkably. Consumers are being given more say in upstream activities, and can engage in more and more steps during the production process. Under the old model, large-scale production is the norm, where companies design and manufacture products on their own, and consumers can only choose what they want from a range of finished products. With companies now better understanding what consumers really want, they offer more and more product categories to provide consumers with more choices, but this often results in overstocking.

New models such as e-commerce and livestreaming enable companies to more directly and accurately assess customer demand and promptly adjust the size of their production runs to avoid overstocking. Companies can even plan how many resources they will need in advance to avoid overcapacity.

In the future, consumers will become directly involved in design processes. They will be able to express their opinions and even make design decisions. For example, modular designs in the flexible manufacturing process can allow consumers to mix and match and decide on the form factors or style of the products they want. Companies only start production once the customer has made their choice. This means the whole production model will become truly personalized. As modular manufacturing offers more options, consumers will be given more freedom to choose exactly what they want, ultimately leading to a fully personalized production model.

Snapshot from the future: ICT-powered flexible manufacturing

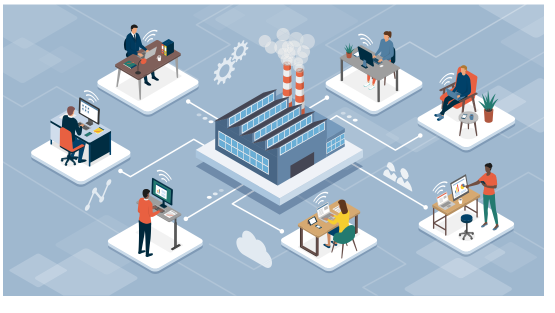
To respond to changing market conditions and set themselves apart in the face of fierce competition, companies must take the initiative and embrace new production models. That's why an increasing number of companies are looking to concepts like flexible manufacturing.

Flexible manufacturing is an advanced production model characterized by on-demand production. It helps companies become more flexible and enables them to rapidly respond to ever-changing market demand. In addition, flexible manufacturing shortens the R&D cycle, cuts R&D costs, and ensures equipment is not left idle, while reducing inventory risks and speeding up capital turnover. Therefore, it allows companies to seize market opportunities and grow sustainably. Flexible manufacturing involves the following areas:

Flexibility of product design and production line planning: After receiving an order for a new category of product, companies need to quickly conduct R&D and design, and rapidly adjust factors such as production line equipment, working procedures, processes, and batch size. This is where ICT comes in. Simulation, modeling, VR, and other ICT technologies can be used to simulate the entire new manufacturing process. This will reduce the cost of new product development and design, and support more accurate adjustment cost projections and capacity projections.

Flexibility of process: In flexible manufacturing, companies can design products based on the personalized needs of customers, or invite customers to directly participate in product design (e.g., using modular systems to enable customers to define what a product will ultimately look like). Both models require an intelligent scheduling system. The system makes automatic adjustments and provides an optimal production plan based on known features such as the factory's production capacity, order complexity, and delivery deadlines.

After a company receives an order, the scheduling system will automatically identify all universal components, custom components, and procedures and materials required to manufacture these components. By coordinating production tasks and the provisioning of materials and tools, this scheduling system maximizes the productivity of all equipment and workers in the factory so that no component will become a bottleneck in order delivery.



Flexibility of equipment: As the number of customizations and small-batch orders increases, factories must be able to switch between production processes in real time. Conventional manufacturing equipment can generally only be reconfigured by trained engineers using specific programming devices and languages. This makes switchover processes time-consuming, and does not support the kind of rapid responsiveness that companies need. In the future, ICT technologies such as visual programming, natural language interaction, and action capture will help factories reprogram equipment quickly and easily. This will help companies promptly meet consumer demand for flexible manufacturing.

Flexibility of logistics: One of the keys to flexible manufacturing is modularization, through which a large number of finished components are manufactured. This requires automated ICT technology to effectively manage warehousing and logistics, which prevents omissions and other errors in the shipment process. Take furniture producers as an example. With large-scale customization, every board, decorative strip, and handle may need its own identification code or radio frequency identification (RFID) tag to facilitate automated packing and loading, and to support traceability throughout the whole transportation and distribution process.

Traditional manufacturing followed a “product > place > people” model that forced sales to start with the site of production. As manufacturing becomes flexible, we can reverse that model to “people > place > product” so that production is based on demand, or even reduce it to a “people > product” approach. This will create a new, truly “people-centric” production model.

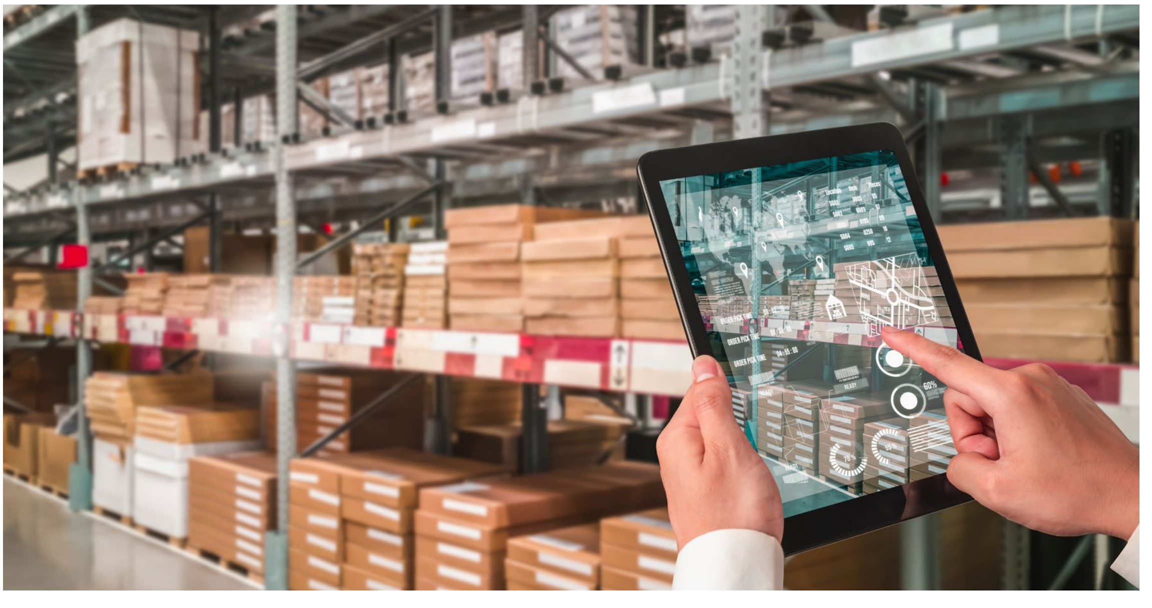
Direction for exploration: Resilient and intelligent supply chains that help enterprises respond to crises

Recent years have seen frequent black swan events, which pose new challenges to the traditional supply chain. Facing constant uncertainty, companies want to consolidate their supply systems to enhance their resilience and ensure business continuity. A global supply chain survey conducted by Allianz Research found that 94% of companies reported disruptions to their supply chains because of COVID-19, and 62% of companies said they were considering looking for new suppliers in the long term[9]. More and more companies regard building a resilient and intelligent supply chain as one of their most important strategic priorities.

Snapshot from the future: Supply chain visualization powered by digital technologies

Supply chain visualization is about using ICT technology to collect, transmit, store, and analyze upstream and downstream orders, logistics, inventories, and other related information on the supply chain, and graphically display the information. Such visualization can effectively improve the transparency and controllability of the whole supply chain and thus greatly reduce supply chain risks.

Supply chain visualization supports the tracking of materials and equipment in upstream activities. Logistics information is displayed in real time, including information on packing, goods logged in, goods logged out, and inspections; goods can even be traced throughout the production process.



With supply chain visualization, the operation data of various transportation vehicles in the logistics system is also available, and the status of these vehicles can be displayed in real time. GPS, AI, 5G, IoT, and other technologies are used to monitor the transportation process and the status of goods when they are in transit. There is a visualized scheduling center that enables the consolidation or splitting of orders at any time, and the optimization of transportation resources and routes. This enables companies to detect and rapidly respond to any logistics emergency by promptly adjusting logistics routes to ensure the timely and safe delivery of goods.

A remote monitoring system monitors the environment in warehouses in real time. This system uses various sensors to graphically display operations and maintenance (O&M) information such as temperature, humidity, dust, and smoke. This allows the timely detection of any signs of fire or water leakage, which enables prompt intervention and prevents material losses. Goods can be tracked in real time as they are logged into and out of warehouses. With the movement of goods, IoT, RFID, and QR code technologies are used to automatically identify and register goods, and the warehousing status data of goods can be accessed remotely in real time.

Snapshot from the future: From supply chain to supply network

In the traditional supply chain model, each link on the chain depends on the previous link delivering as expected. Every link could be a bottleneck that prevents the normal flow of goods down the chain. For example, if the supply of an upstream raw material provider is disrupted, downstream manufacturers will definitely be affected, resulting in inefficient operations or even a standstill for the entire supply chain[10].

With the adoption of ICT technologies such as cloud computing, IoT, big data, and AI, the supply chain will transform into a supply network. In this network, the upstream materials required by every link have multiple alternative sources, and they can be sourced through multiple routes. A multi-contact collaborative supply ecosystem will be created by enhancing the internal and external interconnectivity of enterprises. The failure of any single link will not result in paralysis across the whole supply network.

Conclusion:

New productivity will reshape production models, and enhance resilience

By 2030, digital technologies will be transforming companies. Technologies such as AI, sensors, IoT, cloud computing, 5G, and AR/VR are poised to become new drivers of productivity. They will help make up for labor shortages, so that companies can seize new business opportunities and expand their possibilities.

Huawei predicts that by 2030, every 10,000 workers will work with 390 robots, and the number of VR and AR users will reach 1 billion. Also, one million companies are expected to build their own 5G private networks (including virtual private networks). In addition, cloud services are forecast to account for 87% of enterprises' application expenditures, while AI computing will account for 7% of a company's total IT investment.

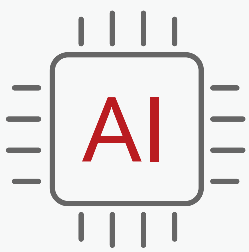
In the future, product design, process design, equipment functions, logistics, and distribution will all be reshaped to become more flexible and serve new people-centric production models. As 3D printing advances and becomes more widely adopted in commercial settings, mold manufacturing, production line adjustment, and many other activities can be eliminated. This will give consumers a much bigger voice in the design and production process, and brand-new personalized production models will be created. Powered by digitalization, supply chains will be visualized and expand into supply networks. This will enable companies to become more resilient than ever and more capable of responding to volatile markets.

Huawei predicts that by 2030,

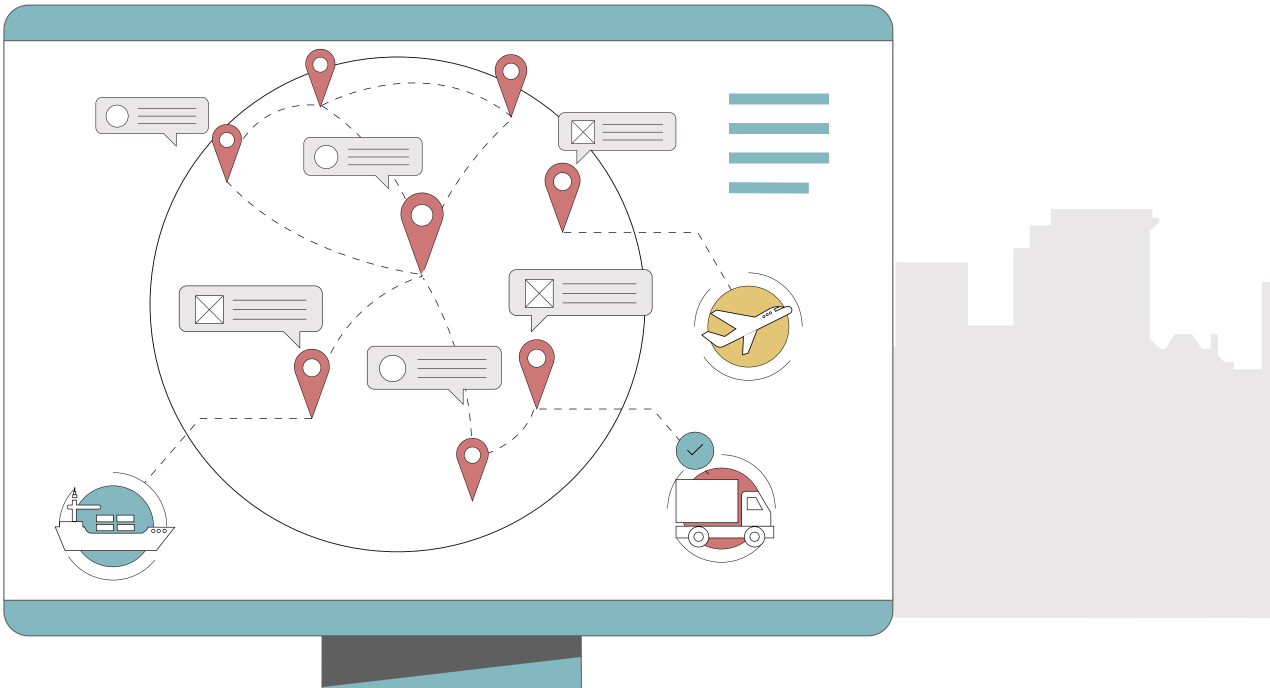
Every **10,000** workers will work with **390** robots.

1 million companies are expected to build their own 5G private networks (including virtual private networks).

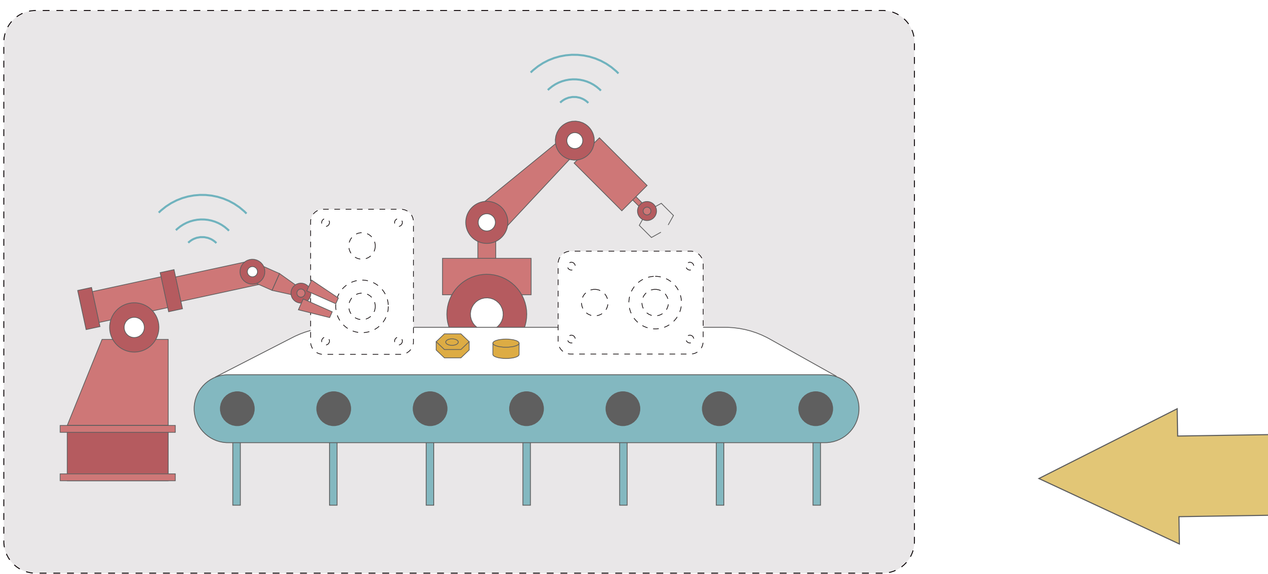
Cloud services are forecast to account for **87%** of enterprises' application expenditures.



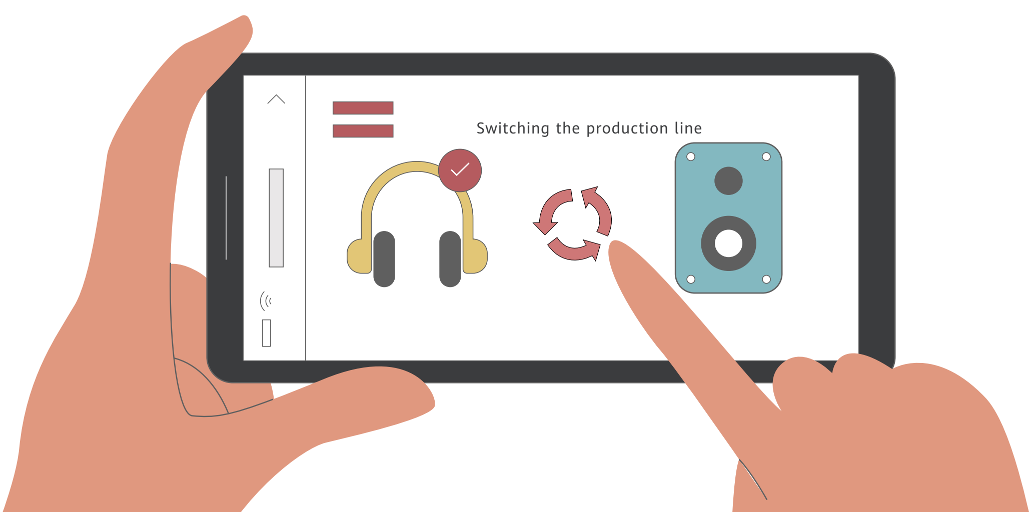
AI computing will account for **7%** of a company’ s total IT investment.



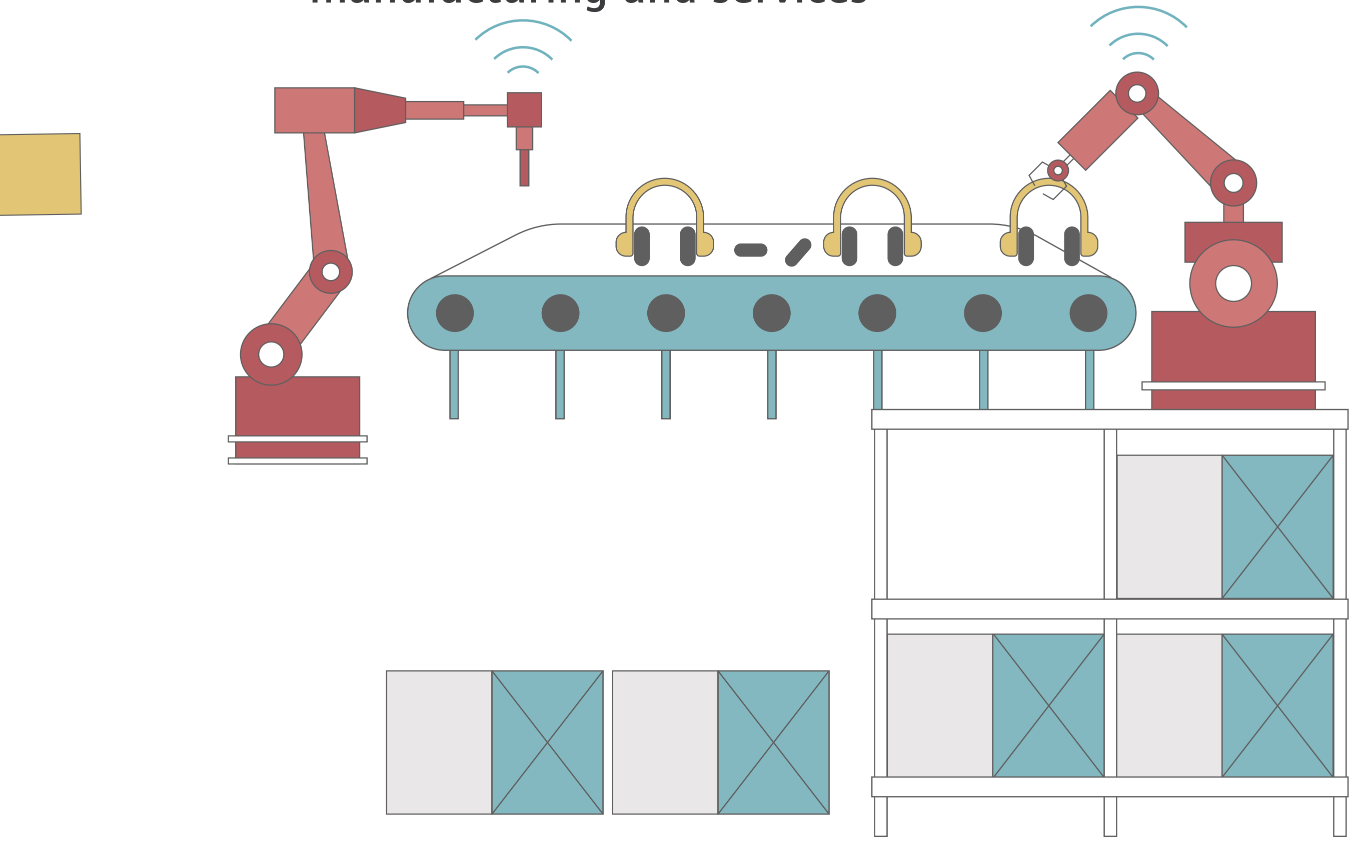
Intelligent supply chains

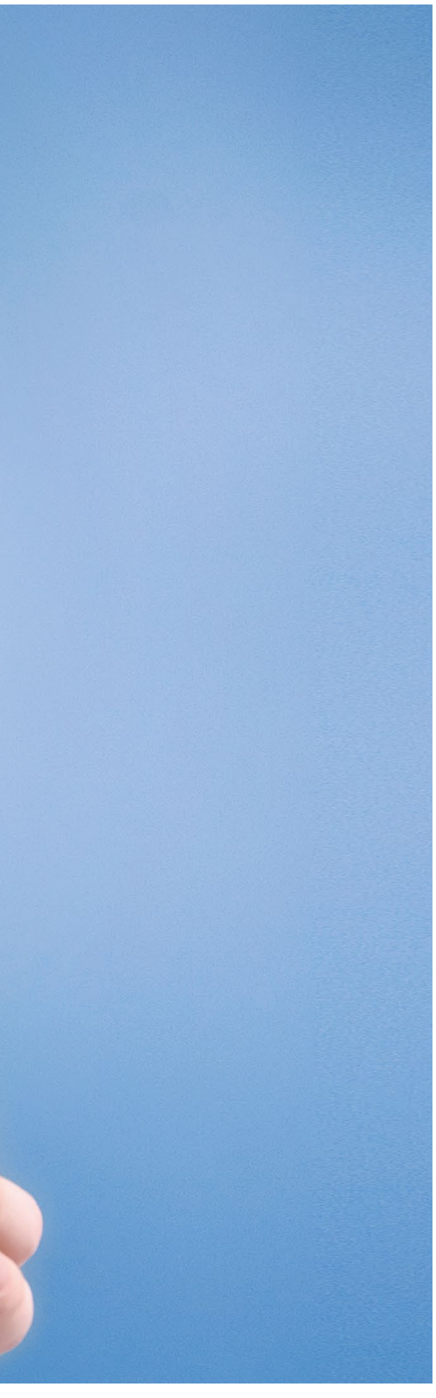


New production models



Unmanned operations in manufacturing and services





Energy

Intelligent, Green Energy for a Better Planet

Climate change is becoming more serious every day. The past decade (2011–2020) was the warmest on record, with the global mean temperature 1.2°C above pre-industrial (1850–1900) levels[1]. Global temperatures have now risen to the point where it has begun to threaten human society. Global carbon dioxide (CO_2) concentrations have reached new highs[2], with globally averaged mole fractions of CO_2 at 410 parts per million (ppm). This means that CO_2 now makes up more than 0.04% of the atmosphere. The oceans are absorbing around 23% of anthropogenic emissions of CO_2, and the resultant acidification of the oceans impacts oceanic organisms and ecosystems[3].

In a large region of the Siberian Arctic, the temperature in 2020 was 3°C above average. In Verkhoyansk, Russia, the mercury climbed to 38°C, setting a new record high temperature for the Arctic Circle[4]. Over the 2010–2019 period, weather-related events triggered an estimated 23.1 million displacements of people each year[5].

Climate change is taking its toll on economy as well. The International Monetary Fund found that for a medium- or low-income developing country with an annual average temperature of 25°C, the effect of a 1°C increase in temperature leads to a growth decrease of 1.2%[6].

Climate change is a global challenge, and many countries have come together to tackle it. At the UN Climate Conference (COP 21) in 2015, parties to the Paris Agreement agreed to intensify efforts to limit global warming to well below 2°C, preferably to 1.5°C, compared to pre-industrial levels, and set the goal of reaching net zero CO_2 emissions globally around 2050. In other words, by the midpoint of this century, the CO_2 emitted by human activities needs to be matched by the CO_2 deliberately taken out of the atmosphere. At the 75th UN General Assembly in September 2020, China pledged to peak its carbon emissions by 2030 and achieve carbon neutrality by 2060.

However, the Emissions Gap Report 2020

by the UN Environment Programme (UNEP) shows that annual emissions need to be 15 gigatons of CO_2-equivalent ($GtCO_2e$) lower than current unconditional Nationally Determined Contributions (NDCs) imply for a 2°C goal by 2030. Limiting the rise in global temperatures to 1.5°C would require an even greater decrease in global emissions by 2030[7].

Concerted efforts are needed to combat climate change and drive the transformation of the global energy mix in three areas: energy supply, consumption, and carbon fixation. On the supply side, renewable energy should be used wherever possible as a cleaner alternative to fossil fuels, for power generation and hydrogen production. This means a shift in the energy production model. It is estimated that the renewable energy share in power generation, which currently stands at 26%, will increase to 42% by 2030[8]. On the consumption side, fossil fuels will need to give way to electricity in the transport, industrial, agricultural, and construction sectors, changing the way energy is used. The share of electricity in final energy consumption is expected to increase from 20% to 30% by 2030[9]. Carbon fixation is another option. If some carbon emissions prove unavoidable, technologies such as soil carbon sequestration and carbon capture and storage can be harnessed, alongside ecological improvement efforts, to remove that carbon from the atmosphere.

As the share of renewables in energy networks continues to increase, challenging the conventional architecture of the energy industry and energy supply chains, a paradigm shift is occurring. With the increasing complexity of energy networks and the increasing digitalization of the energy sector, ICT technologies have become an important part of de-carbonization solutions. The key questions for global warming now are: How can we further increase the share of renewables in the energy mix? How can we adapt to the new energy mix? And how can we fully harness the power of ICT technologies?

Direction for exploration: Renewable electricity generation: Floating power plants

In 2020, the worldwide energy installed capacity from renewable sources increased by 280 gigawatts (GW) or 45%. Of this, 162 GW was contributed by solar power, an increase of 50%, and 114 GW by wind, an increase of more than 90%[10]. By 2050, solar and wind power will account for 60%[11] of the world's total electricity generation. It is estimated that in China alone, the total installed capacity for solar and wind power will reach above 1.2 billion kWh[12] , and 50% of the country's total generation in 2030 is expected to come from non-fossil sources[13]. The German Renewable Energy Federation (BEE), updating its 2030 scenario, calculates that a 77% renewables share of gross electricity demand is required in 2030, in order to achieve climate targets[14] .

The rapid development of inland wind and solar projects is forcing us to confront problems such as shortage of land, distance from electrical load centers, reduced efficiency of solar photovoltaic (PV) systems under high temperatures, and biodiversity loss. A new trend for the future, particularly apparent in island nations, is building wind and solar power installations offshore to take advantage of the excellent geographical features and abundant space of near-shore locations.

Snapshot from the future: Offshore wind, a promising energy source for the future

Some European countries are actively exploring offshore power generation. In 2020, the installed offshore wind capacity in the UK and Germany exceeded 18 GW, accounting for 51% of the world's offshore wind capacity[15]. Denmark is also an active player, with 15% of its electricity generated by offshore wind in 2018. Offshore wind energy still provides only 0.3% of the electricity globally[16], and there is huge room for expansion. Thanks to a large number of technology innovations that have reduced the installation and operating costs of offshore wind farms, offshore wind is set to grow rapidly.

Offshore locations offer higher wind intensity and offshore wind turbines are productive for a greater proportion of the time. Thanks to new technologies, offshore wind turbines can be larger than their onshore counterparts, and consequently have a higher capacity factor.

$$P=1/2\ \rho AV^3\ Cp$$

The equation above is used to calculate the power output of a wind turbine. The generated power, P, is proportional to both the cube of the wind speed, V, and to the swept area of the turbine, A. Offshore wind is better than inland wind, because when wind flows over rough ground surfaces or obstacles, it changes speed and direction. Sea surfaces are less rough and there are fewer obstacles. On average, the wind 10 km offshore is 25% faster than wind at the shoreline[17].

In addition, offshore wind is less turbulent and wind direction is more consistent. As a result, turbines suffer less fatigue, and the service life of offshore wind power equipment is longer.

The swept area of a wind turbine depends on the diameter of the rotor. In 2021, offshore wind turbines with a rotor diameter of 164 meters and a capacity of 10 megawatts (MW) became available. By 2030, an offshore wind turbine is expected to have an average rotor diameter of 230–250 meters and a capacity of 15–20 MW[18]. In contrast, inland wind turbines in 2021 have a rotor diameter of approximately 158 meters and a capacity of 5.3 MW, and are expected to reach a diameter of 170 meters and a capacity of 5.3 MW by 2025[19]. The capacity of offshore wind turbines can be 3 to 4 times[20] greater than that of inland wind turbines.

There are fewer calm periods at sea, so offshore wind turbines can generate power for 3,000 hours a year, compared to 2,000 hours a year for inland counterparts[21], which makes for more efficient use of generator capacity. With technology advances, the capacity factor of offshore wind power can be 40–50%, higher than that of inland wind and twice that of PV systems. In many areas, the capacity factor of offshore wind is close to that of natural gas and coal[22], so offshore wind has the potential to be a baseload technology[23].

Currently, offshore wind turbines are mainly deployed in shallow water areas less than 40 meters deep, within 80 km of the coast. They are fixed by single piles[24]. New floating turbine technologies offer a new alternative with simpler installation and lower cost[25]. Floating turbines can be installed in water up to 60 meters deep. And the new high-voltage direct current

(HVDC) technology offers a more cost-effective solution for transmission at a distance of 80–150 kilometers from the coast. These innovations have greatly expanded the potential of offshore wind.

Other innovations have led to a significant reduction in the cost of offshore wind installations, and the offshore power generation cost in 2040 is expected to be 60% lower than in 2019[26]. In Europe, offshore wind will soon outperform natural gas in terms of cost and be on par with solar PV and inland wind[27]. The Global Wind Energy Council (GWEC) forecasts that global offshore wind capacity will increase from 29.1 GW today to 234 GW by 2030[28].

Annual installations of offshore wind capacity are expected to grow at 31.5% per year over the next five years[29]. The IEA forecasts that offshore wind will become Europe's largest single source of electricity by 2040[30]. This is a boom time for offshore wind power.

Snapshot from the future: Floating PV, the latest trend in Solar PV

According to the Snapshot of Global PV Markets 2021 by the IEA, the total installed capacity of photovoltaics at the end of 2020 was 760.4 GW[31]. In 2020, solar PV accounted for approximately 42%[32] of the total power generation from all new renewable energy sources. Large inland PV power plants are the most common form of PV installation, but there are a number of problems associated with inland solar farms: land acquisition, high costs, and poor performance under high temperatures. Floating PV (FPV) is a new direction for solar PV.

FPV plants can be installed in near-shore marine areas, ponds, small- and medium-sized lakes, reservoirs, river basins, or flooded mining pits. There are three types of FPV installations: thin-film, floating arrays, and submerged. Thin-film PV modules use micrometer-thick solar cells made of non-silicon materials such as cadmium sulfide and gallium arsenide. Thin film is light-weight and does not require the support of a rigid pontoon structure. Submerged PV modules can be supported with or without a pontoon structure. Floating arrays must be supported by rigid pontoons.

Compared with land-based PV (LBPV) systems, installation of FPV systems on water saves land for agricultural use. The lack of obstacles on the surface of the water means less shading loss and less dust. In addition, the natural cooling potential of the body of water may enhance PV performance, due to higher wind speeds offshore, along with the presence of water.

In 2020, a research team from Utrecht University in the Netherlands simulated an FPV system on the North Sea. They found that the apparent temperature at sea was much lower than on land, due to higher relative humidity and higher wind speeds. The average ambient temperature difference between the two locations was 5.05°C, but the apparent temperature difference was nearly twice that, at 9.36°C[33]. This study found that an FPV system performs 12.96%[34] better on average on an annual basis than an LBPV system.

As the technologies mature, rapid growth is anticipated in FPV. On July 14, 2021, Singapore's Sembcorp Industries unveiled a floating solar farm deployed on the Tengeh Reservoir. With 122,000 solar panels spanning across 45 hectares (equivalent to about 45 football fields), the 60 megawatt-peak (MWp) solar farm is one of the world's largest inland floating PV systems[35]. According to Rethink Energy, the global FPV market capacity will exceed 60 GW by 2030[36]. Globally, the estimated potential capacity is 400 GW, meaning that FPV could double the current global installed capacity of solar PV[37].

The floating solar market is set to accelerate as the technologies mature, opening up new opportunities for scaling up global renewables.

Direction for exploration: Intelligent generation-grid-load-storage-consumption through the Energy Internet

The conventional electric power industry is built around large power plants which generate electricity centrally, and a large transmission and distribution network which delivers electricity to consumers. A balance is maintained between generation and demand, and pay-as-you-use is the main charging model.

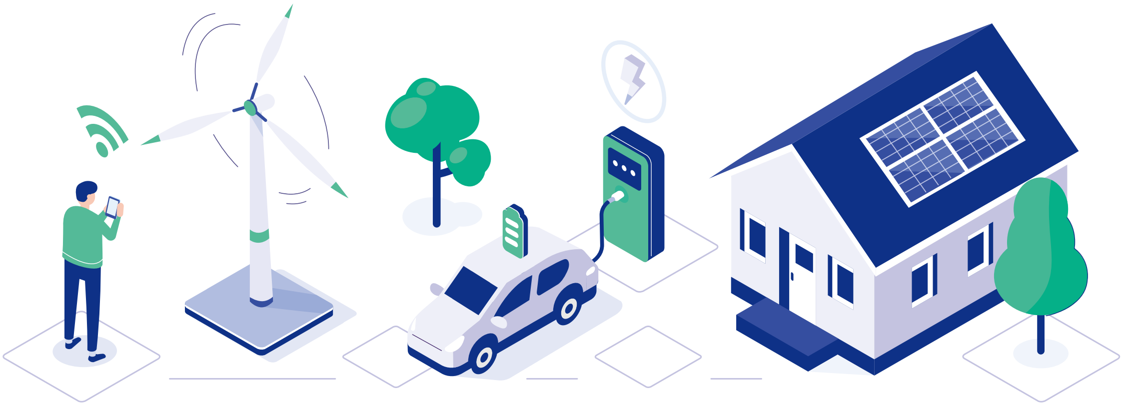
The only flexible variable in traditional power systems is on the production side. To ensure the stability of the power grid, power plants adjust how much power they are generating based on changes in the consumption load. Today, the cost of renewable installation capacity and the Levelized Costs of Energy (LCOE) for renewables is coming down, and renewable energy has become an important power source. Moving forward, distributed power systems will become more widespread, gradually replacing the old, centralized generation model. Electricity plus digital, or "diglectrification," will drive a paradigm shift. As most renewable energy sources are intermittent, future power systems will need to be more flexible in the way they balance supply and demand, and this will require the help of advanced ICT technologies.

Snapshot from the future: Virtual power plants, a paradigm shift for the power value chain

The emergence of virtual power plants (VPPs) is redrawing the boundaries between power producers and power consumers. VPPs are set to reshape the power generation value chain. IRENA defines a VPP as "a system that relies on software and a smart grid to remotely and automatically dispatch and optimize distributed energy resources. In orchestrating distributed generation, solar PV, storage systems, controllable and flexible loads, and other distributed energy resources, VPPs can provide fast-ramping ancillary services, replacing fossil fuel-based reserves."[38]

VPPs aggregate distributed heterogeneous energy sources. Distributed energy sources include innovative renewable energy generation systems, such as rooftop PV plants and small-scale wind power plants. They also include industrial and household energy systems, such as heating, ventilation and air conditioning (HVAC) systems, electric heating pumps, and battery-based hydrogen production systems. To offset the variability of renewable energy generation, VPPs may also be connected to conventional energy sources such as small gas-fired power plants, small hydroelectric plants, and diesel generators. As electric vehicles and household energy storage develop, they will also be incorporated into the heterogeneous energy equipment connected to VPPs.

Commercially, VPPs will leverage economies of scale to realize the commercial model that distributed energy producers cannot achieve alone. In order to participate in the future energy market and generate profits, distributed energy producers should be able to sense market prices in real time. Distributed new energy devices will need to respond to market changes and power grid fluctuations in real time. This requires ICT infrastructure such as interconnected networks and edge gateways or edge computing.



Producers will incur the kinds of transaction costs that come with being part of the market, such as insurance and compliance costs. These additional costs represent a barrier to market entry for distributed energy producers, but by aggregating the large number of distributed energy sources, VPPs reduce costs and generate profits through economies of scale.

VPPs may operate in either grid-oriented or user-oriented models. In the grid-oriented business scenario, VPPs provide services for power grid operators, aggregating power from distributed heterogeneous resources. Typical services include providing frequency response for power grids through aggregated power generation systems, energy storage equipment, and thermal storage facilities. In this case, VPPs treat aggregated distributed resources as a whole and are rewarded for enabling demand-side flexibility in the power grid. In the user-oriented business scenario, VPPs track energy market prices and provide users with paid services such as peak load shifting, reducing their users' power bills. In short, VPPs provide flexibility to power grids by aggregating distributed energy assets, automatically scheduling and managing distributed energy resources remotely, and tracking energy markets in real time. VPPs enable small producers of distributed energy to save costs through less power consumption and profit by delivering generated electricity to the grid. At the same time, they give greater flexibility to innovative power systems based on renewables.

The VPP model requires the collaboration of different players who bring different skills to the table. They include software, new energy, fossil fuel, and electric power companies. A case study saw technology companies partnering with the government of South Australia to install rooftop PV systems and home batteries in more than 1000 low-income homes, and connecting them to form a VPP[39]. The Australian energy market operator, AEMO, released its first review of the VPP model in 2021, arguing that the VPP model provides frequency response following critical grid problems through timely battery charging and discharging, and helps maintain grid stability[40].

In addition to helping stabilize power grids, homeowners who have installed rooftop PV systems and residential batteries have seen their electricity bills fall by as much as 20%[41]. Although there are currently a number of technical and commercial problems standing in the way of the ultimate success of the VPP model, it is expected that VPPs will have a place in the power systems of the future.

Snapshot from the future: Energy cloud as the operating system for the Energy Internet

Conventional energy networks are typically built with centralized architecture. The operator builds up equipment capacity, operates higher voltages, and expands the network to profit from economies of scale. Energy production, transmission, and consumption are separated, and there is no way to implement end-to-end management and scheduling of electricity production, transmission, and consumption. Different energy networks such as electricity, gas, heat, and cooling supplies are separated from each other, which hinders comprehensive energy efficiency. As distributed energy is increasingly deployed, energy consumers that also have production capacity will simultaneously act as producers and consumers, or prosumers, blurring the once clear boundary between energy production and consumption. Demand-side responsiveness is becoming increasingly important. The interconnection of multiple types of energy can also improve comprehensive energy efficiency and contribute to the consumption of renewable energy. There is an urgent need for integrated platforms that address these issues, and energy cloud may be a solution.

Energy cloud is a new multidisciplinary concept which is very much still evolving and crystallizing. It can be understood as the operating system of the energy Internet, and is typically characterized by convergence, openness, and intelligence.

Generation, grids, storage, and consumption of power need to be converged in an end-to-end manner. Generators now include a large number of distributed new energy sources, such as solar energy, wind energy, and biomass, as well as fossil fuel sources such as gas. The most important entity in the grids is the energy router that can direct energy flows free of constraints. Consumers include various industrial, commercial, and household facilities, such as

HVAC systems and electric heating pumps. Storage entities include various fixed energy storage devices at the generation, grid, and consumption sides, plus mobile energy storage devices such as electric vehicles. The energy cloud will also break down the boundaries between electricity, gas, heat, and cooling. By connecting multiple systems of energy sources, such as heat, gas, and cooling, the energy cloud offers a comprehensive converged system that can optimize total energy use through synergy of multiple types of energy.

With the energy cloud, the energy Internet of the future will be a democratic and open system. Energy cloud users will include consumers (e.g. electric vehicle owners and residential power equipment), businesses (e.g. zero-carbon campuses and VPPs), and governments (e.g. zero-carbon cities). The number of users will far exceed that of traditional energy users. In addition, the energy cloud needs to interconnect with third-party systems, such as carbon trading systems. Therefore, the energy cloud should be an open ecosystem. It will offer a variety of open energy data and programing interfaces to developers, enabling them to implement apps for different scenarios. In addition, an energy app store will be built on the cloud to distribute apps to users so that developers can benefit commercially from their work. Open decoupled programing will enable interconnection with third-party ecosystems, such as energy and carbon trading ecosystems, creating the potential for the emergence of new business models for the energy industry.

To enable convergence and openness, the energy cloud must be an intelligent platform. AI algorithms will make energy assets smarter. For example, AI technologies can be used to control the angle of solar panels to increase energy yield. But intelligence will also be built into the fabric of the energy cloud itself. The energy cloud will build data assets based on massive data of distributed energy sources and end-to-end information on energy generation, grids, load, storage, and consumption. It will have a user- and developer-oriented data platform based on data assets and big data modeling capabilities. Algorithms can be used to forecast distributed generation of energy and energy demand based on historical data, to dynamically respond to demand, and to analyze energy market prices in real time. With the support of efficient, intelligent technologies such as AI and big data, the energy cloud aims to enable a frictionless flow of energy from producers to consumers as they demand it. Ultimately, it will create a green, low-carbon, safe, stable, and diverse energy system.

In July 2020, the EU launched a EUR1.8 trillion economic recovery plan. The plan focuses on supporting the EU's green and digital transformation[42]. Green and digital technologies both drive economic transformation. Energy is the foundation of the digital world, and digital technologies will help the energy industry become smarter. Building an energy Internet operating system and promoting the modernization of the energy industry through digital technologies will help reduce emissions across the industry.

Direction for exploration: Efficient power use for ICT: Saving energy and cutting more emissions

According to Shaping Europe's Digital Future, the EU's digital strategy, digital solutions can increase energy efficiency and cut the use of fossil fuels[43] by tracking when and where electric power is most needed. However, the ICT industry itself also needs to undergo a green transformation. It is estimated that the ICT industry accounts for 5–9% of the world's total electricity consumption and more than 2% of total emissions[44]. Data centers and telecom networks need to improve their energy efficiency, reuse waste energy, and increase the share of renewables. The EU requires that all data centers be climate neutral, energy efficient, and sustainable by 2030, and that telecom operators adopt more transparent measures to track their environmental footprint[45].

Snapshot from the future: Low-carbon data centers and sites

According to the IEA, since 2010, the number of Internet users has doubled, global Internet traffic has increased by 12 times, and the electricity consumed by data centers and transmission networks has increased significantly. In 2019, the global electricity demand from data centers was about 200 terawatt-hours (TWh), accounting for about 0.8% of global electricity demand[46]. Data networks consumed approximately 250 TWh in 2019, accounting for approximately 1% of global electricity consumption, with mobile networks making up two-thirds of this figure[47]. Data center power consumption in China alone is expected to exceed 400 billion kWh in 2030, accounting for 3.7%[48] of the country's total power consumption. If data center power usage effectiveness (PUE) improves by 0.1, the result will be 25 billion kWh of power saved and 10 million fewer tons of carbon emissions. If all data centers use green power, carbon emissions will be reduced by 320 million tons each year[49]. Green power and PUE optimization are key measures for low-carbon data centers.

Large ICT companies have been the biggest purchasers of green power, as they strive to reduce carbon emissions in data centers and telecoms networks. Google, Facebook, Amazon, and Microsoft were the world's top four buyers of green power in 2019[50]. Amazon was the world's largest buyer of new energy in 2020, purchasing more than 5 GW, while TSMC and Verizon rose to the third and fourth places[51]. Google has announced plans for "24/7 zero-carbon" global operations by 2030. If successful, this means that the company will not achieve zero carbon by averaging emissions and offsets over a year, but by actually emitting no carbon on an hour-by-hour basis[52]. Facebook plans to achieve net zero emissions across the supply chain by 2030[53]. Microsoft declared that it will achieve negative carbon emissions by 2030 and eliminate all historical carbon emissions by 2050[54].

According to Uptime, the average PUE in data centers around the world in 2020 was 1.59[55]. This means that about 38% of the power drawn by data centers was used for cooling and other auxiliary functions. As an increasing number of high-temperature-proof servers are put into use, cooling using natural air instead of traditional chillers and air conditioners will become possible. This will reduce the energy consumption of

cooling systems, thereby decreasing the PUE. There have been a number of practices in the industry. The cooling system in a data center is powered entirely by seawater[56]. Another data center uses cold outdoor air to ensure its equipment stays at optimal temperatures. A submarine data center keeps its PUE as low as 1.07[57].

In addition to applying renewable energy and free cooling, AI is another effective way to make data centers more efficient and save energy. Sensors in data centers collect data such as temperature, power levels, pump speed, power consumption rate, and settings, which are analyzed using AI. Then, the data center operations and control thresholds are adjusted accordingly, reducing costs and increasing efficiency. AI is used in data center cooling to reduce the energy used for cooling by 40%[58].

China Unicom (Henan) uses Huawei's iCooling@AI solution, which leverages big data and AI technologies to automatically optimize energy efficiency in data centers. It can improve data center PUE by about 8–15%[59]. According to Datacenter Dynamics, the Boden Type Data Center (BTDC), an experimental data center built in Sweden with funding from the EU's Horizon 2020 programme, has achieved a PUE level of 1.01[60] by using AI algorithms to achieve synergy between the cooling system and computing loads, server fans, and temperatures, in addition to environmental cooling. As AI becomes more widespread, heterogeneous computing is on the rise and data center power density is increasing. Large data centers need holistic systems. Balancing power supply, servers, and workloads based on AI algorithms may be the next technological step required to reduce data center system PUE while continuing to support density increase.

In terms of communications networks, in February 2020, the ITU, GeSI, GSMA and SBTi published a science-based roadmap, consistent with the Paris Agreement, to reduce the greenhouse gas emissions of the ICT industry by 45% before 2030[61]. In addition to using green power in data centers, we can also use solar-grid hybrid solutions and a simplified network architecture to build greener telecoms networks with lower carbon emissions.

Telecoms and computing equipment share the same fundamental technologies. Today, Moore's Law is breaking down. Optoelectronic integration is the next step for the industry's structural reform towards higher energy efficiency. The use of optoelectronic hybrid technologies in networks, devices, and chips can continuously improve the energy efficiency of communications devices. The green telecoms networks of the future will be built to support more than 100 times today's capacity, but their total energy consumption will be no higher than that of today's networks. Conventional telecoms networks are defined by their specialist functions, which makes for fragmented operation and maintenance (O&M) and means they cannot keep pace with the latest network automation and intelligence. Networks need to be reconstructed to deliver essential services through a simplified architecture that consists of three layers: basic telecoms network, cloud network, and algorithms. This simplified network architecture will greatly reduce the complexity of the algorithms in autonomous driving networks, reducing the demand for computing, and cut O&M costs, contributing to greener, low-carbon networks.

Using green power, innovative architecture, AI algorithms, and other effective approaches, data centers and communications networks will be more energy-efficient, and will allow us to finally achieve zero carbon.

Conclusion:

ICT makes green energy smarter and the economy sustainable

The world needs to halve its emissions by 2030. On the production side, new energy sources, such as wind and PV, are being deployed at an accelerated pace, providing clean energy alternatives. On the consumption side, electricity's share is increasing, gradually phasing out fossil fuels. The ICT industry is reducing its own energy consumption and emissions, and at the same time, ICT solutions are enabling other industries to cut their carbon emissions. Huawei predicts that by 2030, renewables will account for 50% of all energy generation globally, 3,000 GW of PV plants will be in place, LCOE of PV plants will reduce to US$0.01 per kWh of electricity, and renewable energy will power 80% of digital infrastructure.

In 2030, ICT will make green energy smarter and enable a wide variety of industries to further reduce emissions. ICT will support green, low-carbon transformation and sustainable development for the global economy.

Huawei predicts that by 2030,

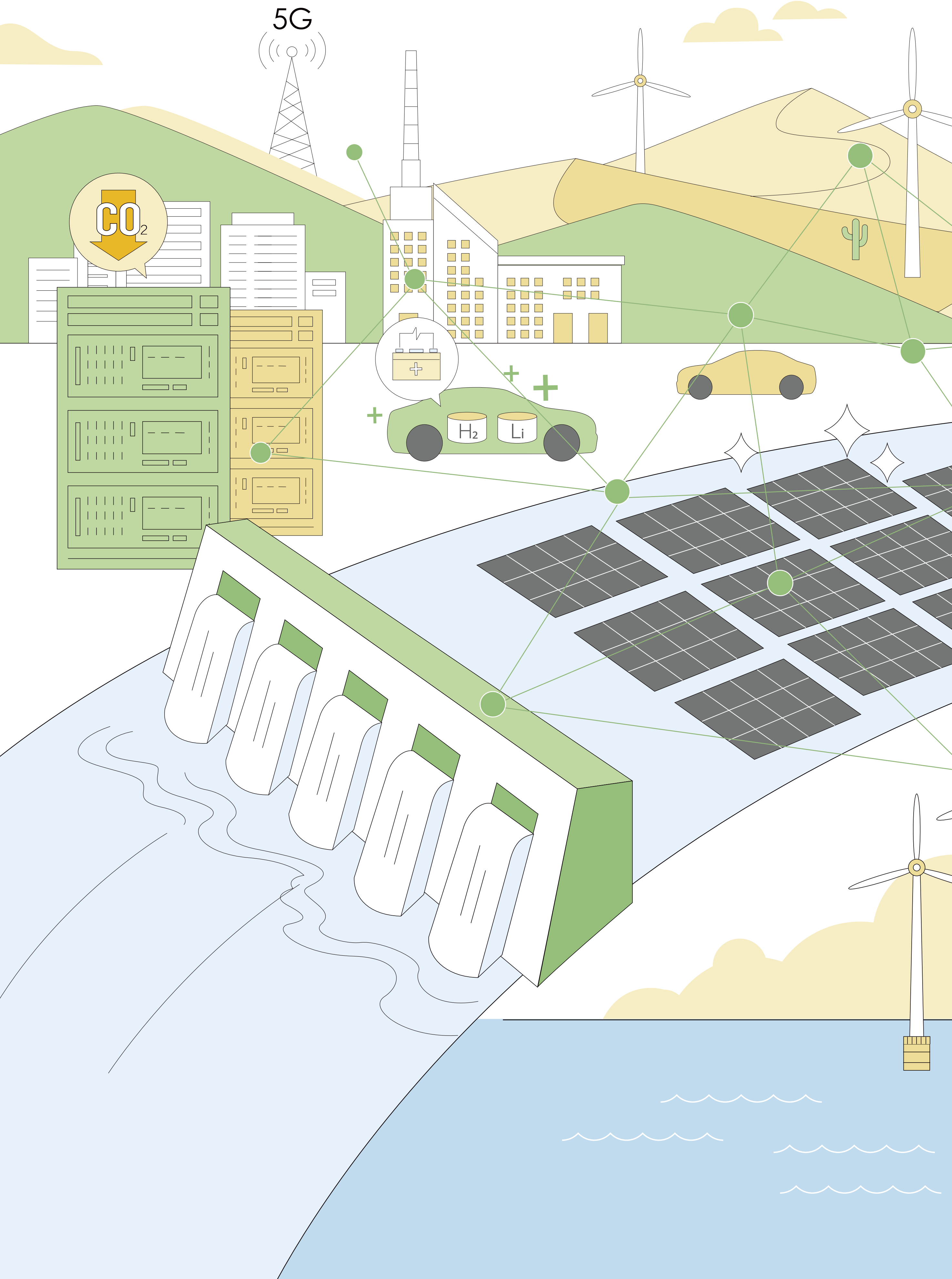
Renewables will

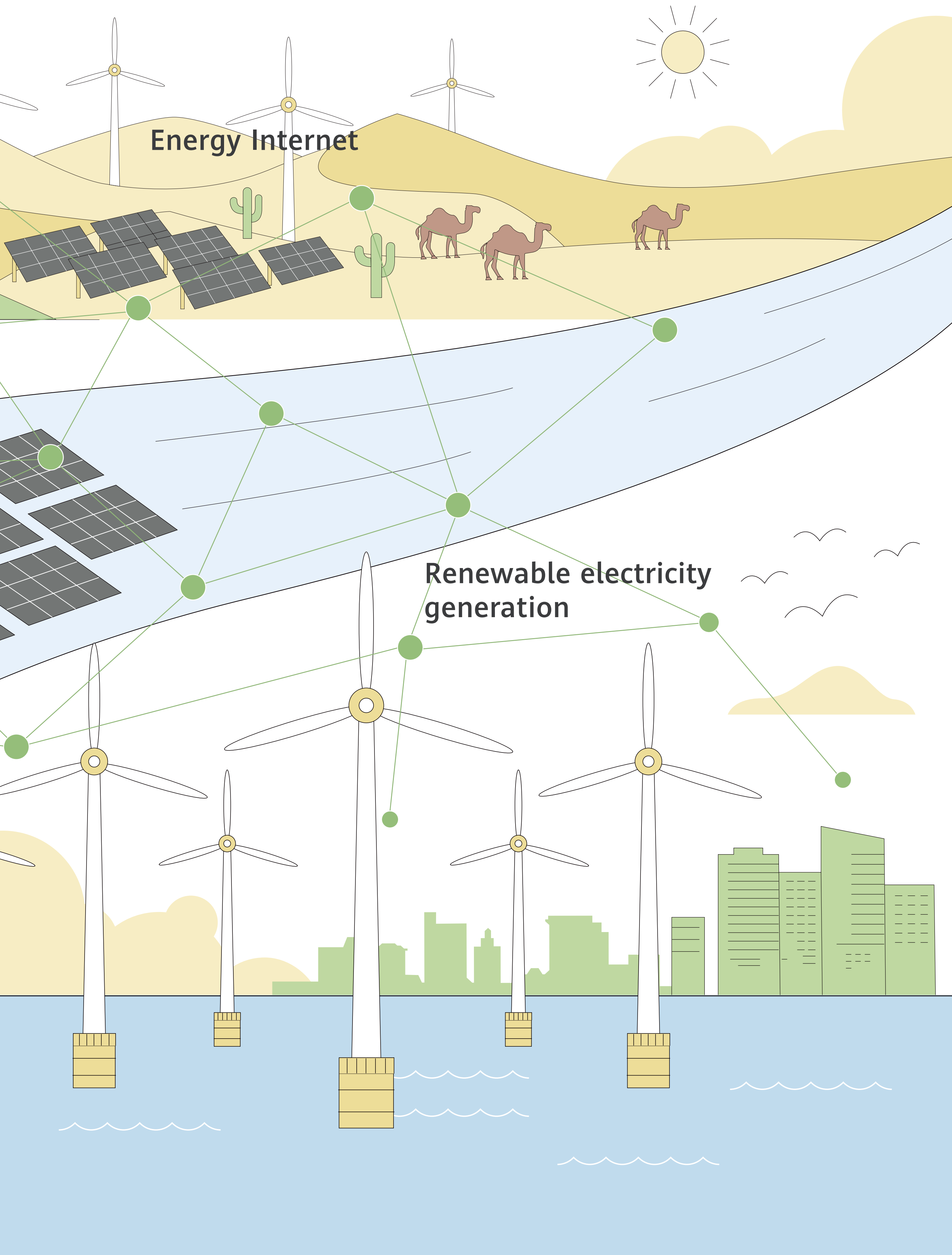
account for **50%**

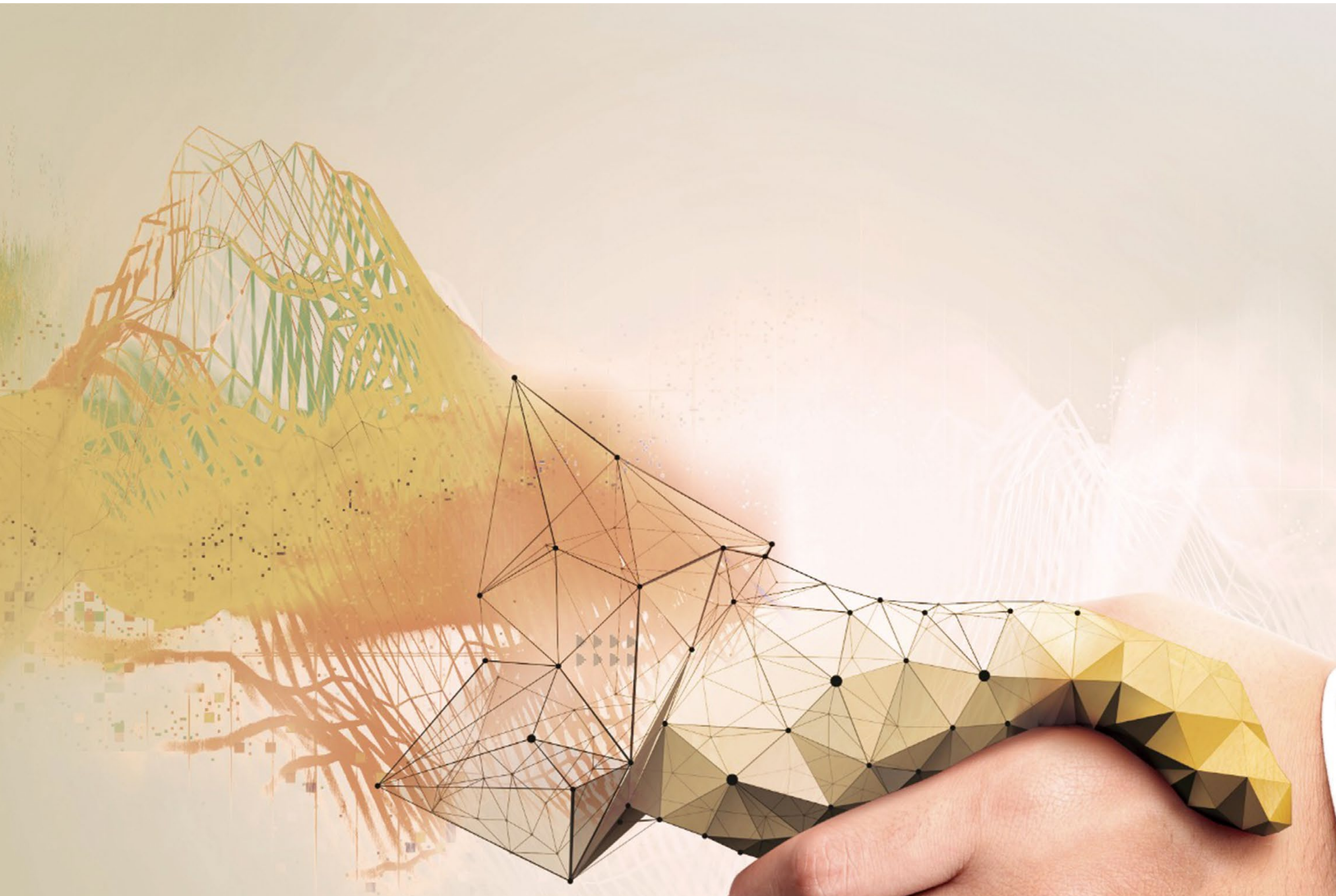
of all electricity

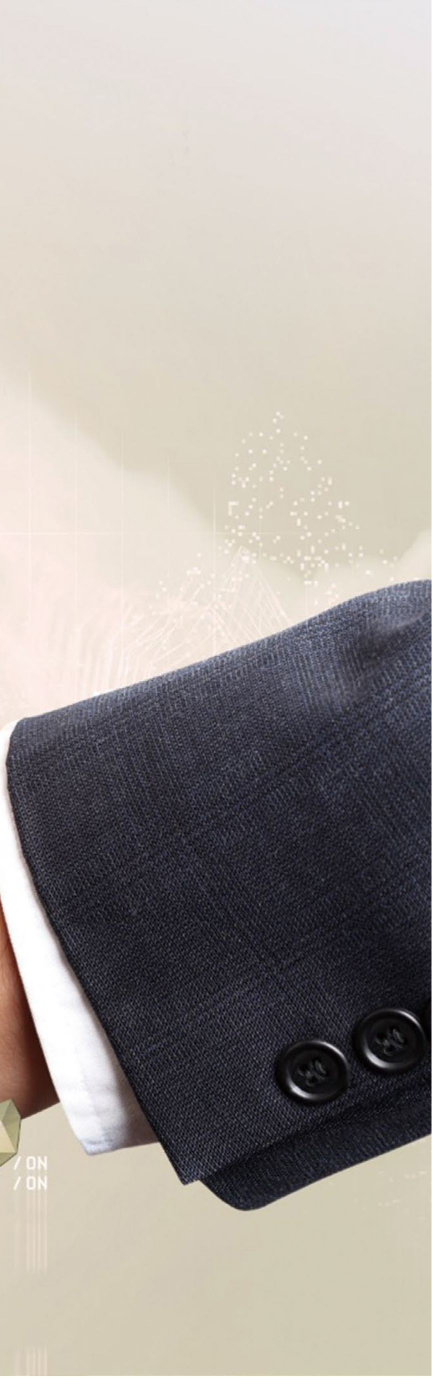
generation globally.

