

## IT titans, 6 major industries under spotlight at Computex 2024

Max Wang, Taipei;  
Jessie Shen, DIGITIMES Asia

The highly anticipated Computex 2024, the apex event of the technology sector, is taking place in the Nangang Exhibition Hall from June 4 to June 7.

This year's Computex has been exceptionally busy. Aside from the CEOs of significant worldwide firms such as Nvidia, AMD, Qualcomm, Intel, and Supermicro, who delivered presentations at the conference, this year's show hosts 1,500 industry professionals. 4,500 booths are showing off their research and development breakthroughs, innovative technologies, and solutions in fields such as artificial intelligence, advanced communications, future transportation, virtual reality, sustainable energy, and innovation.

### AI-capable devices

Observing the technology industry's development trends this year, we see that generative AI technology is maturing and spreading from the cloud to edge applications. This year, the technology industry has mostly used AI as the foundation of business and product marketing.

In addition to AI PCs, AI mobile devices and wearables, AI display terminals, smart interconnected robots, and other new products, there is a growing demand for efficient processors,

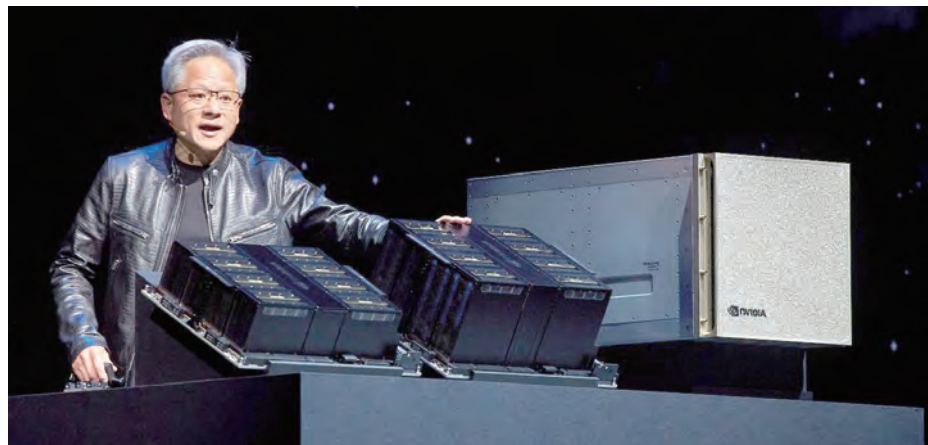
large-capacity memory, high-speed transmission components, and so on. AI PC and AI mobile phone shipments are forecast to surpass 45 million and over 100 million units, respectively, in 2024.

Furthermore, as generative AI introduces a new set of novel applications, the demand for computation in data centers has increased, making AI servers with new cooling architecture, higher energy efficiency, and optimized cabinet design solutions a market favorite. At this year's Computex, Hon Hai (Foxconn), Quanta, Wiwynn, Inventec, Compal, and many other manufacturers are demonstrating their latest AI server solutions.

### EVs

Products and solutions pertaining to EVs are another focus during this year's Computex. The market demand growth rate for EVs this year may be lower than in 2023, despite being influenced by the base period increase, changes in various nations' policies on electric vehicle subsidies, and the expected deferral of Europe's fuel vehicle phaseout plan.

The International Energy Agency (IEA) reports that global EV sales reached 14 million units in 2023, expanding at a pace of 35% annually and making up more than 18% of all automobile sales. Sales are predicted to increase to 17 million in 2024, a 20% increase over the previous year and 20% of all car sales.



AI has become the key spotlight of Computex 2024 with Nvidia CEO Jensen Huang leading the crowd

Furthermore, the global total of public charging piles was nearly 4 million in 2023. The number of new public charging piles will reach 1.62 million this year, including 960,000 DC piles and 660,000 AC piles, as EV sales continue to rise, according to MIC statistics.

The growing demand for electric vehicles has prompted a number of companies to showcase their newly developed products and services related to EVs at this year's Computex.

DFI, Delta, and other companies are showcasing their AI smart charging pile solutions at the event. In addition, Pegatron, Carota, Altek, PSMC, Clientron, MSI, and Sky Genie Auto are presenting their smart driving platform information system, smart cockpit vehicle system, fleet management solution, motor electronic control system, OTA test equipment, enhanced ADAS assistance system, Instrument Control Unit (ICU) and digital instrument, vehicle information security verification, and

autonomous driving solutions.

### Other themes

As 5G development progresses, many telecom operators have shifted to the architecture of next-generation communication technologies including B5G, 6G, Wi-Fi 7, FWA, and low-orbit satellites. They are actively planning and demonstrating their capabilities at this year's Computex.

This year's Computex innovation and start-up exhibition area is noteworthy, as it features the participation of 400 startups from over 30 countries. These companies are engaging in discussions on many issues including AI, green technology, smart transportation, and semiconductor applications.

In addition, many companies are showcasing innovative green energy and environmental protection technology, including renewable energy systems, intelligent energy-saving equipment, and solutions for environmental monitoring.