Efficient power use for ICT




Energy Internet
Renewable electricity generation





Digital Trust

Technologies and Regulations Shape a Trusted Digital Future

Collaboration drives progress, and the basis for collaboration is trust. In the business world, all interactions are ultimately based on trust, from market research to customer engagement, enterprise operations and management, and the supply chain. Now, digital technology is reshaping these interactions and new concepts such as the metaverse are emerging. Building digital trust has become one of the most important strategic goals for companies and other organizations. Interactions between organizations, between organizations and customers, and within organizations, are migrating to the digital world ever more quickly. Valuable digital assets are generated during these processes. However, the basis of these interactions, i.e., the trust, will be lost if information security is compromised or private information is leaked. As a result, business operations, business value (e.g., brand and market value), reputation, and public credibility will all be at risk.

Digital trust is a complex system that covers a range of areas, including privacy, security, identity, transparency, data integrity and governance, and compliance[1]. Therefore, the realization of digital trust must involve different dimensions and use a variety of tools, such as blockchain, privacy-enhancing technology, and artificial intelligence (AI). New technologies and new rules will help shape a trusted digital future.

Direction for exploration: ICT enables digital trust

For both organizations and individuals, digital assets are enabling unprecedented efficiency and convenience. However, digital assets also create high risks of information theft. Their security and integrity rely on technologies such as risk prediction, digital ledgers, encryption, and digital authentication. Research into ICT technologies for digital trust will help support data sharing and data transactions that are traceable and verifiable, with well-defined data ownership. This will help organizations and individuals fully realize the value of their data, effectively manage digital assets, and better protect core data.

Snapshot from the future: Smart contracts on the blockchain

Many companies are looking for solutions that enable contracts to be drafted and executed efficiently, and provide automated performance and neutral supervision. In 2020, the average value of a dispute in the construction sector was US$54.26 million, and the average length of a dispute was 13.4 months[2]. Not only do these disputes cause financial loss, they also disrupt the operations of the companies involved.

Smart contacts, executed on a blockchain, have been one of the most exciting areas of development in recent years.

Smart contracts were first defined by Nick Szabo in 1994 as a set of promises, specified in digital form, including protocols within which the parties perform on these promises. Thus, a smart contract is a special type of agreement that creates, verifies, and enforces performance of obligations. However, smart contracts were only a theoretical concept until the introduction of blockchain, because the technology to create them did not exist.

Blockchain-based smart contracts contain terms expressed in a digital form on a blockchain, and the recording and processing of these terms are completed on the blockchain. The code of the contract contains the obligations of the parties and all information relating to the transactions. The obligations are performed automatically when the set conditions are met, and no third parties are required. Blockchain technology allows information to be recorded and distributed, which ensures that the entire process, from contract storage and access to performance, is transparent, traceable, and non-tamperable. In addition, the decentralized technology that underpins smart contracts can help companies reduce operational costs, improve contract performance efficiency, and prevent third-party interference, which would make transactions more accurate and reliable. However, these same features of blockchain also create challenges to the widespread adoption of smart contracts. For example, if there are errors in the code of a smart contract, the errors cannot be corrected.

The decentralized nature of smart contracts means that they are often not legally enforceable, leaving them governed only by their own code.

Smart contracts have huge market potential in logistics, e-commerce, finance, insurance, and other sectors. According to Capgemini Consulting[3], smart contracts may help US consumers save US$480 to US$960 per mortgage loan, and enable banks to cut costs in the range of US$3 billion to US$11 billion annually by lowering operational costs in the US and European markets. Consumers in the US and EU could save US$45 to US$90 per year on their motor insurance premiums, and insurers would reduce the cost of settling claims by US$21 billion a year globally.

Snapshot from the future: Using AI to identify fraud and maintain organizational reputation and credibility

AI-based behaviors are increasingly human-like. Therefore, some people may use AI for deception. For example, AI-based audio and video can be used for fraud. As reported in the Wall Street Journal[4], the executive of a UK-based energy firm thought he had received a call from the CEO of the firm's parent company and sent US$243,000 to a Hungarian supplier as directed. However, after investigation, the company's insurance firm found that this was a fraudulent request that used AI to mimic the CEO's voice, and the money had been moved to Mexico and other locations. The loss was ultimately borne by the insurance firm.

In 2020, the UK's Channel 4 used deepfake technology to create a "deepfake" queen for an alternative Christmas message on Twitter[5]. This video was criticized by some media commentators and members of the public as

“disrespectful”. A report from the Brookings Institute suggests that deepfake technology will erode trust in public institutions[6].

These stories suggest that it is difficult for humans and traditional technologies to distinguish fake video and audio created using digital technologies from legitimate audio and video. However, the solution to this malicious use of AI may come from AI itself.

Neural networks for deep learning can be used to analyze natural language and images, and even to understand video and audio. This can allow us to distinguish between authentic and deepfake videos. AI can be used to detect differences between videos, or even slight differences in audio waves. It can thus identify whether a video or audio was composed using AI. Machine learning and API technology can be used to automate defense. In particular, discriminator algorithms and causal inference models can be used to automatically detect, assess, and remove fake information on the Internet, and trace it back to the data source to provide evidence for the prosecution of digital crimes.

Snapshot from the future: Privacy-enhancing computation

In the era of big data, data is the new oil. But unlike oil, data will not be depleted. The value of data will be realized again and again in different scenarios and regions by all kinds of enterprises and organizations. However, data sharing presents new challenges to security and privacy. Data mining and analytics driven by machine learning is becoming increasingly prevalent. Sectors such as finance, healthcare, and retail in particular need to guarantee data privacy as they seek to mine data, obtain its value, and share it for collaboration. As data analytics and data warehouse environments become increasingly complex, traditional data desensitization technologies are no longer sufficient. Therefore, privacy-enhancing computation (PEC) technologies are being explored as an alternative.

PEC technologies are data security technologies used to protect and enhance privacy and security during the collection, storage, search, and analysis of private information. PEC supports efficient, high-quality services by

protecting personal data from abuse, while allowing effective use of the data, and realizing its business, scientific, and social value. PEC technologies are being explored in the following areas:

Differential privacy: Random noise is injected into the database to be mixed with personal data, while statistical estimation can still be performed using the data. This method guarantees personal privacy even when the data is shared, because the original data has been scrambled.

Homomorphic encryption: This technique allows users to perform computations on encrypted data without decrypting it. When data is homomorphically encrypted, the computations on the data are also in an encrypted form. When decrypted, the output is identical to the answers that would have been obtained if the computations had been performed on the unencrypted data.

Federated learning: This method allows data to stay in companies' local servers for machine learning. Separate learning models are built after encrypted samples have been aligned, and a virtual joint model is developed based on these models. The performance of this joint model is almost identical to a model trained on data directly gathered in the conventional way.

In addition to the above-mentioned technologies, PEC technologies include a trusted execution environment (TEE), zero-knowledge proofs, k-anonymity, and l-diversity. In the future, PEC will be supported by more algorithms and widely used in more applications, helping us to find the right balance between privacy and data value.

Direction for exploration: Rules redefine digital trust

Technical measures cannot completely eliminate the risk of information leaks, cyber fraud, or other behaviors that compromise digital trust. Rules and regulations are needed for a trusted intelligent world. Personal data security is not only about individual rights; it is also an essential part of national digital strategies. Large platforms with data traffic advantages may grow into digital giants and abuse their position to collect, use, and distribute consumers' personal information. This trend will deepen the distrust between companies and their customers, and will intensify unfair competition, which will be detrimental to communities as a whole.

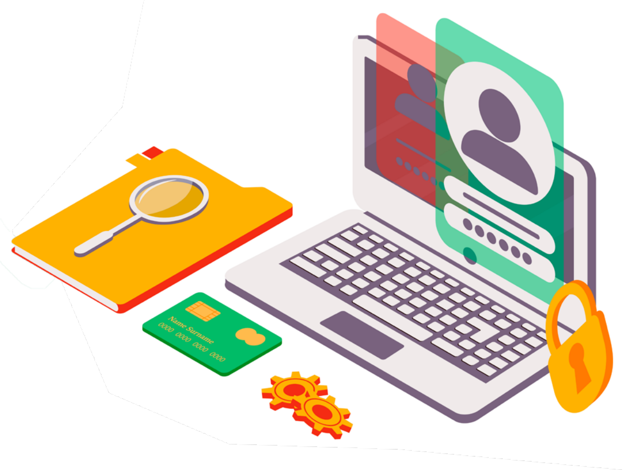
Snapshot from the future: New mechanisms for collecting pesonal information online

More and more laws and regulations concerning the over-collection of data have been passed in recent years. In the context of big data, a fair digital strategy would contain optimized mechanisms that balance the privacy of individuals and the interests of data users creating value with consumer data. The level of control data subjects have over their own personal information will be further enhanced while preserving the conventional approach of obtaining informed consent. In 2021, China promulgated its first Personal Information Protection Law. This law emphasizes multiple basic principles for protecting personal information, including openness, transparency, knowledge of purpose, and minimization. In the future, regulatory frameworks will be further refined so that users will have more knowledge and control over the ways in which their data is collected and used, and the associated risks.

Snapshot from the future: Data protection rules in international markets

The General Data Protection Regulation (GDPR) is currently the most stringent privacy and data security law in the world. It was drafted by the EU and took effect on May 25, 2018. There have been more than 281,000 data breach notifications since it went into effect[7]. Statistics show that by September 2, 2021, there had been 841 fines worth more than EUR1.287 billion, and the largest fine was EUR746 million[8].

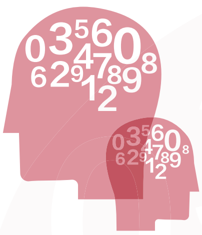
Other countries have also passed laws and regulations on data protection. In 2020, the US published its Federal Data Strategy 2020 Action Plan, which includes the goals of protecting data integrity, conveying data authenticity, and ensuring data storage security. The UAE passed a Data Protection Law for specific sectors, and the Privacy Act 2020 is in effect in New Zealand. On August 20, 2021, China passed the Personal Information Protection Law, which will take effect on November 11, 2021.



Snapshot from the future: The global rise of antitrust action in the data domain

In 2019, the US Department of Justice launched sweeping investigations into the biggest tech giants, suggesting that they had monopolized the market, suppressed competition, and violated user privacy. On May 27, 2020, Japan passed the Act on Improving Transparency and Fairness of Digital Platforms, which was designed to regulate specific digital platforms and enforce obligations to the public on those platforms. On January 19, 2021, the 10th amendment to the German Act Against Restraints of Competition entered into force. The amendment expands the scope of competition law to prevent companies from abusing their positional advantages. In 2021, the Anti-Monopoly Committee of the State Council of China published the Guidelines for Anti-monopoly in the Platform Economy. These regulations represent a global trend towards antitrust action in the data domain.

As antitrust laws are further modernized and adopted, data users and third-party companies will be granted more data rights against industry giants. This will help develop a healthy digital trust ecosystem, and prevent large platforms from committing digital security violations or engaging in other behaviors that compromise fair competition, such as illegally obtaining, abusing, and trading personal data.



Conclusion:

Building an intelligent world with digital trust together

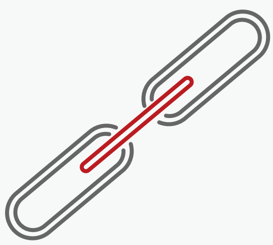
Looking forward to 2030, we will be using technologies such as blockchain and AI to better protect user privacy and data assets; to fight against digital fakes and fake news more effectively; and to reduce the risk of fraud and data theft. PEC and other technologies will enable data sharing with secure encryption, maintaining the value generated by data flows while protecting privacy. Huawei predicts that by 2030, privacy-enhanced computing technologies will be used in more than 50% of computing scenarios, and 85% of companies will have adopted blockchain technology.

At the same time, digital security laws and regulations, such as the GDPR, and the rising tide of antitrust action in the data domain will be seen in more countries. This will help build trust between individuals and organizations, and help organizations maintain legal compliance in terms of digital trust.

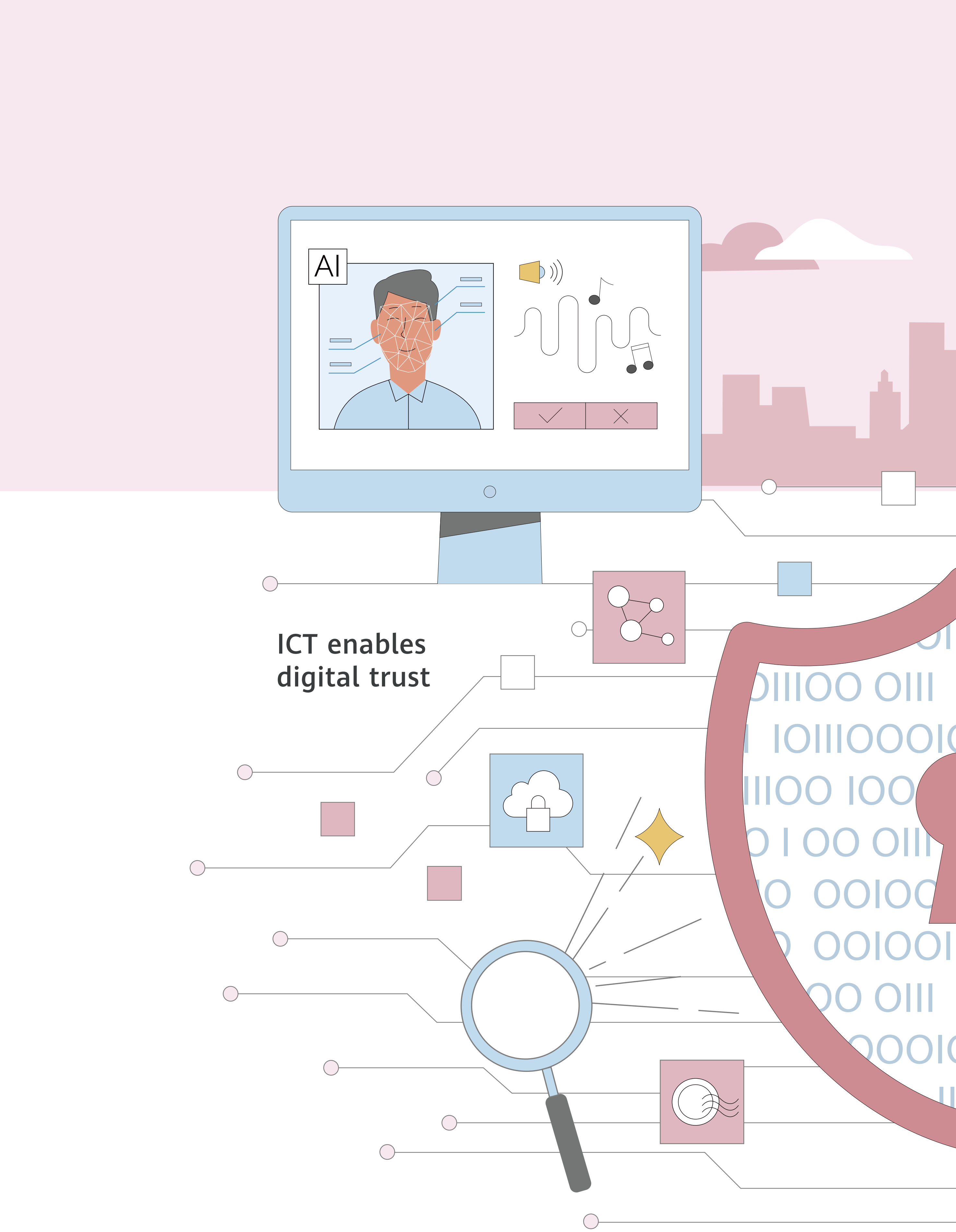
A robust digital trust ecosystem needs the cooperation of multiple parties. Companies will need to maintain their own legal compliance and manage the compliance of their partners, and they may also need to work with regulators to fight against information security violations and data monopolies, as well as to protect user data security. Companies should also proactively support public education and training to help build digital skills and awareness of data and privacy issues in the community. Working together, we can build an intelligent world with digital trust.

Huawei predicts that by 2030,

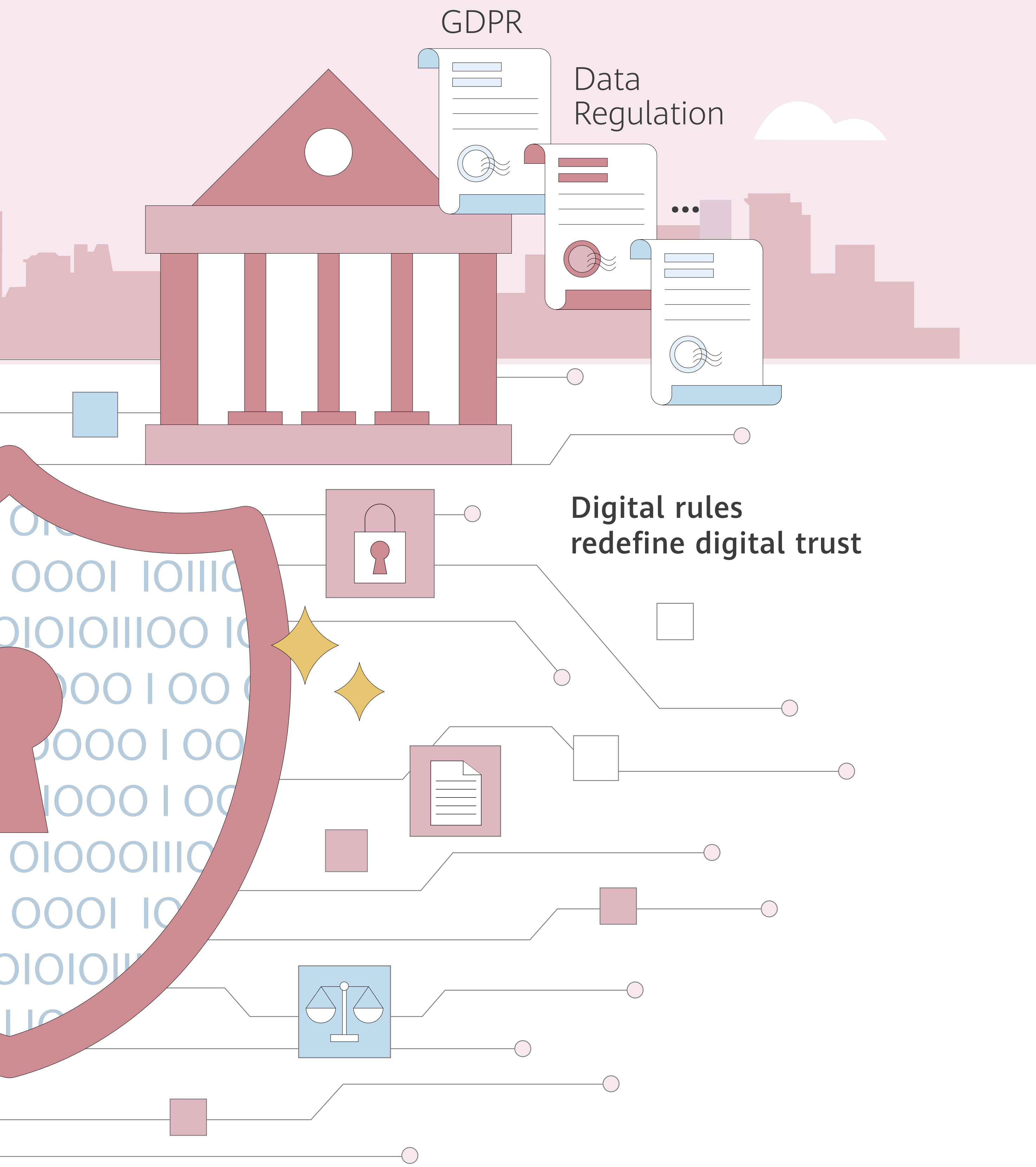
Privacy-enhanced computing technologies will be used in more than **50%** of computing scenarios.



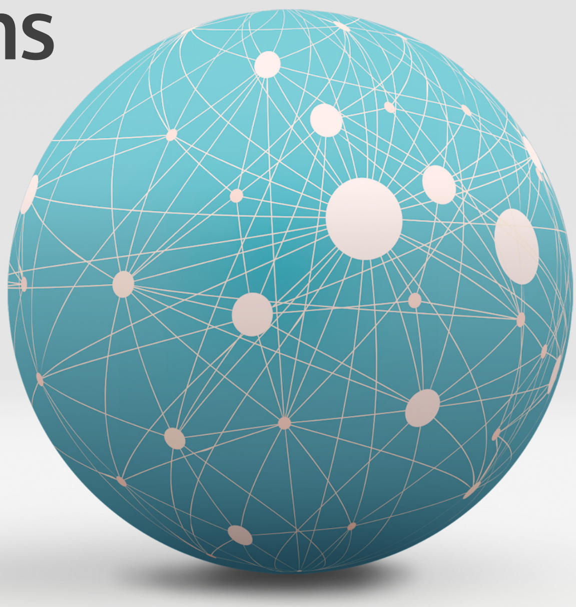
85% of enterprises will adopt blockchain technology.



ICT enables digital trust


GDPR
Data
Regulation
...
Digital rules
redefine digital trust

Communications Network 2030



Industry Trends

Going intelligent has become the general direction that the world is heading in over the coming decade. China, the EU, and the US have all published their new visions for this area.

In its Outline of the 14th Five-Year Plan (2021–2025) for National Economic and Social Development and the Long-Range Objectives Through the Year 2035, China prioritizes industry intelligence as an important area of development, and sets clear development goals for industries including manufacturing, energy, agriculture, healthcare, and education, as well as for government management.

In its 2030 Digital Compass plan, the EU articulates the following targets: By 2030, 75% of European enterprises will have taken up cloud computing services, big data, and Artificial Intelligence (AI), and more than 90% of European SMEs will reach at least a basic level of "digital intensity". To achieve these targets, the EU announced an increase in investment into energy and digital infrastructure.

In its Vision 2030 report, the US National Science Board (NSB) recommends increasing investment in data, software, computing, and networking capabilities over the next decade in order to help maintain the US's competitiveness in the digital economy.

The intelligent development of industries first requires companies to upgrade their networks. In its Industrial Internet Innovation and Development Action Plan (2021–2023), the Chinese government put forward the following measures: (1) Accelerate the network-based development of industrial equipment, drive the upgrade of enterprise Intranet, and promote the integration of information technology (IT) networks and operational technology (OT)

networks to build industrial Internet campus networks. (2) Explore the deployment of new technologies such as cloud-network synergy, deterministic networking, and Segment Routing over IPv6 (SRv6). In its Digitising European Industry platform plan, the EU considers nanophotonics, AI, 5G, and Internet of Things (IoT) to be key enablers of future industrial networks, and plans to increase investment in these technologies in order to stay ahead in the future.

As industries increasingly adopt intelligent technologies, leading telecom carriers around the world are taking action and beginning to explore how they can fully unleash the potential of connectivity in this process. For example:

- China Mobile has unveiled a "5G + AICDE" development strategy, where AICDE stands for AI, IoT, cloud computing, big data, and edge computing.
- China Telecom has set out the goal of building an integrated cloud-network architecture by 2030.
- China Unicom published its CUBE-Net 3.0 strategy, which articulates a new development direction that combines connectivity, computing, and intelligence.
- In its outlook for 2030, Deutsche Telekom aims to become the leading digital enabler in the B2B market, providing comprehensive network, IoT, cloud, and digital services.

A survey conducted by GSMA shows that B2B, cloud, and IoT services that target industry, finance, health, energy, and agriculture will be the most promising areas for carriers worldwide to fully unleash the potential of their connectivity portfolios.

In the world of 2030, many amazing things that we can only dream of today will be a reality.

With highly sensitive biosensors and intelligent hardware connected through broadband networks, we can obtain and track the indicators of our physical health in real time, and securely store massive amounts of health data in the cloud. This will allow us to proactively manage our own health and reduce our dependence on doctors, thus improving our health and quality of life.

New technologies, such as home broadband that supports speeds of over 10 Gbit/s and holographic communications, will enable more intuitive human-machine interactions. An air-ground cubic network will connect all means of transportation, facilitating easy, smart, and low-carbon travel. Sensing technology, 10-gigabit wired and wireless broadband, inclusive AI, and applications that target numerous industries will be available everywhere, allowing us to build urban digital infrastructure that improves the quality of city life.

With Harmonized Communication and Sensing (HCS), automation, and intelligence technologies, we will be able to efficiently protect our environment. New types of labor, such as collaborative robots, automated mobile robots (AMRs), and digital labor, can be adopted in tandem with the industrial Internet to increase accuracy and decrease costs throughout the whole process from demand to production and delivery, while also improving the resilience of the manufacturing industry.

Energy IoT can be integrated into smart grids to form a green energy Internet and fully digitalize all activities, including generation, grid, load, and storage. Zero-carbon data centers and zero-carbon communications sites may soon become a reality. We can also guarantee digital security and trustworthiness by combining blockchain, digital watermarking, AI-driven anti-counterfeiting, privacy-enhancing computing, and endogenous network security.

In 2030, communications networks will evolve from connecting billions of people to connecting hundreds of billions of things, and face many challenges along the way.

First, the scale of communications networks will continue expanding. This means network management will become even more complex, so networks must become more intelligent. Over the next decade, how can we innovate

in software technology to prevent operation & maintenance (O&M) costs from rising in step with the continuous expansion of network scale? This poses a daunting challenge.

Second, IoT scenarios such as unattended operations in industrial and agricultural settings and self-driving vehicles will require carriers to further improve the coverage, quality assurance, security, and trustworthiness of their networks. Over the next decade, how can we innovate in protocols and algorithms to enable networks to carry multiple types of services while meeting the requirements for high quality and flexibility? This will be a very challenging task.

Third, although Moore's law has held true for decades, the semiconductor industry is now struggling to maintain that pace of improvement, and new technologies like quantum computing are not yet mature. Meanwhile, demand for computing power, storage capacity, and network energy efficiency continues to grow, and these factors are increasingly becoming bottlenecks. Over the next decade, how can we innovate in fundamental technologies to build a green, low-carbon network and increase network capacity by dozens of times without increasing energy consumption? This is another extremely challenging task that lies ahead of us.

Communications networks are one of the major forces driving the world forward. The development of communications networks kicked off during the first Industrial Revolution and, unlike traditional industries, it still shows no signs of slowing down after nearly two centuries. In fact, the pace of development of communications technologies has been particularly rapid in recent decades. Both the evolution from 2G to 5G and the shift from the asymmetric digital subscriber lines (ADSL) to gigabit optical home broadband took just 30 years. Over the next decade, we will witness the emergence of new use cases and scenarios for communications technologies and fully embrace an intelligent world.

Future Network Use Cases

Since Samuel Morse invented the electric telegraph in 1837, communications networks have come a long way, moving from connecting individuals and homes to connecting organizations. In today's environment of diverse and rapidly changing services, it takes continuous innovation for communications networks to keep up with the needs of customers. To meet the rich and diverse business needs that will arise in the intelligent world of the next 10 years, communications networks will need to go beyond connecting individuals, to connect multiple perception, display, and computing resources related to each individual. In the near future, networks will have to connect home users as well as home appliances, vehicles, and content resources, while organizations will expect networks to do more than just create connections between employees – they must also connect an organization's machines, edge computing nodes, and cloud resources.

The scope of network connections is expanding, business needs are changing, and the industry has reached a consensus that, over the next 10 years, networks will evolve from 5G to 5.5G/6G, from F5G to F5.5G/F6G, and from IPv4/Multiprotocol Label Switching (MPLS) to IPv6+, and the autonomous driving network will evolve from L2 to L5. In addition, new use cases will continue to emerge.

Next-Generation Human-Machine Interaction Network: A Human-centric Hyperreal Experience

In a world of cold machines, it is up to human beings to adapt to the machines. With the wide use of the automobile, we learned to work with pedals and a gearstick. In the PC era, we learned to use the mouse and keyboard. In the smartphone era, we learned to use touchscreens.

However, with sufficiently advanced levels of intelligence, it is possible to turn this paradigm on its head and have machines adapt to the needs of their human users. Intelligent machines (e.g.,

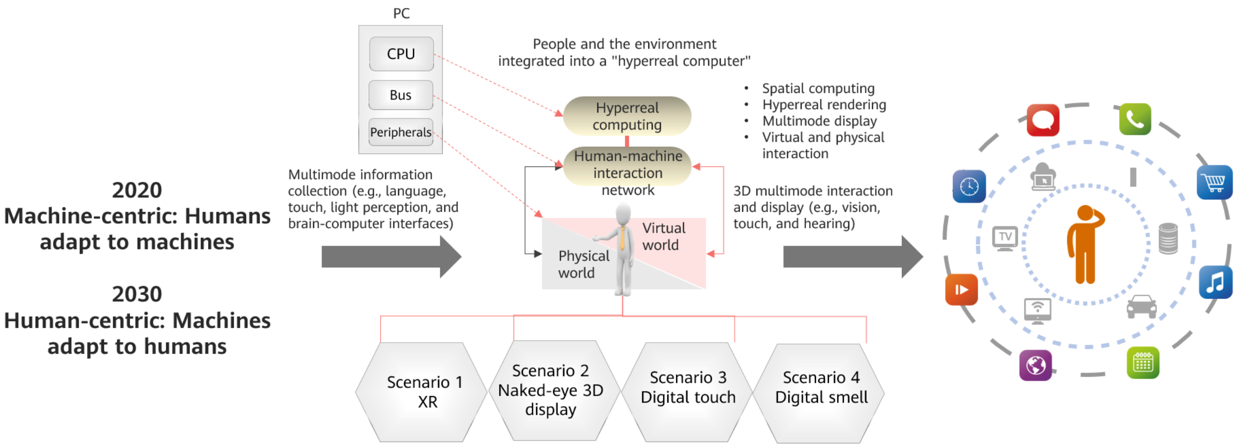
smart screens, smart home appliances, intelligent vehicles, and smart exoskeletons) will be able to understand natural language, gestures, and eye movement, and even read human brain waves, enabling more intuitive integration between the virtual and physical worlds and bringing a hyperreal sensory experience to human-machine interaction. (Figure 1 Hyperreal human-machine interaction experience)

Over the course of the coming decade, communications networks must evolve to support brand-new human-machine interaction experiences such as XR, naked-eye 3D display, digital touch, and digital smell.

XR: An Intuitive Interaction Experience Through a Perfect Synthesis of the Virtual and Physical Worlds

Virtual Reality (VR) is about rendering packaged digital visual and audio content. Augmented Reality

Figure 1 Hyperreal human-machine interaction experience



(AR) refers to the overlaying of information or artificially generated content onto the existing environment. Mixed Reality (MR) is an advanced form of AR that integrates virtual elements into physical scenarios. eXtended Reality (XR), which covers VR, AR, and MR, is a catchall term that refers to all real and virtual combined environments and human-machine interactions generated by computer technology and wearables. Characterized by three-dimensional environments, intuitive interactions, spatial computing, and other features that set it apart from existing Internet devices, XR is considered the next major platform for personal interactions.

In 2020, due to the impact of social distancing caused by COVID-19, demand for VR games, virtual meetings, and AR-assisted temperature taking increased exponentially. The number of active VR users on the US video game digital distribution service Steam doubled. Some manufacturers have unveiled more portable AR-enabled contact lenses, which are expected to go to market in about two years. With the wide adoption of 5G, Wi-Fi 6, and fiber broadband, all of which can deliver gigabit speeds, XR services are set to boom over the next decade. Huawei predicts that the number of VR/AR users is expected to reach 1 billion by 2030.

In its Virtual Reality/Augmented Reality White Paper, the China Academy of Information and Communications Technology (CAICT) divides the technical architecture of XR into five parts: near-eye display, perception and interaction, network transmission, rendering processing, and content creation. The white paper also predicts the development stages of XR. The CAICT's conclusions have, to some extent, been endorsed by the ICT industry. (Table 1 Network requirements of XR services)

Currently, XR is still at the stage of partial immersion. Today, a typical XR experience involves 2K monocular resolution, 100°–120°

Table 1 Network requirements of XR services

Technical System	Technical Index	Partial Immersion 2021	Deep Immersion 2022–2025	Full Immersion (XR) 2026–2030
Near-eye display	Monocular resolution	2K	4K	8K
	Field of view (FOV)	120°	140°	200°
	Pixel per degree (PPD)	20	30	60
	Varifocal display	No	Yes	Yes
Content creation	360° panoramic resolution: Weak interaction	8K	12K	24K
	Gaming: Strong interaction	4K	8K	16K
Network transmission (Average value)	Weak interaction (Mbit/s)	90	290	**1,090**
	Round-trip latency: Weak interaction	20	20	**20**
	Round-trip latency: Strong interaction	5	5	**5**
	Transmission medium	Wired/Wireless	Wireless	
Rendering processing	Rendering computing	4K/90 FPS	8K/120 FPS	16K/240 FPS
		/	Fixation point rendering	
Perception and interaction	Eye interaction	/	Eye tracking	
	Voice interaction	Immersive sound	Personalized immersive sound	
	Tactile interaction	Tactile feedback	Refined tactile feedback	
	Mobile interaction	Virtual mobility (Movement redirection)	High-performance virtual mobility	

FOV, 100 Mbit/s bitrate, and 20 ms motion-to-photon (MTP) latency. If all content is rendered in the cloud, 20 ms of MTP latency is the threshold above which users start to report feelings of dizziness.

We predict that XR will reach the stage of full immersion by 2030, by which time it will be supported by 8K monocular resolution, 200° FOV, and a gigabit-level bitrate. If all rendering is still conducted in the cloud, MTP latency will need to be kept below 5 ms. If technology is developed to support the local rendering of environment-related content that could easily make users dizzy, the latency will be specifically linked to the types of content. For content that requires only weak interaction (such as a streamed video), 20 ms of MTP latency is acceptable. For content that requires strong interaction like games, less than 5 ms MTP latency will be needed.

Therefore, to support the development of XR services over the next 10 years, networks must have bandwidth of at least 1 Gbit/s and latency of less than either 5 ms or 20 ms, depending on the scenario.

Naked-eye 3D Display: A Brand-new Visual Experience Through Lifelike Image Reproduction

The implementation of naked-eye 3D display involves three major phases: the digitalization of 3D objects, network transmission, and optical or computational reconstruction and display.

There are two types of naked-eye 3D display technology: light field display (through lenslets) and the use of spatial light modulators (SLMs).

Light field display leverages the binocular parallax to create 3D visual effects. It uses parallax barriers, lenticular lenses, and directional backlight, all of which impose fairly inflexible requirements in terms of viewing angles. Their large-scale adoption would require real-time capturing of user location and dynamic adjustment.

Table 2 Network requirements of naked-eye 3D display

Technical System	Technical Index	Lenslets (2021–2025)	SLMs (2025–2030)
Maturity prediction		Large-scale deployment and high maturity	Sporadic application
Display	Size	70-inch screen	10-inch to 70-inch screens
	Resolution	16K	16K
Network transmission	Bandwidth	Around 1 Gbit/s	10 Gbit/s–1 Tbit/s (4K, 60 frames, and 10 Gbit/s are required for objects with a size of 10 x 10 cm.)
	Round-trip network latency	Weak interaction: 20 ms Strong interaction: 5 ms	Weak interaction: 5 ms Strong interaction: 1 ms
	Transmission medium	Wired/Wireless	
Interaction design	Voice interaction	Location tracking and spatial sound	
	Gesture interaction	Gesture recognition	
	Mobile interaction	Location tracking and spatial computing	
Availability		Audio: 99.9% Video: 99.999%	

References: IEEE 1981.1 Tactile Internet and Digital Holography and 3D Display

An alternative approach would be to use SLMs. An interferometric method is used to store all amplitude and phase information of light waves scattered on the surface of a 3D object in a recording medium. When the hologram is irradiated with the same visible light, the original object light wave can be reproduced thanks to diffraction, providing users with a lifelike visual experience. (Table 2 Network requirements of naked-eye 3D display)

In recent years, naked-eye 3D display featuring light field display has developed rapidly, in step with the development of user location awareness and computing technologies. Some manufacturers are already showcasing their products. We predict that a large number of use cases will emerge in the entertainment and commercial sectors by 2025. This type of 3D display requires gigabit-level bandwidth and real-time interaction. In strong interaction scenarios, the network latency must be less than 5 ms, and commercial applications will require network availability of 99.999% (this means annual downtime must be less than 5 minutes and 15 seconds).

Over the past several years, breakthroughs have also been made in holographic technology, which is based on optical reconstruction. Product prototypes have been developed with a thickness of 10 cm and a projection size of around 100 cm2. We predict that these small-scale holographic products will become commercially available at exhibitions, for teaching purposes, and as personal portable devices over the next 10 years. They will require bandwidth of around 10 Gbit/s, latency of no more than 5 ms or as low as 1 ms, and network availability of more than 99.999%, the same as that required in commercial settings. True-to-life holographic products will require higher bandwidth (over 1 Tbit/s), but we do not expect them to be ready for large-scale commercial deployment by 2030.

Therefore, the naked-eye 3D display products coming to market over the next decade will need to be supported by networks capable of delivering 1–10 Gbit/s bandwidth per user, latency of 1–5 ms, and 99.999% availability.

Digital Touch: Tactile Internet Made Possible Through Multi-dimensional Sensory Interaction

In IEEE's tactile Internet architecture, digital tactile technology is divided into three layers: user layer, network layer, and avatar layer. The user layer enters information such as location, speed, force, and impedance. After being digitalized over the network, the information is converted into instruction data and provided to the avatar layer. The avatar layer then collects tactile, auditory, and proprioception data and provides the data to the user layer through the Internet to inform users' real-time decision making.

Digital tactile technology has two interaction modes. The first is machine control. Use cases include remote driving and remote control. The second is hyperfine interaction, and use cases include electronic skin and remote surgery. (Table 3 Network requirements of digital touch)

Machine control has numerous use cases in industrial settings, and has high requirements for network availability (above 99.999%). Some industries even require availability to reach 99.99999%. The required bandwidth is generally less than 100 Mbit/s, and the maximum permissible latency varies from 1 to 10 ms, depending on the specific circumstances.

Electronic skin powered by flexible electronics in hyperfine interactions has the most development potential. Electronic skin integrates a large number of high-precision sensors such as pressure and temperature sensors. According to a study by the University of Surrey in the UK, each square inch of electronic skin will require bandwidth of 20 to 50 Mbit/s, meaning that an average hand would require bandwidth of 1 Gbit/s. The wearers of electronic skin won't all be humans; intelligent machines present another class of potential users. The user layer may perform analysis, computing,

Table 3 Network requirements of digital touch

Interaction Mode	Direction of Traffic	Traffic Type	Reliability	Latency (ms)	Bandwidth
Machine control	User-Avatar	Touch	99.999%	1–10	2 Mbit/s
	Avatar-User	Video	99.999%	10–20	1–100 Mbit/s
		Audio	99.9%	10–20	512 Kbit/s
		Tactile feedback	99.999%	1–10	20 Mbit/s (100 DOFs)
Hyerfine interaction	Avatar-User	Tactile feedback	99.999%	1–10	1–10 Gbit/s (Electronic skin)
Active cognitive capability: The network layer also needs to support services such as dynamic performance monitoring, task awareness, and 3D mapping.					

Reference: IEEE 1981.1 Tactile Internet

and decision making based on the massive amounts of data collected by the electronic skin on the avatar layer to control the avatar layer. The user layer can also be directly connected to humans through brain-computer interfaces or myoelectric neural interfaces to deliver an immersive remote interaction experience. We predict that network bandwidth of 1 to 10 Gbit/s will be required in hyperfine interaction scenarios.

Therefore, to support digital touch, networks will need to deliver 1–10 Gbit/s bandwidth per user, availability greater than 99.999%, and latency below 10 ms, or as low as 1 ms in certain use cases.

Digital Smell: Internet That Enables Us to Smell Through Deep Sensory Interaction

Among our five senses, two of them – touch and taste – require direct contact, while three – sight, hearing, and smell – do not. Of the latter three, smell involves the deepest interaction.

Digital smell includes three technical phases: odor perception, network transmission, and smell reproduction.

There have been some use cases for odor perception, such as using composite materials to form a barcode, which can generate chemical reactions according to the odor and create color changes. The relationship between the barcode and odor can then be identified through Deep Convolutional Neural Network (DCNN) algorithms. Use cases can be found in specific scenarios like detection of dangerous goods and detection of food freshness.

There are already some commercial odor reproduction products available in the industry, such as smelling generators for VR games, which use five odor cartridges and selectively release odors from the cartridges. They emit scents such as the ocean, gunpowder, wood, and soil, deepening the immersion of the gaming experience. However, some research reports suggest that the future of smell in VR won't rely on these odor cartridges, but will instead work through brain-computer interfaces to enable people to sense odors more directly and accurately.

The combination of odor perception (using electronic noses) and odor reproduction can help create an Internet that enables us to not only hear and see, but also smell. It is not yet clear what kind of network bandwidth and latency this function will require, but the computing requirements are already relatively well understood.

In a nutshell, the next-generation human-machine interaction network will support brand-new experiences including XR, naked-eye 3D display, digital touch, and digital smell. Making these technologies work will require networks capable of delivering bandwidth of 10 Gbit/s and 99.999% availability, with latency as low as 1 ms for some use cases.

Networks That Deliver a Consistent Experience for Homes, Offices, and Vehicles: The Third Space with the Same Broadband Experience

When we envision the future of self-driving cars, the most appealing feature for many is that we will be able to enjoy the immersive entertainment, social, and work experience we get at home while on the go. Multi-screen collaboration, 3D display, and holograms will all be used both at home and in cars. 8K and 16K smart screens will be gradually adopted at home and MR will be widely used in cars.

With 5G, F5G, and Wi-Fi 6, mobile and fixed broadband basically enters the gigabit era at the same time, making it possible to deliver the same level of experience to users regardless of whether they are at home, in the office, or on the go.

In the future, self-driving cars will become the "third space" beyond homes and offices, and users will enjoy the same broadband service experience in all three scenarios. (Table 4 Network requirements for delivering a consistent experience at home, in the office, and on the go)

Over the next decade, common home and office services will include smart screens, multi-screen collaboration, 3D, holographic teaching, and XR. As true-to-life holographic meetings will not be widely adopted by 2030, the mainstream broadband requirements of home and office services will stand at 1 to 10 Gbit/s of bandwidth and less than 5 ms of latency. In the future, home and office networks will not only provide seamless broadband coverage, but also support brand-new scenarios such as working from home, premise security, and robotics. Based on HCS capabilities, home networks will be able to sense user locations, indoor space, and environment security, and create a more user-friendly living and work environment for people.

Services like multi-screen collaboration, 3D, holographic teaching, and XR will also be available in our self-driving cars. Over the next decade, their key requirements for network bandwidth will be 1 to 10 Gbit/s, and latency requirements will be less than 5 ms. As autonomous driving will require vehicle-road collaboration, it will require network

Table 4 Network requirements for delivering a consistent experience at home, in the office, and on the go

Scenario Type	Commercial Deployment	Home			Vehicle
		Service	Peak Bandwidth	Round-trip Latency	Service
Cinema	Within 10 years	16K video (180-inch screen)	1.6 Gbit/s	50 ms	1.6 Gbit/s, 20 ms (16K XR)
Gaming	Within 10 years	360° 24K 3D VR/AR	4.4 Gbit/s	5 ms	4.4 Gbit/s, 5 ms (24K XR)
Holographic teaching	Within 10 years	10-inch hologram	12.6 Gbit/s	20 ms	12.6 Gbit/s, 20 ms
Holographic meeting	Within 10 to 20 years	True-to-life hologram (70-inch)	1.9 Tbit/s	1–5 ms	12.6 Gbit/s, 1–5 ms (Miniature hologram, 10 inch)
Autonomous driving	Within 10 years	Home robots	10-cm positioning	99.999% availability	5–20 cm positioning Availability: 99.999% to 99.9999%

availability greater than 99.999% and positioning precision of 10 cm.

If networks are to meet the needs of these new technologies and provide a consistent experience across our three spaces (home, office, and self-driving cars), we will need to build new network capabilities that deliver the high bandwidth, high availability, and low latency required.

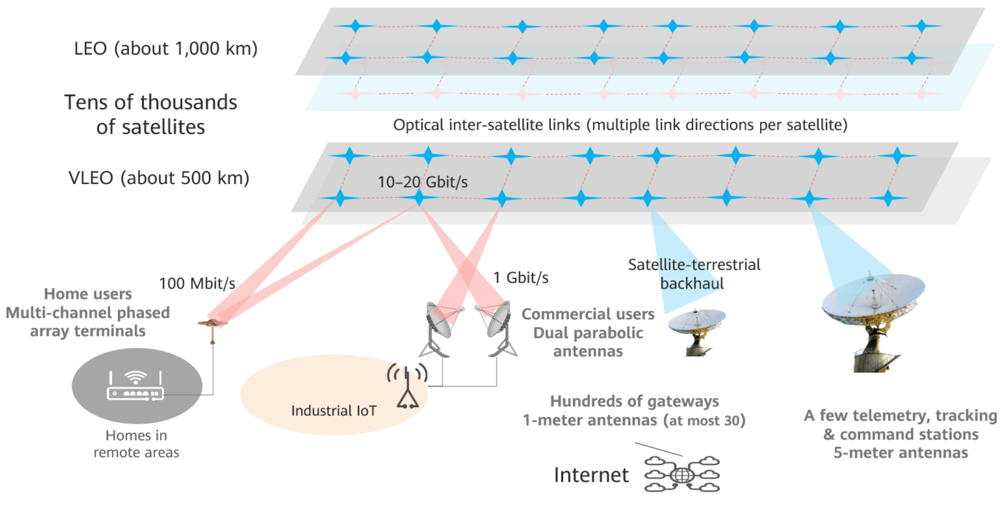
Satellite Broadband Internet: Continuous Broadband Coverage from Ground to Air

Over the next 10 years, connected drones will be more widely adopted, helping create new markets worth tens of billions of US dollars. We will see an increase in intra-city passenger transport in the skies above our cities, supported by tens of thousands of low-earth orbit (LEO) broadband satellites. Satellite broadband will be commercially available on a large scale, which will help turn space tourism and deep-sea exploration into popular leisure activities. Broadband will become an indispensable part of our life; the increased diversification of our leisure activities and the growing demand for unmanned operations in intelligent industry and agriculture mean that broadband will be needed everywhere, from land to sea and sky. A comprehensive broadband network from ground to air will deliver new experiences to people and power the full intelligent transformation of industries. (Figure 2 Satellite broadband network)

Due to spectrum resource constraints and communications disruptions, the actual peak capacity of a single LEO satellite is about 10–20 Gbit/s. Suppose a global satellite network is supported by 10,000 satellites distributed on multiple orbital planes from very low earth orbits (VLEOs) to LEOs, and each satellite maintains links with satellites around it in all directions using over 100 Gbit/s lasercom. The actual effective capacity of the satellite network will be around 100 Tbit/s, considering at least half of the areas passed over by satellites are areas where demand for broadband is minimal (e.g., seas and deserts). Outside of areas covered by cellular networks, satellite broadband providers can use multi-channel phased array antennas to deliver hundred-megabit broadband to consumers and use dual parabolic antennas to deliver gigabit broadband to enterprise customers. They could also transmit data over optical inter-satellite links to hundreds of gateways around the world where they can connect to the Internet. Such a satellite network would be equivalent to a quasi-4G network, providing three-dimensional global coverage and latency of less than 100 ms.

Currently, terminal antennas for LEO satellite

Figure 2 Satellite broadband network



broadband are large, meaning they are ill-suited to mobile scenarios for individuals. Current satellite broadband networks mostly serve homes in remote areas, enterprises, and ships. Some carriers have combined satellite broadband, which is used for backhaul, with cellular networks and WLAN networks on the ground to provide both broadband and narrowband coverage for villages or enterprises in remote areas. With wider adoption of satellite broadband, we may see it applied to mobile scenarios (terminals) such as connected cars and small personal devices. Satellite broadband will deliver a seamless, continuous broadband experience beyond home Wi-Fi and urban cellular networks, meeting the network requirements of people and things.

Industrial Internet: A New Type of Network for Intelligent Manufacturing and Human-Robot Collaboration

The industrial Internet is a new type of infrastructure that deeply integrates ICT into the industrial economy and fully connects people, machines, things, and systems. For industries, this means the birth of a brand-new manufacturing and service system that covers entire industry value chains and paves the way for digitalization, network-based operations, and the intelligent transformation of all industries. An industrial Internet system consists of four key components: industrial control, industrial software, industrial network, and information security. The industrial network is the foundation of the entire system.

Traditional industrial networks are built based on the ISA-95 pyramid model. This architecture was introduced more than 20 years ago and is a manufacturing system centered on human management. However, the development of intelligent manufacturing requires a new architecture that will facilitate human-robot collaboration.

The new architecture will be built upon three equal elements – humans, robots, and an intelligent platform (cloud/edge computing). Private industrial communication buses will be replaced by universal industrial networks and open data layers that support real-time data transmission. The intelligent platform will aggregate data collected from humans and robots for real-time analysis and decision making and support effective collaboration between humans and robots.

Huawei predicts that the total number of global connections will reach 200 billion by 2030, including about 100 billion wireless (cellular) connections (including passive cellular connections) and about 100 billion wired, Wi-Fi, and short-range connections. In industrial settings, the billions of connected devices will include not only pressure, photoelectric, and temperature and humidity sensors, but also a large number of intelligent cameras, intelligent cars, drones, and robots. Industrial networks, currently characterized by a fragmented landscape of different narrowband technologies, will adopt universal broadband technologies.

Universal industrial networks will erase the technical boundaries between consumption, office work, and production. These networks will support multiple types of services using deterministic broadband networks and slicing technologies, such as 5G, Time Sensitive Networking (TSN), IPv6+, and industrial optical networks, allowing companies to connect any workforce and migrate all consumption, office work, and production elements to the cloud.

Universal industrial networks will enable on-demand data sharing and seamless collaboration between office and production systems within a company, between different companies in the same industry, and even between the related services of different vertical industries. They will support broadband-based interconnectivity and multi-cloud data sharing of any workload.

Universal industrial networks will also be smarter than ever, facilitating the movement of data in boundary-free and mobile scenarios across industries and across clouds. They will support intent-driven automated network management and AI-based proactive security and privacy protection, ensuring service security and trustworthiness at any workplace.

An enterprise usually has multiple types of services, so a universal industrial network must ensure the availability, security, and trustworthiness of services. For example, smart healthcare involves services such as remote diagnosis, monitoring & nursing, and remote surgery; a smart grid involves video-based inspection, grid control, and wireless monitoring; and smart manufacturing involves factory environment monitoring, information collection, and operation control. (Table 5 Network requirements of intelligent enterprises)

Based on the typical bandwidth and latency requirements of each service and forecasts on the number of devices used by enterprises in 2030, we predict that a medium- to large-sized enterprise will require network bandwidth of 100 Gbit/s and the maximum bandwidth per user will reach 10 Gbit/s. Acceptable latency will vary greatly from one use case to another, from as low as 1 ms to as high as 100 ms. In addition, it will be necessary to ensure the security and trustworthiness of industrial networks.

Computing Power Network: Orienting Towards Machine Cognition and Connecting Massive Amounts of User Data and Computing Power Services at Multiple Levels

The social value of communications networks is reflected in the services they support. In the past, networks helped establish communications channels between people by providing communications services. Today, with smart devices and the cloud connected to networks, more diverse content services are provided through communications networks.

The networks we use today are designed for human cognition. For example, the frame rate for motion video (typically 30 frames per second [FPS]) is chosen based on the human ability to perceive motion, and the audio data collected is compressed with mechanisms that take advantage of the masking effects of the human cognitive system. For human perception, such encoded audio and video can be considered high quality. However, for use cases that require beyond-human perception, the level of quality may be far from enough. For

Table 5 Network requirements of intelligent enterprises

Industry	Service Type	Network Requirements of Services															
		Number of Connections Per Enterprise	Service Availability (requirements per user or per service)										Security		Trustworthiness		
			Bandwidth Per User (Mbit/s)					Latency (ms)									
			B1	B2	B3	B4	B5	T1	T2	T3	T4	T5	S1	S2	M1	M2	M3
			1–10	10–20	20–50	50–100	>100	50–100	20–50	10–20	5–10	<5	Logical Isolation	Physical Isolation	Visualized	Manageable	Operable
Smart healthcare	16K remote diagnosis	10					1 Gbit/s										
	Monitoring & nursing	2K															
	Holographic remote surgery	5					10 Gbit/s										
Smart grid	Video-based inspection	-															
	Grid control	-															
	Wireless monitoring	-															
Smart manufacturing	Factory environment monitoring	100															
	Information collection	10K															
	Operation control	1K															

Reference: CAICT, Research Report on Industry SLA Requirements for 5G End-to-end Network Slicing

example, robotic monitoring systems will need to detect anomalies by listening to sounds beyond the human audible frequency range. In addition, the average human response speed upon seeing an event is about 100 ms. Therefore, many applications have been designed based on this latency. However, for certain applications that are beyond human usage, such as emergency stop systems, shorter response time is required.

The Innovative Optical and Wireless Network Global Forum Vision 2030 and Technical Directions states that compared with today's networks that are designed for human cognition, future networks designed for intelligent machines such as XR, machine vision, and self-driving vehicles will have enhanced performance in four dimensions:

- Cognitive capacity: Systems will be able to capture objects in the physical world more finely, precisely, and in a multi-sensory manner. For instance, in manufacturing monitoring systems, motion capture at 120 FPS will detect anomalies that would otherwise be undetectable.

- Response speed: Systems will be able to respond to the status change of a controlled object within 10 ms.

- Scalability in computing: Systems will be able to accommodate varying and uncertain workload while achieving high resource utilization, through methods such as dynamic linear scaling of computing resources.

- Energy efficiency: Energy efficiency can be greatly improved if enterprises eliminate on-premise computing resources and adopt a cloud-based model. Moreover, energy efficiency will be further improved with an event-driven approach where a system is deployed on a serverless computing platform.

Intelligent machines will create more accurate data. For example, network clocks and geolocation stamps can be used for precise modeling of the physical world in a digital twin system. This will lead to a shift in data processing and computing, from today's Internet platform-centric model to a data-centric model, decoupling data, computing, and communications.

The network infrastructure designed for machine cognition should satisfy the following requirements:

- Accommodating the collection and transmission of massive amounts of data, having an ultra-low latency, and supporting a very large number of subscribers.

- Managing publishers' data generation and injection based on the overall condition of the system and the importance of the data.

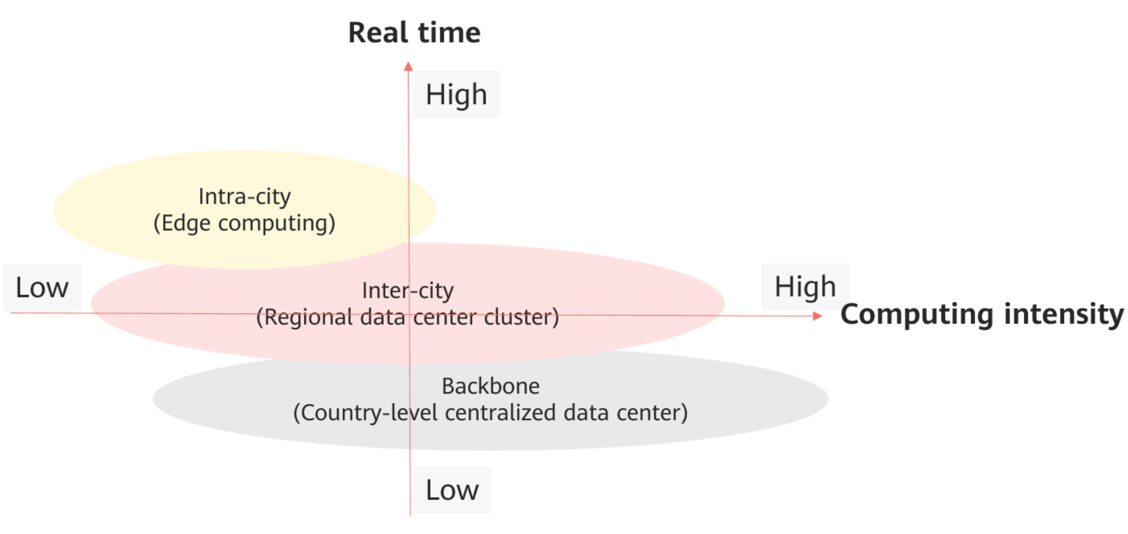
- Supporting the storage and sharing of data among communications and computing nodes in the network.
- Supporting precision time and geolocation stamping.
- Providing strong protections for data security, privacy, and integrity.
- Providing a data brokerage between IP and non-IP nodes, with the data brokerage being accessible through multiple networks.

As the miniaturization of chips approaches its physical limits, the computing industry can no longer rely on Moore's law for rapid development. Economic bottlenecks are encountered when manufacturing CPUs with more than 128 cores in smart devices. In addition, due to bandwidth costs and latency, cloud data centers may not be able to satisfy the massive amount of time-sensitive service processing required by intelligent systems. Machine cognition requires a new type of network in which data analytics and processing can be performed on the edge, and not all data needs to be transmitted to the central cloud.

In the future, the cloud, edge, and devices will be connected, and computing workloads will be apportioned to one of three levels (distributed edge nodes in a city, regional data center clusters that cover multiple cities, or backbone centralized data centers) in real time based on their latency thresholds. In use cases that can tolerate latency of about 100 ms, data may be sent to a centralized data center. In use cases with lower latency tolerance (from 10 ms to as low as 1 ms), computing will be performed in a regional data center cluster or at the edge. (Figure 3 Three levels of computing resources for machine data services)

Computing efficiency and reliability are correlated with network bandwidth, latency, security, and isolation. Therefore, computing and networks should be coordinated. Major carriers have articulated a new business vision for computing and network convergence services based on a new concept of "computing power network". They aim to connect diverse computing power in the cloud, on the edge, and across devices to implement on-demand scheduling and sharing for efficient computing power services at multiple levels. The computing power network represents a significant shift in network design, from focusing on human cognition to focusing on machine cognition.

Figure 3 Three levels of computing resources for machine data services



The Chinese government released the Guiding Opinions on Accelerating the Construction of Collaborative Innovation System of National Integrated Big Data Centers, which states: "With the acceleration of digital transformation and upgrade in various industries, the total volume of data being created by society as a whole is growing explosively, and the requirements for data resource storage, computing, and applications are greatly increasing. Consequently, there is an urgent need to promote an appropriate data center layout, balance between supply and demand, green and centralized development, and interconnectivity. We should build a new computing power network system that integrates data centers, cloud computing, and big data, in order to promote flows and application of data elements and achieve green and quality development of data centers." In addition, the document proposed that "as data centers should be developed on a large scale in a centralized and green manner, network transmission channels between national hubs and nodes should be further streamlined to accelerate the program of 'Eastern Data and Western Computing' and improve cross-region computing power scheduling."

To support proactive development of computing power network standards, ITU-T has launched the Y.2500 series of computing power network standards, with Y.2501 (Computing Power Network – framework and architecture) as the first standard. This series of standards will be compatible with a raft of computing power network standards developed by the China Communications Standards Association (CCSA). Many carriers have incorporated the computing power network into their 6G and future network research. The computing power network will be a key scenario for communications network evolution over the next 10 years.

Cognitive Network: Evolution Towards Advanced Levels of Intelligence

In the academic community, technological advances are often personified to help us understand them more easily. In 1877, German philosopher Ernst Kapp first put forward the concept and theory that "tools and instruments produced by human hands conform to already existing organic structures" in the Elements of a Philosophy of Technology. In 1964, in his book Understanding Media, Marshall McLuhan proposed that mechanical technologies extended our bodies in space and that electric technology extended our central nervous system. In 1995, in The Global Brain Awakens, Peter Russell stated that the various connections were making the earth an embryonic brain and the earth is awakening. The analogies used to describe the tools and technologies humans have created have shifted from body to nerve and to brain as the digital world has advanced towards advanced levels of intelligence.

Communications networks have existed in one form or another for about two centuries. Today, the telegraph and analog telephone networks of the 19th century have long since been superseded by more advanced digital networks.

Over the past 50 years, mobile communications, optical communications, and data communications networks have evolved continuously to remain relevant. These network types, together with optical cables, equipment rooms, and sites, form a robust network trunk.

The biggest evolutionary step taken in the last decade is the development of what the academic community could understand as being the network's nervous system. The human nervous system comprises basic systems that can initiate automatic stress responses and feature closed-loop control, as well as more advanced systems that support thinking, analysis, and cognition (i.e., the brain). The current shift from the software-defined network (SDN) to the autonomous driving network (ADN) is analogous to the evolution of a basic nervous system for networks.

Over the next 10 years, the network nervous system

will evolve along the following two tracks. The first is HCS, e.g., wireless sensing, Wi-Fi sensing, and optical sensing. The second is the development of a "brain" – the digital twin of the physical world that allows autonomous inference and decision making. This is how networks will evolve towards advanced levels of intelligence and feature cognitive intelligence.

Cognitive intelligence is both an engineering and a mathematical issue. To qualify as having cognitive intelligence, a system must be able to sense status changes internally and externally in real time and manage itself through autonomous analysis and prediction.

The construction of cognitive intelligence consists of two dimensions: time and function.

Time: A network may predict changes in the future (T3) by learning historical information (T1 and T2). For example, an L5 ADN will be able to accurately predict performance degradation based on historical performance records and warnings.

Function: A network may predict changes in its functions (information C) by learning function-related information in different environments (information A and information B). For example, a cognitive wireless system may predict user service changes by identifying changes in user locations and channels, and in cyber security, a network may predict changes in the security situation based on detection of abnormal behaviors at the packet level.

The concept of cognitive networks is not new. Some renowned universities, research institutes, and companies have been doing related research for years, but few breakthroughs have been made. Cognitive technology was first used in wireless networks. In 2004, the IEEE established the IEEE 802.22 Working Group and developed the first wireless standard based on cognitive technology. Recent years have seen AI breakthroughs in many industries. Self-driving cars have already driven millions of miles in the real world, completely autonomously. In production quality control, AI vision has significantly cut inspection times. In agriculture, the efficiency of intelligent apple pickers is more than double that of human workers. The communications industry is also exploring the use of AI in networks. We hope that over the next 10 years, the combination of AI and digital twin technology will lead to breakthroughs in cognitive networks, and prediction and judgment of network status will be significantly improved though analysis and inference of multi-dimensional data.

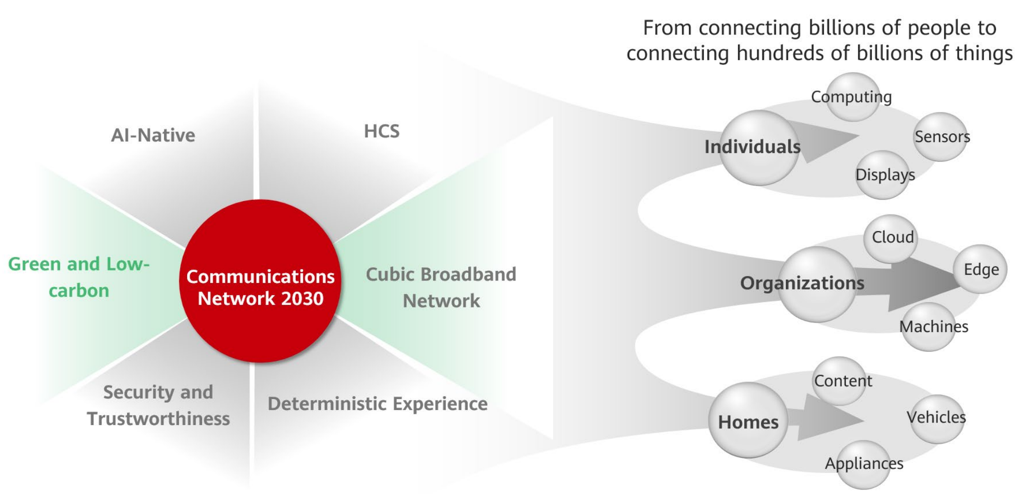
As part of the digital world that is about to awaken, communications networks will have both harmonized sensing capabilities and cognitive intelligence, evolving to a higher form with a robust trunk, sensitive perceptions, and an agile "brain".

Vision for Future Networks and Their Defining Features

Vision for Future Networks

Future networks won't just connect billions of people; they will connect hundreds of billions of things. We envision those connections as being supported by green and cubic broadband networks that are AI-native, secure, trustworthy, and capable of providing deterministic experiences and HCS. (Figure 4 Vision for the communications network of 2030)

Figure 4 Vision for the communications network of 2030



Defining Features

The communications networks of 2030 will have six defining features enabled by 15 key technologies, and each key technology will rely on research on multiple technological fronts. (Figure 5 Defining features of the communications network of 2030)

Cubic Broadband Network

The coming decade will see continuous improvement in network performance. Today's gigabit access enabled by 5G, F5G, and Wi-Fi 6 for homes, individuals, and organizations will evolve toward 10 gigabit capacity enabled by 6G, F6G, and Wi-Fi 8. Huawei predicts that the average monthly data use on wireless cellular networks per person will increase by 40-fold to 600 GB in 2030. In addition, gigabit or higher fiber broadband household penetration and 10 gigabit fiber broadband household penetration are expected to reach 55% and 23%,

Figure 5 Defining features of the communications networks of 2030

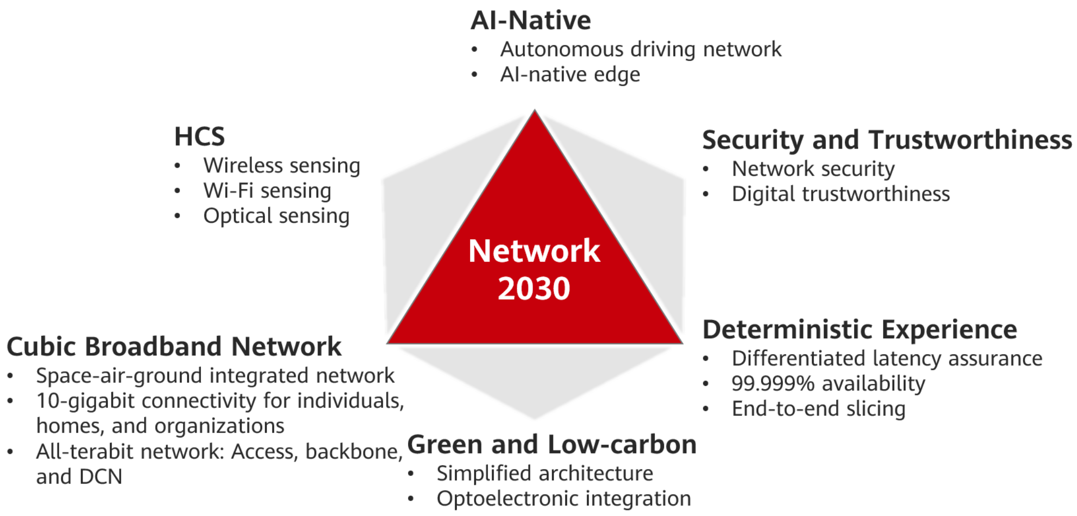
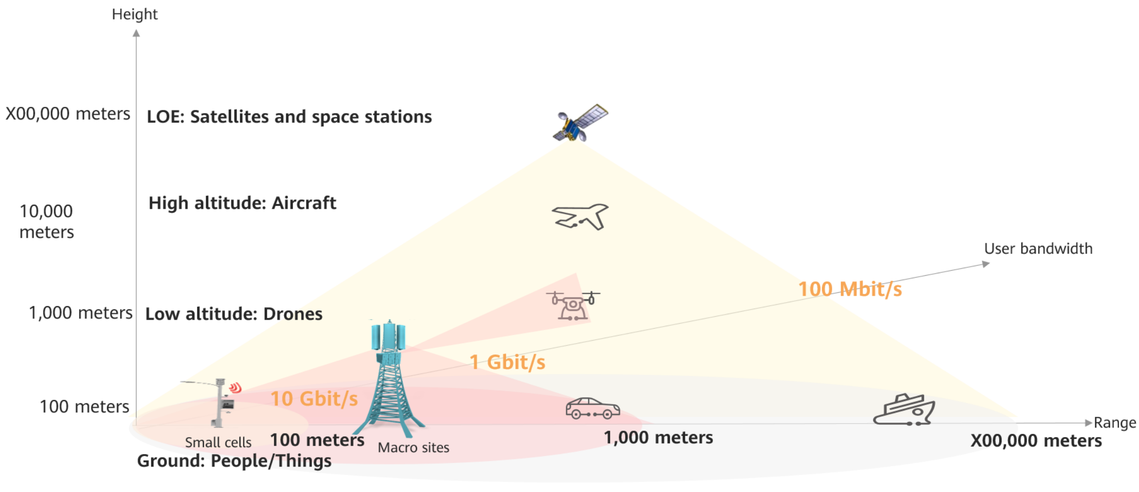


Figure 6 Cubic broadband network



respectively, and the average monthly fixed network data usage per household is forecast to increase by 8-fold to 1.3 TB. Network ports will be upgraded from 400G to 800G or even 1.6T, and single-fiber capacity will exceed 100T. In terms of coverage, network construction up until now has focused on connectivity on the ground, but in the future, we will see the construction of integrated networks connecting the ground, air, and space.

1) Space-Air-Ground Integrated Network: Seamless and Continuous Broadband Experience

Moving forward, broadband coverage will extend beyond the ground, encompassing the air and even space. These networks will connect devices at various heights, such as drones flying less than 1 kilometer above the ground, aircraft 10 kilometers above the ground, and LEO spacecraft hundreds of kilometers above the ground. A cubic network will consist of small cells with a coverage radius of 100 meters, macro sites with a range of 1–10 kilometers, and LEO satellites with a range of 300–400 kilometers, which will provide users with continuous broadband experiences of 10 Gbit/s, 1 Gbit/s, and 100 Mbit/s, respectively. (Figure 6 Cubic broadband network)

For satellite-terrestrial access, devices must be able to easily access terrestrial and space networks. Before we can make that happen, there is still a lot of research that needs to be done. For example, we need to develop:

- New air interface technologies with deep fading, large latency, and highly dynamic performance

- Intra-satellite and inter-satellite beamforming technologies that can evenly allocate active user equipment (UE) to different beams for load balancing and more efficient resource utilization

- Anti-interference technologies for higher spectrum multiplexing rates

- Frameworks that support fast decision making in response to massive volumes of switching requests and complex switching, as well as mobility management frameworks for limited numbers of ground stations

Inter-satellite transmission requires satellites at different orbital heights to form multi-layer constellations, with each layer networking through inter-satellite links. Inter-satellite links are established between satellites in the same orbit, on the same layer, and on adjacent layers, forming a cubic space network. Inter-satellite links will use lasercom and terahertz technologies to support bandwidth of more than 100 Gbit/s. This will require research on adapting industrial products to aerial settings, making phased array antennas more compact, and enabling dynamic laser tracking and pointing.

The network management and control domain consists of an operation and control center, network management center, earth station, and integrated core network. It performs the tasks of satellite network management, user management, and service support. In this domain, we need to research new dynamic routing protocols between ground-based earth stations and constellation networks, and hyper-distributed convergent core networks that support intelligent switching of space-air-ground integrated networks.

2) 10-Gigabit Connectivity for Individuals, Homes, and Organizations

Fiber networks are expected to be widely deployed globally over the next 10 years, transforming today's gigabit connections for individuals, homes, and organizations into 10-gigabit connections.

To deliver 10-gigabit home broadband, 200G passive optical network (PON) technology will likely be used for optical access. The coherent detection technology typically used for wavelength division multiplexing (WDM) will be used in the PON field, which will significantly improve receiver sensitivity and support modulation formats with higher spectral rates, such as quadrature phase shift keying (QPSK) and 16-quadrature amplitude modulation (16-QAM), to achieve higher data rates.

To deliver 10-gigabit broadband for individual users, mobile network research needs to focus on flexible use of the sub-100 GHz spectrum bands and continuous evolution of massive multiple-input multiple-output (MIMO). 3GPP Release 16 has defined two frequency ranges, FR1 and FR2, for 5G new radio (NR), covering all spectrum bands for International Mobile Telecommunications (IMT) between 450 MHz and 52.6 GHz. Research for Release 17 is still underway, and one important focus of this research has been the use of spectrum above 52.6 GHz for 5G NR. This points to industry consensus on fully utilizing spectrum below 100 GHz for 5G.

To make 10-gigabit campus networks possible, more research is needed on next-generation Wi-Fi technologies that support millimeter-wave and high-density MIMO. Theoretically, Wi-Fi 7 standards that are currently being defined should be able to support 10-gigabit user access. With wireless air interface technology approaching Shannon's limit, further evolution of Wi-Fi and mobile technologies will require more spectral resources, which are scarce. This has prompted industry-wide discussions about the feasibility of converging Wi-Fi 8 and 6G.

3) All-terabit Network: Access, Backbone, and DCN

Taking into account the growing broadband requirements of individuals, homes, and enterprises, as well as the need to connect people and things, future access network equipment will need to support terabit-level interfaces. Backbone equipment will support 40–100 Tbit/s per slot and data center equipment 400 Tbit/s per slot.

By 2030, there will be broadband networks that can achieve terabit-level transmission speeds in many parts of the network, from access and backbone to data center networks. These will mostly serve the world's largest cities – those with populations of 10 million or higher.

In the terabit era, datacom equipment will need to have Ethernet interface technology that supports speeds of 800 Gbit/s or even 1.6 Tbit/s to meet service development needs. Unlike 200G and 400G Ethernet, 800G Ethernet is a nascent technology that has yet to be standardized. From a technical standpoint, there are two routes that will take us to 800G: continuing evolution of existing pluggable optics modules and the adoption of new co-packaged optics (CPO) modules. Both module types will have a place in the future market, but pluggable optics modules with a capacity of over 800G are expected to encounter power and density problems, so CPO modules will likely become the preferred choice.

In addition, enabling long-distance transmission capacity of more than 100 Tbit/s per fiber will require technical breakthroughs in backbone WDM equipment, including materials science breakthroughs in high-baud electro-optic modulation and the development of new optical amplifier technology that goes beyond C-band to L-band and S-band.

Deterministic Experience

The ability of communications networks to provide deterministic experiences is key to supporting online office and learning, as well as meeting the security and reliability needs of production environments.

1) 100 ms, 10 ms, and 1 ms Latency Assurance for Differentiated Service Requirements

Over the course of this decade, the Internet traffic model will undergo a fundamental shift from today's top-down content traffic generated primarily from online services, retail, and entertainment to bottom-up data traffic from pervasive intelligent applications deployed across various industries. Intelligent machines will generate massive amounts of data, and this data will need to be processed in data centers. This decade will see a push toward the coordinated development of electricity and computing power to enable society-wide green computing power. Therefore, the networks of the future will need to be able to support more centralized operations of data centers. That will entail meeting differentiated latency requirements, with the acceptable latency for backbone, inter-city, and intra-city network services being 100 ms, 10 ms, and 1 ms, respectively. In addition, networks will need to schedule resources in real time at the network layer based on service attributes in order to make computing power greener and more efficient.

In addition to meeting differentiated latency requirements at the network architecture and system levels, the industry also needs to research end-to-end deterministic latency.

Real-time wireless access services require high instantaneous rates over the air interface. However, due to the spectrum constraints caused by the multiplexing of multiple pieces of UE on a single carrier, it is difficult to guarantee real-time performance. Moving forward, multi-carrier aggregation technologies need to be developed so that carrier configuration is decoupled from transmission, improving the bandwidth of services under latency constraints on multi-band carriers.

For cloud-based wireless core networks, real-time operating systems (OSs) are needed to enhance deterministic scheduling frameworks and ensure real-time service performance.

The optical access networks we have today feature PON technology, which is based on time division multiplexing (TDM). PON uses uplink burst to prevent collisions, making it ill-suited to scenarios requiring low latency. Frequency division multiplexing (FDMA) needs to be explored to allow concurrency of multiple optical network terminals (ONTs) and guarantee low latency by addressing fundamental issues.

For wide area networks (WANs), the current best-effort forwarding mechanism needs to be changed, protocols at the Physical (PHY) and Medium Access Control (MAC) layers need to be improved, and new technologies such as time-sensitive networks (TSNs) and deterministic IPs need to be integrated to ensure on-demand, end-to-end latency.

2) End-to-end Slicing: Logical Private Networks and Services That Are More Adaptable to Vertical Industries

End-to-end slicing provides vertical industries with customized private network services that run independently and are isolated from each other. This is a key area we can work on in order to serve vertical industries. End-to-end slicing is a network virtualization technology with Service Level Agreement (SLA) assurance. Through network slicing, different logical or physical networks can be isolated from the network infrastructure to meet the SLA requirements of different industries and services. Types of slicing include wireless slicing, transport network slicing, and core network slicing. When a carrier provides a slice to a customer, the carrier also provides end-to-end management and services.

Wireless slicing: It can be further classified into hard slicing and soft slicing. Hard slicing is achieved through resource isolation, such as through static resource block (RB) reservation and carrier isolation for specific slices. Soft slicing is achieved through resource preemption, such as QoS-based scheduling and dynamic RB reservation. Currently, the bitrates of different network slices can be guaranteed based on priorities. The next step in the development of network slicing is to explore the most appropriate wireless protocols for the PHY, MAC, Radio Link Control (RLC), and Packet Data Convergence Protocol (PDCP) layers. For example, we could have a PHY layer with a low-latency coding scheme for slices that support ultra-reliable low-latency communication (URLLC) services, or a MAC layer with an optimized hybrid automatic repeat request (HARQ) mechanism.

Transport network slicing: This is achieved through physical isolation or logical isolation. Physical isolation technologies can be optical-layer hard pipes, which carry different services through different wavelengths or through the optical channel data unit-k (ODUk) within a single wavelength. Flexible Ethernet (FlexE) at the MAC layer is also used to isolate services by scheduling timeslots. Logical isolation technologies mainly include SRv6 Slice-ID, traffic engineering (TE), and virtual private network (VPN) at the IP layer. Logical service isolation is implemented through labeling and network equipment resource reservation. Further research is needed in the industry to explore the integration of technologies such as congestion management mechanisms, latency-oriented

scheduling algorithms, and highly reliable redundant links for FlexE, TSN, and deterministic networking (DetNet). This can deliver bounded latency and zero packet loss for physical slicing, as well as low-granularity FlexE interfaces.

Core network slicing: In 5G standalone (SA) architecture, microservices are the smallest modular components of core network functions. In the future, microservices will need to be flexibly orchestrated into different slices based on service requirements, and flexibly deployed in different parts of the network based on differentiated latency and bandwidth requirements.

End-to-end management and services: 3GPP has defined an end-to-end network slicing management function (NSMF), which streamlines network slice subnet management functions (NSSMFs) to enable end-to-end automatic slicing. This can facilitate elastic slice service provisioning and capacity expansion or reduction. Moving towards 2030, the SLA awareness, precision measurement and scheduling of slicing need to be further researched in the industry to achieve automated closed-loop slicing control. In addition, customers in vertical industries must be able to flexibly customize slicing services on demand. More efforts are needed to study how to meet industry customers' Create, Read, Update, and Delete (CRUD) requirements for slices, and how to coordinate the configuration of slices, private networks, and edge services.

3) 99.999% Availability for Industry Production Control Systems to Enable Enterprises to Migrate All Systems to the Cloud

Traditional enterprise management and production systems are human-centric and built based on the ISA-95 pyramid model, such as enterprise resource planning (ERP), manufacturing execution system (MES), supervisory control and data acquisition (SCADA), and programmable logic controller (PLC) systems. As enterprises become intelligent, these systems will be built on human-thing collaboration, and we will see the wide adoption of a new flattened architecture for cloud, edge, things, and humans.

Currently, enterprises are primarily migrating their ERP and MES systems to the cloud, which do not have real-time requirements and require the availability of the cloud and network to be just 99.9%. By 2030, however, enterprises will be migrating all of their systems to the cloud, including systems that require greater than 99.999% availability for the cloud and network (and edge), such as SCADA and PLC.

Moving forward, improving radio access network availability will be a major area of research. 5G can already meet the basic reliability requirements of URLLC scenarios such as ports and mines, in which availability has reached 99.99%. In the future, AI technologies will be introduced to improve the availability of mobile networks to 99.999% by better predicting channel fading characteristics, identifying envelope channel changes, increasing the number of URLLC connections supported by a single unit of spectrum, and enabling intelligent prediction, interference tracking, and end-to-end collaboration.

AI-Native

1) Autonomous Driving Network (ADN): Continuously Evolving Toward L4/L5 Advanced Intelligence

The ADN represents an advanced stage in the development of the network nervous system. Based on data- and knowledge-driven intelligent, simplified networks, the ADN is automatic, autonomous, self-healing, and self-optimizing. It enables new services that offer optimal customer experiences, autonomous O&M, and the most efficient resource and energy utilization.

Currently, the development of the ADN has reached the L2 to L3 stages, with partial and conditional autonomy. The system uses AI models to enable closed-loop O&M for specific units in specific external environments. In the future, the ADN will continue to evolve toward advanced intelligence to achieve closed-loop autonomy for multiple services throughout the lifecycle in more complex cross-domain environments. (Table 6 Levels of the ADN)

Research aimed at supporting the evolution to the L4/L5 ADN should focus on the following key directions.

The management and operation layer: This layer unifies data modeling to decouple data from functions and applications and ensures data consistency across layers. On this layer, the network's digital twin is built to analyze and manipulate the physical network through simulation. In this regard, research should focus on the following technologies:

- Objective-based adaptive decision-making architecture: The traditional architecture centered on function implementation will evolve into an objective-based decision-making architecture, capable of handling complex and unpredictable environments. Key challenges in this regard include resolving potential conflicts

Table 6 Levels of the ADN

Level	L0: Manual O&M	L1: Assisted O&M	L2: Partial Autonomy	L3: Conditional Autonomy	L4: High Autonomy	L5: Full Autonomy
Service	N/A	Single use case	Single use case	Multiple use cases	Multiple use cases	Any use cases
Execution	Manual	Manual/Autonomous	Autonomous	Autonomous	Autonomous	Autonomous
Awareness	Manual	Manual	Manual/ Autonomous	Autonomous	Autonomous	Autonomous
Analysis/ Decision	Manual	Manual	Manual	Manual/ Autonomous	Autonomous	Autonomous
Intent/ Experience	Manual	Manual	Manual	Manual	Manual/ Autonomous	Autonomous

Reference: TMF 2020

between multiple objectives of the system, increasing the environment's predictability, and ensuring collaboration between different autonomous systems, or between autonomous systems and humans.

- Model-driven and data-driven hybrid architecture: The model-driven architecture requires detailed risk analysis and identification of harmful incidents in the design phase. Its advantages include being trustworthy, explainable, and applicable to critical tasks. In the data-driven architecture, machines will gradually replace human operators for situation awareness and adaptive decision making, and become able to handle complex and uncertain scenarios. This will mark the first step toward the ADN. The advantage of this architecture is its high performance, and its disadvantages include poor explainability, limitations of training sample space, and the fact that it is currently only applicable to non-critical tasks.

- Semantics-based intent: In an ADN, autonomous systems interact with each other through intent-based interfaces in a simplified manner, and differentiated internal implementation processes are shielded from the outside, which enables an out-of-the-box feature. Autonomous systems are decoupled from each other by focusing only on achieving the objectives, regardless of the implementation methods. There are four types of intent: user intent, business intent, service intent, and resource intent.

- Network digital twin: In terms of data awareness, research on high-performance networks should strive for near-zero-error measurement. At the modeling and prediction layer, a high-precision approximate simulation model needs to be constructed for research on how to provide high-performance, SLA-supported simulation that has theoretical guarantee based on network calculus and queuing theory. In terms of control management, the issues of resource allocation and optimization of giant network systems need to be resolved by exploring the theory of fast and slow control structure.

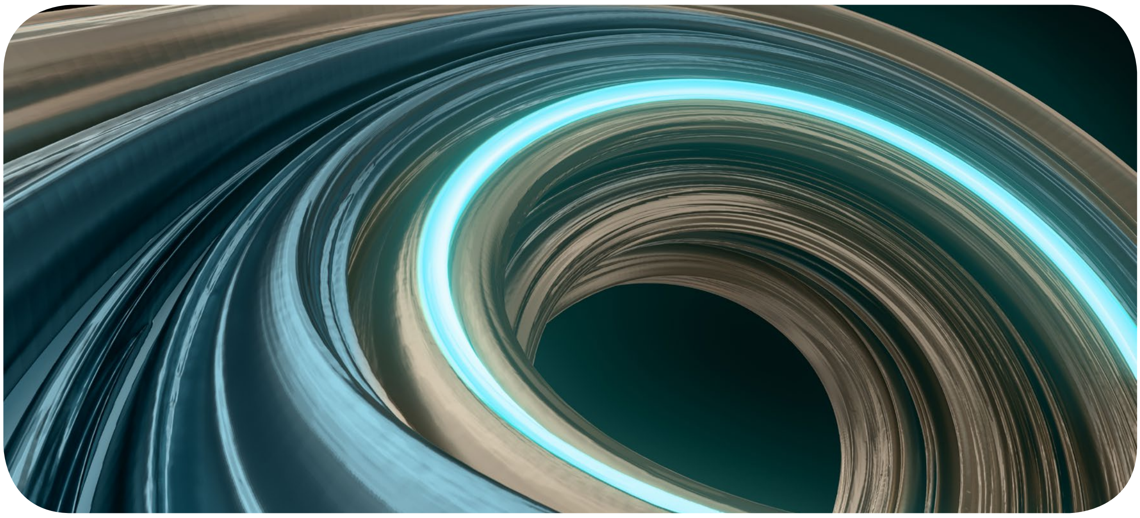
The NE layer: AI should be used not only for O&M, but also for reconstructing NE algorithms and functions so as to deliver AI-native NEs. Through AI-based, real-time analysis and processing of real-time status data of NEs, the ADN can dynamically compensate and optimize parameters, improve the algorithm accuracy of network equipment, and achieve intelligent ultra-broadband, such as cognitive wireless networks and cognitive optical networks. To this end, the equipment's computing power needs to be improved tenfold.

In addition to the advancement of software systems, building L4/L5 capabilities into the ADN also requires that network architecture, protocols, equipment, sites, and deployment solutions be simplified, so as to offset the complexity of network connectivity with a simplified architecture.

2) AI-Native Edge: Reconstructing the Intelligent Edge with Cloud Native and AI Technologies

Within the architecture of the communications network of 2030, the cloud core network will build an AI-native edge by combining the flexibility and openness afforded by a cloud-native architecture and the service-aware capabilities of AI.

The AI-native edge needs to support AI-based service awareness capabilities. Networks for individual consumers will provide efficient encoding and decoding, optimized transmission, experience assurance, and coordinated scheduling capabilities for full-sensing, holographic communications services.



Private networks for industries can enhance the scheduling framework, and provide service assurance for various industries based on deterministic operating systems. For example, during machine vision processing, the MEC-based 5GtoB + AI inference service uses the AI-powered image feature recognition function on the edge to reduce the bandwidth requirements of the backbone network and improve real-time service performance.

The AI-native edge needs to support mesh interconnection and horizontal computing power scheduling. As networks connect to multi-level computing power resource pools, they should be able to sense various resources in order to use computing power efficiently.

To develop computing power awareness, the first thing to do is explore how to measure and model the computing power requirements of AI services. There are various types of computing chips on a computing power network, such as CPUs, GPUs, ASICs, TPUs, and NPUs. The computing power of each type of chips needs to be accurately measured in order to identify the service types to which they can be applied.

Second, computing nodes of a computing power network need to send their computing power resource information, computing power service information, and location information to network nodes. To enable the network to sense multi-dimensional resources and services such as computing power and storage, new computing power routing control and forwarding technologies need to be developed. These could include IPv6+-based computing power status advertising, computing power requirement awareness, and computing power routing and forwarding.

Third, in addition to being able to sense computing power, networks should also be able to flexibly adapt to different IoT device scenarios. Huawei predicts that IPv6 adoption must exceed 90% by 2030 to ensure all things that can be connected are connected. It is thus necessary to develop innovative technologies for hierarchical IPv6 address architecture and ultra-large-scale high-speed addressing and forwarding. These technologies should be compatible with both traditional IP networks and lightweight protocols, so as to ensure the global accessibility of data and computing.

HCS: A New Area Emerging from Communications Technologies

From 1G to 5G, communications and sensing have been independent of each other. For example, a 4G communications system is only responsible for communications, and a radar system is only responsible for functions such as speed measurement, sensing, and imaging. This separation wastes wireless spectrum and hardware resources, and the separation of functions often results in high latency for information processing.

As we approach the 5.5G or 6G era, the communications spectrum will expand to include millimeter wave, terahertz, and visible light. This means the communications spectrum will soon overlap with the spectrum previously reserved for sensing systems. HCS facilitates unified scheduling of communications and sensing resources. Technically, HCS can be broken down into the following three types:

Wireless sensing: HCS is one of the three new use cases proposed for 5.5G, particularly in scenarios such as connected vehicles and drones. With Release 16, precise positioning functionality can already achieve meter-level accuracy for commercial use cases, and future releases are expected to hone this accuracy further, to the centimeter level. As wireless networks move toward higher frequency bands, such as millimeter wave and terahertz, HCS will be applied in areas such as smart cities, weather forecasts, environmental monitoring, and medical imaging.

Wireless HCS technology is still in its infancy and more research is needed in the industry for basic theories such as optimal compromise. As of now, the potential of channel modeling above 0.3 terahertz remains untapped. More research needs to be devoted to far-field and near-field terahertz propagation modeling; spatial target reflection, scattering, and diffraction modeling; and spatial sparsity sensing modeling. In addition, more research is needed on high-performance, low-power radio frequency (RF) chips and components, the structure of super-large terahertz array antennas, and efficient distributed cooperative sensing algorithms such as active radar illumination, environmental electromagnetic control, multi-point cooperative transmitting and receiving, target imaging, scene reconstruction, and channel inversion.

Wi-Fi sensing: IEEE 802.11bf defines Wi-Fi sensing standards applicable to indoor, outdoor, in-vehicle, warehouse, and freight yard scenarios, among others. It covers functionalities such as high-precision positioning, posture and gesture recognition, breath detection, emotion recognition, and perimeter security. Moving forward, more research needs to be directed at both the PHY layer (i.e., new signals, waveforms, and sequences) and the MAC layer (e.g., compromise between measurement result feedback and sensing precision for sensing scenarios based on channel state information [CSI] or signal-to-noise ratio [SNR]). Synchronization and coordination between nodes for single-, dual-, and multi-station radar systems is another problem to address. The last issue concerns collaborative sensing across multiple protocols, including 802.11az, 802.11be, and 802.11ay.

Optical sensing: Optical sensing can be divided into fiber-based sensing and laser radar ("lidar") sensing. Fiber-based sensing is more often seen in energy, electricity, government, and transportation sectors where it is used to sense changes in temperature, vibration, and stress to inform fire monitoring and warning, equipment and pipeline fault diagnoses, and environmental and facility stress monitoring. Lidar sensing is more commonly seen in homes and vehicles, providing functions such as spatial environmental sensing, high-precision positioning, and posture or gesture recognition. Currently, fiber-based sensing tends to have a high false alarm rate in complex environments. More research should be directed at reducing the false alarm rate by introducing AI and big data analytics. For lidar sensing, the 3D panoramic modeling algorithm technology needs to be improved to enable multi-radar coordinate system registration based on lidar sensing data.

Huawei predicts that by 2030, 10 gigabit Wi-

Fi network penetration in enterprises will reach 40%, and F5G private network penetration in medium/large enterprises will reach 42%. In addition, the penetration of 5G private networks in medium/large enterprises will reach 35%. While providing broadband services for enterprises, communications networks will use HCS capabilities to gather static information (e.g., spatial environments, communications blind spots, and obstacles) and dynamic information (e.g., positions, motion tracks, postures, and gestures of people, and the movement of vehicles and objects) to perform data modeling. Coupled with simulation technologies based on the idea of digital twins, the data can help identify and predict changes, empowering numerous industries. HCS represents a new frontier of communications technologies and has huge development potential.

Security and Trustworthiness: A Six-Layer Framework for a New Security Foundation

Today, the very concept of security is changing. It is no longer about centralized protection or a bolted-on feature. Rather, there is a new expectation that security should be an endogenous feature of a network and part of the network's architecture. Beyond that, as we move from the consumer Internet to the industrial Internet, networks require not only security but also trustworthiness.

Security and trustworthiness cover six layers: trustworthiness of components (chips and operating systems), equipment security, connectivity security, management security, federated trustworthiness, and data trustworthiness. Equipment security, connectivity security, and management security fall under network security, while component trustworthiness, data trustworthiness, and federated trustworthiness fall in the trustworthiness realm. The two focus on different aspects but interact in many ways. Ensuring security and trustworthiness requires systemic efforts, involving hierarchical security and trustworthiness technologies such as cross-platform trustworthiness operating systems and chips, endogenous network security, cloud security "brain", multi-intelligent-twin and cross-domain trustworthiness federation, and differential data privacy processing. (Figure 7 Six-layer network security and trustworthiness framework)

Component trustworthiness: Credible data sources are the basis for security and trustworthiness. The Trusted Execution Environment (TEE) at the component (chip and operating system) level is a widely recognized and used solution. Moving forward, chip-level trustworthiness computing technologies will be introduced to network elements (NEs). This will help build a secure and trustworthy running environment for software and hardware based on the underlying NEs, thereby enabling level-by-level verification of chips, operating systems, and applications to ensure data authenticity.

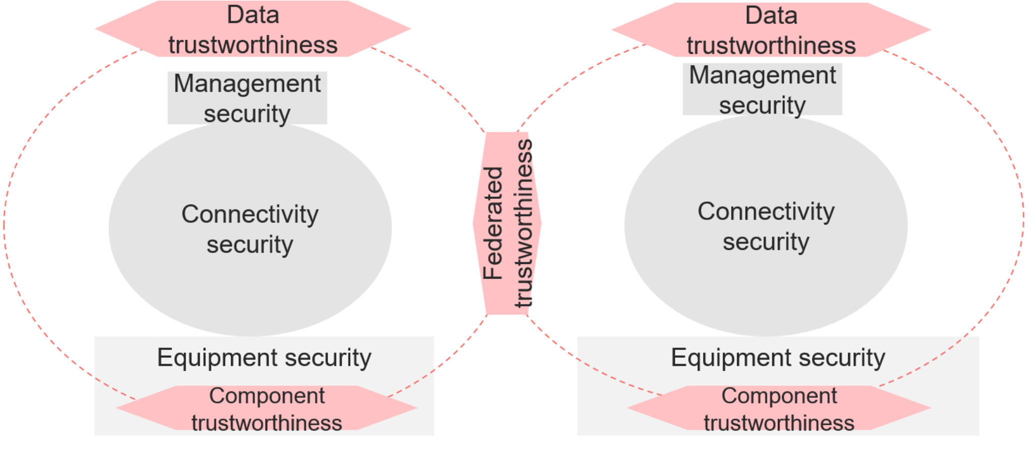
Equipment and connectivity security: Communications protocols and network equipment can be modified to embed trustworthiness identifiers and password credentials in IPv6 packet headers. Network equipment can verify the authenticity and legitimacy of requests based on identifier authentication, preventing identity theft and spoofing and building fine-grained access authentication and source tracing capabilities.

Management security: First, future networks need to adopt a service-based security architecture that integrates cloud, network, and security, so that security functionalities are provided as components and microservices, and can be centrally orchestrated and agilely deployed. Second, as the user base grows and complexity increases, security policies are growing exponentially to the point where the conventional manual approach to planning and management can no longer keep up. More research is needed on traffic and service self-learning and modeling technologies, model-driven risk prediction and security policy orchestration technologies, and security policy conflict detection and automatic optimization technologies.

Federated trustworthiness: To meet the security and trustworthiness requirements across networks and clouds, blockchain technology will be used to build a trustworthy service system for basic digital resources (including connectivity and computing) for future networks. Distributed accounting, consensus mechanisms, and decentralized key allocation will help ensure the authenticity of resource ownership and mapping relationships and prevent anonymous tampering, illegal hijacking, and other security and trustworthiness issues.

Data trustworthiness: Networks process user data at user access nodes and service-aware nodes. Therefore, user data passing through the network must be made opaque to the network, so as to ensure user information security. Research should go into technologies that enhance encrypted transmission of user IDs and communications data, as well as pseudonymization and homomorphic encryption technologies that make user information fully

Figure 7 Six-layer network security and trustworthiness framework



invisible to the network.

Green and Low-carbon

1) Simplified Architecture: Low Carbon Realized by Simplifying Foundation, Cloud, and Computing Networks

Traditional networks are divided by technical specialty, resulting in the fragmentation of O&M services. This model is increasingly difficult to adapt to the development of automated and intelligent networks. In the future, networks need to be reconstructed based on the nature of the services they carry, building a simplified three-layer network architecture consisting of foundation, cloud, and computing networks.

The foundation network is used for connectivity at the equipment port level. It provides access (wired/wireless), switching, and core networks from end to end, based on the 100% fiber-to-site optical foundation that supports optical cross-connect (OXC) or ROADM. The foundation network provides high-bandwidth, low-latency, and high-reliability broadband services, and enables green, low-carbon networks with simplified O&M based on all-in-one full-spectrum antennas, fully converged core networks, and simplified protocols.

The cloud network is used for connectivity between the cloud and devices at the tenant level, and is overlaid on the foundation network using end-to-end slicing technology. It enables agile and open virtual networks that provide SLA assurance, and uses a network for multiple purposes to increase network utilization and save network energy.

The computing network is used for connecting data and computing power at the service level, and provides computing power routing services and trustworthiness assurance for data processing. It is constructed based on distributed and open protocols. Through flexible scheduling of data, the computing network enables green, centralized multi-level computing power infrastructure that has a reasonable layout.

The three networks are interdependent. The computing network depends on the cloud network to enable agile building of virtual pipes and open interfaces that can be provisioned on demand, so as to provide real-time, elastic connections between data and computing power. The computing network also needs the support of the foundation network to enable its most important features: low latency and high bandwidth.

2) Optoelectronic Integration: Profoundly

Changing the Architecture and Energy Efficiency of Communications Network Equipment

In the communications network industry, optical technologies have traditionally been relatively independent from other specialized technologies such as wireless communications and datacom. However, as networks develop toward higher bitrates, higher frequencies, and greater energy efficiency, traditional electronic technologies will soon encounter sustainable development bottlenecks, such as in distance and power consumption. The solution to this is optoelectronic integration.

In the next decade, the development of new products, such as optical input/output chips and CPO, will improve electronic components' high-speed processing capabilities and reduce their power consumption. Coherent optical technologies will be applied to extend the transmission distance of high-speed ports on datacom equipment. New types of antennas that directly connect to optical fibers will be used to reduce the weight and power consumption of base stations. Microwave communications will be superseded by laser communications to support high-speed data transmission between LEO satellites. To meet the communications requirements of underwater mobile devices, wireless coverage will be replaced by visible light which achieves higher penetration than radio waves. Due to its higher transmittance, far infrared light technology will be used to detect brain waves more accurately.

Optoelectronic integration is the way forward for structured improvement of equipment energy efficiency. CPO chips based on optical buses are expected to be in commercial use by 2025. Some academic institutions are researching optical cell switching technology that could potentially replace electrical switching networks. Equipment-level optoelectronic integrated products using optical buses and optical cell switching technology are expected to be developed by 2030. Further into the future, chip-level products that combine optical computing, optical RAM cores, and general-purpose computing cores will also emerge.

Optoelectronic integration technology at the network, equipment, and chip levels can

continuously improve the energy efficiency of communications equipment, and meet the green network objective of increasing network capacity without increasing energy consumption.

3) Summary and Technology Outlook

By 2030, we will be living in a multi-network and multi-cloud world. Billions of people and hundreds of billions of things will be connected to an intelligent world of hyperreal experiences where multiple clouds coexist, including public, industry, and telecom clouds. Connections will be supported by cubic networks consisting of 10-gigabit personalized home networks, 10-gigabit industrial campus networks, 10-gigabit individual networks, and global satellite networks.

In future communications networks, energy efficiency will be continuously improved through optoelectronic integration at the network, equipment, and chip levels in the foundation network. The cloud network will use end-to-end virtual slicing to connect the breakpoints of specialized networks on top of the foundation network, so as to provide differentiated capabilities guaranteed by SLAs for different tenants. The computing network will provide high dynamic connectivity between data and computing power through innovation in IP network protocols, meeting the requirements of intelligent services. Green, low-carbon networks will be enabled by a three-layer simplified network architecture and three-layer optoelectronic integration.

Future communications networks will support deterministic service experiences critical to the intelligent transformation of industries. Users will be connected to multi-level computing resources: 1 ms latency will be guaranteed for data transmission within cities, 10 ms latency within city clusters, and 100 ms latency through backbone networks. The networks will also provide greater than 99.999% availability, and develop secure, trustworthy network capabilities to support the migration of all systems to the cloud across industries.

Future communications networks will support AI-native. By combining NE status data with AI and innovating in algorithms, the networks will approach the theoretical limit and turn non-determinism into determinism, improving network performance. With the combination of network O&M data and AI, big data analytics, and closed-loop optimization, the networks' automation and all-scenario service capabilities will be comprehensively improved. With the AI-native edge, the networks will also be able to sense diversified service requirements of various industries, thereby improving service experiences.

Future communications networks will support HCS. Wireless, optical, and other multimodal sensing technologies will allow networks to collect environmental data and combine it with digital twin technology to provide industries with the brand-new service capabilities enabled by HCS.

Over 20 years ago, IP technology started reshaping the forwarding architecture of communications networks. Over 10 years ago, cloud technology began to profoundly influence the network management control architecture. Over the next 10 years, AI will be embedded into all layers of the network architecture, driving the networks to evolve toward advanced intelligent twins. To support the development of intelligent networks in the future, networks' computing capabilities will be enhanced, and optoelectronic integration will be adopted to enable green, low-carbon communications networks.

In conclusion, the architecture of the communications network of 2030 will evolve towards cubic broadband networks, deterministic experience, AI-native, HCS, security and trustworthiness, and green and low-carbon networks.

Recommendations

William Gibson, famous science fiction novelist and author of Neuromancer, once said, "The future is already here. It's just not evenly distributed yet." AR, the key technology for integrating the virtual and real worlds, was invented by the Royal Navy 60 years ago, and used for the sighting devices of fighter aircraft. Later, MIT established in the 1980s the Media Lab, which is dedicated to changing the way humans interact with computers and delivering personalized digital experiences.

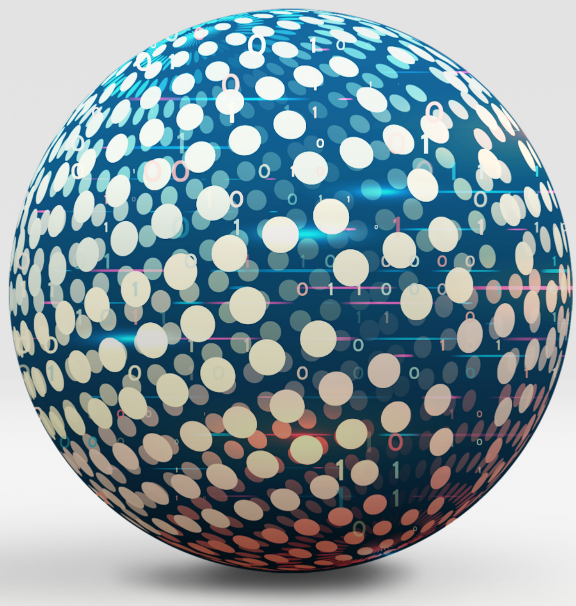
Communications technology and computing technology share the same origin. Less than five years after IBM launched its first personal computer in 1981, the world's first router was invented. Compared with computers, the main distinguishing features of communications equipment are enhanced optical and wireless functions, and network protocol interfaces.

Cloud, AI, and optical, the three key technologies influencing the development of future communications networks, are also reshaping the computing industry. While we may be more familiar with cloud and AI, optical technologies have also been profoundly influencing the computing industry over the past decade. Currently, the industry is focusing on two research areas of optical computing. One is replacing electronic components with optical components to develop optoelectronic integrated computers. The other is using optical parallel processing to build an optical neural network which will increase computing power by 100 times while consuming very little power. The application of optical technologies in computing will also play a part in realizing a green, low-carbon network architecture.

Currently, we cannot find an accurate keyword to represent the target network. 6G/F6G may be the keyword based on the improvement of network capabilities from ubiquitous gigabit networks to 10-gigabit cubic networks. Industrial Internet may be the keyword based on the shift of network application scenarios from consumer Internet to industrial Internet. At the same time, computing power network may be the keyword based on the shift in the nature of services from human-oriented cognition to machine-oriented cognition that supports massive amounts of user data and multi-level computing power services. In addition, optical network may be the keyword based on the evolution of the underlying technology from electronic technologies to optical technologies. The cognitive network or digital twin network may be the keyword based on the improvement of network intelligence from L3 to L5 ADN.

The next decade in communications networks will open up huge space for imagination while also bringing an abundance of uncertainties. All players in the industry need to work together to explore new technology directions and jointly make the vision for the communications network of 2030 a reality.

Computing 2030



Foreword

A decade ago, humanity generated just a few zettabytes[1] of data every year, and mobile Internet, cloud computing, and big data were still in their infancy. Today, these technologies are profoundly changing our world, and computing is playing an unprecedented role.

By 2030, we will be producing yottabytes[1] of data every year. The amount of general computing power in use will increase tenfold, and AI computing power will increase by a factor of 500[2]. The digital and physical worlds will be seamlessly converged, allowing people and machines to interact perceptually and emotionally. AI will become ubiquitous and help people to transcend human limitations. It will serve as scientists' microscopes and telescopes, enhancing our understanding of everything from the tiniest quarks to vast cosmological phenomena. Industries already making extensive use of digital technology will now use AI to become more intelligent. Computing energy efficiency will increase, bringing us closer to low-carbon computing, so that digital technologies can become a tool for achieving the global goal of carbon neutrality.

In the next decade, computing will help us move into an intelligent world – a process of the same epochal significance as the age of discovery, the industrial revolution, and the space age.

Macro trends

After half a century of development, computing has become deeply integrated into every aspect of our work and lives. In the next decade, computing will become the cornerstone of the intelligent world and continue to support economic development and scientific advances.

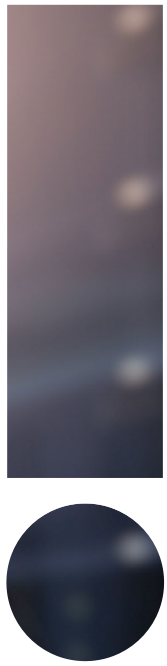
Looking ahead to 2030, many countries and regions, including China, the EU, and the US, will prioritize computing in their national strategies. China's 14th Five-year Plan and Vision 2035 define high-end chips, artificial intelligence (AI), quantum computing, and DNA storage as technologies of strategic importance for the country. The EU's 2030 Digital Compass: the European Way for the Digital Decade lays out a plan whereby, by 2030, 75% of EU companies will be making full use of cloud, AI, and big data, and the EU will have its first homegrown quantum computer. The US has reintroduced the Endless Frontier Act, which authorizes the government to legislate and make grants to promote US research in areas such as AI, high-performance computing, semiconductors, quantum computing, data storage, and data management technologies.

In 2030, the digital and physical worlds will be seamlessly converged. People and machines will interact with each other perceptually and emotionally. Computing will be able to simulate, enhance, and

recreate the physical world. Hyper-real experiences will drive computing to the edge, and necessitate multi-dimensional collaborative computing between cloud and edge, between edge and edge, and between the digital and physical worlds. AI will evolve from perceptual intelligence to cognitive intelligence and develop the capacity for creativity. It will become more inclusive and make everything intelligent. As the boundaries of scientific exploration continue to expand, the demand for computing power will increase rapidly. Supercomputers that can perform 100 EFLOPS[2] and a new, intelligent paradigm for scientific research will emerge. In the push toward global carbon neutrality, computing of the future will be greener, and service experience will get better.

The semiconductor technologies that computing relies on are approaching their physical limits, and this will spark a golden decade of innovation in computing. Innovation in software, algorithms, architecture, and materials will make computing greener, more secure, and more intelligent. It is estimated that by 2030, global data will be growing by one yottabyte every year. Total general computing power will see a tenfold increase and reach 3.3 ZFLOPS, and AI computing power will increase by a factor of 500, to more than 100 ZFLOPS[2].

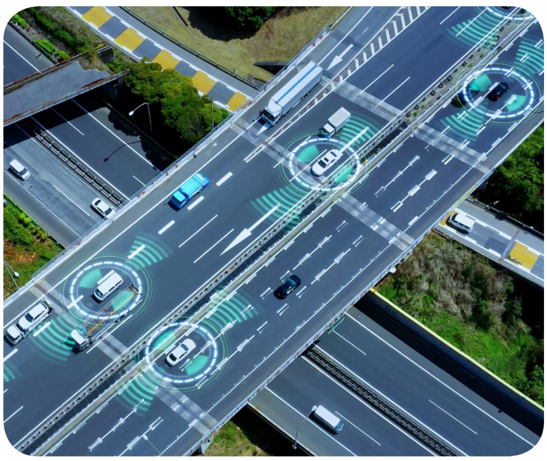
Future computing scenarios



Smarter AI

AI-enabled smart transportation

By 2030, the number of electric vehicles, vans, heavy trucks, and buses on the road worldwide is expected to reach 145 million. Today, all these means of transportation run up against the limited capacity of our road networks. Intelligent transportation is the key to solving this problem.



There will be a wide range of intelligent transportation use cases that use cameras, radars, and weather sensors to collect various types of data. At the edge, data will be read to identify vehicles, traffic accidents, road conditions, and more, and to generate a multidimensional representation of a stretch of road. In the cloud, a digital twin of roads across the city will be produced, constituting a multidimensional representation of real-time and historical road conditions. Policy-based computing on the cloud can help generate different commands for every vehicle and every road, and manage vehicles and traffic signals.

The sheer volume of data involved means that the bottleneck to be addressed is not the capacity of our roads, but the capacity of our computing networks. Suppose a vehicle runs two hours a day on average. For each running vehicle, the compressed data uploaded per second will increase from 10 KB today to 1 MB in 2030, meaning that for every 100,000 intelligent connected vehicles, about 720 TB of data will need to be transmitted every day. The data generated by each running vehicle will need to be frequently exchanged between the vehicle itself and the city.

With the help of intelligent transportation infrastructure that can store and analyze such massive amounts of data, urbanites can look forward to quicker daily commutes (15–30 minutes shorter, on average), less frequent traffic accidents, and vehicles with lower carbon footprints. Increased computing power will boost transportation safety, efficiency, and user experience, facilitating socioeconomic development.

AI-enabled autonomous vehicles

L4 autonomous vehicles will be commercially available on a large scale, and data will be continuously sent to the digital twin. AI learning and training will continue in the digital world, so that AI models will become smarter and eventually outperform humans in coping with complex road conditions and extreme weather. In time, AI will even make L5 autonomous vehicles a reality. The computing power required for intelligent driving will far outstrip what Moore's Law can provide. The corner case library will continue to expand and the demand for computing power will increase. By 2030, an L4 or higher-level autonomous vehicle will require computing power of 5,000 TOPS.

The training of AI models will involve introducing unsupervised or weakly supervised learning into closed-loop data, and using images and visual information obtained from vehicle snapshots to support automatic, unsupervised, video-level AI machine learning and training. Autonomous vehicles demand device-cloud computing. In the future, a vehicle manufacturer will need at least 10 EFLOPS of computing power on the cloud.

Smart cities

Urban areas make up 2% of the world's land surface, and are home to more than 50% of the world's population. Cities consume two thirds of the world's energy and are responsible for 70% of global greenhouse gas emissions (including over 25 billion tons of carbon dioxide). Smart city governance will be the way forward for cities that want to achieve sustainable development. IoT sensors will collect the environmental data that is needed to support the operations of smart cities. In the future, every physical object will have a digital twin. Digital cities made up of digital buildings, digital water pipes, and other infrastructure will be a powerful tool for intelligent urban management. Smart city governance will aggregate 100x more data than conventional city governance and make our cities more efficient.

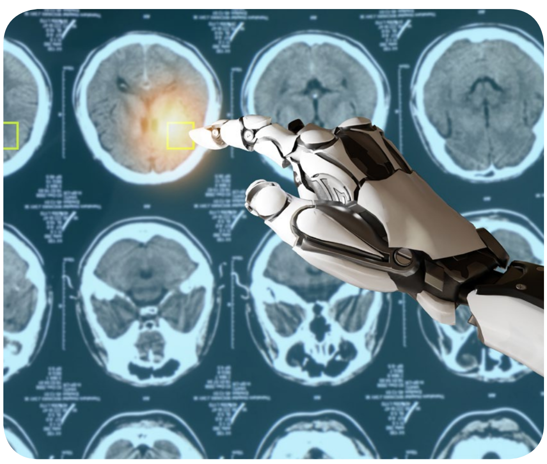
The data storage and analysis capabilities of smart energy infrastructure will make it possible to manage urban energy supply and demand in one system, and to schedule urban energy more efficiently through real-time data processing. For example, a real-time energy efficiency map can be drawn based on urban energy consumption data. This will help dynamically monitor energy usage and ensure targeted energy scheduling, which will cut average electricity consumption in peak hours by more than 15%.

The quality of public services like meteorology, oceanography, and earthquake prediction can deeply affect the life of each resident in a city, and these services rely on massive data computing and processing. With a greater volume and diversity of urban and natural environment data, smart public services will help better predict the impact of weather, oceans, and earthquakes on urban life, making cities more resilient to extreme events. With these smart public services, residents can gauge the impact of climate or emergency events on themselves and their communities using the push messages tailored to their geographic locations.

Data will be at the core of efficient operations of smart cities. How can we effectively manage and use the massive data generated? This is a question we must answer if we want to promote the development of smart cities.

More inclusive AI

AI-enabled precision medicine



In the healthcare sector, AI is already able to automatically identify tiny lung nodules, saving doctors a lot of time compared to conventional identification with the naked eye and manual tagging. AI will play a bigger role in more complex consultations. It will be deeply integrated into the diagnosis process, providing explainable diagnoses and predicting outcomes. The future will bring a new model of healthcare in which AI will provide solutions, and the role of doctors will be to check and approve them. The World Health Organization estimates there will be a shortage of 18 million healthcare professionals by 2030, and AI offers a viable solution to this problem.

AI-enabled drug screening

AI will become more transparent. It will not make judgments inside a black box. Instead, it will show the reasoning behind its conclusions so that we can understand its thinking process. Greater transparency will allow AI to play a greater role in more domains and help us perform more complex tasks, such as screening antiviral drugs. AI will be able to tell us why the drugs are selected, instead of just giving us a list of drugs selected. Results on their own, without the decision-making processes, cannot help us make informed decisions.

AI-enabled personalized education

The process of AI training is also a process of better understanding ourselves. AI makes it more important to understand human intelligence and how the human brain works. This will in turn push humans to rethink and reform education[3]. AI of the future will change our learning and cognition processes. For example, AI instructors will analyze students' behavior, habits, and abilities in detail and then develop personalized teaching content and plans. This will help students acquire knowledge more easily and realize their full potential.

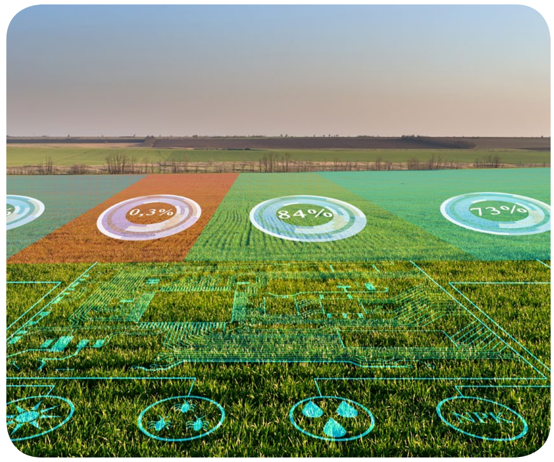
AI will be integrated into every aspect of our lives. It will allow us to analyze, create, and study more efficiently, and open up high-quality resources to many more people. AI will make services like precision healthcare, creative design, cultural education, elderly care, community services, and autonomous driving more inclusive.

Deeper perception

By 2030, there will be 200 billion connections. Hundreds of trillions of sensors will be collecting information about the physical world, including temperature, pressure, speed, brightness, humidity, and chemical concentration. Turning this basic data into sensory information to give robots vision, hearing, taste, smell, and a sense of touch will require deeper perceptual capacities. Issues of data quantity and latency mean that the process of computing for generating sensory information must be completed at the edge. The edge will therefore need to be able to intelligently process data, which would include simulating how the human brain processes information. In the future, a large amount of perceptual computing will be completed at the edge, where about 80% of data will be handled.

Perceptual intelligence makes the gathering and analysis of vast flows of data possible. It will enable more industries to perceive their work, and to build digital twins in the cloud. Digital twins remain in constant balance with their physical models, and support digital innovation.

Intelligent agriculture



In the future, an intelligent space-air-ground integrated network will be built and continuously optimized for remote sensing and monitoring of agricultural information. Advanced information technologies, such as the Internet, the Internet of Things (IoT), big data, cloud computing, and AI, will be deeply integrated with agriculture. This will create a brand-new model of agricultural production that features agricultural information sensing, quantitative decision making, intelligent control, targeted investment, and personalized services. Applications like smart fields, smart greenhouses, smart farming, smart planting, and spraying drones will have increased demand for edge AI computing. Intelligent agricultural sensing and control systems, intelligent agricultural machinery, and autonomous field operation systems will be deployed. These will promote the development of e-commerce, food source tracing and anti-counterfeiting, tourism, and digitalization in the agriculture sector. Agriculture will become more digital, connected, and intelligent.

Intelligent control of equipment

AI will be increasingly adopted in enterprises' production systems. It will support every aspect of company operations, improving workflows, staffing, and coordination across different departments and different sites. Over the next decade, AI will bring massive improvements in quality and cost savings in critical production processes. With the support of AI, manufacturers can achieve intelligent operations and management, massive data analysis and mining, and lower-latency diagnosis and warning.

The Made in China 2025 plan has a target of universal adoption of AI in key manufacturing sectors, with 50% reductions in operating costs, production time, and defects in showcase projects. In deep learning use cases, such as bearing fault diagnosis, steel furnace thermal anomaly detection, and power device overhaul, factories can use AI to diagnose problems and send warnings faster, detect production problems more efficiently, and shorten order delivery cycles.

Production robots

Workers who once operated machines in harsh environments will be able to operate them remotely. More non-operational tasks of enterprises will involve AI. Humans and machines will seamlessly collaborate with each other. AI will reshape enterprises' business operations at every level, from product design, production, and sales to enterprise architecture, employee hiring, and training. For example, enterprises will use AI to analyze factors such as economic development and current events and assess their own growth and trends across the industry. They will then optimize their production plans and create new solutions as input for decisions on new product concepts. AI will play an especially important role in flexible manufacturing that meets personalized needs. It can design customized products and even generate new product designs based on demand changes and product usage. We project that by 2030, there will be 390 robots per 10,000 workers. These robots will be able to accurately understand people's instructions, sense the environment, and provide recommendations.

Lights-out factories, with no human workers at all, are already in widespread use. AI robots are busy on production lines and in logistics, freeing humans from repetitive, boring tasks. In the future, machines will help humans handle dangerous jobs in harsh environments, even in highly variable scenarios. People will no longer need to operate machines onsite. Instead, they will be able to command the machines remotely from the safety and comfort of a control room.

In the mining industry, for example, China has set the goal of achieving intelligent decision-making and automatic collaborative operations by 2025 in large coal mines and mines where severe disasters have occurred in the past. Key roles down in the mines will be assumed by robots, and few, if any, actual workers will have to work underground. The longer-term goal is to build an intelligent coal mine system featuring intelligent sensing, intelligent decision-making, and automatic execution by 2035.[4]

AI will make enterprises more intelligent as it will play a greater role in creative work, rather than just operational work. AI will be more deeply involved in our thinking process and will better interact with people while showing the reasoning behind its conclusions. It will become more reliable and take on a bigger role in complex fields that require high-quality decisions, such as finance, healthcare, and law. In the next decade, AI will continue learning about the physical world and will become smarter. AI will move beyond well-understood scenarios and play a bigger role in empowering humans to do better in more complex tasks. AI will help people transcend human limitations.

An experience beyond reality

Intelligent interaction in living spaces

The AI of today has already helped people complete tasks that were impossible in the past. For example, we can use the cameras on our phones to identify plants and obtain information about their habits and how to grow them. Robots are helping humans perform better. For example, exoskeleton robots can help patients recovering from accidents. Home robots can perform intelligent work like keeping the elderly company and doing household chores. It is estimated that more than 18% of homes will use intelligent robots by 2030.

When AI participates in human thinking and creation, it must be able to explain its thought processes in terms that people can understand. This means that AI needs to be able to use natural language to articulate the logic behind its recommendations. AI will make a leap from perception to cognition, and from weak AI to strong AI.

AI has already made initial attempts at poetry writing and painting. The AI of the future will be able to perform more complex creative work, like film making, art, and industrial design. AI will provide highly customized content services, so that people can get a tailor-made painting or movie at any time. When watching a movie, the audience will be able to decide how the story goes. Based on audience choices, AI will analyze potential storylines and develop the video in response. Each viewer will experience the movie differently, making the content richer. It will also be possible for people to supply a theme and let a creative AI fill in the blanks. This will inspire our creativity and add another layer of richness to our lives.

AR/VR in living spaces

Data will create many digital spaces, such as virtual tourist attractions, holographic conferences, and virtual exhibitions. These digital spaces, together with the physical world, will form a hybrid world. Virtual tours can give us a true-to-life experience of scenery on the other side of the world. They will also allow us to sit side by side and talk with friends thousands of miles away, or have wide-ranging conversations with luminaries of the ancient world. The way people communicate with other people, communities, nature, and machines will be revolutionized, and our ways of living, work, and study will be redefined. It is estimated that by 2030, more than 30% of businesses will operate and innovate digitally, and there will be one billion augmented reality (AR) and virtual reality (VR) users.

Virtual world / Metaverse in living spaces

The seamless convergence of the digital and physical worlds requires the ability to accurately perceive and recreate the physical world, and the capacity to understand user intentions in the hybrid world. The demand for a hyper-real experience means that computing will be brought closer to the edge. Multi-dimensional collaborative computing is required between cloud and device, device to device, and between the virtual and physical words. The physical world will be modeled and mirrored on the cloud, and following a process of computing and the addition of virtual elements, will be recreated digitally. Edge devices will be able to hear, see, touch, smell, and taste, and real-time interaction between people and devices will be possible. Multi-dimensional collaborative computing will change a user's environment into a supercomputer that is able to compute environment information, identify user intentions, and display a virtual world using technologies such as holography, AR/VR, digital smell, and digital touch.

More precise exploration into the unknown

The "high-performance computing (HPC) + physical models" approach has been widely applied in many scientific domains. As humans continue to study quantum mechanics, life sciences, the Earth's atmosphere, and the origins of the universe, our cognitive boundaries will continue expanding to embrace phenomena at both the subatomic and cosmological scales, in which the distances can be as short as 10^{-21} m, or as vast as 10^{28} m. The amount of data and computing that scientists have to process will grow exponentially. The amount of computing power available in the digital world determines how deep and how broad we can explore in the physical world.

CERN, the European Organization for Nuclear Research[5], built a computing pool by connecting supercomputers located worldwide. Scientists used this pool to analyze nearly 100 petabytes of data generated by its Large Hadron Collider (LHC), and ultimately proved the existence of the Higgs boson in 2012. The CERN plans to use the High-Luminosity LHC (HL-LHC), a major upgrade of the LHC, by the end of 2027, which will be able to produce more than 1 billion proton-proton collisions per second. The amount of data to be computed will be 50–100 times greater than that used to prove the existence of the Higgs boson, and zettabytes of data will need to be stored. By 2030, computing will help scientists solve basic problems in more domains.

Environmental monitoring

Environmental protection is a top priority for humanity. New technology and equipment will be powered by AI to ease environmental problems such as the greenhouse effect, soil desertification and salinization. Models built based on big data will help predict the results of different management measures, which can be fed back to algorithmic models to come up with better governance models, like accurately locating pollution sources and predicting pollution diffusion.

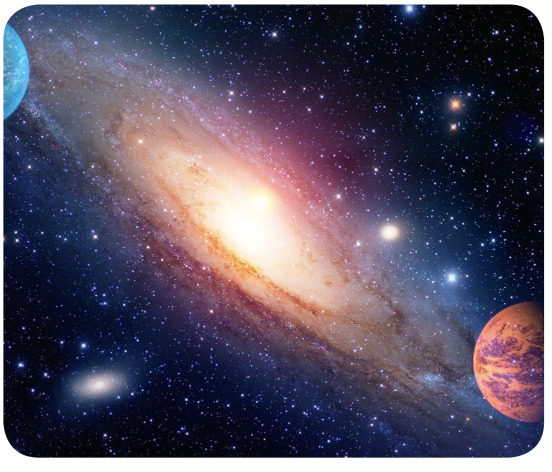
Weather forecasting

Future weather forecasts will use more complex dynamic numerical models to predict the weather more accurately. Potential applications include weather radar quality control, satellite data inversion and assimilation, as well as weather and climate analyses (e.g. short-range and imminent weather forecasts, probability forecasts, typhoon forecasts, extreme or catastrophic weather warnings, storm environment feature classification, and environmental forecasts). Take short-range local weather forecasts as an example. Torrential rainfall in a short period of time is an extremely destructive phenomenon, but it is difficult to forecast when it will happen, because it requires massive amounts of data and huge computing power. If we were to increase the granularity of weather forecasts from the current 10 x 10 km to 1 x 1 km, that would increase the amount of data and computing power needed by two or three orders of magnitude. It is expected that by 2030, with the emergence of supercomputers that can perform 100 EFLOPS, more accurate climate simulations and weather forecasts will be possible.

Seismic and ocean prediction

In the future, AI will be used to monitor earthquakes and estimate the focus of earthquakes in real time, which will make prediction much more accurate. It is very time-consuming to calculate the focal mechanism (also called a fault-plane solution) based on seismic records. Ever since seismologists began calculating fault plane solutions in 1938, focal mechanism parameters have been a huge challenge. AI can effectively solve this complex computing problem. Seismic data can be used to train AI neural networks, which can make prediction systems more accurate and reliable. This will further drive breakthroughs in earthquake prediction.

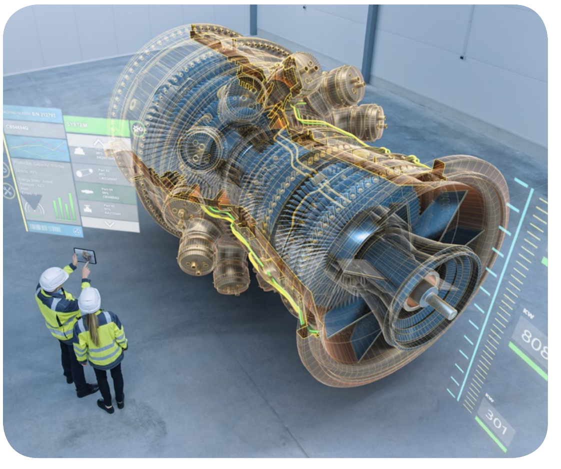
Exploring the structure of the universe



The large-scale structure of the universe is one of the most important current fields of science. Scientists are studying the formation and evolution of cosmic structures over time, to find answers to questions about the composition of the universe, the process of cosmic evolution, dark matter, and dark energy. The conventional method is to use a supercomputer to calculate the evolution of various large-scale structures in the universe based on our current physical theories, and then compare the results with observed data. This, however, requires accurate calculations for hundreds of thousands—or even millions—of cosmological objects. As of today, there are two trillion galaxies and countless planets in our observable universe. Even if we were to pool all of the world's computing resources together, it would still be impossible to complete the calculations.

More accurate simulation of the real world

More precise wind tunnel simulation



Computer wind tunnel simulation is now an important test method for high-speed vehicles such as aircraft, high-speed trains, and automobiles. However, due to the huge amount of computing required for these simulations, the testing system needs to be broken down into sub-systems like tire and engine, and then further divided into even smaller systems to get precise simulation results. This will pose new challenges in verifying whether system design meets requirements. As computing power will increase by 2 to 3 orders of magnitude in the future, wind tunnel simulation is expected to be used in larger sub-systems, or even for the entire system.

AI-enabled research on new drugs

When it awarded the 2013 Nobel Prize in Chemistry to three scientists "for the development of multiscale models for complex chemical systems", the Nobel Committee stated that, "Today the computer is just as important a tool for chemists as the test tube. Computer models mirroring real life have become crucial for most advances made in chemistry today."

Quantum mechanics/molecular mechanics (QM/MM)[6] modeling is one of the most reliable methods for simulating the catalytic mechanisms of enzymes. The high-precision QM model is used in core regions of the enzyme, and the low-precision MM model is used in peripheral regions. This approach combines the accuracy of QM and the fast speed of MM. To use this model to simulate the growth and reproduction of 0.2-micron Mycoplasma genitalium cells over a period of two hours would take the supercomputer Summit[7] one billion years. For more complex studies of thinking, memory, and behavior in the human brain, vastly more computing power would be needed. To predict the response of the human brain to a particular stimulus, it would take Summit 10^{24} years to simulate one hour of brain activity[8].

Turing Award winner Jim Gray divided scientific research into four paradigms: experimental, theoretical, computational simulations, and data-intensive scientific discovery. As we continue with research in dynamically complex fields such as biology, material science, chemistry, and astronomy, it will be increasingly difficult to make progress relying on traditional computation methods. The curse of dimensionality may occur as the number of variables and degrees of freedom increase, and this means that the demand for computing power will increase exponentially.

AI will provide a new solution to the curse of dimensionality and a new path for scientific research. Using conventional methods, it would take scientists several years to analyze the folding structure of a single protein, but with the help of AI, scientists are able to learn the 18,000 known protein structures and produce simulations with atomic levels of precision for unknown protein structures within just a few days. This kind of research is giving us new ways to discover therapies that could prevent and treat cancer, dementia, and other diseases caused by changes in the structure of proteins in cells. The winners of the 2020 Association for Computing Machinery (ACM) Gordon Bell Award[9] simulated a system of more than 100 million atoms using AI. The system was more than 100 times larger than current models and the time-to-solution was 1,000 times faster. This project has brought accurate physical modeling to larger-size material simulation[10].

The scientific computing of the future will rely on a combination of data, computing, and intelligence, which will give rise to new paradigms for scientific research. AI will study existing knowledge, analyze, and draw new conclusions. Online iteration, combined with traditional modeling methods, will speed up scientific exploration and further expand people's cognitive boundaries.

Data-driven business innovation

Computing-enabled data value mining

Cloud computing and big data are now the foundation for digitalization in any industry. They are driving the digitization processes that are making many industries more efficient. A key feature of digitization is that it improves the matching of producers to consumers. Examples include e-commerce platforms and online-to-offline (O2O) models.

10-fold increase in the demand for new services

Full-stack, serverless device-edge-cloud computing will become a key technology for enterprises to modernize and go digital and intelligent. Programming languages, language runtime, as well as application scheduling, operations, and O&M based on the cloud-native computing model will be the foundation for building modern full-stack serverless software. This will allow all applications to be migrated to the cloud, and will result in tenfold gains in performance, efficiency, and cost reduction.

More efficient operations

More efficient resource utilization

The wide adoption of cloud allows companies to use computing resources more easily and quickly. New computing technologies will give companies access to these resources in smaller packages, available more quickly. This will reduce waste in the way companies use these resources. For example, before the cloud, central processing units (CPUs) were used only 10% of the time. Containerization raised this indicator up to 40% or higher. In the future, the wide adoption of new resource management technologies will reduce waste by 50% or more.

Software-defined operations

IT is now one of the core components of any operational system. Internet companies use a DevOps[11] model and are becoming more agile and efficient. By 2030, companies in the manufacturing sector will achieve highly efficient software-defined operations in their more complex value chains.

The industrial Internet will connect the supply chain, manufacturing, maintenance, delivery, and customer service processes. All companies will form a value network that spans the globe. The digital transformation inside a company will expand into an improvement of entire industries, which will translate into greater synergies. And the dependence on data will change: from a company being highly dependent on its own data to being dependent on data from up and down the value chain, or even from other industries.

Companies of the future will use software to manage complex cross-organizational coordination and to define their own operations. For example, they can use technologies like robotic process automation, no-code/low-code development, and AI-supported natural language programming to invoke the capabilities of robotic automation software, obtain required services, and orchestrate business processes. This will mean that even personnel without much expertise can improve processes and fix problems on their own.

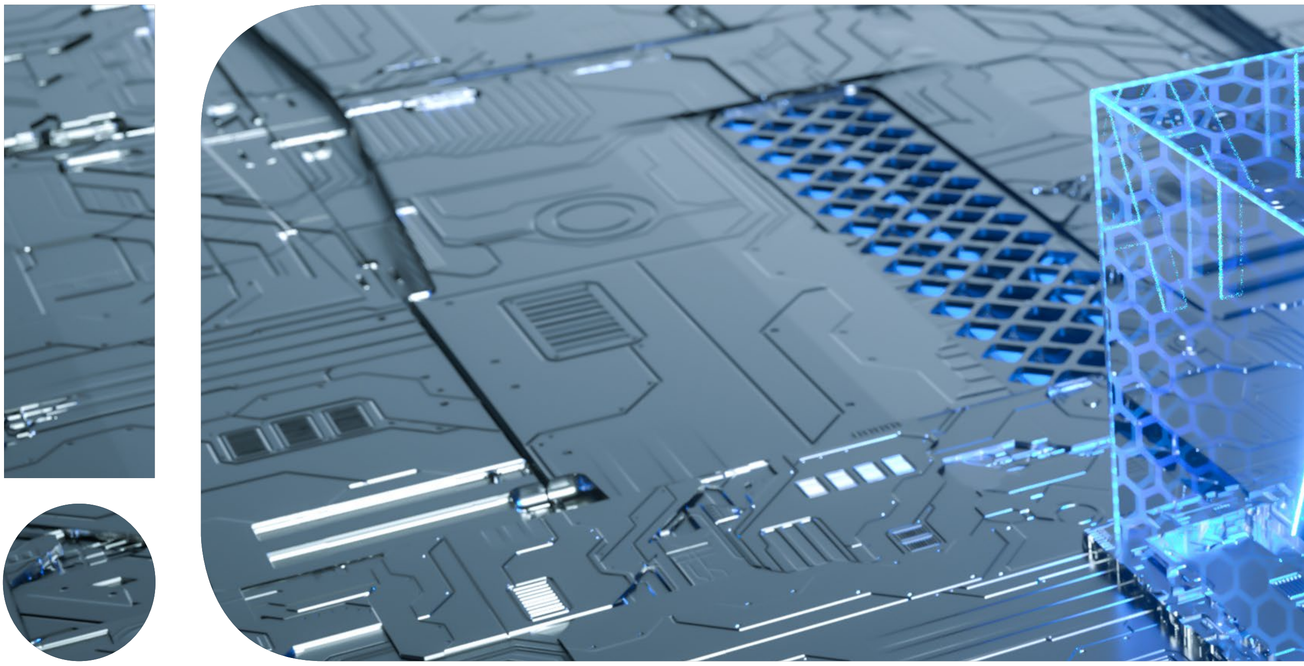
Low-carbon data centers

By 2030, data centers (DCs) will deliver a 100-fold increase in computing power while achieving low-carbon operations, giving companies access to green computing resources.

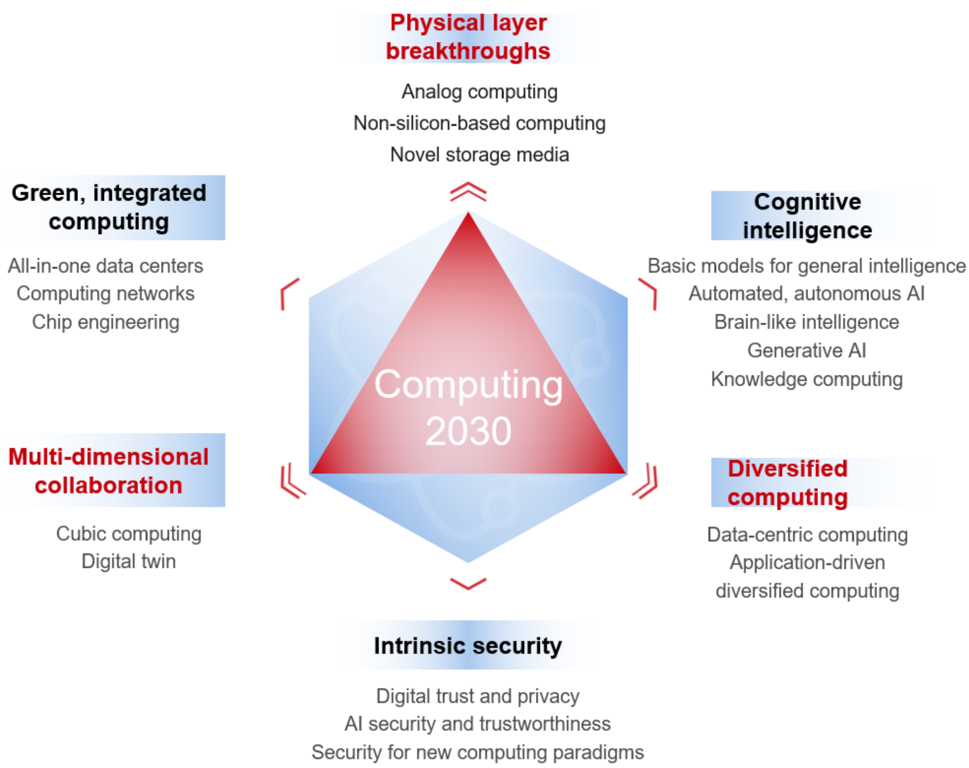
New computing architectures will massively boost energy efficiency. In a conventional computing

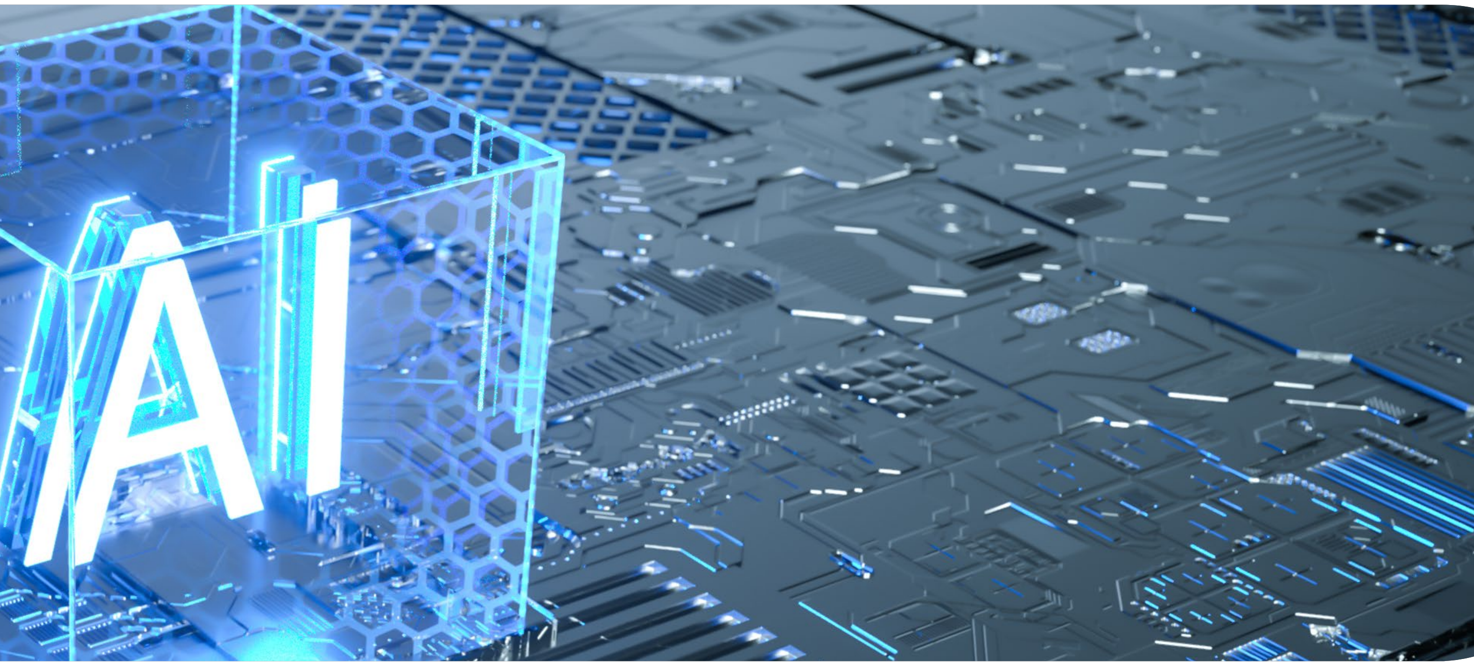
process, more than 60% of the energy is used shuffling data. The data centers of the future will make computing tens of times more energy efficient. Analog computing such as quantum computing and analog optical computing will be important sources of computing power, driving energy efficiency indicators up exponentially.

In the push toward carbon neutrality, data centers will be positioned near energy resources and near areas with high computing demand. This will change the computing architecture on a larger geographic scale. Computing networks can balance the needs of green energy, latency, and costs and achieve optimal global power usage effectiveness (PUE) and cut carbon emissions. Tasks like AI model training or gene sequencing can be done in places with abundant green energy sources and low temperature while tasks like industrial control and VR/AR can be performed in places that are closer to customers' production environments.



Vision and key features of Computing 2030





Cognitive intelligence

AI is evolving from perceptual intelligence to cognitive intelligence. Cognitive intelligence is an advanced stage of AI evolution, at which machines are given the capabilities of data understanding, knowledge representation, logical reasoning, and autonomous learning. It will make machines powerful tools for humans to become more capable and change the world. In the evolution toward cognitive intelligence, semantic and knowledge representation and logical reasoning are important means of cognition, and multimodal learning is an important way to realize information fusion and collaboration. By building large-scale multimodal basic models, AI systems can learn converged representation of multiple types of information to establish multimodal transfer and concordance. This improves an AI system's ability to perceive and understand complex environments, thereby enabling AI applications to work in different environments and on a wide range of different tasks.

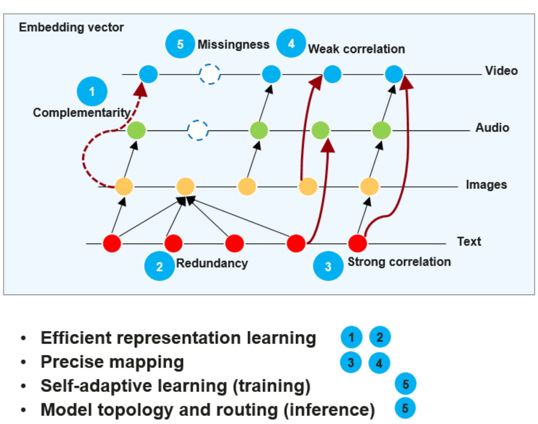
Basic models for general intelligence

AI's ongoing evolution from perceptual intelligence to cognitive intelligence: AI has delivered computational intelligence and perceptual intelligence; it is now on the way to developing cognitive intelligence. Machines have strengths in computing speed and storage. Today, deep learning and big data analytics are enabling machines to perform certain tasks through vision, hearing, and touch, similar to how a human being would. Cognitive intelligence will allow machines to understand and reason like humans. When machines have these abilities, they will become powerful tools that help humans to understand and change the world.

Improving the ability of machines to generalize in the process of solving problems is an important evolutionary path from weak AI to strong AI. AI systems will be given the ability to solve multiple different problems using large-scale, domain-general basic models that can generalize from one situation to another; from one modality to another; and from one task to another.

Multimodal learning is an important approach to building basic models: In multimodal learning, data heterogeneity is the first problem that needs to be solved, which creates a number of challenges: (1) How can complementarity and redundancy in multimodal data be used for representation learning? (2) How can the strong and weak correlations between these representations be processed to produce relational vector maps between modalities? (3) During adaptive learning and multimodal transfer for model training, how can we keep model accuracy within an acceptable range when one piece or a type of data is missing in a certain modality? (4) During inference, when one piece or a type of data is missing in a certain modality, how should model topology and routing adapt for maximum inference gains?

Based on progress to date, we expect that multimodal models will become capable of multimodal, self-supervised learning and the transfer of generally-applicable knowledge. This means that tasks in different domains can be approached using the same multimodal framework.



Breakthroughs in multimodal learning will require advances in the following key areas:

First, the technology to tag training data to associate captions, audio, video frames, etc. Second, multi-stream codecs from single-modal pre-training models to multimodal association coding, which enables multimodal learning with weak information association, with the decoder providing support for cross-modal transfer and generation. Third, self-supervised learning technology, involving semantic alignment and inter-modal predictions between text, speech, vision and other modalities. Fourth, technologies for downstream task fine-tuning that support multimodal semantic understanding and multimodal generation. Fifth, multimodal models that are smaller.

Automated, autonomous AI

Deep learning has not yet successfully developed beyond the stage of supervised learning. Data cleansing and tagging, and the design, development, training, and deployment of models, all require extensive manpower. Development in domains such as transfer learning, few-shot/zero-shot learning, self-supervised/weakly-supervised/semi-supervised/unsupervised learning, and active learning, will eventually drive AI to reach autonomy, eliminating our dependence on manual training, design, and iteration of models. AI autonomy will make models more homogenized, with the same model serving multiple purposes. The

amount of data learned will increase without manual intervention. Models will learn to pick up and train on new data as they operate, improving their capabilities in the process. The scaling of data and the prevalence of online learning will lead to more centralized model production. Industry applications in multiple domains will converge to a handful of or even a single ultra-large model.

However, there are still some major challenges that the developers of autonomous AI must overcome:

1) Training signals can be incorporated online in a self-supervised fashion, so that feedback is available during inference, not just during the training phase.

2) At present, a model's learned representations are formed without constraints. The representations that result from different training sessions may be radically different even if they are of the same model structure. Models need to overcome the problem of catastrophic forgetting, so that learning can be carried out continuously, and training and inference can converge into a single process.

3) Models manually designed for different tasks need to be replaced by models that can learn to encode for different tasks and switch between different modalities in context and on demand.

Brain-like intelligence

Current deep learning technology is largely data-driven and relies heavily on large quantities of labeled data and powerful computing. Backpropagation training algorithms need continuous enhancement in terms of model robustness, capacity to generalize, and interpretability. Drawing on and imitating the way biological neurons work, brain-like intelligence creates digital neurons with richer functionality and promises to enable learning that is event-triggered, uses pulse encoding, and is coordinated both temporally and spatially. Using neurodynamic principles, brain-like computing can deliver both short-term plasticity and long-term memory, and is capable of adaptive adjustment and learning in open environments. Inspired by the sparse connectivity and recursive form of the biological brain, no computation will be performed without pulses, which greatly reduces energy consumption. In the next five to ten years, if breakthroughs in related technologies are made, brain-like computing is likely to begin to outperform other models, as well as consuming less power, in many computing tasks, and be applied in smart devices, wearables, and autonomous vehicles.

At present, brain-like systems are still inferior to the deep learning systems in terms of learning efficiency and computing accuracy, because our understanding of the human brain's learning mechanisms is too shallow. Research in this field will need to advance in two major areas. First, from the bottom up, the systems can simulate pulses in the biological brain, and use neuromorphic chips to recreate neurons and synapses at scale, which should support low power and low latency in time-dependent applications. Second, from the top down, more comprehensive theories of neurodynamics and cognition are needed from a functional perspective, which can then be applied in combination with AI to achieve more robust and general intelligence.

Generative AI

Generative AI powers automated content production: It allows computers to abstract the underlying patterns related to a certain input (such as text, audio files, and images) and use it to generate expected content. Generative AI is used in identity protection, image restoration, audio synthesis, and antimicrobial peptide (AMP) drug research, among other fields.

Generative AI generates data that is similar to training data, rather than simply replicating it, so it can incorporate human creativity into processes of design and creation. For example, a game generation engine can generate 3D games to test the vision, control, route planning, and overall gaming capabilities of an intelligent agent, in order to accelerate the training of the agent. In the development of generative AI applications, the key objective is generation models that are capable of evolving and dynamically improving over time.

The field of generative AI is facing the following challenges:

1) Some generative models (such as generative adversarial networks, or GANs) are unstable, and it is difficult to control their behavior. For example, generated images may not be sufficiently accurate; they may not produce the desired output; and the cause cannot be located.

2) Current generative AI algorithms still require a large amount of training data and cannot create new things. To address this, algorithms capable of self-updating and evolving are needed.

3) Malicious actors can use generative AI for spoofing identities and can exploit vulnerabilities in AI tools to conduct remote attacks, resulting in serious threats to online information security such as data breaches, model tampering, and spam.

Knowledge computing

The industrial application of AI needs the ability to make high-quality decisions based on expert domain knowledge across multiple disciplines. A complete technical system is needed for knowledge extraction, modeling, management, and application. In the next decade, knowledge computing will make a leap forward: In knowledge extraction, the data source will not only include text and structured features, but also complex and multi-level knowledge, which includes several areas of research such as multi-modal knowledge alignment, extraction and fusion, complex-task knowledge extraction, and cross-domain knowledge extraction.

Knowledge modeling will move from developing scenario-specific, atomized, automated, and large-scale knowledge graphs to integrating these scenario-specific graphs into general knowledge graphs. The applications of knowledge will develop, from simple query and predictions to high-order cognitive tasks such as causal reasoning, long-distance reasoning, and knowledge transfer.

The application of knowledge computing will require breakthroughs in algorithms for massive retrieval of sparse information, capture of dynamic-length knowledge, knowledge attention, and large-scale graph computing. The training schema for cognitive intelligence will require advances in high-frequency knowledge retrieval during training and inference and feature enhancement based on knowledge combination. In terms of computing, it will be necessary to solve a number of problems such as training and inference for high-frequency random retrieval, high-speed data communication, and some graph computing puzzles such as random walk and structural sampling.

Intrinsic security

The migration of computing resources to the cloud has gone beyond traditional security boundaries. Traditional add-on security based on the division of trust and untrust zones cannot withstand new types of attack. In order to protect users in an evolving threat landscape, security must become intrinsic. Specifically, that means:

- Security must be an intrinsic capability of a system and a basic feature of chips, firmware, and software.

- Security should be ensured throughout the entire data processing lifecycle (including storage, computing, and transmission), to defend against all kinds of attack.

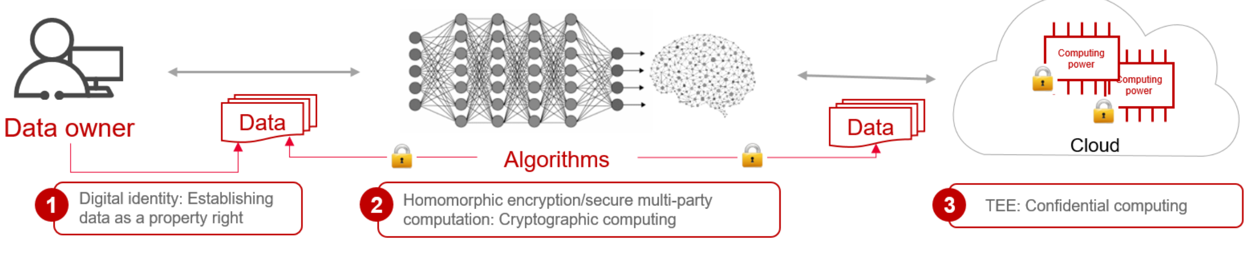
- A hardware-based root of trust is essential. Due to the system access control model, security functions must be implemented based on the highest hardware privilege in order to provide reliable security services on the operating system and applications. In addition, hardware acceleration can improve the performance of security services.

- The principle of open design should be adopted, which means the security of a mechanism should not depend on the secrecy of its design or implementation. Security services should be made open source. This way, service software can embed security into itself based on its own architecture pattern to ensure service security.

Digital trust and privacy

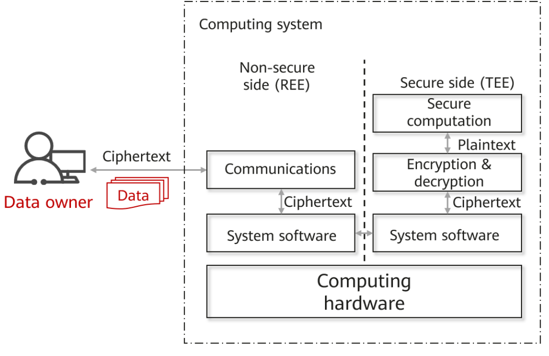
Data processing, in essence, is the process of computing data using algorithms. If all the three elements – computing power, data, and algorithms – are controlled by the data owner, data security and privacy are not really an issue. However, during cloud computing, these elements are often separate. Algorithms and computing power are provided by computing service providers, while users (i.e., data owners) need to upload data to the cloud for processing. Even if users trust the computing service providers, they don't trust the computing service provider administrators who have access privileges. Therefore, the major security challenge of cloud computing lies in protecting user data and privacy. To address this challenge, digital trust systems need to be rebuilt.

Governments worldwide have enacted data protection laws, providing a legal basis for rebuilding digital trust systems. Digital identity and privacy computing are key technologies in this rebuilding process. Digital identity is the



basis for establishing data as a property right, while privacy computing can be used for data analysis and processing without compromising data.

1) Hardware isolation technology that is based on trusted execution environments (TEEs) can be used to process sensitive data. However, the completeness of hardware security isolation mechanisms cannot be mathematically proven, so it may be hard for the mechanisms to prove their own innocence, and security vulnerabilities may exist. On the other hand, TEEs have a smaller impact on performance than cryptographic technology. In the future, privacy computing based on TEE technology will be widely adopted in public cloud, Internet, and major enterprise services. It's expected that TEE technology will be used in more than half of all computing scenarios by 2030.



2) Homomorphic encryption and secure multi-party computation are considered to be the most ideal privacy computing technologies because it is possible to verify their security level mathematically. However, both of these technologies come with a significant performance cost (their processing is over 10,000 times slower than conventional computing). Significant performance improvements must be made if these technologies are to be applied in real-world scenarios. Approximation algorithms are maturing, and homomorphic encryption and secure multi-party computation technologies have already been applied in face authentication, the sharing of health data, and other specific domains. In the future, further breakthroughs based on hardware will be made in these technologies, which are expected to be commercially used in scenarios that require high security, such as in finance, healthcare, and other security-conscious sectors.

3) Multi-party computation is built on the sharing of secret slices between multiple parties. Cryptographic methods like zero-knowledge proofs come with a high performance overhead. However, TEE technology can greatly improve the performance of multi-party computation, while being used to enable the sharing of secret slices between multiple parties. In addition to that, security can be proved mathematically based on TEEs. So this technology is expected to be used in various scenarios in the future.

AI security and trustworthiness

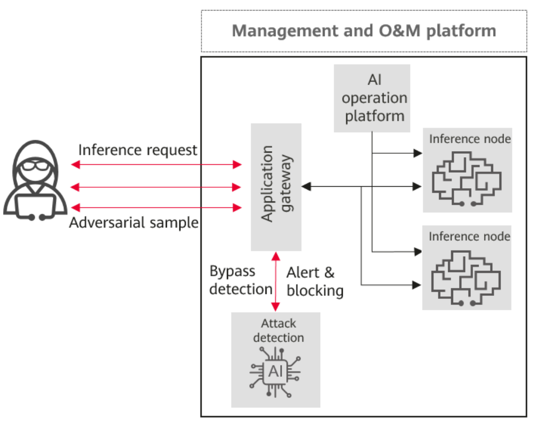
As AI applications become more popular, especially in fields like healthcare and autonomous vehicles, AI-related security challenges are increasing. AI models and training data are core assets of AI application providers. If not properly protected, they may be maliciously recovered and can be used to trace back to the data subjects. In addition, AI models are vulnerable themselves, resulting in more and more evasion and poisoning attacks on AI models. Attacks on AI models in key fields can have serious consequences. As concern about AI increases, there are challenges regarding AI ethics and forensics that will have to be overcome.

To address these challenges, all participants in the AI ecosystem must work together to ensure AI regulatory compliance and governance. They also need to adopt innovative technologies to trace the responsibilities of multiple participants, so as to support responsible AI.

1) Protection of AI models and training data: Encryption, mandatory access control, security isolation, and other mechanisms must be implemented to ensure security of AI models and training data throughout the data lifecycle, from

collection and training, to inference. The major challenge lies in encrypting the high-bandwidth memory data of neural network processing units (NPUs) in real time while ensuring no performance loss. In the future, breakthroughs need to be made in high-performance and low-latency memory encryption algorithms and architecture design for a hardware memory encryption engine on NPUs to provide full-lifecycle protection.

2) AI attack detection and defense: Adversarial sample detection models should be implemented to better identify physical and digital evasions and other attacks on AI models, block attack paths, and prevent misjudgment when AI models are attacked. The main challenge lies in continuous adversarial training against new types of attacks. In the future, independent security products and services to defend against AI attacks will emerge.

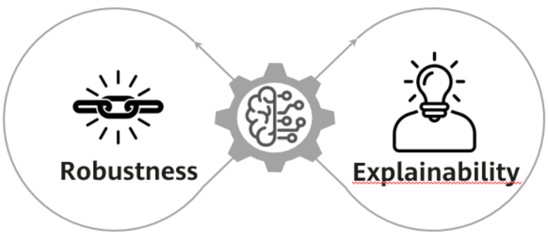


3) In addition to defense against known attacks, the security of an AI model itself must be enhanced to avoid the damage caused by unknown attacks. This can be achieved by enhancing model robustness, verifiability, and explainability.

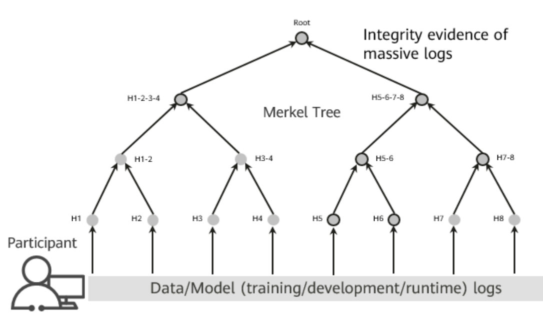
Adversarial training is one of the key technologies for improving the security of AI models. Regularization of models and generalization of adversarial samples are key technologies that need to be improved. Adversarial robustness is expected to increase from current low levels to 80%.

Effective formal verification methods will be available to prove the security of small AI models. However, the formal verification of large models still faces huge challenges.

The ability of AI models to justify their decisions will be vital to minimizing legal or logical risks. Moving forward, an explainable model can be built through explainable data before modeling. Currently, linear models are basically explainable, but there are still huge challenges to be overcome in making non-linear ones explainable. It's still hard to make AI models explainable globally, which means that making some layers of network models visible and explainable may remain the most technically feasible approach for a long time to come.



4) AI models should also be continuously monitored and audited to comply with AI regulations, and blockchain and other related technologies can be used to ensure reliable audit results and real-time tracking of issues.



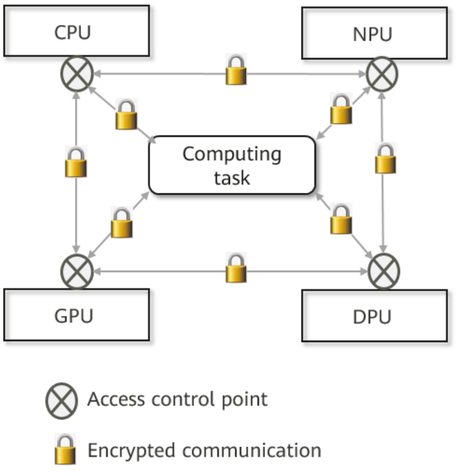
Security for new computing paradigms

In data-centric computing scenarios, computing power extends beyond CPUs, and in particular computing power is moved to memory, using the processing in memory (PIM) technology. This causes the failure of traditional memory encryption mechanisms, making it impossible to deploy hardware-based privacy computing technologies. Even if data is encrypted at the application layer, data is still processed as plaintext, which means privileged users and processes cannot be prevented from data breaches. The only solution for this scenario is to deploy cryptography-based privacy computing technologies (e.g., homomorphic and multi-party computation) to build users' trust in computing service providers.

In data center scenarios where diversified computing power is provided, the migration to the cloud is blurring the boundaries of security, leaving traditional security approaches that were based on security boundaries out of date. That's where the Zero Trust Architecture [12] model comes into play. This architecture addresses the security challenges of untrusted environments by enhancing access policies, proactive monitoring, and encryption. The Zero Trust Architecture model and diversified computing power together plot out the path of security technologies for diversified computing.

1) Security + in-network computing architecture: The Zero Trust Architecture model erases the old boundaries of security, so it employs a finer-grained access control mechanism to support dynamic authentication and resource access policies. That means software implementation consumes a large amount of CPU resources. However, an in-network computing architecture that uses hardware acceleration mechanisms for regular expressions can make policy execution 10–15 times more efficient.

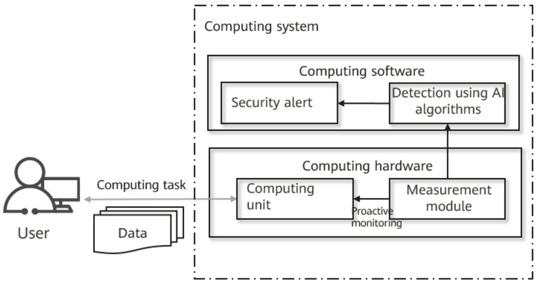
2) Security + diversified computing architecture: A Zero Trust architecture assumes that the network environment is untrusted. It requires encrypted communication throughout the network, including between compute nodes and data centers. Therefore, each xPU in a diversified computing architecture is required to implement the high-performance hardware encryption engine that supports post-quantum encryption algorithms to withstand potential quantum attacks.



3) Security + data-centric peer-to-peer computing architecture: In a data-centric peer-to-peer computing architecture, high-performance SCM will connect with the memory bus in the system. There are increasing risks of data and privacy leakage, as no mechanisms are in place to encrypt residual data in the memory after a power off. Ensuring data security in a data-centric peer-to-peer computing architecture will be a new challenge. For example, in a distributed cluster system where memory is shared across hundreds of compute nodes, it's challenging to protect data without greatly impacting bandwidth performance (keeping the impact close to a theoretical limit that is less than 3%).

4) DC-level dynamic measurement and proactive monitoring: Current computing platforms are generally unaware of the computing tasks running within systems. Even if the systems are attacked, the platforms cannot effectively distinguish malicious behaviors from normal computing tasks. In data centers, we are still facing many challenges in terms of detecting behavior of computing tasks in the system, so that they can measure system status proactively and monitor computing tasks, to detect and defend against potential malicious behaviors adaptively, thereby assuring computing power

security.



Green, integrated computing

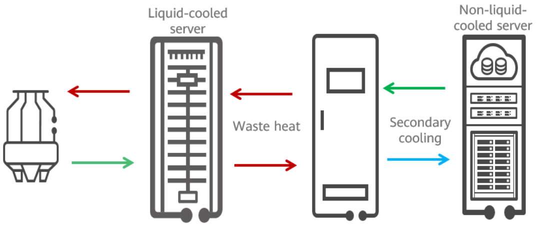
Data centers currently account for about 1% of global electricity consumption. The total energy consumption of general computing has been doubling every three years. The push toward global carbon neutrality will drive a 100-fold increase in computing power while increasing energy efficiency. Ongoing improvements in chip packaging and chip architectures are increasing computing power density and energy efficiency. Co-packaged optics can reduce losses in high-frequency data exchanges. All-in-one data centers will use AI to coordinate power supply, servers, and workloads to achieve an optimal PUE. The ultimate goal is to reduce the PUE to less than 1. Computing networks will connect distributed data centers that provide equivalent services while respecting differences in latency, cost, and green power use, achieving a globally optimal PUE and lowering carbon emissions.

All-in-one data centers

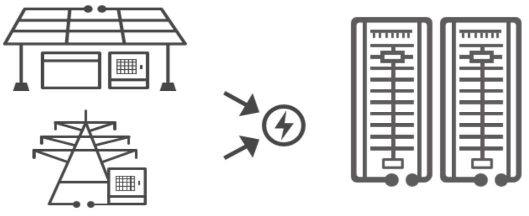
1) DC-level full-stack, converged architecture

Rapid development of compute-intensive technologies such as AI, supercomputing, and cloud computing will enable large data centers to accommodate millions of servers. This will create challenges such as end-to-end heat dissipation, hardware configuration and resource utilization, and unified O&M for millions of central nodes and massive numbers of edge devices.

All-in-one data centers will consume megawatts of power, so we will need to continuously increase their energy efficiency in order to deploy them at scale. Air conditioner-free and chiller-free data centers are now common and liquid cooling technologies are seeing wide adoption. Reuse of waste heat from liquid cooling for heating, secondary cooling, and power generation has become a new growth opportunity in the industry. New technologies are being improved and put into commercial use. As a result, the PUE of some data centers is approaching 1.0, and some are expected to achieve PUE below 1.0 in the foreseeable future. As chip manufacturing and packaging technologies continue to advance, the heat flux density of chips for compute-intensive tasks such as AI and high-performance computing will exceed 150 W/cm^2, and may even go beyond 200 W/cm^2. Native liquid-cooled chips are emerging. With wider adoption of AI, we will see full-stack, automatic, coordinated optimization at the DC level, from power supply and cooling to chip working modes, based on service scheduling and workload features.



Power for data centers needs to be delivered on shorter and more efficient supply paths. New packaging technologies such as 2.5D, 3D, and wafer-level chip (WLC) will enable kiloampere-level chip power supply, which will require new processes, components, and topologies. Power fluctuation due to overclocking and heavy, dynamic loads will require us to rethink server power supply design. Liquid cooling is more complex than air cooling, meaning more difficulty during the construction of equipment rooms, server production, installation, and O&M. It also demands higher skills in data center personnel. Core components such as cold plates and coolants need to be improved in terms of processes and reliability if they are to be deployed at massive scale.



The temperature rises inside 3D chip packages are higher than existing packages. 3D packaging is responsible for nearly 50% of the temperature rise along the heat dissipation path. Therefore, 3D packaging will present new heat dissipation challenges. The thermal resistance of thermal interface materials (TIMs) and cold plates will need to be reduced by 50%, and achieving this will require innovation in materials and processes. Large chip packages like WLC will also require advances in cold plate assembly, coplanarity, and reliability. One viable heat dissipation solution is integrating the chip packaging technology and the liquid cooling technology. With the TIM layer removed, the coolant comes in direct contact with the die inside the chip package. However, this will give rise to reliability issues such as long-term erosion and corrosion, and challenges related to heat dissipation on the surface of the die, jet uniformity, and package sealing.

Waste heat can be reused much more efficiently when water temperatures are high, but for efficient cooling and high chip performance, coolant water temperature must not be too high (not higher than 65°C). Low water temperature presents challenges for data center heat reuse systems. Waste heat reuse in secondary cooling is expected to be in large-scale use by 2025. However, the current efficiency of power generation from waste heat is less than 5%. Large-scale adoption will require breakthroughs in key technologies, such as new power generation materials with high ZT values. In addition, stable heat sources are required for waste heat reuse. The temperature of the liquid-cooled return water depends on chip workloads. Therefore, service scheduling, workload control, and coolant flow control will be needed to help provide stable heat sources for the waste heat reuse system.

Data center-level full-stack energy efficiency optimization will require open interfaces to monitor and control cooling towers, water pumps, coolant distribution units (CDUs), uninterruptible power supply (UPS), electricity meters, and servers. Developing the specifications of these interfaces will be another challenge.

Flexible hardware configuration: As service types and processor platforms become increasingly diversified, IT resources in cloud computing and 100 EFLOPS supercomputing data centers will see a dramatic rise in both scale and complexity. There will be a gradual evolution from the current server-based delivery model to a component-based one. As a result, resource utilization will increase from the current 30% to 70%. To support automated O&M and component-based supply, specifications must be developed for hardware form factors and software and hardware interfaces.

Automated, intelligent equipment O&M: With millions of servers deployed in data centers, automation can improve the efficiency and accuracy of construction and O&M by orders of magnitude. Large numbers of nodes are being deployed at the edge, and automating their integration will spare us corresponding increases in labor and operation costs that edge deployment would otherwise bring. Automation will also improve our ability to troubleshoot edge systems. AI and big data will help make better informed decisions; learning algorithms and dynamic adjustment of hardware and software configurations will increase IT resource efficiency and energy efficiency. Incidents like the COVID-19 pandemic will require data centers to support contactless delivery and O&M.

As Industry 4.0 and AI continue to develop, automation technologies are rapidly maturing. Intelligent and unmanned adaptive data centers (ADCs) will be deployed widely, making automatic and dynamic matching between data centers and service workloads a reality.

Computing networks

1) Cross-region distributed super data centers

The central idea of a computing network is to use new network technologies to connect computing center nodes distributed across different geographical locations. The purpose of such a network is to achieve real-time awareness of the status of computing resources, to coordinate the allocation and scheduling of computing tasks, and to transmit data, so that the system as a whole forms a comprehensive network that senses, allocates, and schedules computing resources across the board. Through this network, computing power, data, and applications will be aggregated and shared.

Computing centers have multiple layers and management domains. Different computing centers differ greatly. The types of applications deployed, datasets stored, and computing architectures may vary from site to site. Management policies, billing standards, and carbon emissions standards may also vary. If we are to build computing networks, there are several things that need to be sorted out first: coordination between different computing centers; a trusted transaction and management mechanism for computing power, data, and applications; and unified standards. The ultimate goal is to build computing architecture that is open, energy efficient, and delivers high resource utilization.

2) Converged applications will form a digital continuum

Hyperscale AI models, the explosive growth in the volume of data, and the increasing requirements for precision and speed in scientific computing will require massive computing power and new applications. The distributed applications of the future will integrate real-time and non-real-time data processing, model training and inference, simulation and modeling, IoT, and information physics to form a "digital continuum". This will solve the problems that individual computing centers find hard to solve. For example, a digital meteorological model, which combines neural networks and real-time data, can provide short-term and imminent weather forecasts at high frequency and high resolution, bringing tangible benefits to our everyday lives. Distributed large-scale models can use the resources of multiple computing centers to speed up model training. New applications will support the connectivity between different computing centers and between computing centers and edge computing facilities. Computing centers will no longer be independent systems; instead, each center will be a node in an interconnected computing network. In order to meet the computing and data processing requirements of complex applications, users from multiple organizations can share computing power and data distributed across multiple computing centers.

3) Collaborative scheduling for cross-domain computing centers

Multiple computing centers distributed at different geographical locations will be connected to support new distributed converged applications. Training hyperscale models will require the resources of multiple computing centers, and complex converged applications may also rely on the computing power and datasets of different computing centers. Application diversity, resource heterogeneity, and inconsistent management strategies will all pose new challenges to the scheduling system. The scheduling system needs to be aware of the computing power and storage resources required by applications; it will need to know the locations of data, to reduce data movement overheads; and it will need to understand how applications communicate to reduce communication overheads. The scheduling system also needs to be aware of the availability and heterogeneity of resources in different computing centers in real time, and the network status of different computing centers. In addition, the system needs to make the optimal decisions while taking into account the required cost-effectiveness and energy efficiency, in order to adapt to differences in resource pricing and carbon emission standards that apply to different computing centers. That is, the scheduling system must be capable of discovering resources, aware of the characteristics of applications, aware of software and hardware heterogeneity at computing centers, and aware of local management policies. This will make it possible for the scheduling system to deliver globally optimal efficiency in computing, data movement, and energy use.

Chip engineering

1) 2.5D chiplet packaging and integration technology will continue to improve chip computing power and product competitiveness

The hard dimensional limits on wafer exposure (25 mm x 32 mm for one reticle) present huge technical barriers to increasing total die size and die yield. This issue is impeding efforts to improve chip performance and cut chip costs.

2.5D silicon/fan-out (FO) interposer + chiplet technology can increase die yield and reduce chip costs. Stacking and integration help achieve greater chip performance, and provide better adaptability to different product specifications. In

addition, the energy consumption per bit in 2.5D packaging is just half that of the board-level interconnection solution used in conventional packaging.

As the industry continues to advance and the demand for chips grows, it is estimated that by 2025, the size of a 2.5D silicon/FO interposer will be more than four times that of a reticle, and the substrate is expected to be larger than 110 mm x 110 mm. Larger 2.5D and substrate processes pose engineering challenges in terms of yield, lead time, and reliability. To address these challenges, converged, innovative substrate architectures will be needed.

2) 3D chip technology is expected to outperform conventional architectures by dozens of times

3D chip technologies present significant advantages over advanced 2D/2.5D packaging and heterogeneous integration: better interconnection density, bandwidth, chip size, power consumption, and overall performance. 3D chip technologies will be critical to chip and system integration in key scenarios such as high-performance computing and AI.

3D chip technology will evolve from die-to-

wafer (D2W) to wafer-to-wafer (W2W) and µbumps, and then to hybrid bonding, and finally to monolithic 3D technology. This technology will be widely used in different types of stacking, including 3D memory on logic, logic on logic, and optical on logic, and will gradually extend to multi-layer heterogeneous stacking.

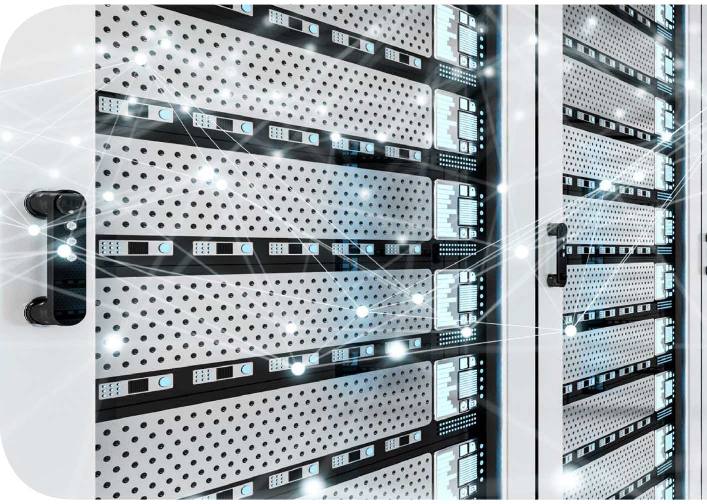
3D chip stacking requires the use of ultra-high-density bonding technology with pitches smaller than 10 µm. 3D chips have significant advantages over 2.5D packaging in terms of bandwidth and power consumption, so power consumption per bit is expected to fall by 90%. Ongoing research is required into technologies for working with smaller through-silicon vias (TSVs), both in materials and processes. One drawback of 3D stacking is that it multiplies local power density and current density, with implications for the system's power supply and heat dissipation paths.

3) Co-packaged optics for Tbit/s-level high-bandwidth ports

Compute-intensive chips (e.g. xPUs, switches, and FPGAs) will deliver increasingly higher I/O bandwidth. It is expected that the port rate will reach terabits per second or higher by 2030. As the speed per channel increases, serial communications at speeds of 100/200 Gbit/s or higher will create challenges in power consumption, crosstalk, and heat dissipation. Conventional optical-to-electrical conversion interfaces will no longer meet the demands of increasing computing power. Co-packaged optics are expected to cut end-to-end power consumption by 2/3. Co-packaged optics can replace pluggable optics and on-board optics, and will become a key technology for higher port bandwidth. If the technology is to be widely adopted, challenges in engineering technologies will need to be addressed, including 3D packaging of photonic integrated chips (PICs) and electronic integrated chips (EICs), ultra-large substrate and optical engine (OE) integration, and chip power density.

4) Power supply for power-intensive chips

The demand for increasing computing power and the development of chiplet technology continue to drive up chip power consumption. The power supply for kW-level chips will no longer be a problem, but more innovative and efficient power supply strategies will be required for 10kW-level wafer-level chips. New power supply architectures such as high-voltage single-stage conversion and switched-capacitor hybrid conversion, combined with engineering technologies such as low-voltage gallium nitride (GaN) power devices and high-frequency integrated magnets, can further improve the end-to-end energy efficiency and power density of board power supply.



High-voltage (48V) direct power supply is key to addressing the problem of chip power supply. To implement this technology, it will be necessary to first develop new materials for substrates and packaging, along with the processes necessary to apply them, which can accommodate the high voltages. Efficient on-chip voltage conversion and core-based power supply are the way forward for research.

5) Chip-level heat dissipation technology

The power consumption of computing chips has risen sharply, and heat dissipation has become a major barrier to further chip performance improvement. There is an urgent need for new heat dissipation technologies and materials. Lidless chips, advanced package- and chip-level liquid cooling, and high-conductivity TIM1 materials that reduce path thermal resistance are expected to provide the heat dissipation capacity necessary for kW-level chips, and even 10 kW-level chips. They will open the way for major advances in chip performance. Dynamic chip thermal management and system-level coordinated heat dissipation will also be key technologies for ultra-power-intensive chips.

Diversified computing

In the future, data will be processed in the right place, using the right kind of computing. For example, network data will be processed on data processing units (DPUs) and neural network models will be trained on NPUs. Computing power will be everywhere. Peripherals such as hard disks, network adapters, and memory will gradually become capable of data analysis and processing. Converged applications call for a unified architecture for diversified computing. Currently, tools from different vendors are siloed from each other, greatly hindering the development of diversified computing.

Data-centric computing

1) Symmetric computing architecture (in-memory data processing)

In Von Neumann architecture, data needs to be moved from storage to the CPU for processing, and this movement of data consumes a large amount of computing power and energy in the system. In addition, numerous memory, storage, and transport formats need to be converted back and forth during data processing and exchange, which consumes a lot of CPU time and leads to low energy efficiency. At the same time, data volumes are mushrooming, and hardware deployment cannot keep pace. This will exacerbate existing issues related to input/output (I/O), computing power, and networks. Such issues slow down data migration, hinder processing efficiency, and affect a system's overall energy efficiency.

These issues can be properly addressed with a symmetric computing architecture that supports memory pooling. Under this architecture, unified memory semantics will be used to process and exchange data throughout the data lifecycle, and even ensure that all data is processed in the memory. This architecture can eliminate

the need for format conversions, improve data migration speeds, expand the memory available for applications, and ultimately enhance the entire system's data processing capability. This will be one of the major approaches to faster computing. Building this architecture will require breakthroughs in multi-level memory architecture, large-capacity non-volatile memory, and other key technologies.

2) Ubiquitous computing (intelligent peripherals)

In the future, a diverse range of xPUs will provide different types of computing power. In addition, we believe that an architecture with ubiquitous near-data computing will be a way forward. Under this architecture, data will be processed in the right place with the right kind of computing power, which will help reduce data migration and boost overall system performance.

Ubiquitous near-data computing may involve the following directions:

(1) Near-memory computing. In current systems, the effective bandwidth available for data migration is limited by the bandwidth of the external bus. In the future, multiple concurrent programmable computing units will be added to the dynamic random access memory (DRAM) control circuit, and the DRAM array structure will be optimized to improve concurrent internal data access. This will multiply effective bandwidth for data computing in the DRAM, and help overcome the bandwidth bottleneck caused by the memory wall.

(2) Near-storage computing. Currently, a fixed data acceleration unit (such as a compression engine) can be added to a solid state drive (SSD) controller specifically to process data. In the future, multiple operator engines in the SSD controller could be invoked on demand through application programming interfaces (APIs). Coupled with compilers, this approach can support more flexible offloading of compute workloads, and improve the energy efficiency of data computations in general scenarios.

(3) Computing using memory based on SmartNIC, which will evolve to a DPU-based, data-centric computing architecture. In the future, in-network computing power will be flexible and programmable, existing within open, heterogeneous programming frameworks, for a service-driven in-network computing paradigm. This will support acceleration across the board, including storage, security, and virtualization, and will greatly improve the performance of distributed applications, such as HPC and AI convergence, big data, and databases. Fine-grained dynamic scheduling and efficient

interaction of all computing resources in data centers will become possible.

3) Computing using memory

Computing using memory is a tight coupling between processing and storage units, which allows storage media to function as both a storage unit and processing unit. This erases the boundary between computing power and storage, effectively overcoming the power wall and the memory wall. This technology is expected to be at least 10 times more energy efficient than traditional Von Neumann architecture.

Computing using memory based on mature memory technologies like static random-access memory (SRAM) and NOR flash is expected to be in commercial use on a large scale within two to three years. This technology will make AI inference and operation on devices and the edge 10 times more energy efficient. Computing using memory, powered by new non-volatile memories like resistive random-access memory (ReRAM), phase change memory (PCM), and magnetoresistive random-access memory (MRAM), is still in the experimental phase, but given their high performance and low energy consumption, they have the potential to be used in data centers in the next decade.

Breakthroughs in the following areas will also be required before computing using memory can become commercially available on a large scale.

Computational precision: Computational noise and issues of component consistency and stability can cause computational errors, so computing using memory is less precise than conventional computing systems. Therefore, algorithms will need to be optimized to account for the kind of compute circuit on which they are running.

Software ecosystem: Computing using memory is a type of data-driven computing. Neural network models need to be deployed on the right storage units, and the entire computational process will be controlled through data flow scheduling. This necessitates the development of more intelligent, efficient, and convenient data mapping tools.

System architecture: Computing using memory, powered by new non-volatile memory, uses a calculation method that multiplies matrices by vectors. Today, these systems are often used in specific machine learning applications (e.g., neural network inference and training), and it is difficult to extend them to other use cases. In addition, they cannot cooperate with existing storage systems to efficiently process data. To overcome these challenges, a full-stack design that facilitates synergy between storage devices, programming models, system architecture, and applications will be essential to ensure that the architecture of computing using memory works for general purposes.

4) Buses: From board-level buses to DC-level buses

With the exponential growth of computing power and data, large, centralized data centers that focus on AI, HPC, and big data will become more important. Compared with intra-node buses, the networks connecting entire data centers has a huge latency, bandwidth gap, and heavy network software stack overheads. All of these features degrade computing power. Lightweight software stacks, with high bandwidth, low latency, and memory semantics, exist at the board level, and will be extended to the entire data center through the memory-semantic bus. This will enable optimal performance and energy efficiency for the entire data center.

For memory-semantic buses, the biggest challenge lies in building open, equal, interoperable buses, interfaces, and protocol standards. This helps prevent the fragmentation of standards for computing system buses, which would only hinder advances in computing performance and large-scale computing.

Application-driven diversified computing

The next generation of computing systems will bring a new paradigm, characterized by domain-specific hardware, domain-specific programming languages, open architectures, and native security architectures.

1) New paradigm for scientific computing

With breakthroughs in AI computing methods and AI computing architectures, a new paradigm is emerging in scientific research, in which machine learning is combined with first-principles-based physical modeling. In the next decade, intelligent scientific computing will be involved in every aspect of scientific research and technological innovation. The effort to efficiently integrate AI algorithms with scientific computing presents unprecedented challenges and opportunities.

- In terms of the fundamentals, there are challenges regarding the computational frameworks and mathematical methods of the new computation approach. There is a need for new frameworks and approaches that ensure a given problem can be effectively solved using AI. That is, the mathematical methods and frameworks must ensure computability, learnability, and interpretability. Therefore, over the next decade, hardware and software infrastructure must be built based on mathematics and AI research and provide appropriate implementation, assessment, and testing systems.

- In terms of data, a large number of different data sources are required to boost scientific research, engineering, and manufacturing using AI. First, different fields of scientific research rely on different sources for their data. These data sources may include instruments, simulations, sensor networks, satellites, scientific literature, and research findings. Currently, there are still great challenges to overcome regarding the availability and shareability of this data. Second, there are challenges in using AI to generate effective data that is based on physical principles and complies with basic laws of physics (such as symmetry and conservation laws). To address these challenges, scientists from different domains, AI experts, mathematicians, and computer scientists need to work together.

2) AI enabling intelligent storage

Storage systems are now expected to address loads of increasingly diverse and complex service requirements and to offer simplified system management and O&M.

Storage systems of the future will be able to use AI to proactively manage and respond to their internal and external environments,

to learn continuously, to be workload-aware and adaptive, and to automatically optimize themselves to deliver gains in resource allocation, cost, performance, reliability, usability, etc. In addition, manual O&M will need to evolve to automated intelligent O&M using AI.

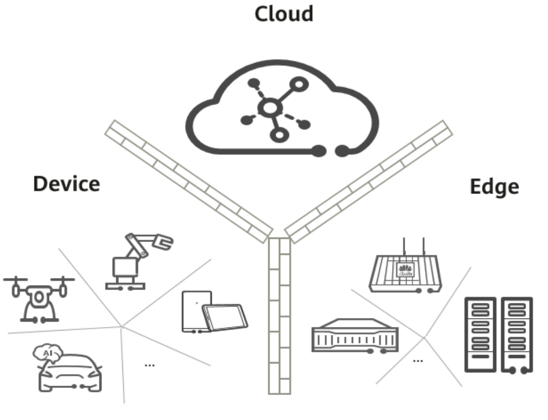
Progress has already been made in the application of AI in indexing, automatic optimization, and resource allocation in storage systems. However, breakthroughs in the following four areas are still needed:

- **Workloads:** The impact of I/O workloads on system performance needs to be modeled to identify the key indicators and factors affecting module performance, to accurately assess system performance, and to simulate real-world service scenarios.
- **Data:** Data distribution, data lifecycle, and data context need to be perceived so that systems can improve data access performance, reduce the consumption of resources by back-end garbage collection, and improve data reduction ratios.
- **Systems:** Rules and patterns need to be identified based on past data, computing tasks need to be arranged and scheduled efficiently, and systems need to be optimized during runtime to improve system parameters and resource allocation, reduce system power consumption, and ensure that fluctuations in system performance are controllable and do not undermine reliability.
- **Operations:** Automated O&M is needed to eliminate the need for manual work; faults need to be automatically analyzed to identify root cause; and any system suboptimality needs to be detected, prevented, and rectified automatically.

Integrating top-down load modeling and bottom-up adaptive learning to support intelligent storage has become an area of interest. A great deal of current research is aimed at developing intelligent storage systems featuring automatic performance optimization, automatic QoS control, intelligent data awareness, self-learning of rules and policies, intelligent scheduling, low-power controls, simplified planning and configuration, prediction of system issues, and automatic root cause analysis.

Multi-dimensional collaboration

Computing and storage infrastructure are distributed in different locations on the cloud, edge, and devices. Such infrastructure can be horizontally and vertically coordinated to complement each other and enable cubic computing. This addresses problems such as poor service experience, uneven distribution of computing, low utilization of computing resources, and information silos.



Multi-dimensional sensing and data modeling enable the physical world to be mirrored, computed, and enhanced to form digital twins. With light field holographic rendering and AI-assisted content generation, the digital world is precisely mapped to the physical world. Multi-dimensional collaboration between time and space, as well as between virtuality and reality, enables seamless integration of the physical and digital world.

Cubic computing

1) Edge computing

The world of the future will be an intelligent one in which everything is connected. As 5G technologies mature and see increased application, edge computing will be widely deployed in the ICT industry. It is expected that the global edge computing market will be worth hundreds of billions of US dollars by 2030, but at present the value of this market is US$10 billion. To apply edge computing on a large scale, we must first confront challenges in areas like edge intelligence, edge computing network, edge security, edge standards, and open ecosystems.

Edge intelligence: Intelligent upgrades of vertical industries like manufacturing, power grids, city administration, transportation, and finance are important drivers of the exponential growth of edge computing. Development kits for basic AI capabilities, such as incremental learning, transfer learning, device optimized model compression, and inference scheduling and deployment, are needed to solve common issues encountered by many industries currently undergoing intelligent transformation. A development kit is needed to address common issues unique to intelligent manufacturing. This industry is characterized by samples or images with complex backgrounds and low contrast, small size training samples, and weak supervision. Development kits should also be developed for other industries, to form a comprehensive set of software development kits (SDKs) for application enablement.

Edge computing network: Future service demands will drive edge devices to support a greater range of services. As such, these devices will need to be mobile, low-power, and smaller, but computing, storage, bandwidth, and latency will become bottlenecks. Holographic and multi-dimensional sensing services require 100 times more computing power than is currently available, storage capacity will need to expand by 100 or even 1,000 times, and network bandwidth will need to increase to tens of terabits per second. Industries such as intelligent manufacturing, intelligent power grids, and intelligent transportation require millisecond-level deterministic latency. To meet the demands of edge acceleration, offloading, and performance breakthroughs, we need hyper convergence of computing, storage, and networking, with efficient use of diversified computing. This will pose new challenges to edge software and hardware architecture.

Edge security: Edge devices are physically closer to attackers. Being located in complex environments, edge devices are more vulnerable to attacks from physical hardware interfaces, southbound and northbound service interfaces, and northbound management interfaces. Data is often a core asset of users, so data loss or theft may cause significant losses to users. It is estimated that 80% of data will be processed at the edge by 2030. It is thus paramount to strengthen security and privacy protection during data collection, storage, processing, and transmission at the edge. In addition, the security and privacy of core assets such as edge applications and models must be strictly protected. Data silos caused by data privacy protection must be prevented as this would make it difficult to fully unleash the potential value of data and AI algorithms in sectors such as healthcare, finance, and industry.

Edge standards and open ecosystems: Edge devices for different industry applications differ greatly in computing power, functions, software and hardware formats, and interfaces. Proprietary software and hardware solutions and interface protocols from different vendors are often not interoperable, which greatly hinders the adoption of edge computing. The edge computing system, software and hardware frameworks, and related interfaces and protocols need to be standardized, and corresponding test and acceptance standards need to be established for better interoperability between edge devices, software, and protocols. In addition, open ecosystems need to be built for each industry to attract investment from more vendors and partners.

2) Multi-device collaboration

Animals like ants and bees create swarm intelligence through collaboration. The multi-device collaboration technology aims to achieve similar breakthroughs to improve the problem-solving capabilities, overall performance, and robustness of multi-device systems.

Multi-device collaboration takes various forms, such as task sharing, result sharing and intelligent agents. In task sharing, devices collaborate by performing subtasks of a particular task. In result sharing, devices collaborate by sharing parts of the results. The processing capability of each device at any given moment depends on the data and knowledge that the device owns or receives from other devices. In the form of intelligent agents, devices collaborate on the basis of independence and autonomy.

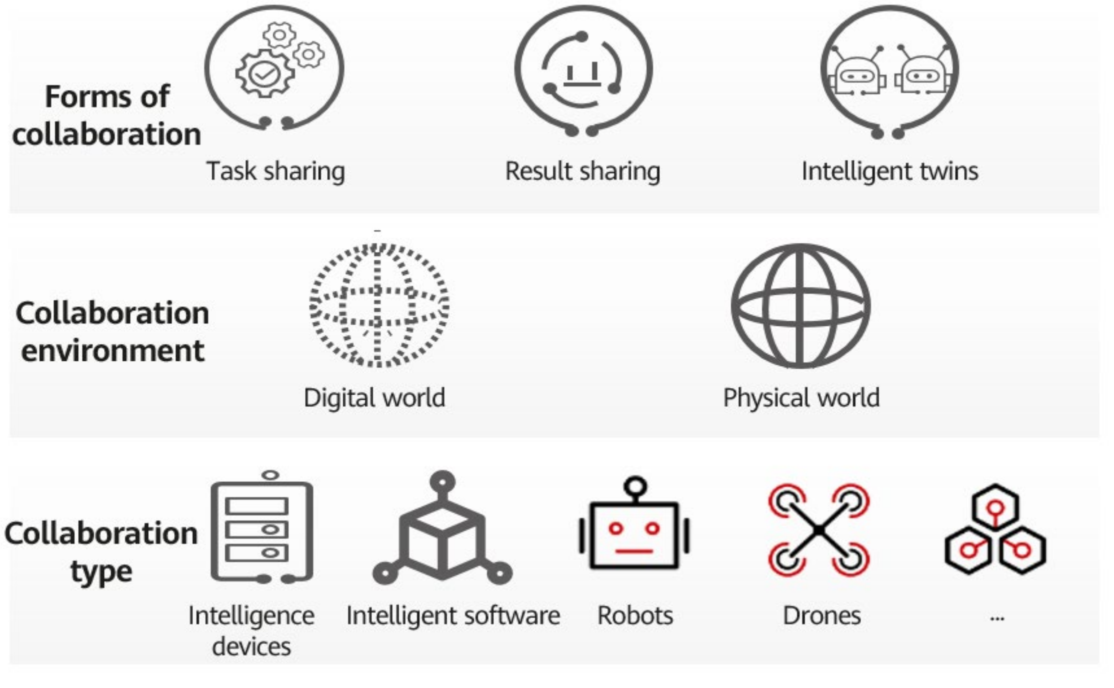
Effective multi-device collaboration requires solving problems related to cooperation and conflict resolution, global optimization, and interaction and collaboration consistency.

Cooperation and conflict resolution: A deadlock or livelock may occur during multi-device collaboration. Deadlocks make devices unable to perform their respective next-steps, and livelocks make devices work continuously without making any progress. Coordination mechanisms and algorithms are critical for preventing deadlocks and livelocks in interactive processes.

Global optimization: It is difficult to achieve global optimization when multiple devices are collaborating based on local information. However, collaboration based on a global view often means large communication traffic, which can overburden the system. Efficient and secure acquisition of high quality and reliable global situation estimations determines the efficiency and effectiveness of multi-device collaboration.

Interaction and collaboration consistency: Each device obtains information from other devices through network communication and adjusts its own state. In practice, because the connectivity between multiple devices is unreliable or there are barriers to communication, collaboration consistency issues may arise. Therefore, the ability to address such issues determines the robustness of a multi-device collaboration system.

Multi-device collaboration systems will gradually



evolve from simple cooperation and connection to autonomous swarm intelligence.

3) Device-edge-cloud computing

AI and emerging data-intensive applications, such as intelligent manufacturing, intelligent cities, smart inspection, and intelligent transportation, are developing rapidly. The need to improve application experience, such as by reducing latency, reducing bandwidth costs, and enhancing data privacy protection, drives the development of device-edge-cloud computing. To develop an integrated computing architecture, the following challenges need to be addressed.

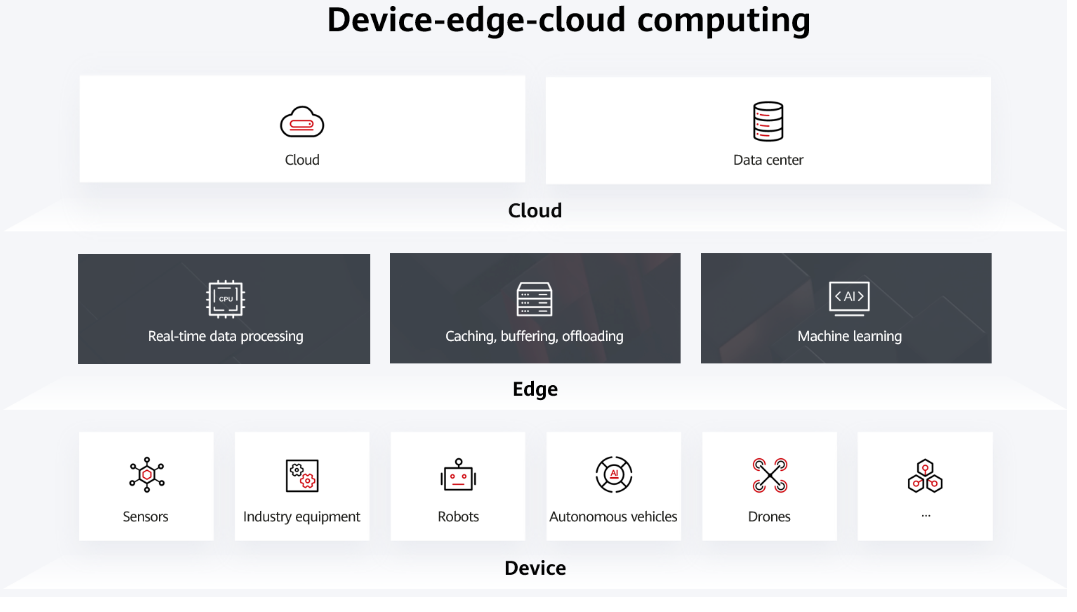
Task collaboration: How should a computing task be divided into multiple subtasks? How should subtasks be deployed and scheduled on the device, edge, and cloud? Where should a subtask be performed (on the device, edge, or cloud) and when? The migration of computing subtasks across clouds, clusters, and nodes is also a challenge.

Intelligent collaboration: The model of training on the cloud and inference at the edge is moving toward device-edge-cloud collaborative training and inference. Challenges in the following areas need to be addressed to achieve device-edge-cloud synergy: precision and rate of convergence of collaborative training; latency and accuracy of collaborative inference; and data silos, small sample sizes, data heterogeneity, security and privacy, communication cost, and limited device/edge resources.

Data collaboration: Data is the basis of intelligence. Diversification and heterogeneity pose challenges for data access, aggregation, interaction, and processing.

Network collaboration: As the scale of the device-edge-cloud computing network grows, access by a large number of devices and subnets brings great challenges to device, network, and service management. We need solutions for the challenge of ensuring reliable real-time connectivity.

Security and trustworthiness: How can security and privacy be ensured when edge devices and their data are connected to the cloud? How can the cloud protect itself from edge-side attacks? How can the data sent from the cloud to the edge be protected?

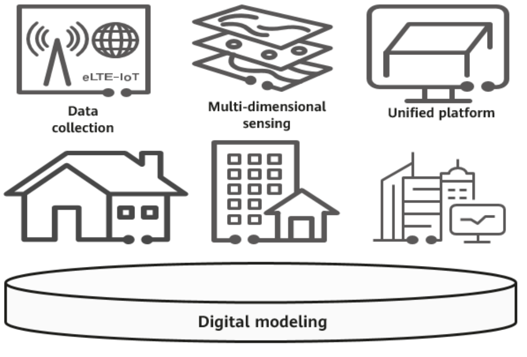


Digital twin

1) A unified digital twin platform is the way forward

Under the digital tide of various industries, such as smart factories, smart cities, and virtual social media, there is no unified platform for creating personalized digital twin systems. This platform needs to focus on the unification of data formats and development tools of 3D models, and provide diversified computing power and storage space required for modeling large amounts of data.

2) Multi-dimensional sensing and digital modeling technology



The physical world of the future will have a digital twin. These two worlds will be seamlessly converged and work in tandem to improve the efficiency of product design, product manufacturing, medical analysis, and engineering construction. The process of mapping the physical world to its digital twin will face numerous challenges, such as multi-dimensional sensing, 3D modeling, and light field data collection and storage.

Multi-dimensional sensing: Massive amounts of data on the physical world, including images, videos, sounds, and temperature, humidity, and mechanical records are collected and stored. The acquisition, processing, and convergence of data with more dimensions requires high-resolution sensing, object location, imaging, and environment reconstruction, and the amount of data generated in this process is even larger. The process of screening, preprocessing, modeling, and simulation of such massive amounts of data relies on powerful computing and the deep integration of multiple disciplines, such as artificial intelligence, cognitive science, control science, and materials science.

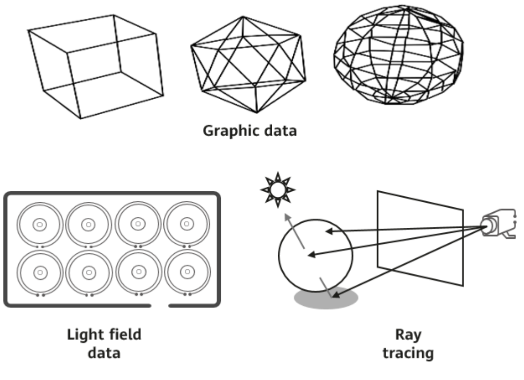
3D modeling will require 100 times more computing power. 3D modeling, which is based on images and video streams of different angles and massive amounts of data collected by array cameras and depth cameras, requires huge computing power. The volume of high-precision data collected by a 100 plus-channel camera array is 100 times higher than that of 2D images. The resolution will increase to 8K and the required computing power per channel will see a 4-fold increase. The required computing power for modeling is 100 times higher. Managing this massive amount of multi-dimensional data and transforming it into a 3D model is a big challenge. In addition, in the consumer market, depth information of images can be obtained using the 3D camera on a phone, and medium- and low-precision modeling based on the depth information can be performed on the phone. The 3D camera of a phone is usually a binocular camera, structured light camera, or time-of-flight (ToF) camera. A unified, efficient, and economical software and hardware system for 3D modeling is required for high-level and consumer-level modeling, the digital transformation of various industries, and the flourishing of the digital twin industry.

AI-enabled material generation in digital modeling: Powered by AI image recognition technology, intelligent generative algorithms, and strong AI computing power, digital models can automatically recognize the physical properties of images, such as metalness, roughness, reflectivity, refractivity, and surface normal vector. This would then generate materials like we see in the real world in the form of a 3D model. To support this process, a unified and open material description language is needed to exchange 3D graphic data between different industries.

100 times more light field data will make compression a key technology: Light field camera arrays will collect 100 times more image and video stream data, which will then be used for synthesizing 3D video streams and light shading in rendering. Such massive amounts of data mean that data storage and processing will be a huge challenge. Breakthroughs in fast compression and storage of light field data are therefore essential, as these are the key to subsequent rendering and imaging.

3) Light field holographic rendering technology

Breakthroughs in visual and interactive technologies need to be made for a digital twin display system to provide users with the same experience as they have in the physical world. Most products currently on the market have deficiencies in rendering quality, fidelity, and rendering delay. Real-time ray tracing and zero-delay transmission can directly improve user experience and are key technologies for photo-realistic authentic rendering. Advanced rendering such as ray tracing requires 10 times more computing power than traditional rendering. Utilizing storage to replace computing can meet part of the demand for computing power while reducing latency, but this would necessitate greater storage space. Moving forward, cloud-based holographic rendering of light fields will be an important area of research.

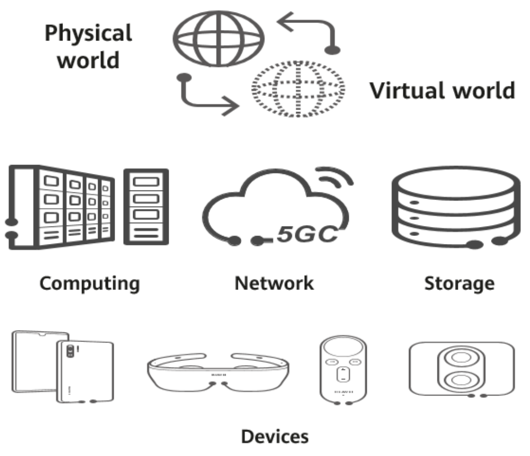


Advanced rendering technology will deliver a 64-fold increase in resolution: The mainstream technology of holographic rendering of light fields has evolved from rasterization rendering to much more advanced rendering technologies such as ray tracing. In scenarios such as gaming and extended reality (XR), a near-real experience can be made possible with 16K binocular resolution, 120 frames per second, and a latency of no more than 8 ms. Strong interaction services use 64 times more computing power and require a latency of 5 ms. These services need breakthroughs in key technologies such as 3D modeling, material generation, and ray storage. Device-edge-cloud computing clusters can provide converged computing power for rendering, AI, and video streaming. When these compute resources are combined with content creation software for advanced rendering, near real-time and high-performance rendering solutions can be created.

AI-based content generation: AI can enable 3D modeling, automatic material generation, super resolution, and noise reduction. AI technologies such as generative adversarial network (GAN), natural language processing (NLP), and natural language generation (NLG) will generate 3D images of avatars and allow them to have vivid expressions and engage in natural language conversations. This will greatly aid communication between people in different parts of the world. AI content generation will also be used in industrial design, XR content creation, and special visual effects.

4) Interaction between the physical and digital worlds for hundreds of millions of users

Allowing hundreds of millions of users in the physical world to interact with digital twins places high demands on computing, storage, and network bandwidth. This is because it requires a large amount of state queries and message transmission. When people and things can interact with each other at latencies less than 5–10 milliseconds, the bandwidth reaches hundreds of Mbit/s per user, and the required computing power increases to tens of TFLOPS per user, network-edge-cloud collaboration and real-time data processing and transmission for hundreds of millions of users will be possible, but this is a very challenging goal.



Physical layer breakthroughs

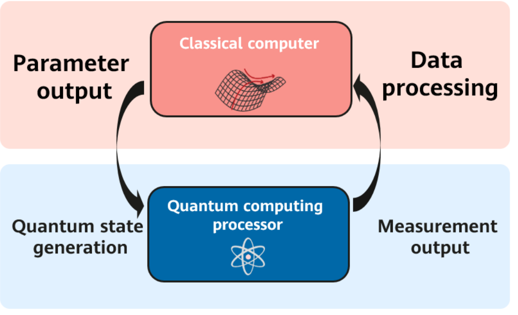
Both academia and the industry are exploring potential breakthroughs at the physical layer, including analog computing, non-silicon-based computing, novel storage media, and optimized chip engineering, to keep improving the energy efficiency of computing and storage density. For example, quantum computing offers exponential advantages over traditional computing in data representation and parallel computing. Analog optical computing consumes little power yet achieves high performance for certain computing tasks. 2D materials and carbon nanotubes have high carrier mobility and shorter channels, and are expected to replace silicon. Significant breakthroughs have been made in ferroelectrics, phase change materials, and device structures, resulting in significant improvement of storage density and read/write performance. Multi-layer and multi-dimensional optical storage has huge potential for long-term storage of cold data. Breakthroughs in DNA storage will need to be made. These breakthroughs in key technologies at the physical layer will revolutionize computing and storage.

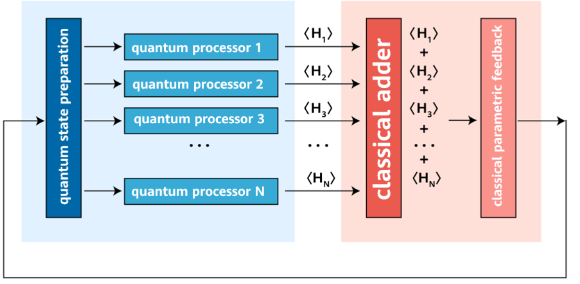
Analog computing

1) Quantum computing: A technology of strategic importance for the future of high-performance computing

Quantum computing is undergoing rapid progress in engineering, and a chip with more than 1,000 qubits is expected to appear within the next five years. Quantum computing is now in an era of noisy intermediate-scale quantum (NISQ). The most feasible path forward is building a hybrid computing architecture that combines the accuracy of classical computers and the performance of quantum computing. This hybrid computing architecture will be used in quantum chemical simulation, quantum combinatorial optimization, and quantum machine learning, as those are the three scenarios that have the greatest commercial potential. Quantum chemical simulation can provide new computing power for research and development of pharmaceuticals and new materials. Quantum combinatorial optimization, where combinatorial optimization problems are encoded as quantum dynamics, can be used to optimize logistics scheduling, route planning, and network traffic distribution. Quantum machine learning will provide a new path for accelerating AI computing.

The focus of the next decade should be on developing a dedicated NISQ-based quantum computer, while continuing to increase the number of qubits in a single quantum chip, prolong coherence time, and enhance fidelity. More efforts should be made to optimize the interconnection of quantum chips to enhance system scalability, so that sufficient computing power will be available to solve those complex problems. At the same time, we also need to make quantum computing more fault tolerant, improve system reliability, optimize quantum algorithms for different application scenarios, and improve the quantum software stack, while reducing circuit depth and complexity. These





are part of the broader efforts to bring NISQ-based quantum computing to commercial use. However, building a universal quantum computer will be a long, challenging process.

2) Analog optical computing: Competitive in certain complex computing tasks

Light propagates at a high speed with negligible power consumption. In certain optical systems, mathematical models are used to describe their associated physical phenomena, such as interference, scattering, and reflection. Certain computing tasks can be accomplished by utilizing the physical characteristics of light, such as amplitude and phase, and the interactions between light and optical devices. In addition, as a boson, a photon allows parallelism in degree of freedom, such as wavelength division multiplexing, mode division multiplexing, and orbital angular momentum (OAM) multiplexing. Multi-dimensional parallelism is an important direction forward for optical computing.
Early breakthroughs of optical computing are expected to appear in convolution computing, Ising model solving, and reservoir computing, followed by application in signal processing, combinatorial optimization, sequence alignment, and AI acceleration.

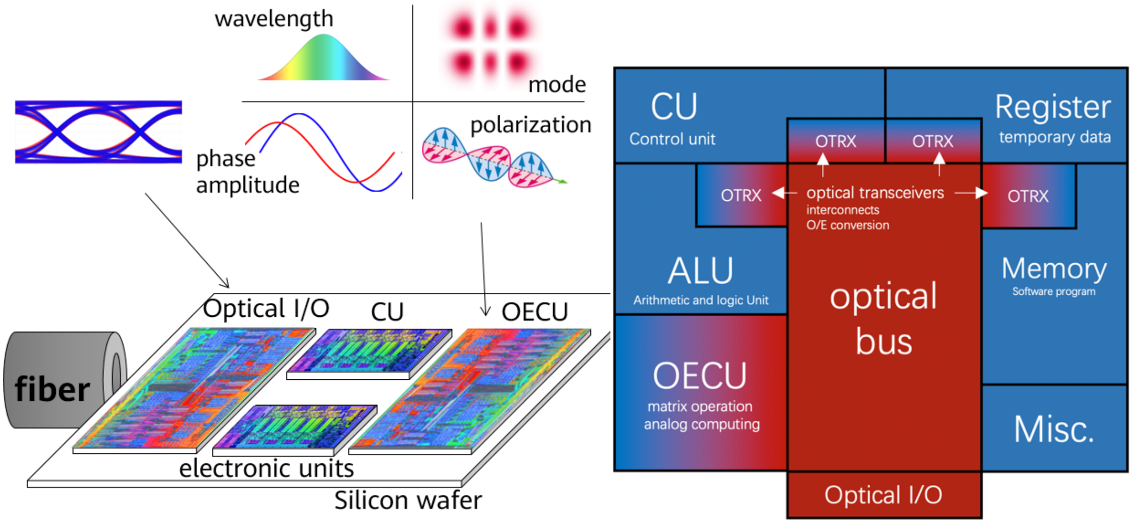
There are still formidable challenges for the commercial application of optical computing, such as insertion loss, noise control, heterogeneous integration, and co-packaging of electronic and optical devices. The drive circuits used in optical computing also need to be further integrated with optical chips to reduce power consumption and area. As optical computing and electrical computing each have their own advantages, optoelectronic hybrid computing architecture is a promising direction for future development.

Non-silicon-based computing

1) 2D materials: A potential material to extend Moore's law

2D materials offer several advantages, including shorter channel length, high mobility, and the possibility of heterogeneous integration, and are expected to be used as transistor channel materials to extend Moore's law as far as 1 nm technology node. In addition, 2D materials with ultra-low dielectric constants can be used as the interconnect isolation materials of integrated circuits. 2D materials are expected to be first adopted in domains such as optoelectronics and sensors, and eventually in large-scale integrated circuits and systems.

At present, 2D materials and relevant devices are still in the basic research stage, and many of the necessary breakthroughs in materials, devices, and processes have yet to be made. Over the next five years, we need to realize industrial-grade wafers made of 2D materials and constantly improve their yield. In addition, we need to keep optimizing the electrode contacts and device structures to improve the comprehensive performance of 2D transistors. Once these improvements are made, 2D materials are expected to be applied in large-



scale integrated circuits within ten years.

2) Carbon transistors: The most promising technology to extend Moore's law

Carbon nanotubes have great potential in both high performance and low power consumption because of their ultra-high carrier mobility and atomic-level thickness. In cases of extreme scaling, carbon nanotube transistors are about 10 times more energy-efficient than silicon-based transistors. Carbon nanotubes are expected to be commercially used in biosensors and radio frequency circuits in 3 to 5 years.

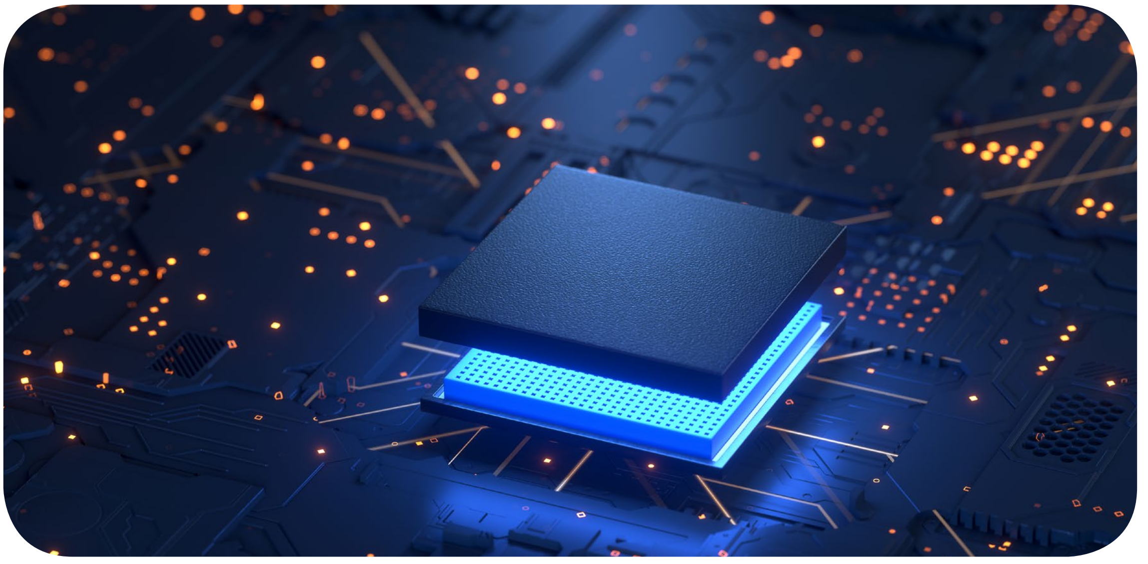
The next five years will see more efforts invested to improve the fabrication process of carbon nanotube materials, reduce surface pollution and impurities, and improve material purity and carbon nanotube alignment. In addition, the contact resistance and interface state of these devices need to be optimized to improve injection efficiency. Supporting electronic design automation (EDA) tools also need to be developed. Small-scale integrated circuits can be used to verify end-to-end maturity of carbon-based semiconductors, which are expected to be initially applied to flexible circuits. Looking ahead to the next decade, when carbon-based semiconductors are scaled down to the level of advanced silicon-based processes, there will be opportunities for large-scale application of this technology in high performance and high integration scenarios.

Novel storage media

While traditional storage mainly uses magnetic media, it is predicted that by 2030, 72% of enterprise storage, including both primary and secondary storage, will be based on all-flash. Furthermore, 82% of enterprise service data will require backup. Because of the differences between hot, warm, and cold data throughout the lifecycle, the evolution of storage media will diverge in two directions: higher speed with better performance, and massive scale at lower cost.

1) Novel media for memory

Currently, hot data is stored in SSDs and transmitted to DRAM for processing, because the latency of DRAM can be up to 1,000 times lower than that of SSDs. However, physical conditions limit DRAM from further density or voltage expansion. Therefore, neither SSDs nor DRAM are the best options for hot data storage. There are now many novel media technologies for memory, such as PCM, MRAM, ferroelectric RAM (FeRAM), and ReRAM, and those media



outperform DRAM in performance, capacity, cost, lifespan, energy consumption, and scalability. They also support byte-level access and persistence, making data migration unnecessary. Eventually, they will become commonly used media for hot data storage, but for now they face two major technical challenges:

Capacity: The total amount of hot data in 2030 will be equal to the total amount of the data stored on SSDs today. The capacity density of hot data media needs to be increased by at least ten times to reach the current level of SSDs, which is 1 TB/die. Such media should also support on-demand expansion unrestricted by processors, memory interfaces, network latency, and bandwidth. Media such as FeRAM, ReRAM, and MRAM face structural and material challenges.

Energy consumption: In the global push toward carbon neutrality, there is considerable pressure to reduce the power consumption of storage media for massive amounts of hot data. Resistor-based data storage technologies such as PCM and ReRAM require high data write voltages and therefore consume more power. The operating voltage of FeRAM, however, is relatively low, and its power consumption per bit is just one tenth that of ReRAM and MRAM, and a mere hundredth of that of PCM, making FeRAM the most promising candidate.

2) High-density NAND flash media

In the future, most hot data will be generated from warm data, which means warm data will become the largest reservoir of hot data. Therefore, warm data media must balance performance, capacity, and cost. NANDs will replace hard disk drives (HDDs) as the primary storage medium for warm data and are evolving towards multi-level cells and 3D stacking. The biggest challenge is to expand the capacity and reduce the cost of NANDs while achieving the same level of performance and lifespan as quad-level cells (QLCs).

Performance and lifespan of multi-level cells: For every additional bit a cell stores, the voltage needed for the data doubles, reducing read/write performance and lifespan by several folds.

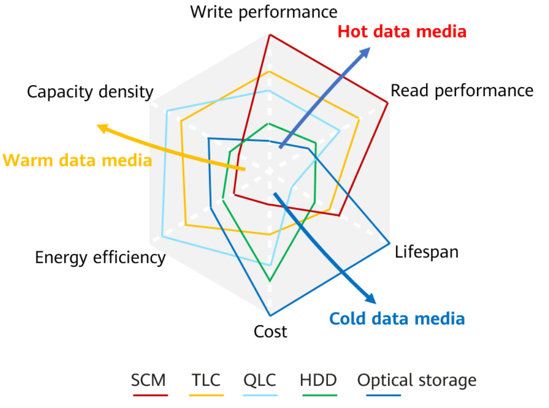
3D stacking process: At present, no mainstream 3D NAND SSDs contain more than 200 layers, but by 2030, we are likely to see products that stack close to 1,000 layers, and the aspect ratio of dielectric through-silicon vias will reach 120:1 (more than double the current level), making processing much more difficult.

3) Optical storage

In the future, the amount of cold data requiring long term storage will increase from 1.2 ZB to 26.5 ZB, and their retention time will grow by 5–10 times. At the National Archives Administration of China, for example, the retention time of key file data has been extended from 100 years to 500 years, and the amount of cold data that needs to be stored is expected to grow from 100 PB to 450 PB. Traditional hard disks and tapes can no longer meet such requirements. With the ongoing research on codec algorithms as well as the read/write mechanisms of transparent materials such as quartz glass and organic glass, optical storage will become the leading storage medium for massive cold data. Before that, however, two challenges must be overcome:

1. The service life of optical storage media needs to be extended tenfold and adapted for use in various complex and harsh environments.

2. Compared with Blu-ray, future optical storage media are expected to have ten times the capacity, perform ten times better, and be available at 1/5 the current cost.



Acknowledgments

During the drafting of this Computing 2030 report, we received invaluable support from Huawei's own team and external consultants. More than 300 experts and professors participated in the discussions that led to this report, contributing ideas and sharing their vision of Computing 2030. We would like to extend our special thanks to them.

(Contributors listed in alphabetical order)

André Brinkmann (Professor, Johannes Gutenberg University Mainz)
Bill McColl (Former professor at the University of Oxford)
Chen Wenguang (Professor, Tsinghua University)
Feng Dan (Changjiang Distinguished Professor, Huazhong University of Science and Technology)
Feng Xiaobing (Professor, Institute of Computing Technology, Chinese Academy of Sciences)
Gan Lin (Assistant Professor, Tsinghua University)
Guan Haibing (Changjiang Distinguished Professor, Shanghai Jiao Tong University)
Guo Minyi (Professor, IEEE Fellow, member of the Academia Europaea, Shanghai Jiao Tong University)
Jarosław Duda (Assistant professor, inventor of asymmetric numeral systems-based compression algorithms, Jagiellonian University)
Jia Weile (Associate professor, Institute of Computing Technology, Chinese Academy of Sciences)
Jin Hai (Changjiang Distinguished Professor, IEEE Fellow, Huazhong University of Science and Technology)
Jin Zhong (Professor, Computer Network Information Center, Chinese Academy of Sciences)
Miu Xiangshui (Changjiang Distinguished Professor, Huazhong University of Science and Technology)
Onur Mutlu (Professor, ACM Fellow, IEEE Fellow, ETH Zurich)
Pan Yi (Professor, Fellow of the American Institute for Medical and Biological Engineering, Foreign Fellow of the Academy of Engineering Sciences of Ukraine, member of the UK's Royal Society for Public Health, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences)
Shu Jiwu (Changjiang Distinguished Professor, IEEE Fellow, Tsinghua University)
Sun Jiachang (Professor, Institute of Software, Chinese Academy of Sciences)
Tian Chen (Associate professor, Nanjing University)
Tian Yonghong (Professor, Peking University)
Wang Jinqiao (Professor, Institute of Automation, Chinese Academy of Sciences)
Wu Fei (Professor, Zhejiang University)
Xie Changsheng (Professor, Huazhong University of Science and Technology)
Xue Wei (Associate professor, Tsinghua University)
Yang Guangwen (Professor, Tsinghua University)
Zheng Weimin (Professor, academician of the Chinese Academy of Engineering, Tsinghua University)

Digital Power 2030

The next energy transformation: Low-carbon and sustainable energy

Scaling back fossil fuel consumption and greenhouse gas emissions has become an urgent task for the world

Coal, oil, and electricity have been widely used since the 18th century. They played pivotal roles in the first and second industrial revolutions, helping civilization transition from an agricultural society into the industrial economy. As a cornerstone of global economic development, energy has consistently driven social progress, poverty reduction, and improvement of peoples' livelihoods.

However, human activities have clearly impacted the planet's ecology with greenhouse gas (GHG) emissions reaching record highs in recent years. According to the United Nations' Intergovernmental Panel on Climate Change (IPCC), human activities produce about 23.7 billion tons of carbon dioxide (CO_2) every year, of which around 20 billion tons are produced by burning fossil fuels. As a result, the amount of CO_2 in the atmosphere now is 27% higher than its average level over the past 650,000 years. The large amount of coal burnt during the industrial revolutions has resulted in a spike in CO_2 levels, putting our ecosystems at unprecedented risk and contributing to severe ecological and economic imbalances. This has driven people to discuss how to use fewer fossil fuels to reduce GHG emissions.

Fortunately, a clearer consensus has been reached among the scientific community and governments on climate change, and the Paris Agreement specifies that our most important goal in the global fight against climate change

is to achieve carbon neutrality by the middle of the century. Countries around the world are taking action. By the end of 2020, 44 countries and economies had officially declared their own carbon neutral targets. Some of these countries and regions have already achieved their targets, and others have already incorporated their targets into public policy and legislation.

The energy development strategies and practices adopted by major economies around the world have proven that reducing reliance on fossil fuels is one of the best ways to achieve carbon peak and neutrality goals. This requires countries to step up efforts to develop renewable energy while simultaneously improving energy efficiency and reducing overall consumption of fossil fuels.

Multiple countries have put forward targeted energy reform and GHG control goals. For example, China's National Development and Reform Commission and National Energy Administration have released the Energy Production and Consumption Revolution Strategy (2016–2030), which specifies that by 2030, China's new energy demand will mostly be met by clean energy. The strategy proposes reducing total energy consumption to at most 6 billion tons of coal equivalent (tce), with non-fossil fuel only making up about 20% of the total primary energy supply by 2030. China has also pledged to achieve CO_2 emissions peak by 2030, if not sooner.

The EU's 2030 climate and energy framework aims for net GHG emissions reductions of 55% compared to 1990 levels and an increase in renewable energy consumption to 38–40% by 2030. The US government has also pledged to achieve a 50–52% GHG emissions reduction from 2005 levels by 2030, and one of the most important steps to achieve that goal is to require the US grid to get 80% of its electricity from emission-free sources by 2030.

Sustainable, renewable energy will play a vital role in driving sustainable development of the world economy

Population growth and national industrialization have driven energy demand to unprecedented levels. Since commercial oil drilling began in the 1850s, experts estimate that we have harvested more than 135 billion tons of crude oil, with that figure increasing every day. Currently, global annual consumption of primary energy amounts to about 14 billion tons of oil equivalent, of which more than 85% is fossil fuels.

This means that fossil fuels will soon run out. According to BP, we will run out of global oil, gas, and coal resources in 50, 53, and 134 years, respectively, if our current extradition and consumption patterns do not change. This makes the development of renewable energy sources imperative to ensure sustainable development. UN Secretary-General António Guterres himself said, "Renewable energy is crucial for building a sustainable, prosperous and peaceful future" at the High-level Dialogue on Energy in March 2021. He also noted, "The year 2021 must be a historic tipping point towards sustainable energy for all."

According to Goal 7, set in the United Nations 2030 Agenda for Sustainable Development that was adopted at the seventieth session of the United Nations General Assembly: By 2030, ensuring universal access to affordable, reliable and modern energy services; by 2030, increasing substantially the share of renewable energy in the global energy mix; by 2030, doubling the global rate of improvement in energy efficiency; by 2030, enhancing international cooperation to facilitate access to clean energy research and technology, including renewable energy, energy efficiency and advanced and cleaner fossil-fuel technology, and promote investment in energy infrastructure and clean energy technology; and by 2030, expanding infrastructure and upgrading technology for supplying modern and sustainable energy services for all in developing countries, in particular least developed countries, small island developing States, and land-locked developing countries, in accordance with their respective programs of support.

Countries around the world are making the development of renewable energy an important part of their future energy strategies. Many countries have formulated specific strategies, plans, targets, regulations, and policies to support the development of renewable energy.

South Korea, for example, has unveiled a new long-term plan for renewable energy that increases the share of renewable energy sources in the electricity supply. According to this plan, by 2034, all coal-burning power plants in South Korea will be shut down and renewable energy will make up 40% of the country's total energy supply, up from their current 15.1%.

France's National Energy Plan 2030 notes that the country will continue to increase the share of renewables, especially wind power, in their power generation in order to achieve an energy transformation. France plans for renewables to account for 40% of its total power generation, with wind power accounting for 20%.

Germany plans to increase the percentage of renewables in their energy mix to 30% from their current 18%. Chile formally launched a green

hydrogen strategy in November 2020 to promote transformation of the country's energy structure that plans to reduce the amount of coal power to 20% by 2024 while gradually increasing hydropower, wind, photovoltaic (PV), and biomass power generation. Their ultimate goal is to increase the proportion of renewables in the country's total energy supply to 70% by 2030, and shut down all coal-burning power plants by 2040. Brazil has also continued to introduce policies and measures to provide funds and policy support for the development of PV infrastructure and projects. By 2035, total investment in the Brazilian electricity industry will exceed US$30 billion, 70% of which will be used to develop renewable energy technologies such as PV, wind, biomass, and ocean energy.

Cost-effective wind and solar power will drive an energy revolution

The cost of renewable electricity generation is dropping rapidly, even if fossil fuels still dominate the global electricity supply. Coal remains the biggest source of electricity by far, supplying 37% of all electricity, with gas coming in second supplying 24%. This is because fossil fuels are cheaper than other sources of energy. If we want to transition to a truly decarbonized energy system that primarily relies on renewable energy, we must ensure that renewable energy is cheaper than fossil fuels.

The global renewable energy industry has emerged as a promising new market in recent decades. Many countries have made wind and solar power generation part of their new energy strategies, and invested significantly into supporting R&D and industrial development in these areas. Driven by technological innovation, wind and solar power generation is also growing increasingly affordable. Oxford University's Max Roser found that the Levelized Cost of Energy (LCOE) of large-scale terrestrial PV plants was US$0.36 per kWh of electricity in 2009 and that within just one decade that price had declined by 89% to US$0.04 per kWh of electricity.

However, electricity from fossil fuels, especially coal, is not getting cheaper. Coal-burning power plants have a maximum efficiency of 47%, often leaving little room for substantial efficiency improvements. The price of electricity from fossil fuels is also not only based on the cost of technology itself but, to a significant extent, the cost of the fuel. The cost of coal that power plants burn accounts for around 40% of total costs. This means that even if the cost of constructing a power plant declines, the price of the electricity it generates will not continue to drop until it reaches a certain point.

However, each time the cumulative installed capacity doubles, the price of solar modules declines by 20.2%. Solar power has already become much more affordable in recent years, and this trend will continue as new PV module technologies and processes mature.

In addition to these cost benefits, wind and solar power generation is more flexible than traditional fossil fuel plants. Resource endowments have long influenced domestic energy development and utilization. However, as wind and solar are becoming new renewable energy sources, they can transcend the limits of resource endowments and produce electricity anywhere as long as their relevant requirements are met.

For example, distributed PV has attracted investors from many industries due to its low investment threshold. As wind and solar power generation becomes more affordable and flexible, more users are willing to use distributed PV systems in campuses, industrial complexes, and businesses, changing how energy is produced and utilized around the world. By the end of 2020, the global installed capacity of wind power and solar power had exceeded 650 GW and 750 GW, respectively.

Offshore wind power, for example, is an important type of wind power that occupies no land space. The power generated from offshore wind turbines is directly delivered to coastal load centers nearby, avoiding the waste that long-distance transmission causes. Because of this, we are currently seeing a shift from onshore wind farms to offshore wind farms.

By 2024, distributed PV systems will account for nearly half of the entire solar power market, with industrial and commercial distributed PV systems holding a major market share. Floating PV plants have become popular in many regions because they offer larger power generation capacity, no land requirements, and lower impact on water bodies. Over 60 countries have been pushing for wider adoption of floating PV plants, with total power output from floating PV plants expected to reach over 60 GW in the next five years. We expect that the global demand for fossil fuels will also peak in next five years as wind and solar power become more affordable and the installed capacity spikes.

Power electronics and digital technologies are a key driver of energy transformation

Power electronics technologies ensure security and control during energy system transformations

Power electronics technologies play a key role in electricity generation, distribution, transmission, and consumption. As more electricity is generated from renewable energy sources such as wind and solar, energy transformation efforts will focus on building an energy system that will be centered on electricity, connected to power grids, and based on power electronics devices. The inclusive interfaces, fast response times, and high conversion efficiency of power electronics devices are already being widely implemented in electric power generation, transmission, and consumption.

a) In electric power generation, power generating systems using renewable energy, such as wind and solar, cannot directly transmit electricity to local grids like conventional electric generators. Power from these renewables first needs to be converted into frequency-adjustable AC using power electronics technologies to meet the grid transmission requirements. For example, PV inverters and wind converters can adjust voltage waveforms through power electronics switches to enable the transmission of renewable electricity to local grids, making power generation more efficient.

b) In electric power transmission, intelligent power electronics devices can significantly improve long-distance power transmission performance, power flow distribution, and reliability of power supply. This makes electrical grids more secure, thereby making power transmission over large-scale grids more secure, reliable, and efficient.

c) In electric power distribution, large numbers of distributed power supplies, microgrids, and flexible loads are being connected to power distribution networks, increasing the requirement for plug-and-play power supply and the overall amount of reactive power in the transmission lines. Problems such as voltage spikes and harmonic distortion are also becoming increasingly serious. There are limited ways to improve the power quality and supply stability of traditional distribution networks, meaning these networks alone can no longer meet user requirements. Power electronics devices, such as multi-functional power electronics transformers, DC circuit breakers, and DC switches, can instead be used to guarantee the power quality of different load categories and meet tailored electricity needs.

d) In electric power consumption, demand for DC power and proactive source-load interactions are increasing due to the application of distributed power supply and energy storage devices, and the emergence of new types of facilities, such as data centers, communications base stations, electric vehicle (EV) charging stations, computers, and LED lights. Switching power supplies and switchgears with high efficiency, high power density, high reliability, and low cost are meeting the increasingly diverse personalized needs of users and the demand for quality assurance of electric power.

Demand for new types of power semiconductor devices is set to skyrocket. Energy systems of the future will need to better utilize renewable energy resources, meaning the bar for energy transmission and control subsystems is going to rise significantly in terms of safety, efficiency, and intelligence. We will need entirely new electricity transmission and distribution networks designed specifically for renewables, more efficient terminal systems that work better with other subsystems like distributed power supply and energy storage, and more comprehensive service systems that are integrated with information systems. Changes introduced by these new systems will need to be managed, compensated, or controlled by power electronics devices, which currently rely on silicon-based components to a large extent.

However, the reality is that silicon-based components are going to hit a wall soon. The physical properties of silicon mean that there will no longer be a way to further improve performance. Many are already struggling to further reduce the energy use of silicon-based components. These components simply won't be a good fit for generating, transmitting, consuming, and absorbing clean energy at scale in future energy systems.

Third-gen power semiconductor chips and components, based on silicon carbide (SiC), stand out for their high voltage, high operating frequency and temperature, and high speed switch. These SiC components have delivered a huge boost to the reliability, availability, energy density, and energy conversion efficiency of power electronics devices while simultaneously reducing overall cost and energy loss. SiC components are ideal for sectors with high requirements on energy conversion efficiency, such as electricity generation based on renewables (e.g., solar and wind), ultra-high-voltage direct current electricity transmission, new-energy vehicles, rail transportation, industrial power supplies, and home appliances.

The next decade will see increased efforts to adopt and further innovate third-gen power semiconductor devices. The substrates used

in SiC components are currently four to five times more expensive than those used in silicon components, but their costs are expected to break even by 2025. By then, high uptake of SiC components in new-energy vehicles, industrial power supplies, and other domains will help drive down costs. New technologies that will enhance the performance and reliability of SiC components are also expected to emerge. These trends will prime the SiC sector for explosive growth and market development. Yole Développement predicted that the market for SiC components will expand rapidly, from US$600 million in 2020 to US$10 billion by 2030. It is estimated that by 2030, over 70% of solar inverters will use SiC components, compared to the current 2%. By then, SiC components will most likely be found in more than 80% of charging infrastructure and EVs, and be widely used in the power systems of communications networks and servers.

Digital technologies drive intelligent transformation of energy systems for greater value creation

Energy systems will soon become more distributed, thanks to the rapid increase in the number of new renewable energy facilities (e.g., wind and solar) and the increasing flexibility of applications that support these systems. Energy systems of the future will be decentralized, just like a nebula, with ecosystems made of numerous distributed energy applications. Large power plants, campuses, buildings, households, EVs, and countless other facilities will also develop their own energy systems. These distributed energy systems will not be sustainable if they rely on traditional models. Intelligent connectivity and control powered by digital technologies will make energy systems highly intelligent and connected, which in turn will make them safer and more stable, efficient, affordable, and flexible. They will then be better positioned to reduce carbon emissions and more effectively generate clean energy.

Advances in emerging technologies, particularly 5G, cloud, AI, big data, and IoT, are ushering all sectors of society into a new digital era. This will be an era where all things can sense, connect, and work intelligently. This vision of ubiquitous connectivity and pervasive intelligence is already becoming a reality. The following new digital technologies are being adopted in the energy sector at an increasing pace and will soon become game changers:

Networking: Low-power wide-area networks (LPWANs) are quickly gaining commercial popularity around the world. With wide coverage, low latency, and massive connectivity, 5G is ideal for IoT applications and is permeating a growing array of scenarios that require on-demand, intelligent connections between people, machines, and things.

Information processing: Information perception, knowledge representation, and machine learning technologies are progressing by leaps and bounds, driving IoT's ability to intelligently process data to levels we have never seen before.

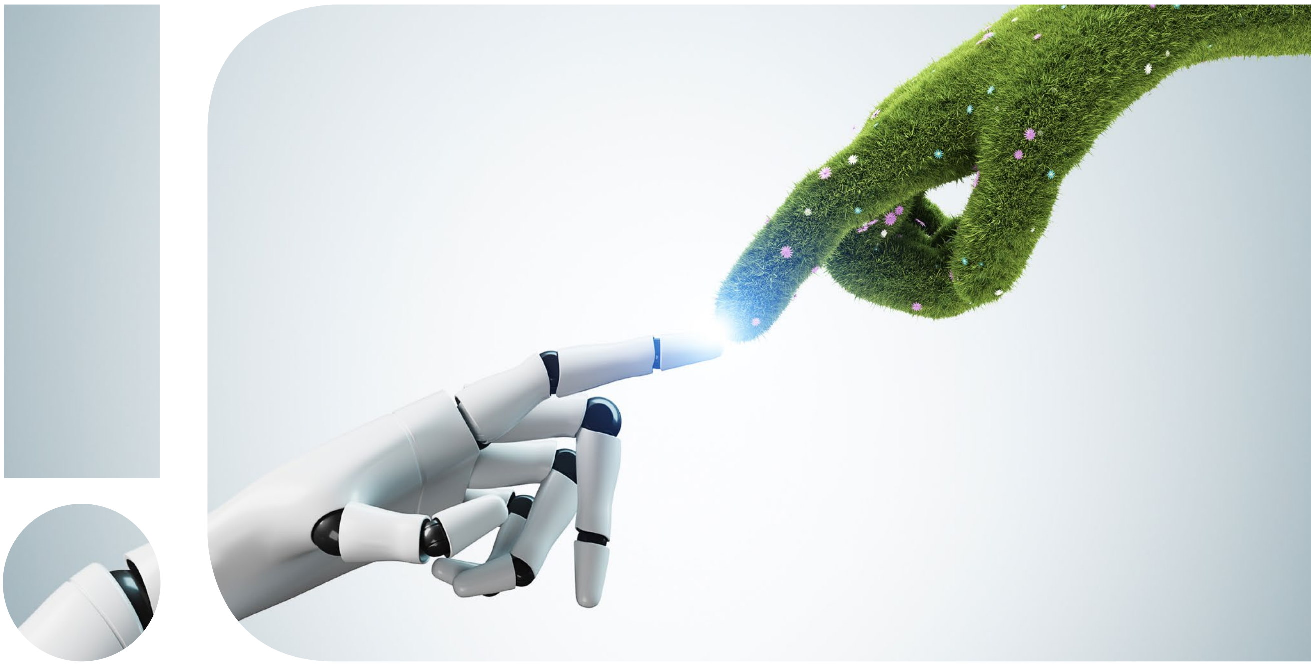
IoT virtual platforms, digital twins, and OSs: Widespread adoption of cloud computing and open-source software is reducing barriers to entry for those who hope to play a part in the energy sector. Cloud computing and open-source software are also boosting the popularity of energy system OSs and digital ecosystems.

By making energy systems more intelligent, we are moving energy supplies closer to users for better experiences. As distributed energy systems continue to grow in popularity, users will become prosumers – those who both consume and produce energy.

Highly intelligent energy systems can flexibly determine when to switch from generating electricity (when energy prices are high) to storing electricity (when energy prices are low). These energy systems can also adjust energy generation based on the load, and vice versa. This means the systems will be able to transfer energy flow to and from each other across time zones and distances.

EVs will be able to double as storage facilities that feed electricity back into the power grid during peak hours to help meet demand. Data centers will be able to provide heating by reclaiming huge amounts of the residual heat they produce. Communications sites will be even more versatile, supporting vehicle-to-infrastructure systems and intelligent "brains" for cities. Intelligent sockets for homes will become endpoints that detect, meter, and trade electricity. Distributed energy, energy storage, and the electricity spot market will thrive. There will be an untold number of prosumers, and they will enable energy systems to respond to demand and provide value-added services.

Digital technologies play an essential role in supporting the source-load interaction of energy systems by enabling high-speed and high-frequency computing. Digital technologies are also weaving their way into the fabric of many different sectors. As this happens, data centers, communications base stations, and other types of facilities used for transmitting, computing, and mining information flow will grow in number and consume more energy. This brings us to the question: How can new ICT energy infrastructure reduce its own energy consumption and emissions? We will address this question below in this report.



The era of digital power: Digital and intelligent transformation for integrated information and energy flows, and green and low-carbon operations

In the next decade, renewables like wind, solar, and hydro power will replace fossil fuels as the main sources of electricity. Electrification of consumer terminals on power grids is also increasing. Technologies for EVs, hydrogen energy, energy storage, heat pumps, and thermal energy storage are advancing rapidly. Transportation, heating, and other energy-consuming systems are quickly shifting away from diesel, petrol, natural gas, and coal and towards electricity. Energy systems will soon be embedded with more advanced plug-ins, and be underpinned by an energy cloud OS that integrates information flows and energy flows. Connecting electricity production and consumption will enable two-way, Internet-based interactions between all different types of industry players that handle everything from energy sources and power grids to load management, energy storage, and consumption.

Transforming energy systems will open up enormous opportunities for innovation in technology, business models, and operations in the energy sector. Electricity generation involving renewables (e.g., solar power) in particular will have many opportunities. The same is true in the green mobility sector, which will be driven mainly by EVs, and other energy-consuming sectors, particularly ICT energy infrastructure.

It's predicted that by 2030, renewables will account for over 50% of all electricity generation globally, LCOE of PV plants will reduce to US$0.01 per kWh of electricity, and over 3,000 GW of PV plants will be in place. The proportion of electricity in global energy consumption will increase from the current 20% to 30%. Over 50% of vehicles sold will be electric. Renewable energy will power more than 80% of ICT energy infrastructure.

PV plants will be grid-friendly, intelligent, and convergent, with lower LCOE

LCOE of PV plants may drop to US$0.01 per kWh of electricity by 2030

LCOE is a measure of the average net cost of electricity generation for a PV plant throughout its lifecycle. It is used to compare the electricity generation costs of PV plants with other types of plants. Under a full-lifecycle investment model, LCOE is determined by a plant's upfront investment, operation & maintenance (O&M) expenses, and the system's useful life. By 2030, LCOE of PV plants is expected to plummet, possibly even down to US$0.01 per kWh of electricity.

PV plants are composed of PV modules and balance of system (BOS) components (such as electrical cables, solar inverters, and wires). Generally speaking, about 45% of a PV plant's investment goes into its PV modules. Over the next decade, this percentage is expected to drop by at least 15 points, because of improved engineering techniques, decreased manufacturing costs, and the ever increasing efficiency of PV modules. This means more investment will go to BOS components and O&M. On top of this, technological innovations will drive up the overall cost competitiveness of PV plants:

a) PV plant systems will be able to support higher voltage. As input voltage increases, so does output voltage. This in turn can reduce line loss in direct current systems and loss in low-voltage transformer winding, significantly increasing the systems' efficiency. In addition, solar inverters and transformers will become more compact, translating into a huge reduction in transportation and O&M workloads. PV plant maintenance will also be automated. Thanks to these trends, by 2030, the system voltage of PV plants will reach 1,500 V or even higher, further slashing LCOE.

b) Solar inverters will deliver higher power density and efficiency because of advanced materials like SiC and gallium nitride (GaN), better heat dissipation in chips, and topology technologies. These materials and technologies increase solar inverters' voltage, operating temperature and frequency, and reduce loss. By

2030, solar inverters will see their power density grow by over 70%.

c) PV plants will use modular, standardized components. For example, solar inverters, power control systems, energy storage, and other key systems will use standard interfaces that allow for flexible capacity expansion and rapid deployment. All internal DC and AC circuit breakers, inverters, controllers, and heat dissipation components will be modularized. This will eliminate the need to enlist experts for maintenance work, slash O&M costs, and enhance system availability. Full modularization at the system and equipment levels will be the future trend.

d) PV plants will be digitalized inside and out. As digital and PV technologies converge, they will make O&M, production, and asset management simpler, more intelligent, and more efficient. With AI, PV plants will transform into intelligent systems. AI will handle the tasks that used to be performed by highly-trained experts, and support autonomous and collaborative optimization inside PV plants. Intelligent tracking algorithms make it possible for PV modules, trackers, and solar inverters to work in tandem to continuously find the maximum power point (MPP) of solar panels, thus maximizing power output. With AI, fault location will be more precise and O&M times can be reduced from months to minutes. Other benefits of AI include higher electricity generation efficiency, better O&M experiences, and greater productivity and safety. By 2030, AI is expected to be used in 90% of PV plants.

PV generators ensure stable operations of power grids through proactive support for frequency and voltage fluctuations

PV power generation technologies have the potential to make power grids more resilient. PV power generation fluctuates wildly over time, so it can only meet the energy dispatching demands of power grids with the support of regular power supply services such as peak shaving and backup. Therefore, as more PV generators are brought into a power grid, the grid itself will become more vulnerable. For example, the power grid's system inertia may drop, and its ability to regulate frequencies and control system voltage may suffer. What's more, the characteristics of faults and oscillations on the power grid may change significantly.

Effectively integrating PV generators into power grids and harmonizing operations is key to incorporating large amounts of renewable energy into power grids and changing the energy mix. In a power grid, fossil fuel power plants and hydropower plants typically use conventional synchronous generators. These synchronous generators utilize mechanical structures to provide stable voltage and frequency, thus facilitating frequency regulation and voltage control. However, as asynchronous generators gradually displace synchronous generators in power grids, the way power systems work will change fundamentally. In response, renewables-based power systems will need to simulate the technical indicators of synchronous generators, in order to proactively support the grid's frequency and voltage fluctuations. The goal will be to help power grids become safer and more reliable.

PV power generation technologies combine power electronics, energy storage, and digital technologies to simulate the electromechanical transients of synchronous generators. When connected to power grids, PV generators have many of the same external characteristics of synchronous generators, such as inertia, damping, primary frequency regulation, and reactive voltage control. As a result, PV generators can offer technical specifications that are similar to the synchronous generators used in fossil fuel power plants. PV power generation technologies can proactively support the operations of renewables-based power systems and make them more grid-friendly. This will help renewables become mainstream and provide a solid technical foundation for incorporating renewables into power grids.

Energy clouds will intelligently converge energy and information flows to synergize generation, grids, loads, and storage

Energy clouds that converge energy and information flows will function as the OS of the digital power industry. They will direct information flows, regulate energy flows, and spark an energy revolution in which bits can be used to manage watts. In the future, electricity will be the main energy carrier in energy systems, and digital and power electronics technologies will be leveraged to transform all aspects of power infrastructure, including power generation, transmission, distribution, usage, and storage. New energy will be observably, measurably, controllably, and adjustably enhanced to address the vulnerability of new energy access systems and increase new energy consumption. Improving the ability to control and regulate extensive terminal systems, like microgrids, integrated energy, and distributed power supply will also enable real-time interaction between power generation units and users. The data generated by these networks will allow power generation units to learn from and adjust to user consumption habits, improving resource utilization. This will improve the quality, safety, and stability of electricity systems.

a) The physical distribution of energy resources is often the inverse of actual energy demand, but the energy cloud will remove these time and distance limitations from energy flows. Take the situation in China, for example. Northwest and Southwest China have abundant wind, solar, and water resources but low demand for power consumption, while Central, East, and South China have high demand but insufficient energy resources. When new energy sources are centrally accessed over local grids, transmission between these regions is further hindered by high randomness and volatility. On the consumer side, the large numbers of user devices and power supplies, such as EVs and distributed power supplies, results in an increased demand for distribution network resources and increased vulnerability in regional power grids. Grids need stronger zoning and interconnection, simplified system operations, and the ability to provide mutual support. Fault isolation also needs to be strengthened to prevent cascading faults that would cause backbone power grids to break down. An energy cloud can arrange for more sharing of distribution network resources with technologies such as active distribution networks and flexible DC distribution networks, making it perfect for scenarios like microgrids, virtual power plants, and integrated energy systems. An energy cloud will improve the digitalization of transmission and distribution networks, make their operations more flexible and adaptable, and enhance overall network control capabilities.

b) In conventional power grids, over the entire process from electricity production to consumption, more than 50% of resources are wasted. An energy cloud will make the connections between electricity production and consumption more resilient by enabling unified management. Synergy between generation, grids, loads, and storage will automate the distribution of integrated energy resources. Regional nodes will be able to be monitored and managed in real time, and regional energy consumption will be equalized to balance production and consumption. In this way, electricity production and consumption are intelligently aligned and collaboratively operated so as to improve resource utilization. For example, optimization algorithms can ensure that solar PV and wind power generation and storage can adapt to their power market, and take into account the local weather forecast and other factors that influence production. Data integration then ensures the optimal combination of power generation. Flexible interconnection and digital control of multiple integrated energy sources will strike a balance between energy supply and demand over larger networks. This will make energy systems more flexible and better able to meet different objectives like cost-effectiveness, carbon emission requirements, and comprehensive energy efficiency. This makes it possible to use a wider range of energy types to meet complementary types of demand. Flexible conversion and integrated demand responses from multiple energy sources will make electric power systems more flexible and ready to take on more renewable energy.

Transportation will be electrified, with sales of EVs catching up with combustion-engine vehicles by 2030

Currently, the transportation industry is highly dependent on fossil fuels like petroleum, and the transportation sector is responsible for about a quarter of all energy-related carbon emissions. In Europe, transportation is the second largest carbon emitter, right after electricity generation. In the United States, transportation is the single largest source of GHG emissions. The transportation industry is primarily centered around four methods of transport – road, rail, aviation, and shipping – each of which has different requirements for green fuel. The rapid advancements in battery technologies and the growth of comprehensive electricity infrastructure have made electricity the most important clean energy alternative in road and rail transportation. Interaction between electricity systems and transportation systems is increasing, and the two are becoming more integrated. EVs, for example involve transportation, power consumption, and energy storage. EVs and their charging or power facilities will become critical hubs for the integration of electricity systems and transportation systems.

Many countries have already begun to promote EV development in recent years. Large-scale investment in electrification, intelligent driving, and Internet of Vehicles (IoV) technologies have greatly improved the competitiveness of EVs. Consumers are also turning to EVs in increasing numbers because of their energy-saving, eco-friendly, intelligent, and high-tech features. Though the global automobile industry as a whole is still suffering from the fallout of COVID-19, EV sales rose by 41% in 2020, with sales exceeding 3 million vehicles, achieving a market share of 4%. EV sales exceeded 1 million in Europe, 1.3 million in China, and 250,000 in the United States. As battery costs decline, performance improves, and autonomous vehicle technologies evolve, EVs are expected

to become as cost effective as combustion-engine vehicles by 2025, and the EV market growth will increase. More than 40 million EVs are expected to be sold globally by 2030, tying with combustion-engine vehicle sales. Adequate charging facilities will be needed to support this development. According to the International Energy Agency (IEA), by 2030, the global number of private charging piles is expected to reach 100 million, collectively delivering a total charging power of 1,500 GW and a total charging capacity of 800 TWh. The number of public charging piles is expected to reach 20 million, with a total charging power of 1,800 GW and a total charging capacity of 1,200 TWh.

New materials and digitalization will redefine the EV experience and safety

Extensive application of wide-bandgap semiconductor materials and digital control technologies will collaboratively be used to help EVs achieve an optimal energy efficiency ratio. As power components, topologies, and control algorithms related to power electronics advance, power devices will deliver record high efficiency. The application of new technologies and materials such as silicon carbide will increase bandgap width almost threefold and electric field strength 15 fold, double electron saturation rates, and triple thermal conductivity when compared with traditional silicon. An E2E architecture covering charging, driving conditions, power transmission, power conversion, heating, cooling, and energy recovery will be constructed and continually upgraded to deliver system-level efficiency optimal for EVs. Powered by digital technologies, intelligent electrothermal synergy, intelligent torque distribution algorithms, and intelligent electro-hydraulic braking distribution will be used to achieve high efficiency at every level from components to systems and from the power domain to vehicle operation. To further save energy and maximize EV range, a hyper-converged and domain-based control architecture will be used to implement multi-energy scheduling through coordinated control of electric energy, kinetic energy, thermal energy, and energy recovery, achieving a high vehicle-level efficiency in all aspects such as power charging, storage, and consumption. Intelligent electrothermal synergy will allow the heat from the motor and inverter to be intelligently recaptured through the heat pump system to the passenger compartment for heating. Drive torque can be intelligently distributed to balance braking ability and energy recovery. Other technologies such as optimal distribution between motor braking and hydraulic braking will also extend EV range.

Digitalization is redefining EV experience. EVs are delivering better driving experiences

than combustion-engine vehicles in terms of acceleration, control, and intelligence as battery energy density increases, battery management improves, and the electric control systems become increasingly precise in calibration. High-power EVs with fast acceleration are becoming more and more common. 300 kW, 400 kW, 600 kW, and 800 kW EVs now regularly outperform combustion-engine vehicles. Distributed electric drives are replacing mechanical limited-slip differentials (LSDs) in combustion-engine vehicles to archive faster acceleration at curves and better off-road driving, making EVs more enjoyable to drive. In terms of innovative intelligent features, the service oriented architecture (SOA) and centralized electrical and electronic architecture (EEA) allow software features to be kept up-to-date over the lifetime of an EV's power domain, and intelligent remaining range estimation lets drivers hit the road without worrying about how much charge they have left. EVs operating in smart track mode can adjust their thermal systems to boost power and control front and rear drive torque to drive for more fun. Intelligent accelerators and converged drive and brake give drivers more control. The motor itself will monitor the tire slip and adjust torque in real time to prevent hydroplaning on wet or snowy roads, greatly improving driver control and safety.

The energy cloud will improve vehicle energy efficiency management at the cluster level and make periodic service experience all-online. In an era of intelligent EVs, the diversity of the market and customers' personalized needs will more heavily influence automotive R&D, launch, and lifecycle requirements. These changes will lead to manufacturing and product service model innovation, driving a digital transformation of the entire automotive industry. For example, digital twin technology for the power domain of alternative fuel vehicles is based on vehicle power system digitalization and IoV technologies. With this technology, real-time operating data of vehicles and their components is continuously collected through sensors and fed to a digital twin created on the cloud. This data is then synchronously fed into a digital model on the cloud. The cloud generates data for real objects monitoring their real-time operating status and power domain status, and then enables real-time interactions to ensure that the EV power system works reliably and efficiently. Cloud computing brings natural advantages in terms of computing power, algorithms, model training, big data storage and analytics, and ecosystem partner participation. Digital modeling on the cloud can therefore be achieved based on the real-time data of vehicle components such as batteries, motors, and electric controllers. The cloud can build fault prediction analysis algorithms, parameters for efficient operating status, device aging models, intelligent fault rectification algorithms, and intelligent calibration algorithms. It remotely diagnoses and rectifies faults, and enables OEMs and service providers to proactively optimize product design based on actual end user requirements. These features improve service efficiency, reduce vehicle use costs, and enhance user experience.

Kilovolt flash charging will be widely applied to supplement energy supplies

As ranges increase and charging becomes easier, more and more customers are choosing EVs. From a technical perspective, EV range is being increased by improving battery energy density, and fast charging is becoming more common by increasing battery voltage for fast charging. Take electric passenger cars, for example. Average battery capacity is expected to increase from 60 kWh to 100 kWh by 2025, and mainstream charging voltage is expected to rise from 500 V to 1,000 V by 2030, bringing the EV industry into the kilovolt era. Each charging gun will be able to deliver 480 kW of power, up from the current 60 kW and charging will speed up — today a full charge takes about an hour, but this could be brought down to less than 10 minutes in a few years. This is almost comparable to how long it takes to refuel a combustion-engine vehicle. We will also have kilovolt-level EV power systems that are intensive, integrated, and well-coordinated, helping to decrease current and reduce energy loss. High-voltage platforms and precise high-rate charge/discharge curve design will enable efficient coordination between charging, discharge in driving, and kinetic energy recovery. High-voltage technologies will be widely used in charging infrastructure systems. For example, high-voltage silicon carbide technology will promote high-efficiency and high-density application and support the evolution of high-voltage platforms. Based on the ChaoJi charging technology roadmap, a 1,000 (1,500) V charging voltage platform will support a maximum charging power of 900 kW. This type of supercharging technology will be widely deployed on intercity highways.

EVs can collaborate with various energy systems to become regulators of energy flows

EVs will become fully involved in interactions with energy systems as important regulators in energy flow control. The large-scale promotion of EVs and renewable energy creates opportunities for vehicle-grid synergy. There is increasing demand for a large number of flexible power sources on the power-generation side and for adjustable load resources on the consumption side. Unlike more common electrical loads such as household appliances, EVs are highly flexible and adjustable. As wireless charging, smart charging, and autonomous vehicle technologies mature and are widely adopted, EV users will be free to choose when to charge, discharge, and swap batteries, participating in the electricity spot market and ancillary service market based on their needs. This will reduce the impact of EV charging on the power grid, provide new resources for the power system to schedule, and avoid a large amount of wasted investment in power grid and power supply.

The number of EVs worldwide could exceed 150 million by 2030. Ideally, by that time, the energy storage capacity could be 40 times as large as the energy storage capacity installed in 2020, with the potential to serve as an adjustable load and a flexible power source. In the future, EVs participating in the frequency-modulation ancillary service market will have higher value. By 2025, EVs will be able to take full advantage of their role as flexible loads and perform orderly charging as a way to contribute to user-side applications such as peak load shaving, distributed PV charging, demand response, peak clipping ancillary services, and spot market balancing. As the cost of power batteries decreases and their service life increases over the course of the decade, EVs will become increasingly capable of serving as distributed power sources. With the support of platforms such as microgrids and virtual power plants, EVs will be able to provide services like frequency modulation and spot power balancing through both charging and discharging.

Charging infrastructure connects vehicles, transportation systems, and mobile lifestyles, as well as diverse energy use scenarios. It is the point of convergence for energy and transportation, in terms of transactions, interaction, behavior, and information. It is one of the important enabling components of the energy cloud. Large-scale construction of charging networks and the development of technologies such as digitalization, IoT, cloud computing, big data, and AI bring about multi-level improvements in intelligence: Intelligent charging infrastructure will make charging networks visible, manageable, controllable, and optimizable, greatly reducing maintenance and operation costs and increasing efficiency and revenue. As a data interface, charging piles can be utilized to build a smart charging network that integrates vehicles, charging piles, power grids, the Internet, and value-added services. This network will leverage charging piles' strengths in terms of scale, integration, data, and connectivity, create multiple new business models, and generate a virtuous cycle of economic and social benefits. Charging piles enable charging facility operators to provide data consulting services to support business

district construction, real estate development, 4S store planning, the second-hand car market, digital payment, and e-commerce operations as a way to monetize, expand sources of revenue, and improve market operation capabilities. For local governments, charging piles can provide data support for urban planning, power dispatching, everyday services, and infrastructure construction, making charging infrastructure an important part of smart cities.

ICT energy infrastructure is going green

A decade from now, there will be hundreds of billions of connections, the total amount of general computing will have increased 10-fold, and the total amount of AI computing will have increased 500-fold. ICT technologies will empower other industries to reduce their carbon emissions, and a global reduction of emissions by 20% is predicted. Construction of related infrastructure will increase. For example, the number of telecom sites will increase from 10 million to 55 million, and the number of data center racks will increase from 4.2 million to 10 million. ICT will account for 4% of global power consumption, up from less than 2% today. Building efficient and low-carbon communications networks and data centers will not just be an operational necessity for enterprises, but their civic duty.

In addition to providing high-quality ICT services, leading operators worldwide have made carbon-neutrality declarations and launched various initiatives. Vodafone and Orange have proposed to achieve net zero emissions by 2040, while Telefónica has brought its carbon-neutrality target forward to 2030. In addition, Google has set the goal to power its operations — across all of its data centers and campuses worldwide — entirely with carbon-free energy by 2030. Microsoft has pledged to be carbon negative by 2030, and to remove all the CO_2 Microsoft has ever emitted, either directly or through electricity use since the company was founded in 1975, from the environment by 2050. The municipal government of Beijing has required that data centers be built with their own distributed renewable energy facilities and be powered 100% by clean energy by 2030. Key players in Europe's cloud infrastructure and data centers have developed a self-regulatory initiative, the Climate Neutral Data Centre Pact. Japan plans for its data center industry to become carbon neutral by 2040. It is predicted that, over the next decade, the ICT energy infrastructure will evolve in a number of directions that will be discussed in the following sections.

Comprehensive architecture refactoring is making ICT energy infrastructure simple, converged, flexible, and efficient

Networks and data centers will become larger and more complex. The ongoing pursuit of simplicity is driving the development of ICT energy infrastructure architecture toward greater convergence. Most of today's telecom sites are built indoors, and traditional air conditioners are used for cooling. The overall energy efficiency of these sites is only 60%. Conventional power supply solutions typically use multiple sets of power supply equipment, each supporting a different voltage system, which complicates deployment. We believe that the form of telecom sites will change dramatically in the next decade. What once filled an equipment room can now be squeezed into a cabinet, and what once filled a cabinet can now be mounted on a pole. Sites are becoming simpler and more reliable, with smaller footprints and lower rent. The way in which data centers are built will also change rapidly. Traditional concrete buildings usually take more than 20 months to build, and concrete is neither environmentally friendly nor recyclable. Prefabricated modular data centers will become increasingly common over the next decade. Prefabrication reduces the use of high-carbon-emission materials, such as concrete, rubber, and rock wool sandwich panels, and greatly reduces onsite construction and maintenance. This way, a data center housing 1,000 cabinets can be built in only a few months, meeting the requirements for rapid service rollout.

In terms of network and data center power supply solutions, power supply link convergence will become a major trend. Adapting to more renewable energy sources, ensuring compatibility with multiple energy supplies, and being ready for smooth evolution are the directions in which the power supply architecture will evolve. Examples include multi-mode scheduling control and management, modular overlay evolution, and the convergence of different services and devices across multiple scenarios. With this converged architecture, power supplies and batteries of telecom sites are being converged into a blade form factor. Power supply, energy storage, temperature control, and power distribution are being integrated into a single module, and on-demand evolution is being enabled to support cross-generational network evolution. All data center power supply links, including transformers, uninterruptible power systems (UPSs), and power distribution, will be converged to reduce footprint. All backup power will be based on lithium batteries to support intelligent collaboration between electric power generation, storage, and consumption, reducing the required configuration capacity of the data center UPS, and reducing the footprint and construction costs of data centers.

Green energy is becoming mainstream in ICT energy supply

The global trend of digitalization is turning ICT into an energy-intensive industry. Driven by the goal of carbon neutrality, ICT infrastructure providers are increasingly turning to clean energy sources such as PV, wind, and hydrogen. With the cost effectiveness and flexibility of these distributed energy sources, over 80% of ICT infrastructure power systems will use distributed green energy over the next decade. As the power consumption of individual telecom sites is low, distributed PV systems are likely to become a major power supply solution, helping to make zero-carbon communications networks a reality. Instead of taking the traditional approach of signing renewable-energy power purchase agreements (PPAs) and purchasing green certificates, sources of clean power will be integrated into the design of data centers themselves. For example, distributed PV plants are built on data center campuses and rooftops, and large-scale terrestrial PV plants, wind farms, and other types of clean-energy structures are built in surrounding areas to directly supply data centers. With lithium batteries replacing lead-acid batteries, the backup power of telecom base stations and data centers will soon be fully lithium battery based. With intelligent regulation, these traditional unidirectional distributed energy systems will participate in the ancillary service market which includes services like power grid peak clipping. This will allow the systems to help address

the unpredictability and intermittency of wind power and solar power. Therefore, the revenue coming from ICT infrastructure power supply will increase, maximizing the commercial value of basic resources. The stability and reliability of the entire energy system will also improve.

ICT energy infrastructure O&M is moving toward full automation

Over the next ten years, ICT energy infrastructure O&M will gradually become more autonomous, thanks to technologies such as neural networks, knowledge graphs, and domain shift. The combination of AI with other technologies will greatly improve O&M efficiency, eliminate a large amount of repetitive and complex manual work, and improve the fault prevention and prediction capabilities of energy infrastructure using big data. Data-driven differentiated service models will enable highly automated and intelligent operations of ICT energy infrastructure, profoundly changing how O&M is performed.

a) From manual operations to autonomous system operations: Operations involving inefficient and repetitive work (for example, configuration delivery, change, and upgrade) will be automated. The role of O&M personnel will change from in-the-loop intervention to on-the-loop management, greatly improving operation efficiency and addressing the pressure of intensive maintenance in the future. This significantly shortens the time needed for construction and service rollout.

b) From manual decision-making to machine-assisted and even autonomous decision-making: Traditional O&M approaches that rely on expert experience will change. Driven by data, AI machine learning can be leveraged to assist in or even automate decision-making under human supervision. This enhances the system's ability to cope with complexity and uncertainty, and greatly improves the response speed, resource efficiency, and energy efficiency of energy infrastructure.

c) From open-loop operation management to closed-loop guaranteed operations: The automation of ICT energy infrastructure will streamline data flows from end to end and make closed-loop autonomy a reality. Closed-loop autonomy will be implemented based on predefined SLA policies during construction, maintenance, and optimization to ensure manageable and guaranteed electricity production and consumption policies. This enables differentiated services for ICT energy infrastructure, more efficient resource utilization, higher revenue, and differentiated business innovations.

Conclusion

How advanced the energy sector has become can be seen in the extent to which renewable energy, digital technologies, and power electronics technologies have developed and converged. In the future electricity-based energy sector, power grids will be like the backbone networks in ICT, power electronics devices will be like gateways, and the energy cloud will be like the OS. The way in which energy flows are processed, moved, and stored will also change. As a result, the large-scale development and utilization of clean and low-carbon energy, wide connection of multi-level energy networks, active and passive participation of multiple loads, and collaborative decision-making and operation of multiple service logics will become a reality. Energy flows and information flows will be deeply integrated and support each other over the next decade. This will be a key transition period for a comprehensive energy transformation that will impact the development of energy landscapes over the next century.

The energy sector is now entering the digital age. Technological innovation regarding information flows and energy flows is becoming increasingly synchronized, shifting the focus of innovation from individual devices and scenarios to systems and industries as a whole. Energy networks are expanding from local networks into global networks, and operations are changing from device-based to cloud-based. Energy systems are becoming visible, measurable, and controllable on a wider scale. The convergence of energy flows and information flows is extending on a larger scale across time and space, further amplifying the value of energy systems. Energy systems are becoming more economical, cleaner, and safer to operate, while new models of electricity production, transmission, storage, and consumption are being promoted. All of this facilitates the deep integration of energy systems with information systems and even commercial systems. An energy system will no longer be a simple standalone energy network; it will also be a critical infrastructure platform that co-exists with other networks, such as transportation networks, carbon footprint networks, and information networks, which are collaboratively controlled across industries. The method and scope of energy cloud management collaboration are not limited to individual devices, systems, or industries in the energy system.

Technological advancement and energy transformation mutually catalyze each other and have profound implications for the future of the energy sector. Only by recognizing the major trends can we better address the challenges of the future, and only by focusing on today's realities can we seize the opportunities of the times. In the emerging digital power era, we must all work together to build new alliances, explore new ways to collaborate across value chains and ecosystems, and contribute to energy innovation and development worldwide. Only together can we drive energy transformation, build low-carbon, electrified, and intelligent energy systems, and make the world a greener, better place for all.

Intelligent Automotive Solution 2030

Foreword

ICT enables an intelligent automotive industry and helps carmakers build better vehicles

The beginning of the 2020s has marked a rapid shift towards more intelligent electric vehicles within the automotive industry. A new era for the automotive industry is just on the horizon, and we will soon see these profound changes affect our daily lives.

There is an industry-wide consensus that vehicles will be more electric and intelligent

Carmakers are embracing this trend by actively adjusting their business strategies and ramping up R&D investment. Many have made transformation a core part of their future development strategy and have already begun to take concrete steps in this area.

Technology and user experience are driving rapid growth in the new energy vehicle (NEV) market

In 2020, NEV companies were able to buck the trend and grow, helping them establish a mid- to long-term competitive edge over other players in the automotive market. This can be attributed to two key drivers – technology and user experience – that NEV companies have been able to leverage

through heavy investment into R&D and closer analyses of user requirements. According to China Association of Automobile Manufacturers, in 2020, 1,367,000 NEVs were sold in China, marking a remarkable 10.9% increase from the previous year, despite the overall 1.9% decline in China's passenger vehicle market due to the pandemic.

Data and software are turning traditional vehicles into intelligent and software-defined vehicles

Data and software support faster iteration of vehicle functions, helping vehicles deliver experiences beyond consumer expectations. New, ever-evolving functions and services are also providing stable revenue streams for carmakers, pushing the industry to move away from product-centered operations towards user-centered operations.

What it means to "build better vehicles" is changing dramatically for carmakers

Users are increasingly focused on intelligent and electric features, rather than the traditional mechanical aspects of a vehicle. To make great intelligent electric vehicles, carmakers need to use digital platforms to achieve faster vehicle development and improve efficiency at lower costs. They also need to enable fast software iteration, ensure vehicle safety and trustworthiness, and address other challenges that consumers might face. These are what it means to "build a better vehicle" in the era of intelligent electric vehicles.

In the future, the market for new intelligent connected vehicle components will be worth trillions of dollars. Huawei hopes to bring its decades of ICT expertise to the automotive industry as a provider of new components for intelligent connected vehicles. As vehicles become more electric and intelligent, Huawei wants to help carmakers build better vehicles.

Macro trends: Cross-sector collaboration for shared success

The automotive industry is changing rapidly, and so are its products and industry landscape. As ICT is integrated into the automotive industry at an increasing speed, cross-industry collaboration becomes increasingly important. Huawei is committed to researching basic ICT technologies and bringing its ICT expertise to the automotive industry through partnerships with carmakers.

Faster industry upgrade brings a bright future for electric and intelligent vehicles

Favorable policies create new opportunities for electric and intelligent vehicles

Carbon neutrality has become a globally recognized mission. Many countries are racing to become carbon neutral. The transportation industry plays a key role in this process as it presents huge opportunities to conserve energy and cut emissions. This in turn makes the NEV industry very promising.

The EU has tightened carbon emissions standards and increased penalties, significantly driving up compliance costs for traditional fossil fuel vehicles. As part of its broader efforts to stimulate the NEV market, the EU also now offers purchasing incentives and tax benefits for those who buy electric vehicles. Similarly, the US has released a 2030 plan for electric vehicles, and is currently accelerating the deployment of

charging infrastructure.

In China, low-carbon vehicles are playing a key role in the government's carbon peak and carbon neutrality ambitions. The transportation industry – the automotive industry in particular – is setting out a roadmap for reaching its carbon peak and carbon neutrality goals. The Chinese government has also tightened its dual-credit policy, which assesses carmakers according to their efforts to cut fuel consumption and produce NEVs. This policy has yielded positive results and stimulated significant investment into NEVs. China's big push for the electrification of public transportation is also driving NEV sales.

Many governments are fostering positive policies and regulatory environments for their intelligent automotive industry through independent research and policy guidelines. China, for example, has introduced many policies on intelligent connected vehicles in recent years, including specifications for vehicle quality and safety, functional safety, cyber security, data security, and road tests, which has facilitated the construction of demonstration zones for new products. Looking ahead, standards and regulation for intelligent vehicles will continue to develop in line with new technological advances. This kind of regulatory environment is critical to commercializing mature technologies and promoting sustainable growth in the intelligent automotive industry.

China's New Infrastructure Plan has laid out requirements for enhancing top-level designs for information, convergence, and innovation infrastructure, while improving underlying infrastructure like 5G, big data centers, artificial intelligence (AI), and charging infrastructure for NEVs. China is also promoting a new development model – the dual circulation model – which aims to create a powerful domestic market while promoting consumption and more room for investment. This new model will allow Chinese carmakers to compete globally and increase internal circulation through more domestic demand.

ICT accelerates upgrades in the intelligent automotive industry

The new vehicle lifecycle sees core functions continuously upgraded. This means vehicles need more sophisticated electrical/electronic (E/E) architecture, system on chip (SoC) computing power, software and data use, and cyber security. This is changing the automotive industry at a fundamental level, as it embraces more advanced ICT technologies and solutions.

Moore's law has long been the golden rule for the semiconductor industry, profoundly influencing the development of PCs, digitalization, Internet, and more for over 50 years. The next 10 years will continue to see Moore's law governing the development of computing power required for intelligent vehicles. Huawei predicts that that a vehicle will require more than 5,000 trillion operations per second (TOPS) of computing power by 2030 to enable the further advancement of telematics applications like intelligent driving, intelligent cockpits, and extended reality (e.g., augmented reality [AR] and virtual reality [VR]).

5G (including 5.5G) promises high bandwidth, low latency, and ultra reliability, making it possible to meet the essential connectivity requirements of intelligent vehicles. By 2030, intelligent digital platforms, powered by emerging technologies like 5G, cloud, big data, Internet of things (IoT), and optical technologies, will connect the physical and digital worlds for vehicles. This will greatly drive innovation and upgrade in the automotive industry.

Changes in supply: Vehicle sales will surpass those of fossil fuel vehicles by 2030

Advanced battery technologies and increasing electric vehicle supply will further drive down the cost of electric vehicles, giving electric vehicles a huge price advantage over traditional fossil fuel vehicles. The boost that China's New Infrastructure Plan has given to advancements in charging and battery swap technologies is making electric vehicle charging as easy as filling up a traditional gas tank.

Carmakers are rapidly expanding their presence in the NEV market. Jaguar will go fully electric by 2025, and other big players like Volvo, Bentley, BYD, and Geely have pledged to switch to fully electric vehicles by 2030. Volkswagen, BMW, and a few others expect half of their global sales to be fully electric by 2030.

According to China's State Information Center, Chinese car brands have managed to keep hold of a solid 35% of China's vehicle market over the past five years. The overall push to increase the quality of domestic manufacturing has also driven

Chinese carmakers to develop high-end brands that prioritize connected, autonomous, shared, and electric (CASE) vehicles. This has resulted in the launch of pure electric platforms, and a surge in in-house development and partnerships that integrate ICT to make vehicles more intelligent.

By 2030, the global NEV market, especially the Chinese NEV market, is expected to grow significantly, with global NEV sales outnumbering those of fossil fuel vehicles.

Changes in demand: Stimulating the intelligent electric vehicle market

User demand for intelligent electric vehicles is on the rise. As electric vehicles will soon cost much less and become more convenient, China is set to benefit greatly from its large consumer market. This gives the nation a base from which it can further develop intelligent electric vehicles. The Chinese market has long been less saturated than more developed markets, and Chinese consumers are also proving to be more receptive to new developments like electric vehicles and intelligent driving.

Due to China's constantly changing demographics, income structure, and consumer purchasing behavior, its consumption structure is also changing at a rapid pace. China will soon become a middle- or high-income economy, and consumption is expected to continue to grow as per capita GDP and household disposable income increases. Consumer distribution is also changing, which means demand is becoming more diversified. China's Post-2000 generation, a Gen Z analogue of true Internet natives, are big fans of new technology and individual expression. They represent a huge engine for growth in domestic consumption. A sliver economy – all economic activities linked to China's older age groups – is also emerging. The two- and three- child policy is also reshaping consumer demand.

Such changes in consumption structure will also affect car consumption, both directly and indirectly pushing China's vehicle market from relying on traditional models to digital models, from commodity-oriented to experience-driven, and from valuing commonality to valuing individuality.

Changes in product attributes: Reshaping the automotive value chain

Key vehicle differentiators: From powertrain and chassis systems to intelligence

As vehicles' power systems become electric, their powertrain and chassis systems will gradually become more standardized. This makes intelligent cockpits, intelligent driving, and other intelligent functions the key differentiators of vehicles. The intelligence level of vehicle cockpits and driving systems will become key factors in users' purchasing decisions. Over-the-air (OTA) updates will be used to deliver superior user experiences and further increase users' uptake of intelligent functions.

Such shifts also present an opportunity for carmakers to expand their hold on the vehicle market. As the laws, regulations, and policies on intelligent vehicles continue to improve and intelligent driving technologies become mature, autonomous driving will enter large-scale commercial use via robotaxis and closed or semi-closed low-speed driving scenarios by 2030 before gradually being implemented in passenger vehicles. In addition, human-machine interaction technologies will continue to advance and the intelligent cockpit application ecosystem will continue to improve, making vehicles an intelligent mobile "third space" outside of home and workplace.

A wider industry: From automotive products to all-scenario mobility services

5G (including 5.5G), IoT, AI, edge computing, and low-carbon technologies are still rapidly

developing, converging, and iterating. They are accelerating the automotive industry's CASE transition. How carmakers will commercialize intelligent vehicles in specific scenarios is becoming increasingly clear.

As intelligent driving technologies continue to improve in different market segments, scenario-specific autonomous driving applications will be become more widely adopted. New forms of autonomous vehicles will emerge for specific scenarios, and the connection between transportation tools across different scenarios will become more seamless. Autonomous mobility services will appear in every link of people's travel. This will fundamentally change how people travel, how people interact with transportation tools, and how transportation tools interact with each other, greatly improving "mobility-as-a-service".

People's basic mobility needs have gradually changed from owning different transportation tools for different scenarios to using integrated mobility solutions for complex mobility scenarios. Many third-party application developers are mobilizing industry resources to develop new service applications for different scenarios, with the purpose of seamlessly connecting different transportation tools and providing end-to-end intelligent mobility services for users. These mobility solutions and services are providing new revenue streams for the automotive industry.

New profit models: From hardware to software and services

As key differentiators for vehicles change and the automotive industry's reach expands, individual intelligent vehicles will become platforms for continuous value creation. This will reshape the automotive industry's standard business model and value distribution pattern.

Carmakers have long profited from one-off deals – multiplying the unit price by the total number of vehicles or hardware units they sell. Software-defined vehicles will turn software and services into new revenue streams for carmakers, and their profits will be determined by software fees

and car parc. Moving forward, data and software will support ongoing, OTA iteration of vehicle functions by allowing carmakers to remotely repair and upgrade products, and improve user experience. This will give users more flexible and operable service models, driving a shift in the industry from product-centered operations to user-centered operations. These revenue streams will also be more stable for carmakers.

The automotive industry as a whole will focus on the new operation and charging model created by intelligent driving as it greatly expands profits for carmakers. In addition, software-defined vehicles will reshape the value chain, and creating more opportunities to unlock value. This will attract more third-party developers and innovators to invest in the intelligent automotive industry, which will help improve the intelligent connected vehicle ecosystem and build a positive cycle of value creation.

Cross-sector collaboration will define the new industry landscape

Carmakers and tech companies work together to maximize their strengths

Intelligent vehicles are the product of multiple industries, built on the integration of core digital technologies (e.g., ICT, software, big data, and AI) and traditional mechanical technologies. Emerging carmakers are the frontrunners in the first phase of the CASE journey. However, other carmakers are also joining the trend and beginning to improve their core competencies in software, electronics, and big data.

Auto underbody solutions are slowly standardized, becoming a shared platform on which other industries can grow. Tech companies, from consumer electronics manufacturers to Internet companies, are taking advantage of this trend and expanding into the automotive industry either on their own or through alliances. These companies have large amounts of capital, strong experience in ICT, significant technological innovation capabilities, and huge brand recognition. Their entry into the industry is driving rapid development in intelligent connected vehicles and pushing the automotive industry towards CASE 2.0.

The automotive industry has been around for over 100 years, and carmakers have emerged as leaders in manufacturing, quality control, safety, and reliability capabilities. Tech companies, on the other hand, have amassed extensive experience and advantages in intelligent technology applications, such as AI algorithms and big data. Software-defined vehicles will significantly change how companies capture value, serve their customers, and build their talent mix. To meet increasing user requirements, all companies along the value chain should become more agile and adapt to this new environment.

Software and hardware will decouple and general platformization and standardization will continue. This means the only way forward will be to foster a more open supply chain system and adopt more flexible vehicle models. Carmakers and tech companies will need to maximize their respective strengths, while also relying on cross-industry partnerships to find new and innovative ways to achieve business and social value.

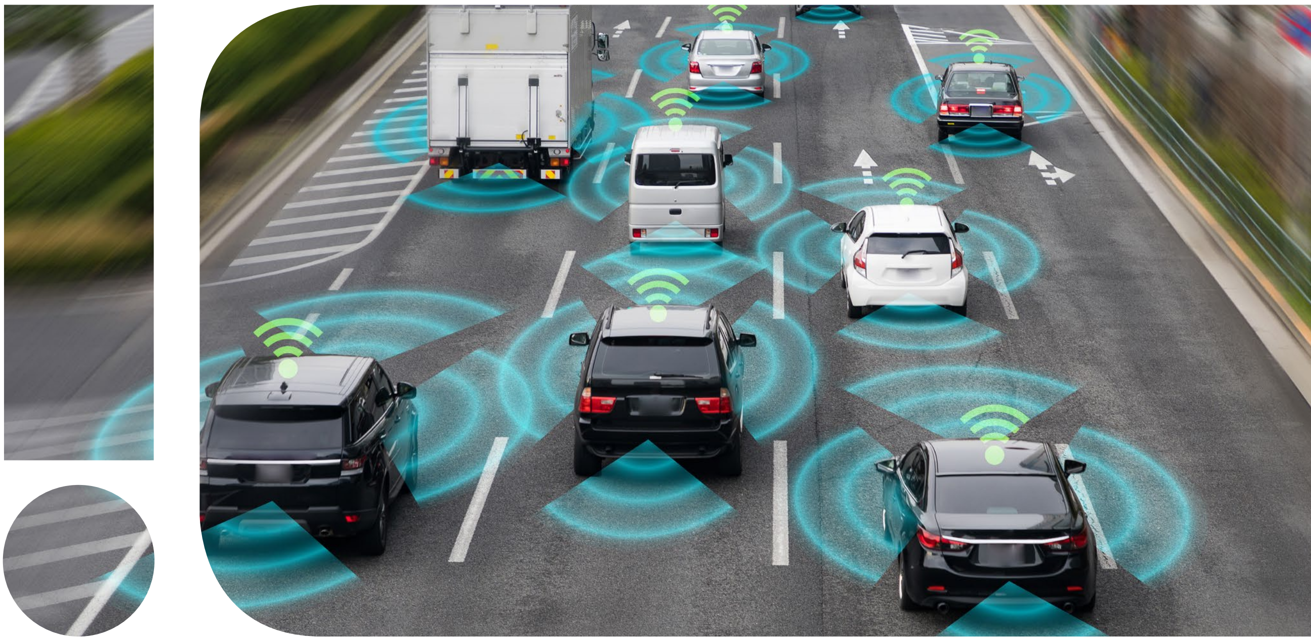
ICT is key to better travel experiences in the growing mobility sector

As the automotive industry's reach expands, demand for transportation and mobility services in new market segments will continue to increase. This will drive exponential growth in the forms and quantity of vehicles, as well as their related infrastructure. More and more traditional vendors are announcing transition plans to become "mobility solution providers", intending to tap into this huge new market. Different players will have different roles to play.

Mobility solution providers provide end-to-end solutions to meet user requirements across different mobility scenarios and control user traffic. Operators in closed scenarios understand operation requirements, customize the form factors of vehicles, and deploy infrastructure in closed scenarios. Carmakers use their existing manufacturing platforms and supply chains to manufacture these vehicles. Tech companies provide solutions like intelligent software and hardware, intelligent driving, and cockpit connectivity and control. Third-party ecosystem developers provide massive numbers of apps and deliver seamless travel experiences to users.

As the forms and quantity of vehicles continue to grow alongside their related infrastructure, we need to connect the vehicles and infrastructure in specific scenarios. Scheduling and connection across different scenarios, data sharing between different vehicles, and scenario-based smart service applications will need to be hosted on a cloud-based "brain". ICT will play a key role in connecting these disparate points and enabling scenario-based digital service sharing.

Intelligent driving powered by ICT will also make travel safer and more efficient. Long-distance driving will become easier and commercial use will become less costly and more efficient. Intelligent connectivity and spaces will make vehicles the real "third space", bringing more enjoyable experiences to users, so that users are willing to spend more time in their vehicles. By integrating people, vehicles, and home scenarios, carmakers aim to improve the in-vehicle user service ecosystem, provide intelligent services, and bring digital to every vehicle for a fully connected, intelligent world.



New scenarios: Bringing digital to every vehicle

As digital technologies are widely adopted and carbon neutrality has become a globally recognized mission, it is becoming an obvious trend that vehicles will become more electric and intelligent. Bringing digital to every vehicle will empower intelligent driving, intelligent spaces, intelligent services, and intelligent operations. This will allow for safer and more efficient transportation, greener and more convenient travel, more fun and intelligent lifestyles, and more efficient and lower-carbon operations.

Intelligent driving: Safer, smoother, and more efficient travel experiences

Intelligent driving can be divided into fully autonomous driving, highly automated driving, and advanced driver assistance systems (ADAS). Intelligent driving scenarios include closed roads, public roads, and other driving scenarios. Autonomous driving will disrupt the mobility industry and fundamentally change people's lives. The earliest applications have been on closed roads like high-speed roads and campuses, and will gradually come to cover public roads like those in urban areas. According to Huawei, autonomous vehicles sold in 2030 will account for 20% of new vehicles sold in China and 10% globally.

In 2030, robotaxi services provided by self-driving fleets are expected to cut labor costs and provide 24/7 mobility services that are more flexible and affordable.

Intelligent driving technologies will be integrated into existing modes of transport to provide safe,

efficient, and affordable mobility service solutions that meet different travel needs and deliver the best possible experience. Mobility resources will be centrally managed and data will be shared in real time, making it possible to build an end-to-end, point-to-point, and door-to-door mobility network. This will in turn help maximize the use of all available mobility resources.

When a user plans a trip, a cloud-based brain weighs up all the possibilities and, based on real-time awareness of the traffic situation, offers the optimal route and mode of transport. Diverse mobility resources will allow users to enjoy efficient, green, and safe travel while maintaining a dynamic balance in urban transportation capacity, contributing to sustainable urban development.

Intelligent spaces: From a flexible mobile space to an intelligent living space that integrates the virtual and physical worlds

Vehicles are no longer just a tool for transport. Their relationships with people and with their surroundings are changing dramatically.

Advanced intelligent driving technology will free commuters to enjoy work, study, entertainment, and much more within their vehicles. When vehicles serve as mobile offices – or even mobile living rooms – it will change how people think about their daily commute.

Powered by human-machine interaction, in-vehicle optical technologies, and immersive AR/VR technologies, intelligent cockpits will become multi-functional units. People will find themselves spending more time in their vehicles even when they do not want to go anywhere. For example, it will not be uncommon to see people sitting in parked vehicles watching movies.

The way we interact with our vehicles is about to experience three major changes. First, a cockpit will no longer be a combination of a steering wheel, an instrument panel, and a screen; it will integrate the virtual and physical worlds. Features such as voice control, facial recognition, and gesture interaction will make interactions simpler, more natural, and more efficient. What's more, it will not be long until brain-computer interfaces start seeing commercial application. Second, our vehicles will no longer passively await our instructions; they will intelligently

anticipate our needs. Technologies such as AI, biometric recognition, emotion perception, and vital sign monitoring will allow vehicles to better understand drivers' behavior, habits, and thinking, and become close partners that truly understand drivers' needs. Third, in-vehicle optical technologies will offer a richer spatial optical experience, and AR/VR technologies will further transcend the barriers of time and space. Such an immersive experience will drive broader and richer vehicle applications in both mobile and static scenarios.

An intelligent vehicle will become a truly intelligent space that integrates the physical and virtual worlds.

- **Driving:** The combination of in-vehicle sensors and wearable devices can accurately monitor drivers' health indicators, recognize fatigue, and send timely reminders to ensure safe driving.

- **Entertainment:** Passengers will be able to have true-to-life experiences of concerts and sports events without having to be there in person. A cinema will no longer be the best place to watch movies. AR technology will make gaming more immersive. Vehicles can become a personal entertainment space, a private cinema, an outdoor cinema, and a preferred place to play games with friends.

- **Mobile office:** Seats can be adjusted and rotated and windows can be used as large projection screens. A conferencing stream on a smartphone can be easily transferred to the vehicle, and the shielding function of the vehicle's sound zone ensures the privacy of the conference. Vehicles will become mobile offices, allowing professionals to get work done on the way to the airport, a restaurant, or their homes.

- **Social networking:** Drivers will not miss the beautiful scenery outside the window. Cameras mounted on the outside of the vehicle can be used to record, edit, and share beautiful moments. Getting stuck in a traffic jam no longer has to be boring. You can watch movies, play games, and make friends with nearby drivers using the head unit. AR/VR brings your friends within easy reach. With separate sound zones created within your vehicle, you will even be able to keep conversations private from other people in the vehicle.

Intelligent services: More intelligent scenario-based services

Digitalization is reshaping the world, and as a result, consumption patterns will change dramatically over the next decade. Services in all kinds of industries will be available online and become more customized, personalized, responsive, and scenario-based. As digital technologies are deeply integrated with vehicles, services will be intelligently and rapidly pushed to users based on the scenario they are in. This will be achieved in three ways.

First, as vehicles become more intelligent, interactions between users and vehicles will inform the services to be provided. Intelligent algorithms can identify, analyze, and understand users' interactions, predict their behavior based on their basic information and historical preferences, and provide the right services. Intelligent vehicles will continuously improve their understanding of users, in order to deliver better services.

Second, intelligent vehicles will make it possible to efficiently and accurately identify user needs in real time. By identifying and analyzing vehicle data, location information, and surrounding environments, intelligent systems can determine what scenarios users are in, proactively predict user needs, and provide the right services.

Third, brand-new, interconnected operating systems (OSs) can create more service scenarios, giving rise to an application ecosystem that is based on new modes of interaction. As the digital world approaches and the digital economy develops, a richer digital ecosystem is emerging to support all scenarios. In the connected world, more services will be provided by intelligent vehicles. The scenario-driven functions and services offered by connected vehicles will become increasingly intelligent, efficient, and convenient.

We can even imagine a situation in which a group of people would want a pizza while driving across town. Mobility-as-a-service providers will provide a shared vehicle that perfectly matches the passengers' preferences, select a high-rated pizzeria located along the planned route, and order the pizza in advance. The restaurant will then prepare the food, which will be collected by a drone. When the vehicle arrives at the designated handover location, its sunroof will automatically open and the drone will lower the food inside. This is a level of service that can only be achieved when every part of the process is seamlessly connected.

Intelligent operations: Autonomous driving is expected to be applied in commercial vehicles first, boosting the productivity of intelligent operations

Commercial vehicles are important tools for transport and the functioning of modern society. Their evolution into intelligent autonomous vehicles can help achieve the goal of carbon neutrality and boost work and operation efficiency while contributing to a more mature intelligent vehicle ecosystem. By 2030, commercial autonomous vehicles will be used on trunk lines and public roads in addition to closed areas and dedicated roads. This will make intelligent operations a reality and greatly increase productivity.

When autonomous vehicles are operated in a closed area, it is possible to enumerate all the scenarios in which a vehicle might find itself and foresee potential emergencies. For this reason, commercial autonomous vehicles will find their first large-scale commercial applications to be in closed areas like ports, mines, farms, campuses, airports, and closed scenic spots. In these areas, intelligent commercial vehicle technologies will not only be applied to transport vehicles; they will also be integrated with operations management systems to build

unmanned manufacturing systems where autonomous driving applications have been integrated into the core production systems that support transportation and distribution. This means intelligent vehicle technologies will be commercially used on a large scale.

By 2030, Vehicle-Road-Cloud collaboration solutions will make it possible for multiple autonomous vehicles to collaborate in closed areas, which means autonomous driving has the potential to be commercialized in vertical industries. Service capabilities, like comprehensive environment perception, global resource scheduling, dynamic service mapping, multi-vehicle cooperative driving, lane-level route planning, coordinated signal control, and service simulation testing, will further streamline service processes, make the collaboration between multiple autonomous vehicles a reality, and increase scenario-based operation and transportation efficiency. All of these will help cut costs and boost productivity.

Cloud scheduling and high-definition (HD) maps will be critical to the service management and scheduling of autonomous vehicles. When intelligent commercial vehicles are used in closed areas, operation managers will use the vehicle cloud control management system to schedule and monitor those vehicles, and support service and safety management using HD maps. In a port, for example, the operation control platform of an intelligent horizontal transport system will be connected to the terminal operating system (TOS), which means the scheduling of autonomous container trucks will be fully integrated into the automatic port scheduling system. This level of integration will help fully automate port operations. In addition, the dynamic layers of HD maps can be updated to reflect changes in the location and status of quay cranes, yard cranes, locking/unlocking stations, and container yards.

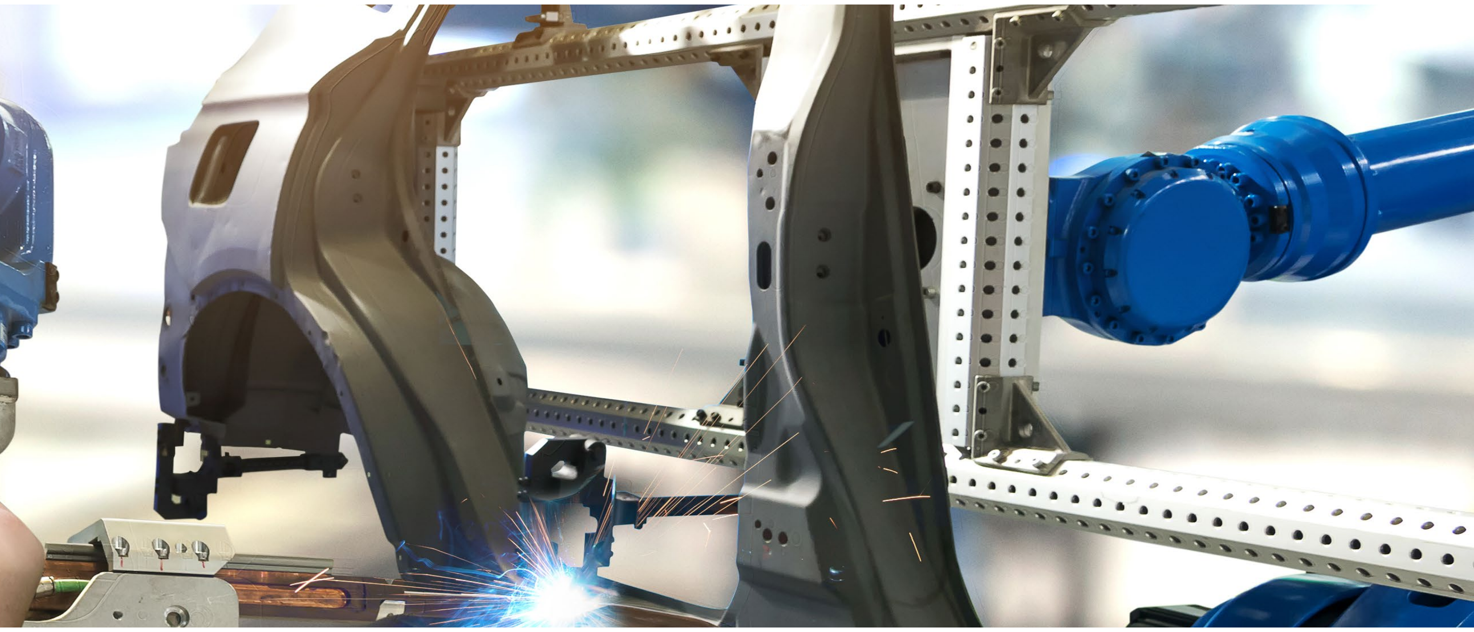
As road infrastructure is upgraded over the next decade, on trunk lines, commercial vehicles will gradually shift from assisted driving systems to fully autonomous driving. As electric vehicles are widely used in urban short-distance transportation, and roadside network infrastructure becomes more intelligent, the penetration rate of intelligent commercial vehicles is expected to rise sharply on more complex public roads, including urban roads. Building on the basic capabilities of autonomous vehicles and their potential commercial application, carmakers can work with ecosystem partners to build more viable intelligent driving applications to overcome the challenging scenarios faced by commercial vehicles.

Technology projection: New components will drive sustained innovation in intelligent vehicles

Evolving to a central computing and communications architecture for software-defined vehicles

Today, most vehicles still use a decentralized E/E architecture. Under this architecture, each separate function has an independent controller, so a vehicle has almost 100 controllers and over three kilometers of wiring. This makes vehicles costly and heavy, and difficult to automate vehicle assembly. In addition, electronic control units (ECUs) are often developed by different vendors, meaning they have inconsistent standards which carmakers struggle to develop new functions or perform OTA updates. In the future, intelligent connected vehicles will be even more complex. The volume of data collected by vehicle-mounted sensors will dramatically increase. This will raise the bar for real-time data transfer and data processing. These trends mean that vehicles' E/E architecture must evolve.

With the rapid development of digital and intelligent technologies, vehicle functions are becoming more integrated and centralized. There is a widespread understanding in the automotive industry that a decentralized architecture is no longer viable, and that it must evolve first into cross-domain architecture, and eventually into a single central computing platform. Vehicle



functions will become applications loaded onto a central processor so that they can share a single set of sensors and actuators. Components will gradually become more standardized. This will help reduce the cost and complexity of new function development in the long run. Meanwhile, domain controllers will evolve through the addition of new software features. By 2030, the E/E architecture of vehicles will be a computing and communications architecture that consists of a central computing platform, zonal control, and high-bandwidth in-vehicle communications.

High-performance central computing platforms power software-defined vehicles

Central computing plus zonal control, either in a hub-and-spokes or ring model, offers architecture stability and functional scalability. Within this architecture, new external components can easily be added from the gateway of the nearest zone, and with pluggable hardware, computing capacity can be upgraded as necessary, enabling simple iteration and upgrades of application software on the central computing platform.

The mobility scenarios are complex and subject to frequent change. New functions such as intelligent cockpits, intelligent driving, and vehicle control are constantly being developed. A high-performance central computing platform is required to support them. This platform can perform several thousand TOPS. It must be based on a powerful SoC with a vehicle-specific operating system, middleware, toolchain, and a centralized architecture. Such platforms will offer a stable architecture for software-defined vehicles, while still allowing room for smooth evolution. The chassis, powertrain, cockpit, and intelligent driving system will each place different demands on the central computing platform in terms of security, latency, dynamic response, and the supporting software ecosystem. A high-performance in-vehicle central computing platform will use hardware virtualization and a central functional safety framework, as well as AI algorithms to deliver the necessary levels of security and ensure that hardware resources are available as required for each domain. The technologies required include:

High capacity processors: SoC will deliver the thousands of TOPS of computing power required by the vehicle, including the chassis, powertrain, cockpit, and intelligent driving system. Key technologies needed to build SoCs include computing in memory and trustworthiness and functional safety islands.

High-speed concurrent processing for guaranteed low latency: High bandwidth is only part of low latency; an even more crucial factor is the capacity to process data in real time. High-speed concurrent processing enables central systems to simultaneously receive data from multiple sources. It prevents data surges, and will enable the vehicle to handle the data generated by an ever-growing number of apps and the ever-increasing demands on the system.

Hypervisor secure partitioning of hardware: The Hypervisor allows one physical server to function as multiple virtual servers and delivers customized functional safety for the different domains of a vehicle. AI engines monitor and forecast the workloads for different virtual partitions and dynamically schedule hardware resources, thereby achieving secure partitioning and load balancing.

Inter-app freedom from interference (FFI): The Hypervisor partitioning function delivers a secure silo of resources for the applications, communications mechanisms, OS, and hardware accelerators. Within the processors, a dedicated safety island provides a safety system that reaches the standards of 3-Level Safety Monitoring. The safety island's intelligent fail safe and fail operational functions enable coordination of the safety responses with other vehicle functions.

Building on a powerful central computing platform, the software-defined vehicle sector will concentrate efforts on agile development and real-time release of new functions to deliver the diverse experiences that users will demand of mobility.

In-vehicle, high-bandwidth, and multi-protocol networks for software-defined vehicles

The centralization of vehicle functions will drive substantial changes in the access approach and method used for communications. Vehicles will be divided into a number of zones, each with its own gateway. Zones will be defined by function, physical position, criticality, and safety. Sensors and actuators will be connected to the nearest access points to transfer data to the backbone network and then to the central computing platform. This approach will reduce the total amount of wiring required, and support the development of new functions. Sensors will no longer be limited to a single function, and actuators will not be bound to a directly connected controller.

By 2030, multiple access protocols will be in concurrent use. Local Interconnect Network (LIN), Single Edge Nibble Transmission (SENT), and Peripheral Sensor Interface (PSI5), though slow, will still be used because of their cost advantages. But ultra-high-definition (UHD) cameras, 4D imaging radars, and high-resolution lidars will require much more bandwidth. According to Huawei, in-vehicle network transmission speed per link will exceed 100 Gbps by 2030. Vehicle Ethernet will become standard, and optical technologies will be widely deployed in vehicles because of their high bandwidth, light weight, insensitivity to electromagnetic interference, and low cost.

Conventional communications technologies are predominantly signal-oriented, using protocols such as Controller Area Network (CAN) and LIN. This approach deeply integrates communications with vehicle component deployment and routing, creating a problem: A change of the transmit/receive nodes will lead to changes of all nodes along the route.

Ethernet communications are service-oriented and can effectively address this problem. When the transmit/receive nodes are changed, no other nodes on the route will be changed. This will:

- Decouple communications from vehicle component deployment and routing, making it easier to scale up.
- Make interfaces standardized and contractual.
- Achieve interconnectivity of in-vehicle services.

Once these technologies are realized, a point-to-point backbone network for software-defined vehicles can be created. The technologies that would be required for this include:

High bandwidth copper communications: Signal attenuation is significant in copper cable over even short distances. Enhanced coding and algorithms will be required for intelligent power distribution and high-speed, high-bandwidth Ethernet transmission (10 Gbps to 25 Gbps). This will provide the high bandwidth required by in-vehicle applications for backbone networks.

In-vehicle fiber communications: For bandwidths over 25 Gbps, copper is no longer an option, because of cost, engineering, and electromagnetic compatibility (EMC) challenges. This makes fiber an excellent solution. Fiber is cost-effective, light-weight, and is not affected by EMC issues. If solutions can be found for vehicle-related problems around temperature, vibration, and service life, optical fiber communications will be widely used in in-vehicle applications and support the evolution to higher-bandwidth communications.

Deterministic latency: Real-time communications protocol stacks as well as time-sensitive network (TSN) protocol suites at the transport layer will need to ensure end-to-end deterministic latency at the microsecond level for vehicles. Transmission policies can be designed for specific service scenarios to meet the needs of different communications functions.

New wireless communication technologies for high-quality in-vehicle connectivity

By 2030, in-vehicle wireless communications will remove all barriers to connection within the vehicle. Any component will be able to connect using sliced wireless capabilities, so that new vehicle applications can call on them as needed. A new air interface will be required to deliver extremely low latencies of less than 20 microseconds for unidirectional transmission, five nines reliability, synchronization accuracy within one microsecond, up to hundreds of connections

and concurrent service provisioning, plus end-to-end cyber security. This is the level of quality required for vehicle connectivity. Service-specific resource management mechanisms will support in-vehicle wireless connections. Wireless slicing will make many things possible, like lossless audio streaming, UHD video apps, and ultra-low-latency interactive games. By taking the collaboration of multiple information domains to the next level, the interior of a vehicle will become an infotainment center offering immersive sound, video, images, and even light applications and sensations.

In-vehicle wireless communications technologies will transform the in-vehicle networking and enable simple upgrades of various vehicle modules. The use of wireless in place of wired connections will address design, production, assembly, and maintenance challenges created by vehicle wiring, and put an end to the highly-coupled architecture that wired connections create. In its place will be a platform + modular communications architecture, which can be replicated across many different models of vehicle. The flexibility of wireless communications allows for a range of different architectures, providing standardized wireless access interfaces. When vehicle-mounted devices become modular, standardized, and plug-and-play, the costs of vehicle development will fall, and smooth and ongoing evolution of the foundational platform will be supported.

Decoupled, service-oriented architectures for software-defined vehicles

Vehicles are now digital, intelligent products. User values, preferences, and needs no longer require vehicles to be a tool for transport; now, as with the phone, users want personalized experiences. Smart technologies, personalized features, and user experience are now the key factors guiding consumers' vehicle choices. At the same time, hardware and the associated technologies that go into vehicles are becoming less easily distinguishable, and carmakers are looking to software and algorithms to create competitive advantages and deliver more value. All industry players are now pursuing software-defined vehicles.

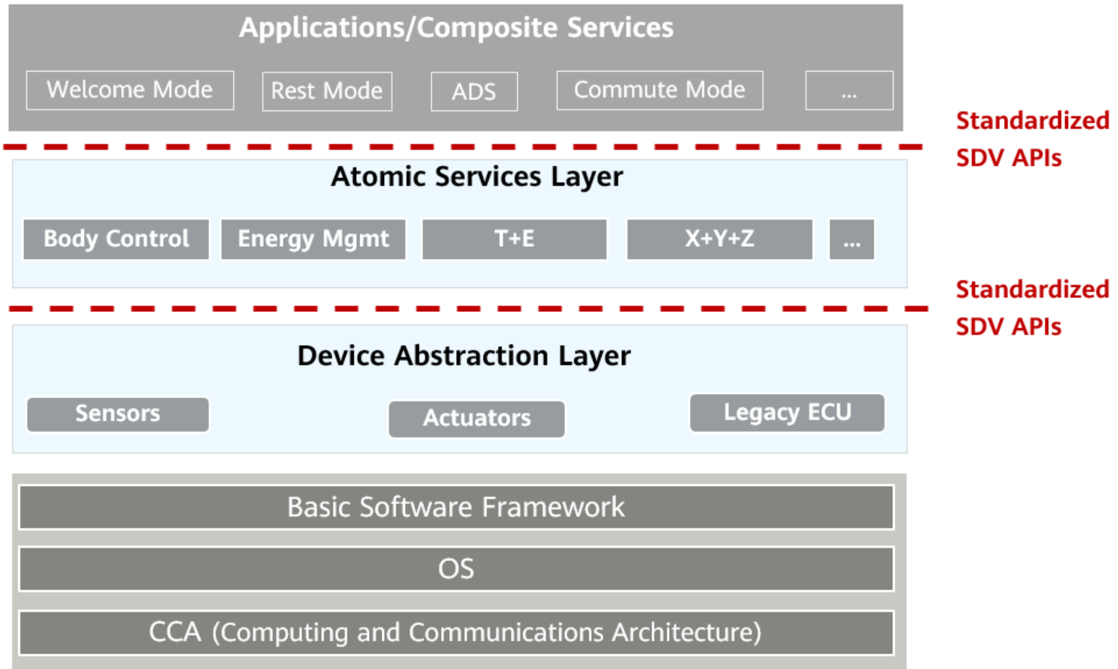
Being software-defined means that software is a key feature in a vehicle's concept, development, testing, sales, and after-sales services. It also means that the entire process will be constantly refined, with refreshed experiences and continuous value creation. A key feature of these vehicles is the decoupling of the software from the hardware. In terms of physical delivery, that means that the hardware and software are delivered separately. In essence, vehicle functionality can be expanded, software can be replicable and upgraded, and hardware can be swapped out or scaled up.

OTA updates can keep software at peak performance; plug-and-play components can be freely added to expand functionality. Flexible, scalable software-defined vehicle platforms will be used to help intelligent vehicles meet the challenges of complex use cases and growing demands on functionality. A service-oriented architecture (SOA), with decoupled layers of software, has also been recognized as the best option for general software architecture. To realize this architecture, the system will need to add a device abstraction layer and a layer of atomic services. (Fig 1 Service-oriented hierarchical software architecture)

Atomic services provide basic service capabilities. By masking differences in hardware, they enable upper-level applications to be replicable and portable across different vehicle models. The device abstraction layer normalizes the capabilities of sensors and actuators, so that software can be decoupled from the hardware, enabling plug-and-play replacement and upgrades of hardware modules.

Services will be decoupled from system design to create basic service units. Each separate hardware component will be abstracted into a standardized service component, and each service

Fig 1 Service-oriented hierarchical software architecture



component will provide one atomized function. These can be called on recursively and combined to produce complex functions, thus reusing the software as much as possible. Ecosystem partners can develop vehicle applications using platform components and standardized APIs that will then be managed by the vehicle platform. The platform will carry out app authentication, granting of access privileges, API call, security checks, and emergency management. Users will be able to choose vehicle apps the same way they do for their mobile phones. This will give users access to new vehicle experiences at very low cost: same vehicle, but a new journey every day! Developers, in turn, will benefit from consumers downloading their apps.

Whole-vehicle coordination and intelligent use of data deliver better safety

As a central computing architecture is more widely adopted, sensors and actuators will be standardized and plug-and-play, reducing the new vehicle development costs and delivering better reliability. Powertrain and chassis systems will be fully wire-controlled, with the control functionality increasingly separated from the mechanical components. The intelligent central computing platform will integrate modules from the engine, electrical system, heat, electronics, and information and communications modules. This will provide well-coordinated management of control functions, thereby delivering higher performance and more precise control. With increased centralized control over the vehicle powertrain and chassis, control latency will also decrease. Controller response time will be 1ms, greatly reducing execution latency. Coordinating control functions across all vehicle systems will continue to increase the value of the intelligent control system over time.

AI is the best way to control the vast streams of data that will come from vehicles and their surroundings. Data can be used to analyze the driving behavior of drivers, and develop a profile of an individual's driving style. If a driver deviates from this profile, the analytics system can issue warnings to the driver or even intervene to prevent accidents. Precision analytics can also help modulate vehicle controls for assisted driving. Using the central processor's built-in understanding of vehicle dynamics, the vehicle can optimize the driver's controls, and work with the driver to deliver a safer driving experience. For example, the vehicle could use its visual sensors and apply its understanding of the vehicle bodywork attitude, to make refined adjustments according to dynamic control principles to deliver improved precision in driving controls. It might notice if the road's surface becomes more slick, if the vehicle's weight is poorly distributed, if the vehicle yaws as it turns, and much more. Intelligent vehicles will sense the road's surface, external obstacles, the behavior patterns of surrounding drivers, and indicator lights, and combine all of this data to recommend corrective actions in the event of an emergency. They can then coordinate the vehicle's systems to significantly improve driving safety.

Intelligent driving: Making autonomous driving a commercial reality faster

Huawei uses sensor fusion technology to deliver superior safety in intelligent driving. Different types of sensors, including lidars, mmWave radars, ultrasonic radars, and cameras, are used to support the fusion and reconstruction of multi-dimensional environmental information.

At the perception layer, Huawei uses HD maps that will cover the entire world and integrated inertial navigation systems to support multi-source high-precision positioning. Coupled with AI chips that provide huge computing power, Huawei's intelligent driving solution can support sensor fusion in urban, highway, and parking scenarios. This further supports autonomous driving in complex road conditions.

Continuous algorithm training for better user experience and autonomous driving

There are still many technical challenges that need to be overcome if we want to make autonomous driving a commercial reality. Given the complexity of corner cases in real-road conditions and how difficult it is to collect long-tail data, perception algorithms, planning and control algorithms, and simulation and training algorithms will be crucial for autonomous driving experience.

If we look at sensor fusion algorithms, there are many technologies that determine a vehicle's ability to perceive and understand its surroundings. These technologies include a vision-based perception framework, lidar point cloud generation and enhancement, lane-level traffic light processing in complex light environments, flashing and fuzzy light source processing, color processing, light source luminance differences, overlapping object recognition, and vehicle interaction prediction.

As perception and prediction both involve uncertainties, the industry needs to further develop core planning and control algorithms that deal with multi-object and multi-stage game-theory decision making, motion planning, human-like decision making and planning models in complex interactions that involve object risks and scenario risks, and extraction and automatic labeling of key scenarios from massive amounts of data.

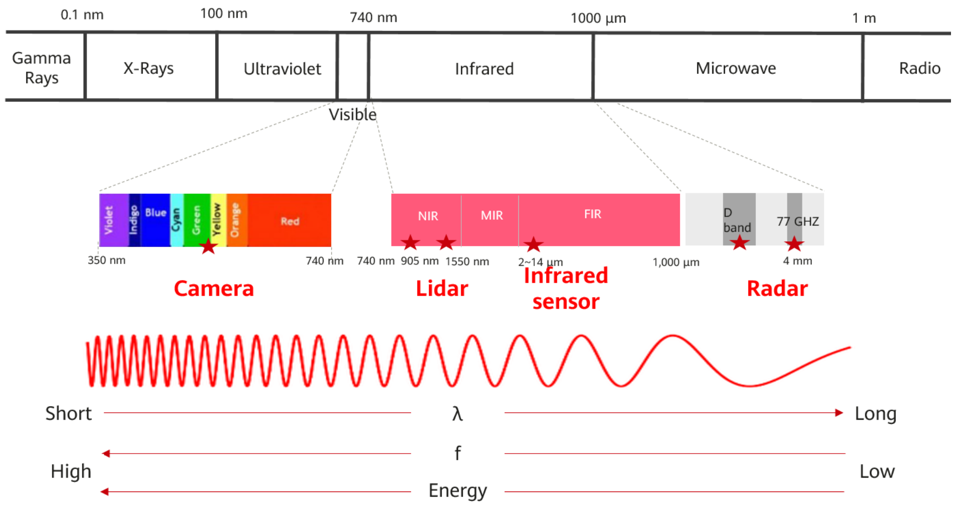
It will also be necessary to build a model for interactions between different traffic participants, create large-scale simulation scenarios, and ultimately form a comprehensive, high-fidelity simulation system.

Developing full-spectrum perception capabilities to make everything sense

As the automotive industry becomes more intelligent, perception systems become increasingly important. One day, they will be a cornerstone of intelligent driving. Ideally, the sensors that enable these systems will reach full coverage and cover all objects, all scenarios, and all weather conditions.

- **All objects:** People, vehicles, obstacles, road

Fig 2 Frequency bands for sensors



facilities, structures, etc.

- **Full coverage:** 360-degree coverage without dead zones.
- **All scenarios:** Highways, urban areas, traffic jams, accident scenes, and construction zones.
- **All weather conditions:** Day, night, rain, snow, fog, strong light, low light, and other harsh environments.

However, the industry still has a long way to go before it can make ideal sensors a reality. To make this happen, perception capabilities need to be built based on all sections of the spectrum. (Fig 2 Frequency bands for sensors)

1) Radars: The shift from the 77 GHz band to D band (110 GHz to 170 GHz) will significantly improve resolution

Radar sensors can perform consistently in all weather conditions because they work on super-long wavelengths. These systems excel in measuring velocity, so they can create unique value in dynamic and static separation and simultaneous localization and mapping (SLAM). However, currently, their poor resolution is limiting their use scenarios.

Radar resolution can be improved by utilizing either ultra-high bandwidth or large-scale antenna arrays. According to existing international standards, the 76 GHz–81 GHz frequency bands are already allocated to automotive radars. This means high-range resolution can be achieved through the higher 4 to 5 GHz spectral band. Angular resolution is determined by antenna arrays, which means the more antennas allocated for transmission and reception, the higher the angular resolution. Current radar systems still use three transmitting antennas and four receiving antennas (a 3T4R antenna array). Huawei recently improved on this by launching a 12T24R antenna array radar. However, the antenna arrays used in wireless communications have already reached 128T128R configurations.

Automotive radar sensors need to be physically small, which means the door-size antennas used in wireless communications systems are unsuitable for automotive applications. These size restrictions and the 77 GHz wavelength mean that antenna arrays for radar sensors would be 48T48R to 64T64R, at maximum. The shift towards higher frequency bands will continue. The D band (110 GHz to 170 GHz)

provides ultra-high bandwidth and has generally not yet been allocated or used for other services. The 140 GHz band is still being researched, but has a relatively suitable atmospheric window, so its propagation is less attenuated. What's more, wavelengths in this band are reduced by half. That means imaging radars that use an ultra-large 128T128R antenna array can be used in a limited space while still delivering high resolutions.

2) Lidar systems are moving from the 905 nm wavelength with time of flight (ToF) to the 1550 nm wavelength with frequency-modulated continuous waves (FMCW). Lidars are being integrated with chips, and high-performance 4D lidars will be widely adopted.

Thanks to their relatively mature components, 905 nm lidars have already been widely adopted and are ready for mass production. From a technical perspective, as the industry moves from analog to digital and from discrete to integrated, the transmitting and receiving components are being arranged in arrays, and core modules will be integrated directly into chips. These trends mean high-performance, low-cost, highly integrated, and highly reliable lidars may be the way forward.

Most light at 1550 nm, a near-infrared band, is blocked by human cornea before it reaches the human retina, which means it does not cause much damage to human eyes. Because of this, the 1550 nm band allows lidars to transmit at higher powers, which can greatly increase coverage. In terms of modulation, FMCW's use for radars can also be applied to lidars. FMCW lidars deliver better performance through high-performance 4D imaging (which can be used to measure velocity), strong anti-jamming capabilities, and higher sensitivity and dynamic range. In addition, FMCW lidars can be mass produced at lower costs, when combined with silicon photonic and optical phased array (OPA) technologies.

However, FMCW technology at the 1550 nm wavelength is far from ready for commercial use and will require concerted efforts from across the entire value chain to develop further. Further exploration of silicon photonic technology in line with Moore's law is one way that can help support FMCW. More complex and discrete optical functions can also be integrated into silicon-based chips, making lidar sensors more integrated, more affordable, and smaller.

3) Cameras will integrate visible light and infrared thermal imaging technologies to work in all weather conditions.

Cameras are a type of passive sensors and are most similar to the human eye. They can sense

surrounding objects through catoptric imaging. As one of the three major types of sensors used in vehicles, cameras are the most critical for identifying elements in a static environment, like traffic lights and road signs.

However, cameras also have drawbacks. The performance and confidence of catoptric imaging suffers at night and in low-light conditions, and heavy rain and snow can impede a camera's line of sight, greatly reducing its visibility. Cameras cannot independently overcome these harsh weather conditions.

However, in the infrared spectrum (in the 2–14 μm band) right next to the visible spectrum, a thermal emission imaging system can be used. Sensors that work on this band have effective night vision, and can detect objects through rain, snow, sand storms, and fog. They even have certain perspectives to further meet the requirements of all weather conditions. Vehicles are now starting to come equipped with infrared thermal imagers that have night vision, but this application still needs a low-cost solution for mass production.

Full convergence: Accelerating innovation in sensor form factors for simplified deployment

As vehicles become more intelligent, the number of sensors per vehicle is increasing dramatically. Vehicles come equipped with many sensors, from 1V and 1R1V to 5R5V and 6R13V3L configurations, and more will be deployed in the future. Here, R refers to radar, V means vision (i.e., camera), and L means lidar.

However, a vehicle's body is limited in size, and the installation and deployment of sensors raise even high requirements for body design, such as strict requirements on fascia, thickness, installation intervals, and flatness. This makes the entire style and design process much more difficult, because designers must balance vehicle appearance and sensor performance.

To make sensors easier to deploy in vehicles, innovation in form factors is a must. Miniaturization of sensors will be the way forward. In addition, sensors will need to be designed to better fit vehicle styles. Integrated designs that consider both sensors and vehicle style can greatly reduce the constraints on the vehicle body. This requires joint efforts from multiple sectors including materials, processes, and engineering.

1) Distributed antennas and central computing

Radars generally come in an integrated design that encapsulates front-end antennas and back-end signal and perception processing units into one box, to support point cloud generation, object detection, and perception processing. As central computing becomes more common, radar signals can be segmented and extended using the techniques already utilized in Huawei's distributed base stations. Conventional monolithic radars can be used only to generate point clouds, while perception processing units can be integrated with domain controllers. Radars that use distributed antennas and central computing deliver better performance, consume less power, and occupy less space than current radars. Solid-state lidars can also be deployed in a similarly distributed manner.

2) Integration into the vehicle body

An alternative solution for separate deployment is integrating sensors with existing components into vehicle bodies. Shark fin antennas already support GPS, 4G/5G, and frequency modulation. Surround-view cameras are also already integrated into rearview mirrors. Similar solutions can be implemented in the future, such as installing lidars into front-view headlights, integrating distributed antennas into glass components and doors, and combining far-infrared sensors with existing cameras. Such combinations can make sensors more adaptable and easier to deploy, but they also raise the bar for sensors. More effort is needed to address heat dissipation, interference, EMC, and other related issues that may arise from such convergence.

Sensors can also be integrated with each other. Low-cost distributed inertial measurement units (IMUs) can be physically integrated with sensors, which can help change sensors' motion compensation from inter-frame compensation into intra-frame (signal-level) compensation and improve the accuracy of sensor attitude perception. This model can help further improve sensors' overall performance and safety in scenarios like vibration, dead reckoning, and slopes.

3) Surface-mount sensors will reshape sensor deployment

Surface-mount sensors are the ultimate vision for future sensor deployments. Such sensors would need to be smaller and flatter, and eventually plug-and-play. Highly integrated chips would be necessary for this solution, as well as:

- **Microlens array technology:** An assembly of precision-manufactured microlens, just a few millimeters deep and high, used to project a tightly-focused beam. This can massively reduce the focal length between light sources and lenses, making flat designs possible. Though this technology is now mainly used in projectors, it also provides a new possible path for miniaturized lidars and surface-mount sensors.

- **Smart Skin (conformal antenna) technology:** An antenna array designed to conform to the shape of the carrier. This allows an antenna array to be directly attached to the surface of a carrier like a Band-Aid, so that the antenna array can be integrated with the platform structure. In order to work on surfaces with different curvatures, antennas need to detect a surface's curvature according to its bending state and automatically correct the phase of electromagnetic waves. This can maximize the gain effect, make sensors more adaptable, and add more flexibility in vehicle style design.

New technologies (e.g., microlens arrays, smart skin/conformal antennas, and silicon photonic technology) create unlimited possibilities for sensor integration in future intelligent vehicles. In the future, sensors will be like a layer of skin attached to the outer surface of intelligent vehicles. To make this vision a reality, all players along the value chain need to work together to advance in multiple fields, including materials, processes, and engineering.

Central computing platforms: Providing

large computing power to support intelligent driving

Powerful computing platforms will provide the fundamental computing power needed for intelligent driving. More sensors, both in type and number (over 50), will be deployed in a single vehicle to enable intelligent driving in all scenarios in complex road conditions. This includes 100-million-pixel cameras, event cameras, 4D imaging radars, high-resolution lidars, ultrasonic radars, infrared detectors, and sound detectors. In addition, all of these sensors will be increasingly accurate. These sensors will generate massive amounts of data that needs to be analyzed and processed in real time.

With improved chip processes, a vehicle will come equipped with a computing power of over 5,000 TOPS and over 3 million Dhrystone million instructions executed per second (DMIPS). When combined with computing in memory technology, the energy efficiency ratio of new vehicles will be further improved. These central computing platforms will provide the computing power needed to enable intelligent driving.

HD maps: New digital infrastructure to enable everywhere intelligent driving

By 2030, next-generation HD maps will become an essential digital infrastructure that enables the widespread deployment of intelligent driving. The shift from navigation maps to ADAS maps makes real assisted driving and energy-efficient driving possible for commercial vehicles. The current evolution of ADAS maps into HD maps will allow intelligent vehicles to hit the road.

The HD maps that we are currently using for intelligent driving provide HD static environment information, like lane-level road networks and basic road facilities, and positioning layer information, like vector features. This information can support high-precision positioning through map-matching, environment perception assistance, and lane-level route planning for intelligent driving vehicles. With these functionalities, HD maps become an integral part of the current intelligent driving system. By 2030, the next-generation of HD maps will cover all road networks, and 100% of roads with physical lane marks will be precisely digitized. Furthermore, lane-mark free roads will be given virtual lane indications. Elements that highly impact driving safety will be updated in real-time, while the dynamic layers of HD maps will be created, including layers for dynamic traffic events, real-time non-line-of-sight traffic participants, and semi-dynamic traffic environments (e.g., road conditions). Reference layers for dynamic behaviors, positioning outputs, and planning results will also be created and regularly updated to support safer intelligent driving.

Huawei will leverage its expertise in AI chips, algorithms, standardization, and ecosystem construction to create a next-generation HD map system that features advanced sensing, AI, and high-reliability positioning technologies to serve as the new digital infrastructure for intelligent driving by 2030.

V2X cloud brain for multi-vehicle collaboration and intelligent driving

Connected infrastructure is continuing to improve, and intelligent driving is already being adopted in more scenarios. The challenge we face is no longer just making single vehicles intelligent. Now, we want to make vehicles that cooperate with each other. This will push commercial intelligent driving and intelligent transportation to the next level.

Ubiquitous V2X connections need to be built to intertwine people, vehicles, and road infrastructure. A vehicle-road intelligent cooperative driving platform needs to be established on the cloud to streamline end-to-end application scenarios. Thankfully, multi-vehicle cooperative intelligent driving will quickly become a commercial reality, thanks to services like comprehensive environment perception, global resource scheduling, dynamic service mapping, multi-vehicle cooperative driving, lane-level route planning, coordinated signal control, and service simulation testing.

A cloud-based brain can integrate all the necessary information elements of people, vehicles, roads, and environments, improving vehicles' ability to perceive dynamic traffic environments. A cloud-based brain can also share driving strategies between different vehicles, allowing vehicles and traffic infrastructure (e.g., traffic lights and signs) to work in synergy. This allows us to optimize overall (not partial) driving strategies and further promote the development of intelligent transportation systems.

Intelligent cockpits: AI speeds up software and hardware upgrade

The vehicle cockpit is a hub where a user interacts with their vehicle. By 2030, more than 90% of vehicles will have an intelligent cockpit and from now onwards, we expect to see a new intelligent mobile device ecosystem gradually take shape.

Whether an intelligent cockpit can create a superior experience depends on its hardware and software, especially the chips and OS. In addition, the advancement of intelligent peripherals and in-vehicle perception algorithms will ensure users always have fresh experiences along the way.

Open cockpit OSs for an ever-evolving application ecosystem with brand-new experiences

Compared with smartphones and other consumer devices, an intelligent cockpit has more peripherals and supports multiple users, multiple concurrent operations, and multiple ways to interact. An intelligent cockpit has many components, including an instrumental cluster and an in-vehicle infotainment system. Its OS must be designed and developed in ways that support interactive experiences and ensure safety.

By 2030, cockpits powered by new quality of service (QoS) technologies and a multi-kernel OS structure for better software-hardware isolation will be able to schedule and deploy resources more flexibly across domains. In addition, convergence of different domains will drive down the cost of system evolution and contribute to a thriving ecosystem.

An ever-evolving application ecosystem is vital for cockpits to bring brand-new experiences to users. This raises the bar for the consistency and stability of cockpit OS's application interfaces.

At present, the cockpit OS market is plagued by the fragmentation and customization of solutions across the industry. For example, when a carmaker is developing a function that needs to work with a camera or microphone, they have to create unique versions of the function for different vehicle models because each hardware platform is different. This fragile approach to software development prevents different hardware components from effectively connecting with each other or sharing software capabilities.

There is yet another challenge for carmakers in software development. The work is outsourced to multiple vendors, with each being responsible for different functions, and redundant copies of the same function may co-exist. This creates chaos in version management, leads to a lot of extra development work, and complicates software upgrade and maintenance.

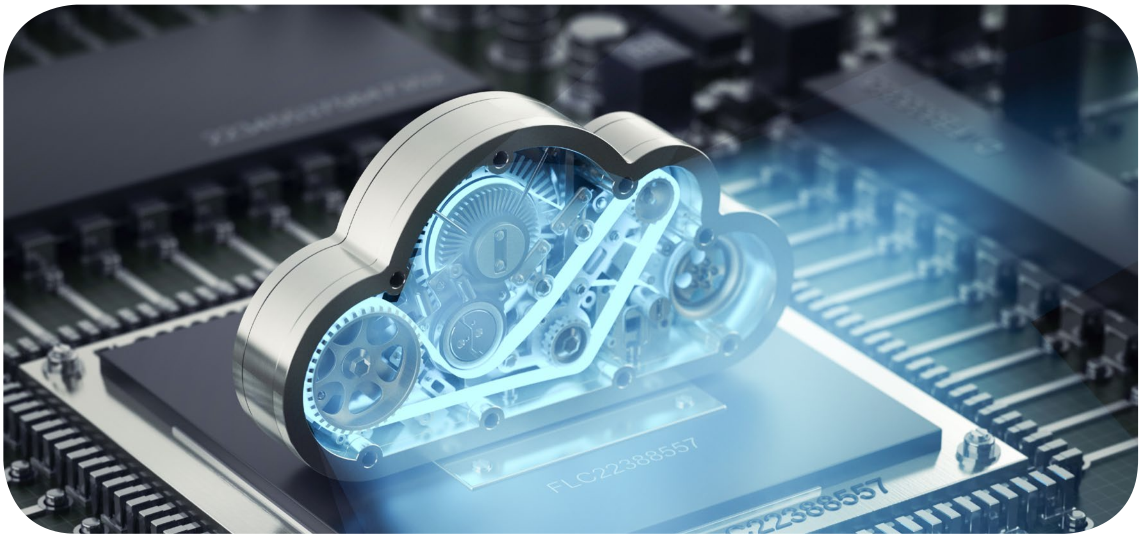
In the future, intelligent cockpit OSs will provide developers with unified APIs that are necessary for a thriving and ever-evolving application ecosystem. Providers of intelligent cockpit solutions will have to upgrade their platforms, open development suites, and other capabilities designed for application developers. In this way, cockpit solution providers will hone their own tech strengths and help developers create, promote, and monetize applications more rapidly. When providers of intelligent cockpit solutions work closely with developers, they can unleash the power of innovation to provide users with more appealing content, services, and experiences.

Intelligent sensing: AI-powered interaction algorithms for an unprecedented level of experience

By 2030, cockpits will have become more intelligent with the addition of new intelligent peripherals (e.g., transparent screens, AR-HUDs, holographic projectors, smart wearables, mmWave radars, and ToF cameras) and AI-based sensing and interaction technologies (beyond just voice and haptic). These peripherals and AI technologies will take the interactive cockpit experience to a new level, in terms of driving safety, in-vehicle communications, infotainment, and other aspects.

Today's in-vehicle voice assistants take the form of 2D virtual figures on the screen. Going forward, they will become 3D holographic robots, supported by mature media-less hologram projection technology. Users will be able to customize the appearance of the 3D robot. The robot will be incredibly expressive both emotionally and in terms of actions. Thanks to AI, the robot will be able to speak naturally, as it will not be subject to very limited programmed phrases.

More advanced hologram technology will enable



a vehicle to project information anywhere within its interior. While you are on a video call, the vehicle will project a hologram of the caller, who "sits" right next to you and talks to you "face to face". While in transit, the vehicle projects information in places that are convenient for the driver to see. Users interact with vehicles just as people naturally interact with each other.

Technologies that support main interaction capabilities, such as voice, visual, and audio, will evolve further. Specifically:

Voice: If we look at driving safety, voice will be the most important interaction capability of intelligent cockpits. Going forward, voice experiences will evolve to be more intuitive, agile, seamless, flexible, and emotionally resonant.

- **Intuitive:** Clear, easy-to-understand content for voice interaction, so that users can act without a second thought

- **Agile:** Efficient execution and timely feedback

- **Seamless:** A unified voice interface that provides seamless access to services across platforms

- **Flexible:** Easy manipulation of the interaction process and retention of the interaction status, and interactions are never abruptly ended

- **Emotionally resonant:** Voice assistants that resonate with users on an emotional level and promote brand image

In the future, voice-based interactive experiences will get even better with improvements in areas like front-end noise reduction, speech recognition (e.g., through visible and conversational assistants), generic voice understanding, and emotion-rich natural conversation.

Visual: The second key way that vehicles interact with users will be through vision. Currently, in-vehicle visual recognition technology is mainly used for the driver monitoring system, cockpit monitoring system, and detection and recognition of human-machine interactions (e.g., through gestures). In the future, more advanced visual recognition technologies will emerge to deliver functions such as detecting living beings in the vehicle, monitoring users' health, and enabling secure payment, entertainment, and integrated audio-video services.

Users will be able to interact with their vehicles in new ways, with vehicles providing more precise and convenient services. For instance, eye tracking technology and AR-HUD will work together to identify objects within eyesight in real time, and then project information about the objects, such as their details and relevant ads. When vision-based gesture recognition technology is used in conjunction with mmWave radars, the accuracy of gesture recognition will be improved and users will be able to smoothly interact with their devices. When there is a lot of background noise, integrated visual-audio technology will read the user's lips and translate lip movements into commands, supporting speech recognition and vehicle control across all scenarios.

Audio: Intelligent cockpits of the future will possess more advanced capabilities in terms of sound pickup and noise reduction, sound zone, and active noise cancelation.

- **Sound pickup and noise reduction:** The quality of in-vehicle phone calls and the accuracy of speech interaction will be improved by filtering out background noise.

- **Sound zone:** By analyzing audio signals before they are sent to the amplifiers, the vehicle can create multiple isolated sound zones, so that all passengers can enjoy immersive, multichannel stereo sound without disturbing each other.

- **Active noise cancelation:** This will continue

to be the next big thing for in-vehicle audio technology over the next ten years. More advanced hardware and algorithms will be needed to thoroughly cancel out the noise generated by the vehicle engine, road traffic, and wind, in order to make driving more comfortable.

Over the next decade, breakthroughs in components, algorithms, and architecture will take in-vehicle audio to a new level, transforming vehicles into mobile entertainment hubs.

Hardware plug-and-play: Standard interfaces keep user experiences fresh throughout hardware lifecycles

Consumer devices like smartphones usually last two to three years, and their software and hardware integration packages are small. Vehicles have a longer lifespan, including a sales window of 5 to 10 years and a useful life of 10 to 15 years. Carmakers tend to research and develop multiple vehicle models at the same time, so they have no choice but to simultaneously maintain a whole range of software and hardware versions.

As novel applications appear, hardware performance needs to keep up. For example, chips and other key components like cameras and displays will have to stay up-to-date throughout a vehicle's lifecycle. Cockpit hardware upgrading will lead to new business models in the post-sales stage.

Inside a vehicle, chips or chip modules must be able to meet the computing needs of software and hardware for the next three to five years. The chip module itself needs to be designed for backward and forward compatibility, so as to ensure easy upgrade (e.g., through pluggable components) and strike a balance between hardware lifetime and computing demand. Hardware plug-and-play is essential for this type of chip module.

When a key peripheral is replaced by a new one, it is necessary to install a new driver to support the new peripheral. Carmakers should aim to make this process as simple as installing a new driver on a Windows operating system. Also, it should allow OTA updates for certain parts of the cockpit OS. To make this possible, unified standards should be established for the interfaces of cockpit hardware, in order to eliminate issues caused by customized interfaces of head units, cameras, displays, HUDs, intelligent seats, intelligent steering wheels, in-vehicle robots, intelligent windows, holographic projection hardware, etc. To standardize cockpit hardware, the automotive industry needs to double down on standardizing communications interfaces, including those for short-distance wireless communications, wired communications, video, and audio. Standardization is key to driving down component costs and cultivating a hardware ecosystem.

Multi-device collaboration through distributed technology for seamless intelligent experiences

Intelligent vehicles are integrated systems, and how they interact with users will involve the broader surroundings. Connection and interaction of devices depends on both general-purpose cloud technology and a distributed software bus. Huawei's HarmonyOS for Automotive provides a distributed software bus to create a seamless experience across nearby devices, making it easier and more comfortable for users to interact with their devices. Huawei's distributed platform enables multiple devices to work together across scenarios, such as transit, office, and home. This creates an intelligent cockpit service system that delivers superb experiences for users in all scenarios, no matter what devices they are using at the time.

Imperceptible sensing and zero-wait transmission are the deciding factors for realizing seamless experience across nearby devices. If we are to meet these preconditions, we must first devote efforts to answering the questions of how different devices should discover and connect with each other, how connected devices should come together to form a network, and how transmission between devices that use different protocols can be realized. The key technologies here include automated device discovery, connection, networking (e.g., multi-hop automated networking and multi-protocol hybrid networking), and transmission (e.g., diverse protocols and algorithms, and intelligent perception and decision making).

A distributed software bus connects different types of devices by using a "protocol shelf" and a software-hardware collaboration layer, despite the different protocols used by the devices. The hub module of the bus analyzes commands to discover and connect devices. The task bus and data bus provide other functions, such as transferring files and messages between devices.

Several preconditions have to be met before an intelligent vehicle and IoT devices can work synchronously to offer an interactive experience. In terms of design logic, it is important that interactive experience of the intelligent cockpit be consistent with that of mobile phones or other such devices. Regarding operating logic, intelligent cockpit applications must provide the unified functions of smartphone applications, and may be designed based on the hardware capabilities of cockpit peripherals. In addition, users want a seamless experience when they switch between cockpit and smartphone applications, especially when it comes to their calendar, navigation, music, video, and conferencing applications.

Ultra-fast connectivity technology for smartphones and head units will be able to help quickly converge the hardware resources, system capabilities, and service ecosystems of phones and vehicles, so that carmakers will be able to leverage the computing power of smart devices (e.g., smartphones) and the services of the mobile Internet ecosystem. The distributed platform for devices will link transit and other scenarios (e.g., office and home) to offer users the best possible experiences across devices and create an intelligent cockpit service system for universal scenarios.

In-vehicle optical applications: Lighting up a new vision for drivers and passengers

Viewing: Panoramic, immersive holograms for eye-opening experiences

Humanity's desire for superior visual experiences knows no limits. As vehicles become more intelligent, their front windshields, side windows, and panoramic sunroofs are quickly becoming displays that can show information through

lifelike holograms. As laser and pixel technologies continue to evolve, they are expanding the roles of headlights, from being mere sources of lighting to projecting 3D information in all directions around the vehicle.

In-vehicle optical applications are designed to create superior visual experiences as they support information display, interaction, and entertainment. In terms of navigation, the windshield can be used to enhance driving safety by displaying essential information and safety warnings like road directions and obstacles. The windshield and even rear side windows can be used for entertainment, serving as holographic screens that offer the kind of immersive 2K/4K viewing experiences that you get at the cinema. Curved panoramic sunroofs can project customized light patterns, to mimic everything from meteor showers and constellations to deep-sea coral reefs. Going forward, intelligent vehicles will also be equipped with headlights capable of wide-gamut, high-pixel projection that will allow users to project and watch movies outdoors.

Vehicles are quickly becoming the third living space for people after their home and workplace. This is why user demand for visual experiences on the go has been increasing. Users expect immersive experiences that deliver images and video with higher resolution and broader view. Users are also looking for new eye-friendly technologies that can help with carsickness. This means interactive functions not only have to create truly immersive experiences, but also help users avoid getting carsick while making long video calls or watching movies. In addition, rear-seat passengers expect optical display technologies to deliver a wide array of entertainment functions without fatiguing their eyes.

Looking to the future, spatial optical technologies – in tandem with human-focused experiences – will reproduce the real world with extremely high-resolution images that are sharper than even the human eye can process. This will require:

- **Wide-view, immersive technology:** Spatial optical technologies like freeform mirrors, diffractive optical waveguides, and polarization beam splitting can be used to project images up to 100 inches in size from a 10-inch display. With directed light field technology, users can watch 3D movies on in-vehicle displays without 3D glasses. Also, directed sound field technology offers amazing acoustics that would previously have been available only to best seats of an IMAX cinema.

- **True-color UHD displays:** Displays will soon come fitted with optical engines for 2K, 4K, and 8K video, and diffuser film displays

based on a micro-nano structure. These UHD displays will significantly enhance pixels and brightness, making the text and images they display so crisp that the pixels will not be visible to the naked eye. With RGB lasers, UHD displays will support DCI-P3 color space or even BT.2020 color space for 8K video, perfectly showcasing the true colors of objects being displayed.

- **Visual health:** Virtual imaging systems can display images at a 3-meter distance from the viewer's eyes, thus eliminating the risk of myopia. Passive cool tint technologies also make it possible for displays to emit zero radiation and reduce the amount of blue light that reaches the eye.

- **Human-focused experiences:** Carsickness is caused by conflicting information the human body receives from eyes and ears when a vehicle is in motion. Staring at an on-board screen in a moving vehicle often exacerbates this problem. When engineering technologies that focus on dynamic human factors are incorporated into on-board screens, this problem can be minimized or avoided completely. Eye fatigue occurs when the eyes' ciliary muscles contract too tightly. The technologies that enable eye tracking, vision screening, and diffuser film displays that automatically adjust the distance between your eyes and the images being projected can all help relax ciliary muscles.

Sensing: Intelligent all-around protection in all weather conditions for greater driving safety

Advances in the field of in-vehicle optical technology will make environment modeling more detailed and comprehensive. Innovative applications like near infrared sensing, thermal optical sensing, and image optical sensing will support more precise environment modeling, enabling drivers to see the invisible. Based on modeling, external environment information collected by sensors will make driving safer and more intelligent while meeting the service needs of drivers and passengers. Sensor-powered headlights can provide accurate information about low-visibility environments (such as when it's rainy, snowy, foggy, or nighttime) to alert drivers to blind spots. When data collected by sensors operating on different frequency bands is pooled together, useful insights can be extracted and used to identify biological factors and risks more accurately than ever before.

As more people tend to spend more time in their vehicles, it is crucial that intelligent vehicles of the future can support potentially life-saving safety functions, such as the ability to monitor the status of drivers, passengers, and the vehicle itself. Light sensing technologies like in-vehicle infrared and optical sensors can be used to promptly detect things like respiratory problems, an irregular pulse, and heart issues, and give timely alerts. These technologies can also identify driver fatigue and issue heads-up alerts to reduce the chance of human errors and accidents. They can even be used to detect whether children or pets are left locked inside a vehicle and automatically trigger solutions.

In the future, optical technologies that operate on different frequency bands will work with image recognition and processing technologies to bring intelligent vision to vehicles, ensuring all-round safety in all weather conditions.

- **All weather conditions:** All objects with temperatures above absolute zero emit infrared radiation over the 8–14 µm waveband. Infrared imaging technology can convert that infrared radiation into a visible picture, helping remove glare and identifying objects in rain, fog, or even total darkness. Conventional imaging technology that works on the 400–700 nm waveband (a waveband that is visible to human eyes) do not currently work in exceedingly dark, rainy, or foggy conditions.

- **All-round protection:** When an object vibrates, the frequency of the light signals it emits changes. Optical Doppler spectrum technology can capture micro-vibrations on a user's skin. Therefore, when combined with infrared thermal imaging technology, optical Doppler spectrum technology can be used to monitor a user's vital signs, including heart rate, respiration, and body temperature. This means the vehicle can issue alerts when the driver is becoming fatigued or if someone in the vehicle needs medical assistance.

Connecting: New interaction methods ensure better driving safety and stronger emotional bonds

In-vehicle optical applications provide new ways for vehicles to interact with their users and the world around them. These applications are also crucial for driving safety. Inside vehicles, augmented-reality head-up displays (AR-HUDs) are an intuitive tool for vehicle-driver interaction. They can directly display information on the windshield, enabling the driver to more easily view information, instead of having to look down at various instruments. The AR-HUDs can display real-time information, such as AR navigation directions, alerts for obstacles, vision augmentation in rain, fog, and darkness, as well as information about nearby gas stations and other services.

Intelligent lighting systems also provide a new way for vehicles to interact with the outside world beyond just their horn and signal lights. When a vehicle is in motion, intelligent lighting can project interactive information onto the road, such as vehicle width, alerts for rain and fog, and night vision, to help drivers make more informed decisions and enhance safety. In addition, intelligent lights can project useful information for pedestrians, such as turn and right-of-way signals. Intelligent lights are also capable of projecting emotions, showing customized information such as patterns, emojis, texts, and weather data, and can even enable other forms of interaction through light shows and concerts.

In the future, a variety of in-vehicle optical applications will create even more ways for intelligent vehicles to interact with people. This can happen through:

- **HUDs:** Currently, AR-HUDs use megapixel-level optical modulation engines and spatial optical technology, but future adoption of dual-focal plane technologies will pave the way for even more advanced multi-layer AR-HUDs that will be able to project dashboard information two to three meters in front of the driver, and navigation and other useful information seven to ten meters in front of the driver. Future naked-eye 3D technology

will also further improve the interactivity and experience of HUDs.

- **Lights:** With megapixel-level modules and optical lens, automotive lights will be transformed into projectors, displaying information such as vehicle-to-vehicle distance alerts and animated greetings. Vehicles that use precision laser lighting and sensing technologies will be able to interact with the environment through methods like dynamic ground projection (dubbed "dynamic light carpets"), to illuminate the surrounding area outside the vehicle and provide centimeter-level precision lighting. This will undoubtedly make driving safer and more fun. In the future, current/voltage modulation will also make it possible to display information in beams, and visible light communications technology will be able to support vehicle-to-vehicle communications.

- **Windows as displays:** Ultraviolet light projectors and fluorescent film glass will turn vehicle windows into colorful, full-size displays where notifications of the driver's intent, ads, and other types of information can be shown to pedestrians.

Intelligent vehicle-cloud services: Helping carmakers become digital, service-oriented companies

In the era of intelligent connectivity, we need smart vehicles, intelligent roads, and also an intelligent cloud. This cloud needs to enable data-driven closed-loop iteration of intelligent driving algorithms and effectively connect people, vehicles, and roads. It will also enable a variety of intelligent connectivity applications, such as intelligent cockpits, intelligent mobility power, intelligent driving, and intelligent transportation. Intelligent vehicle cloud services will be deeply integrated with capabilities at the application layer and become more agile to adapt to a fast changing market. With their leading innovation capabilities, vehicle cloud service providers will establish their unique competitive strengths in the market and help carmakers become digital, service-oriented companies.

Closed-loop management of intelligent driving data on the cloud for faster development and continuous iteration of algorithms

To solve the long-tail problem with intelligent driving, intelligent driving developers need to continuously enrich corner case datasets and simulation scenario libraries for iteration of intelligent driving algorithms. During this process, they need petabytes of data and a huge amount of computing power (more than 1,000 GPUs) for algorithm training, and must simulate driving astronomical distances (as far as 10 billion miles) to validate an algorithm. In addition to large storage capacity and computing power, the iteration of algorithms also requires reliable, secure, and scalable infrastructure services. The conventional model for data center construction puts the costs and O&M responsibility on intelligent driving developers, so we expect that cloud computing technologies will be widely applied in intelligent driving to address these challenges.

Cloud service providers will need to be capable of providing one-stop services on the cloud as this can help address the complex engineering problems of intelligent driving, like data collection, data replay, automatic labeling, identification of corner cases, incremental dataset generation, model management, training task management, model delivery, simulation scenario library building, simulation test, and algorithm adaptation. These tasks often account for more than 70% of the development workload. Therefore, intelligent driving algorithm development should be data-centric, and during this process, data should be decoupled from algorithms. This will help build an open enablement platform that provides a complete and automated development toolchain. Carmakers and developers will then be able to quickly build up their intelligent driving development capabilities, address the engineering complexity in algorithm development, lower technical barriers to entry, and allow for faster algorithm development and iteration. Specifically, intelligent driving developers need to:

(1) Provide scalable, secure, and compliant infrastructure for intelligent driving algorithm development

Hyperscale data storage and computing centers, built based on cloud platforms, can provide the massive uploading capacity, compliant storage services, and massive computing resources needed to handle the massive amounts of data that will soon be generated by intelligent vehicles. Carmakers developing algorithms for intelligent driving will be able to access affordable, scalable, reliable, and secure infrastructure.

(2) Address engineering incoherence and support the DevOps of intelligent driving algorithms

There are two major issues that hinder intelligent driving algorithm development: fragmented toolchains and lack of coherence in the development process. Cloud service providers must develop the ability to provide full-lifecycle services on the cloud, covering everything from algorithm development and testing, to commercialization and optimization. These capabilities should include a complete development toolchain, preset algorithms, and datasets and scenario libraries that are continuously updated, simulation, and validation. Support should be provided for user-defined algorithmic models and heterogeneous hardware. Cloud-based infrastructure and AI capabilities can help ensure a closed-loop intelligent driving development process, from data collection and processing, and identification of scenarios (especially corner cases), to algorithm management, training, simulation, and validation. These capabilities can help carmakers' intelligent driving R&D teams quickly locate problems, optimize algorithms, and conduct testing and validation, allowing for faster algorithm development and iteration.

New cloud-based simulation technologies to accelerate validation and iteration of autonomous driving algorithms

Cloud-based simulation can speed up the testing and validation of intelligent driving. Specifically, intelligent driving developers need to:

1) Recreate a virtual version of real-world scenarios to quickly build a simulation scenario library. The developers can use data collected by

vehicle-mounted sensors, HD maps, and publicly available simulation tools, to simulate complex, real-world road test scenarios on the cloud and quickly simulate complex traffic flows. This whole process will take only a few minutes.

2) Support the validation of multi-vehicle collaboration. Intelligent driving developers need to improve their abilities to validate different intelligent driving algorithms in multi-vehicle collaboration and vehicle-road collaboration scenarios.

3) Play out scenarios devised in a virtual cloud environment on real vehicles on empty roads to better validate algorithms. This can boost efficiency and safety while also providing true-to-life test scenarios.

4) Improve the efficiency of cloud-based parallel simulation. Intelligent driving developers can use container technologies and the resources available on the cloud for large-scale parallel simulation. In a single day, tens of millions of kilometers of driving can be simulated, shortening the algorithm iteration cycle from weeks – or even months – to just days.

HD maps are essential for simulation, whether in a virtual environment that simulates the real world, or a hybrid of virtual and real-world scenarios. Cloud service providers can build their HD map service capabilities in the form of cloud services while ensuring regulatory compliance during data storage and application. They also need to provide detailed coverage, unified standards, layered services, and dynamic updates, in order to support a wide variety of applications, including positioning applications, intelligent driving simulation/operations, intelligent and connected industry parks, intelligent driving, and smart cities.

Vehicle cloud service providers need to provide map data storage and application services based on secure and compliant dedicated cloud infrastructure while meeting applicable regulatory requirements. Leveraging the large storage capacity of cloud, abundant computing resources, and intelligent algorithms, the providers can process road test data in a secure and compliant manner and support third-party partners in intelligent driving development and map data application services.

Huawei hopes to form an alliance of HD map service providers to create greater synergy across the industry. Huawei's HD map cloud services support data access by other map service providers. Following an intelligent inspection of data quality and post-processing, Huawei can establish unified service standards and integrate resources from other map service providers. This helps reduce redundancy in map drawing work and cut costs. Through open and flexible operations, Huawei can provide users with HD map data services that are responsive, up-to-date, and all-encompassing.

Cloud-edge-device collaboration can help dynamically distribute HD map data and update maps through crowdsourcing. With cloud-edge-device collaboration, HD map data can be distributed on demand and flexibly used, facilitating intelligent driving and smart city applications. In addition, collaboration between the cloud and intelligent connected vehicles, other traffic participants, roadside infrastructure, and roadside edge computing helps update the dynamic layers of HD maps in real time, while crowdsourcing enables the update of static layers, helping keep HD maps up-to-date.

A unified data service system to enable carmakers to become digital, service-oriented companies

As ICT technologies are increasingly integrated into the automotive industry, cloud computing is allowing intelligent connected vehicles to provide new functions and experiences. Carmakers rely on cloud-based intelligence to build scenario-specific big data application service capabilities. This is because they need to integrate their underlying business logic with big data and AI when developing applications and services based on their vehicle data. For example, early warnings of battery thermal runaway require a combination of knowledge and technology in both the electrochemical and AI fields. As intelligent connectivity scenarios are further segmented, carmakers need to build a unified data application service system and remain open to collaboration to meet the massive service demand.

Application-centric intelligent vehicle cloud services, built based on the Cloud-native 2.0 architecture, can help carmakers build a unified and complete data service system covering collection, storage, consolidation, computing, management, and services. This system should enable closed-loop data management and continuous iteration of applications. Carmakers can also work with third parties in various domains to provide intelligent applications and services that meet personalized consumer needs, and become digital, service-oriented companies.

1) Full data collection and aggregation based on cloud-native infrastructure

Cloud service providers need to build unified, efficient, and application-centric data service platforms with diverse computing capabilities and a high level of coordination between software and hardware. Such a platform should be able to manage and support heterogeneous hardware, and shield applications from the complexities of the hardware. With open and standardized connected vehicle technologies, cloud service providers will be able to connect massive numbers of vehicles and provide millions of concurrent services on the cloud. This will help build up data channels, and collect and aggregate data from components such as intelligent mobility power, intelligent driving, intelligent cockpits, and intelligent connectivity systems. This will help build a full data lake for intelligent connected vehicles on the cloud and support the creation of digital twins of vehicles.

2) Data value mining and application innovation based on cloud-native intelligent services

There are two key preconditions for enabling intelligent services: The first is the building of full-lifecycle data management capabilities on the cloud, from data collection, transmission, and storage, to labeling, analysis, and application, in order to reduce data governance costs and fully unlock the value of data. The second is the development of a converged data analytics platform with cloud-native technology, which helps transcend data boundaries and enable efficient cross-source and cross-domain collaborative analytics.

As the AI technology and ecosystem on the cloud mature, cloud platforms will continue to enrich algorithm libraries and provide capabilities like automatic labeling and preset algorithms to lower the technical barriers to entry for AI application development. The hyperscale computing power and massive amounts of data available on the cloud can enable data value mining and the innovation of intelligent big data applications. This will then enable the development of intelligent applications like intelligent diagnosis of core vehicle components, vehicle status monitoring, vehicle function preference analysis, and analysis of component working conditions. The data obtained from these applications and services can also help optimize product design and R&D.

3) Build a thriving service ecosystem based on agile and open application development capabilities

Cloud service providers can use cloud-native and microservice technologies to split applications into microservices that can be independently deployed and run. Such high scalability enables applications to have higher availability. It will also be necessary to move from the traditional model, in which application development and delivery are separate, to the integrated DevOps model, as this can speed up application development and iteration.

Cloud service providers can also open up their data service capabilities to enable third-party development of intelligent applications and help carmakers build an ecosystem of applications essential to our everyday lives. This will allow companies to offer data as a service, thereby unlocking the value of data more rapidly. Carmakers and their partners can provide users with rich personalized services and applications, develop new use cases and business models, and explore services as a new source of revenue.

Intelligent mobility power: High convergence, high efficiency, and high voltage

A highly converged and simplified power system with a self-evolving, AI-enabled structure

The development of conventional power system architectures is a lengthy and costly process. With the electrification of more vehicles, more electric vehicle (EV) components need to be integrated to simplify development, adaptation, layout, and evolution.

Most of EV power systems currently in use are three-in-one units that integrate motor control units, motors, and reducers, which, in essence, is an integration of functions and hardware. The EV power system of the future will be a six-in-one unit that integrates the three-in-one electric drive system and the three-in-one on-board charging system. Compared with the traditional separated model, a converged model would be 30% smaller and 20% lighter. A six-in-one unit can be flexibly adapted to simplified vehicle layouts, and its higher level of integration would mean that carmakers can cut spending on component testing and introduction and effectively slash development costs.

Developing a converged power system requires the convergence of multiple components by integrating electric systems, boards, chips, algorithms, and control units. A hyper-converged system would also need higher-level EMC to eliminate any interference that the convergence process may cause. In addition, a more efficient heat dissipation system would be necessary to meet the increased cooling needs of such units. More converged control functions for power systems will also need an architecture shared by in-house AUTOSAR Classic Platform (CP) and the AUTOSAR Adaptive Platform (AP) that supports a domain-oriented control evolution.

Hyper-converged power systems can enable multiple system-level value-added features and satisfy the development requirements of any on-demand functions, such as OTA-enhanced power devices, rapid in-vehicle software upgrades, driving modes updates, continuous optimization of noise, vibration, and harshness (NVH), and wider performance boundaries. All of these would improve the performance of the EV power system over the whole lifespan of a vehicle.

The convergence of multiple control units provides EV power systems with a unified platform that can upgrade intelligently. Iterative component features can be scheduled more efficiently using an AI platform, thereby meeting more complex and intelligent power needs effortlessly.

Thanks to AI, EV power systems will be able to learn, iterate, and evolve on their own. This will be seen in applications such as:

(1) Proactive safety warnings that can be provided one week advance to warn of potential battery thermal runaways, and cloud-based intelligent remote calibration that can enable an EV power system to perform better over the entire lifespan of the vehicle.

(2) Visible, controllable, and predictable EV power system quality that is available throughout the vehicle's entire lifecycle.

(3) AI-enabled EV power system service life predictions that can extend the system's lifetime by avoiding scenarios and operating conditions that could affect service life.

(4) On-demand maintenance or even completely eliminated maintenance, shifting away from the current maintenance paradigm based on mileage and driving hours to greatly improve customer satisfaction.

Bits managing watts: AI-powered optimization maximizes energy efficiency at three levels

Thanks to digital technologies like 5G, AI, big data, and IoT, EV power systems can use "bits managing watts" to create synergies between electric, thermal, and kinetic energy. Simultaneously, an internal AI system can be used to optimize energy efficiency in all working conditions. This would allow the electricity stored in EV batteries to be fully converted into power, helping alleviate range anxiety and offer better travel experiences. Future electric vehicles will need to be efficient at three different levels: the component level, the system level, and the vehicle level.

First, components can be made more efficient through leveraging insulated gate bipolar transistors (IGBTs), silicon carbide (SiC) components, and gallium nitride (GaN) components. Advanced packaging technology can also be used to improve cooling systems, reduce parasitic parameters, and improve the electrical robustness and reliability of power modules. This will make even high-voltage, high-temperature, and high-speed operations more efficient.

Second, systems can be made more efficient by leaning harder into the convergence trend and making EV power systems a coupling component that integrates the mechanical, electrical, and thermal fields, among others. A digital and intelligent development platform can be used for multi-objective and multi-parameter optimization, circuit topology improvement, and algorithm control. Moreover, AI and big data technologies will further improve virtual testing and remote online calibration. Each subsystem within an EV power systems will then convert power more efficiently, increasing the overall efficiency of the complete EV power systems.

Third, vehicles as a whole can be made more efficient by digitally connecting their different system components that are currently separate, including the electric drive, thermal management, steering, and braking systems. This will allow the different components to complement each other in terms of energy efficiency. In addition, bits must be used to manage watts. This can help better coordinate electric, kinetic, and thermal energy, effectively reduce energy loss in non-power systems, and increase energy supply efficiency. For example, excessive heat from the motor can be used to preheat batteries in the winter, eliminating the need for a separate positive temperature coefficient (PTC) heater. The on-board charger (OBC) and the air conditioner compressor can share a high-voltage topology to maximize the utilization of power devices.
In the future, with the support of cloud and AI, vehicle energy consumption management will become more intelligent. Ultimately, optimal energy consumption management policies will be created for the same vehicle in different scenarios and working conditions.

1000 V power systems and 5-minute fast charging

The electric vehicle industry is developing rapidly as consumers become more accepting of electric vehicles. However, charging convenience, battery life, and safety remain the three major concerns that put consumers off from going electric. Among these three problems, charging convenience receives most attention.

If the charging time for a 100 kWh battery pack can be cut from 50 minutes to less than 5 minutes, we would need to increase the charging power of each charging gun from 60 kW to 500 kW. To achieve fast charging under the same power conditions, high-voltage power systems are needed as they can effectively solve the problems that low-voltage systems face, like excessive charging current, high charging facility costs, heat dissipation difficulties, and safety considerations.

The technologies required to create such high-voltage power systems are basically ready for use. One interesting example is the new generation of power semiconductor technologies based on SiC. The critical breakdown field strength of SiC-based power devices is nearly 10 times that of those made using pure silicon, which means that SiC-based devices can easily operate at higher voltages. SiC-based devices of 1500 V or above are already ready for mass production and can be used to develop high-voltage platforms for new-energy vehicles. In addition, the low on-state resistance and the high thermal conductivity of SiC-based devices can boost system efficiency. The development and application of SiC materials provides a solid foundation for the evolution of EV power systems from low-voltage systems to high-voltage systems. High-voltage vehicle components like OBCs, battery management systems (BMSs), and powertrains are also already ready for mass production.

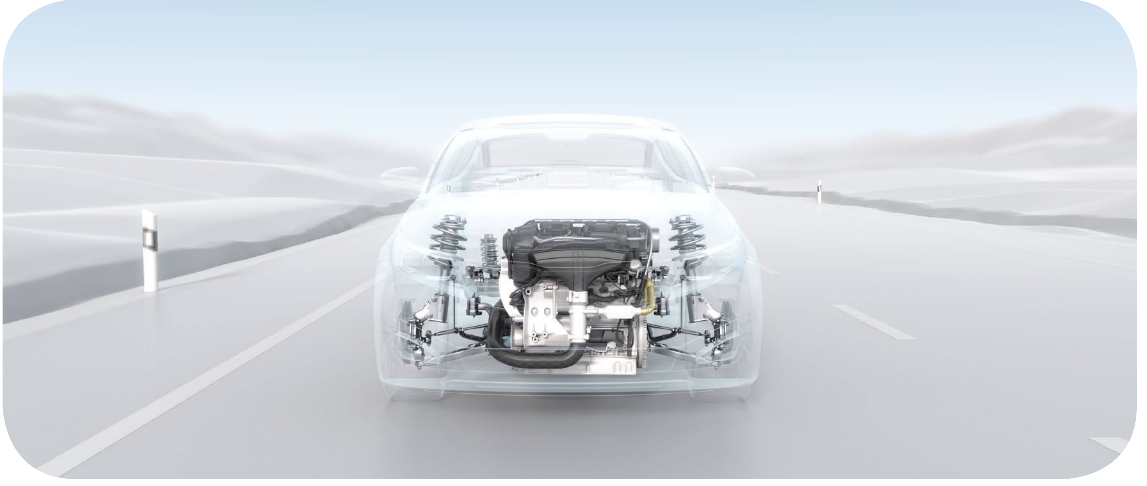
The voltage plateau of EV power systems will further increase from 800 V to 1000 V, charging currents will increase from 250 A to 600 A, and charge rates will rise from 2C to 6C. In addition, the time for the state of charge (SOC) from 30% to 80% will be cut from 15 minutes to 5 minutes. As the voltage of SiC-based power devices continues to increase, EV power systems with higher operating voltages will come more common in the market.

Safety and trustworthiness: Defense in depth through cyber security and functional safety integration

Today's vehicles are increasingly electric, connected, and intelligent, and have complex and integrated E/E systems. The safety of the new functions that come with these new developments is now often a matter of system safety engineering. Finding ways to protect against external attacks, cope with unforeseeable faults and diverse environmental conditions, and ensure safe and reliable mobility has become a common challenge facing the industry. As we embrace intelligent connected vehicles, the notion of vehicle safety and security is also changing. For intelligent connected vehicles, and intelligent driving systems in particular, there are three key aspects to safety and security – functional safety; safety of the intended functionality (SOTIF); and cyber security, data security, and privacy protection.

Functional safety: High-standard safety and reliability in all scenarios and all weather conditions

By 2030, the importance of functional safety will have become more salient with the widespread use of technologies such as intelligent driving, chassis-by-wire, high-bandwidth communications, wireless networking, regional access, SOA design, and central computing and control. In addition, enormous challenges are anticipated for traditional functional safety in terms of security



design and continuous iteration of super software platforms, AI algorithm security, situational awareness and security planning and control in all scenarios and all weather conditions, and security design and verification of fail-operational vehicle E/E systems. However, with continuous improvements in advanced materials, manufacturing processes, architecture designs, as well as multi-level redundancy through the chip, circuit, board, and system, along with a minimum safety system design, intelligent vehicles of the future will meet or even exceed the safety requirements of Automotive Safety Integrity Level (ASIL) D. This means they will be able to function with guaranteed levels of stability and reliability in any scenario and under any weather conditions.

At the architecture and system level, heterogeneous backup of domain controllers will become the standard configuration to address system failure and random hardware failure. This will help ensure mini driving control systems and mini driving functions such as driving, braking, steering, gear shifting, mandatory lighting, critical speed control, gear indication, having at least one door that can be opened, and mandatory end-outline marker lamps. It will also help ensure that emergency processing mechanisms, such as communications and power supply redundancies, are in place for key signals and systems. Vehicles with highly automated systems or more advanced systems must support redundant heterogeneous systems that have the necessary environmental awareness capabilities.

To make this happen, the following technical measures are required:

(1) A fail-operational design for the central computing platform. The central computing platform is a core component, and in order to meet safety requirements, its key hardware, software, and system architecture must be built with a fail operational design:

Reliable architecture: This requires a multi-layer fault monitoring and handling framework; virtual hardware security partitioning and heterogeneous software deployment; redundancy designs for power clocks, communications links, and redundant vehicle controls; and fault cross-check interlocking mechanisms. These measures will help completely resolve common cause failures and cascading failures. In addition, the minimum safety system design helps ensure 100% safety and trustworthiness.

Automotive-grade engineering: Highly-conductive materials, phase change materials, and new liquid cooling technologies can be harnessed to achieve efficient heat dissipation at the chip, component, board, and system levels. High-voltage, high-current switch technology and high-speed cable technology can be used to reduce the bit error rates of high-speed signal transmission. Electromagnetic interference suppression, high-reliability small-scale surge protection, and power management technologies can help improve systems' EMC. Additional advanced processes can also help extend service life and meet safety requirements beyond ASIL D. These processes include design for assembly, vehicular adhesive dispensing, coating, gluing, pipe corrosion and condensation test limits, optical coating, and soldering.

Maintenance-free across the full service life: Based on chip, software, and system failure model libraries, proactive maintenance is performed automatically in real time to predict, prevent, and mitigate 0 km failures.

(2) Architecture-level redundant safety:

Input/Output (I/O) device abstraction and SOA-based security design and FFI design for proximity sensors and actuators; secure, redundant access of important I/O devices and distributed, secure deployment of critical services; real-time status monitoring and fail-safe management for SOA-based service platforms; minimum safety system design; high-level safety with significantly better vehicle availability and user experience

Model-based systems engineering (MBSE) and model-based safety analysis (MBSA) integration:

- Integrated model design, analysis, and verification of system design, functional safety, and execution models;
- Adaptive, safe hierarchical architecture and high cohesion, low coupling of modular components;
- Paced iteration of vehicle software and fast functional safety design, analysis, and verification.

Architecture-level functional safety simulation and verification:

- Vehicle functional safety simulation: Hazard scenario library, fault database, vehicle architecture platform model, simulated external interfaces, vehicle model, driver behavior, fault response model, and operating environment model;
- (Semi-) autonomous functional safety verification;
- Closed-loop functional safety of vehicles as a whole.

(3) Double-guaranteed communications security: This requires a security design with end-to-end communications security, real-time monitoring of network faults, data flow rerouting in the event of network failures, and dual-fed and selective receiving. This helps ensure reliable degradation and safe operations in the event of a network fault. The wake-up function, power supply, and the links of the external sensing input, vehicle control information input, and key actuator output must be redundant.

SOTIF: Turning "unknown hazardous" scenarios into "known not hazardous" scenarios

When fully autonomous driving becomes a reality, the responsibility for driving safety will shift from the driver to the vehicle. Unlike functional safety, SOTIF is mainly about preventing and addressing unreasonable risks caused by functional insufficiencies and reasonable foreseeable misuses, including those caused by inadequate consideration of target scenarios, and diverse environmental factors that may affect sensor performance. Beyond those, SOTIF also needs to constantly explore "unknown hazardous" scenarios and turn them into "known not hazardous" scenarios. SOTIF will be one of the key issues to be addressed before intelligent driving can be fully commercialized.

As intelligent driving functions become more widespread, SOTIF-related issues will be increasingly prominent. The following points are particularly important:

(1) The development of SOTIF standards in all key markets should be closely monitored. These standards must be fully implemented at all stages from system design and specification to verification, validation, and field monitoring. This is key to ensuring end-to-end SOTIF.

(2) Research is needed for SOTIF-related risk analysis and response. This can be done from multiple aspects such as compliance with traffic regulations, determination of operational design domain boundaries, dynamic driving task execution, minimal risk manoeuvre, human-machine interaction, and response to foreseeable scenarios.

(3) Traffic regulations should be digitalized so that they can be understood and followed by autonomous vehicles, allowing them to easily integrate into the larger transportation system.

(4) It is essential to have a safety time domain design that comprehensively considers the total time needed for autonomous vehicles to sense, plan, and act, and maintain a proper time margin. This will help cope with uncertainties, including uncertainties caused by other traffic participants' actions, to guarantee safety on the road.

(5) Sufficient simulation, testing on a closed test track, and road tests are needed, along with enhanced audits and assessments, in order to fully explore potential functional insufficiencies and triggers that can led to hazardous behavior.

(6) During the vehicle operation phase, data analytics and predictive functions should be used to promptly identify potential hazards caused by evolutions in the environment and driving habits, and appropriate measures need to be taken accordingly. Ultimately, the intelligent driving systems should reach a safety level comparable to an excellent human driver, in order to meet user expectations for the safety of autonomous vehicles.

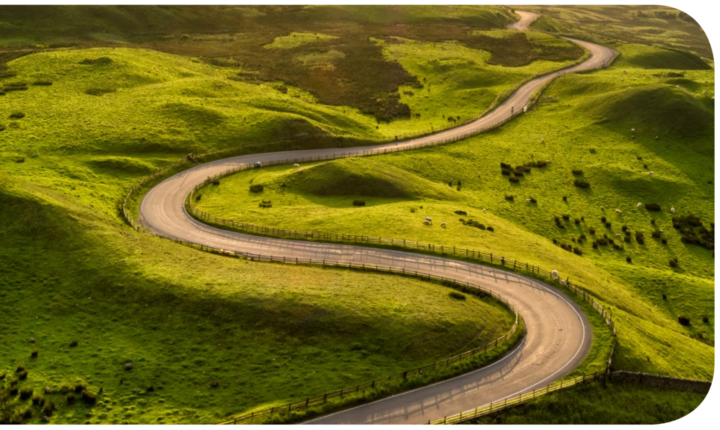
Cyber security, data security, and privacy protection for a comprehensive safety net

Typically when we talk about intelligent mobile devices, the focus is on the safety of the user. What sets intelligent vehicles apart is that designers must take into account the safety of people in and outside the vehicle. Strictly adhering to safety guidelines and ensuring the security of intelligent automotive products and services is a fundamental requirement for ensuring the safety of both people and property.

The automotive industry is embracing the CASE transition. This is expected to create many risks in new cyber security, data security, and privacy. Addressing these risks requires the joints efforts of all stakeholders, including regulators, carmakers, and component suppliers. They need to work together to build a full-lifecycle safety net for intelligent vehicles, in order to address cyber security challenges and ensure the safety and security of intelligent automotive products and services.

1）Cyber security

According to security consulting firm Upstream, in the period from 2011 to 2019, there were 342 attacks targeted at intelligent vehicles. The number of cyber attacks directed at intelligent vehicles has been rapidly increasing. The attack methods have evolved from traditional system hacking through physical contact to remote attacks that require no physical contact at all. Remote attacks account for more than 90% of all cyber incidents, and nearly one-third of them affect vehicle controls. In terms of the distribution of the attacks, the vehicle cloud is the most frequent target, accounting for more than 20% of all attacks. The vehicle cloud, key, and on-board diagnostics (OBD) system interfaces are the points most vulnerable to attack. There's also an emerging type of attack that targets a



vehicle's sensors. In the future, 24/7 connectivity will be a prerequisite for continuous functioning and updating of intelligent vehicles. More and more peripherals and network interfaces will be deployed in the vehicles. Always-online vehicles mean more attack surfaces and more susceptibility to a wide range of attacks that come in different forms, and the consequences of these attacks could be serious.

The next generation of computing and communications architecture must be designed with the following considerations in mind:

(1) As logical functions are centralized in high-performance computing devices, the potential impact of an attack will become wider.

(2) The use of SOA-based services will create challenges related to critical service permission control and communication security.

(3) Open platforms will present persistent security risks for third-party software and hardware.

(4) In addition to conforming to laws and standards, intelligent vehicles of the future will also need to ensure the trustworthiness of both process and results.

The central computing platform is the last line of defense against attacks. The cyber security of the central computing platform needs to be approached from three major directions: platform security, in-vehicle security, and outside-vehicle security. The vehicle connects with the outside world mainly through sensors and network interfaces. Technical tools such as access authentication, intrusion detection and prevention, and AI security and attack defense are required to keep attackers out. Inside the vehicle, computing and control modules are the main targets of attacks. Access control, security isolation, security degradation, and diagnosis and recovery technologies can be used to ensure the system cannot be breached by hackers.

In short, cyber security involves a sophisticated attack-defense system and needs to be addressed on multiple fronts:

(1) Central computing platforms – the lifeline of intelligent vehicles – need to be protected through underlying designs of core technologies (e.g., chips, OSs, and encryption algorithms) and architecture; key technologies such as root of trust, encryption algorithms, trusted computing and OTA, and intrusion detection and isolation; better understanding, detection, and evaluation of AI uncertainty; and enhanced SOTIF-related capabilities.

(2) Carmakers need to work with tier-1 and tier-2 suppliers to build an in-depth defense system across the entire vehicle lifecycle to prevent remote control of vehicles by hackers, and avoid leakage of data stored locally or on the cloud.

(3) Intelligent vehicles need to be resilient. Cyber security and resilience measures should be put in place to allow a hacked system to continue providing stable vehicle control services, or isolate the hacked vehicle control services and perform deterministic degradation to let other unaffected vehicle control services continue to work effectively and safely.

Specifically, the safety of vehicles should be enhanced across their full lifecycle and from different aspects such as cyber security governance, customer requirements, architecture design, protection and defenses, anomaly detection, prompt responses, and fault recovery. In 2030 and beyond, the key technologies for vehicle safety and security will include AI-based on-board intrusion detection systems, third-party software supply chain security and runtime integrity protection, resilient architecture, intrusion tolerance, zero-trust-based continuous authentication and source tracing of third-party devices, and detection and defense of malicious signals from sensors. Resilient architecture and intrusion tolerance depend on the integration of cyber security and functional safety in both the design phase and runtime.

Design phase: In this phase, steps should be taken to prevent known attack types. For known attack types and failure modes, threat analysis and risk assessment (TARA) and hazard analysis and risk assessment (HARA) should be integrated, so as to support comprehensive cyber security and functional safety requirements and avoid direct conflict and overlaps

Runtime: Designers should assume that during runtime, hackers will sometimes be able to successfully breach the system. Based on the system mission, designers should further integrate security detection, functional safety and cyber security risk analysis, functional safety response, and the fixing of cyber security vulnerabilities, in order to prevent functional failures caused by cyber attacks.

2）Data security and privacy protection

Intelligent connected vehicles will exist in a complex connectivity environment consisting of people, vehicles, roads, clouds, and networks. Data is becoming an essential resource, which is not only driving the service innovation and development of intelligent connected vehicles, but is also making data an integral part of our everyday lives. As data security of intelligent connected vehicles becomes increasingly important, corresponding governance systems will be incrementally optimized in all aspects, from data security awareness and data environment security control to data O&M security control, data asset security control, and data application control. Intelligent connected vehicle manufacturers, suppliers, and service providers must all improve data security capabilities. This includes implementing technological protections (e.g., user authentication, data encryption, access control, application management, intelligent anonymization, and network protection), and building systemic data security and compliance throughout organizational and product development, transactions, business commitments, and services. Specifically, the following principles should be applied to the use of data: lawfulness, fairness, and transparency; purpose limitation; data minimization; accuracy; storage limitation; integrity and confidentiality; and availability.

Sensitive personal data should be centrally managed to ensure security. Full-lifecycle privacy protection measures that cover vehicle systems, cloud big data platforms, backend service platforms, and data processing by third parties must be put in place. There are three major technical measures that can make this happen:

(1) Intelligent in-vehicle processing: Sensitive personal data should not be sent to the cloud. Instead, intelligent functions, such as voice and facial recognition, should be realized in the vehicle.

(2) Data invisibility: Technologies such as homomorphic encryption and differential privacy can minimize the use of sensitive personal data.

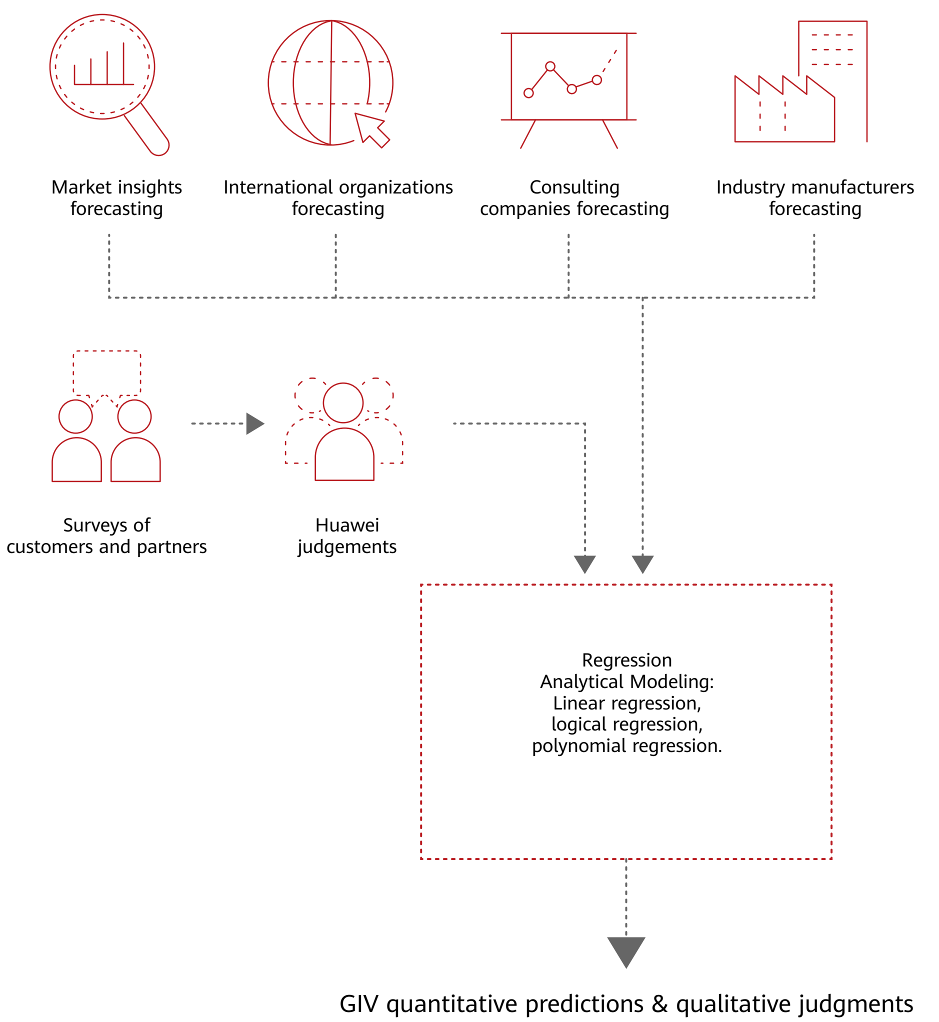
(3) End-to-end privacy: Privacy protection should be integrated into and prioritized throughout the entire data lifecycle from personal data collection and use to its retention, transmission, disclosure, and disposal, and the entire process should be transparent, well structured, strictly controlled, and traceable.

Data Prediction Methodology

Humankind has always been driven to explore, and bring certainty to this uncertain world.

We believe that collaboration is essential for future exploration. Over the past three years, in addition to our own Huawei experts who have been building their own insights and plans for the future, our GIV@2030 research team has also been communicating extensively with industry scholars, customers, and partners. We have studied data, methodologies, and insight reports from prominent players around the world, including the United Nations, GSMA, and other third-party consulting firms.

The GIV@2030 research team used trend extrapolation and time series forecasting to arrive at their findings about the future.



Definitions of the Metrics

Category	No.	Indicator	Definition	Prediction for 2030
Connectivity	1	Global connections	Total number of connected people and things worldwide	200 billion connections
	2	Wireless (cellular) connections	Total number of people and things connected through wireless (cellular) technology worldwide	32.5 billion cellular connections 80 billion passive cellular connections
	3	Gigabit/10 gigabit fiber broadband household penetration	Households using gigabit or higher fiber broadband as a proportion of total households worldwide Households using 10 gigabit fiber broadband as a proportion of total households worldwide	Gigabit or higher fiber broadband household penetration: 55% 10 gigabit fiber broadband household penetration: 23%
	4	4K/8K TV household penetration	Households with 4K/8K TVs as a proportion of total households worldwide	4K TV household penetration: 58% 8K TV household penetration: 17%
	5	IPv6 adoption	Connected devices and applications using IPv6 worldwide	90%
	6	VR/AR users	Number of VR/AR users worldwide	1 billion
	7	Robots per 10,000 workers Household robots	Number of robots per 10,000 workers Households with home robots as a proportion of total households worldwide	390 robots per 10,000 workers Households with home robots: 18%
	8	Intelligent connected vehicle penetration	C-V2X vehicles as a proportion of total vehicles	60%
	9	Average monthly data use on wireless cellular networks per person	Average monthly data (GB) used on wireless cellular networks per person (all owned devices)	600 GB (40-fold increase)
	10	Average monthly fixed network data usage per household	Average fixed network data usage per household, per month	1.3 TB (8-fold increase)
	11	5G private networks	Number of 5G private networks (including virtual private networks) Proportion of medium or large enterprises using 5G private networks	Number of 5G private networks: 1 million 5G private networks penetration in medium/large enterprises: 35%
	12	Gigabit/10 gigabit Wi-Fi network penetration in enterprises	Proportion of enterprises using gigabit Wi-Fi networks and 10 gigabit Wi-Fi networks	Gigabit Wi-Fi network penetration in enterprises: 50% 10 gigabit Wi-Fi network penetration in enterprises: 40%
	13	FTTR/FTTD penetration	Fiber-to-the-room households as a proportion of total households worldwide/Fiber-to-the-desk enterprises as a proportion of total enterprises worldwide	Fiber-to-the-room household penetration: 31% Fiber-to-the-desk enterprise penetration: 41%

Category	No.	Indicator	Definition	Prediction for 2030
Computing	14	General computing power & AI computing power	Total general computing power of global data centers and edge servers Total AI computing power of global data centers and edge servers	General computing power (FP32): 3.3 ZFLOPS (10-fold increase) AI computing power (FP16): 105 ZFLOPS (500-fold increase)
	15	Computing as a proportion of total enterprise IT investment	Investment in general computing and AI computing servers as a proportion of total annual IT hardware investment of an enterprise (The IT hardware investment include investment into data centers, edge, and device hardware)	32%
	16	Cloud services as a proportion of total enterprise application expenditure	Proportion of enterprise application expenditure on cloud services	50%
	17	Global data volume generated annually	Total volume of data generated by all digital devices and digital instruments worldwide each year, including devices, edge, and data centers, as well as newly generated, replicated, or cloned data, back-up data, process data, and cached data.	1 YB (23-fold increase)
	18	Penetration of privacy enhancing computing technology	Proportion of scenarios where the built-in privacy computing capability of the computing system is enabled	50%
	19	Blockchain technology penetration	Proportion of enterprises deploying blockchain technology	85%
Vehicle	20	Electric vehicles as a proportion of new vehicles sold	Electric vehicles as a proportion of new vehicles sold annually	50%
	21	Autonomous vehicles as a proportion of new vehicles sold	Autonomous vehicles as a proportion of new vehicles sold annually	Global: 10% China: 20%
Energy	22	Proportion of renewables in global electricity generation	Proportion of electricity from renewables in the total electricity generated worldwide	50%