KIOXIA

## KIOXIA Memory Maker

Making Flash Memory Solutions for Every Application

X0001, 3F, TaiNEX Hall 2



Global Leader in  
AI-Powered Cybersecurity

## Connecting AI Security

📍 AI Security Cafe  
R103/R104, TaiNEX 1  
(next to 1F Area J Entrance)



Find out more on  
the last page



# AI chips for all: DEEPX CEO Lokwon Kim’s vision to democratize AI technology

News highlights

The applications of AI are becoming increasingly widespread in our daily lives. With the continuous iteration of generative AI and large language models, various innovative applications are emerging. These include, such as smart home appliances, autonomous vehicles, and robots to VR/AR applications.

These edge devices are equipped with AI processors that compute on the device directly, further accelerating the launch of innovative Edge AI applications. However, these devices also face several challenges, including excessive GPU power consumption and high costs limiting the widespread adoption of Edge AI products.

DEEPX, an AI chip startup from Korea, won three CES Innovation Awards 2024 in January for its unique AI chip ultra-gap source technology. The awards were in the categories of Computer Hardware, Embedded Technology, and Robotics, and Computer Hardware & components.

Their NPU processor boasts low power consumption and cost-effectiveness and addresses the issue of insufficient accuracy found in existing NPUs on the market. This technology was recognized by lots of semiconductor companies during CES 2024 and is scheduled for mass market release in late 2024.



DEEPX CEO Lokwon Kim Photo: DEEPX

## Learn from ARM to build DEEPX into a leading On-device AI company

“The main battlegrounds in the AI era will move to the ‘edge’. Just as ARM dominated the CPU market with smartphones, the semiconductor company that dominates the edge will dominate the AI market,” DEEPX CEO Lokwon Kim said.

Kim, a former senior researcher in Apple’s Application Processor (AP) design, returned to South Korea to establish DEEPX after gaining extensive experience



DEEPX won three CES Innovation Awards in 2024 Photo: DEEPX

in semiconductor design at renowned companies like Broadcom, Cisco, and IBM T.J Watson. His goal is to build DEEPX into a leading AI company in the era of on-device AI, reminiscent of how ARM revolutionized the CPU market with its low-power technology.

On-device AI, which processes information within a mobile device without connecting to a server or cloud, is a burgeoning field. ARM broke Intel’s dominance in the CPU market with its efficient, low-power processors, which are now prevalent in smartphones and expanding into PCs and servers. Kim aspires for DEEPX to have a similar impact on the AI semiconductor industry.

Kim identified a weakness in Korea’s semiconductor ecosystem, particularly in system semiconductors. Drawing inspiration from Morris Chang, the founder of TSMC, who returned to Taiwan to establish a leading foundry after learning from the American semiconductor giant TI, Kim saw an opportunity to address these gaps in the South Korean market.

Chang predicted a demand-driven semiconductor market, leading to the creation of TSMC, which produces semiconductors on consignment. Similarly, Kim believes that the opening of the AI semiconductor market offers a chance for innovation and growth.

## Democratizing AI Technology

can utilize.

To pursue this vision, Kim established DEEPX in Pangyo, South Korea’s hub for fabless companies. He aims to create an AI semiconductor company with a unique value, promoting innovation over exclusivity.

DEEPX is positioned not as a competitor to global semiconductor giants but as a complementary force, enhancing the global semiconductor landscape by providing accessible and advanced AI technology.

## Analyzing the Difference Between NPUs and GPUs: DEEPX’s Unique Solution

The necessity for developing specialized AI semiconductors, such as NPUs, stems from the limitations of GPUs, which have traditionally been used for AI computations. GPUs, originally designed to process graphic data, excel at handling large amounts of data simultaneously, making them suitable for AI learning tasks.

However, their high power consumption and operational costs present significant drawbacks. This makes them less ideal for “edge AI,” which involves running AI applications directly on devices like controllers, robots, and self-driving cars - collectively referred to as the “edge.”

NPUs (Neural Processing Units) are modeled after the human brain, offering the benefits of lower power consumption and reduced production costs. However, existing NPUs have struggled with accuracy and support for the latest AI algorithms. DEEPX stands out among fabless companies by addressing the core challenges of AI semiconductors in one comprehensive solution.

Unlike typical NPU vendors that release a single chip, DEEPX recognizes that different electronic devices require varying levels of semiconductor capabilities. For instance, AI for closed-circuit television (CCTV) primarily needs to analyze video, whereas AI for robots involves far more complex computations.

To address this, DEEPX has developed from low-end to high-end performance four chips at once: one that can connect a single electronic device for AI computations, and another that can link three or four devices for broader AI tasks. This universality across devices is a key reason DEEPX received the Innovation Award at CES 2024.

DEEPX’s commitment to high performance at lower power and cost also earned them the Innovation Award. The company’s latest AI algorithm, Yolo7, runs on their semiconductor DX-V1, produced using Samsung’s 28nm process.

This algorithm was previously incompatible with conventional NPUs. In addition, the DX-M1 chip boasts a design area one-third the size of other NPUs, and its manufacturing cost is similarly reduced by one-third. Combining low unit costs with high performance in a low-power NPU, DEEPX’s products are poised to lead the AI semiconductor market.

NPUs are categorized into data center-based NPUs, which handle large-scale inference, and edge-type NPUs, designed for use in electronic devices such as robots, smart cameras, smart factories, consumer electronics, etc. DEEPX targets both data center and edge NPU markets. DEEPX’s NPUs overcome the common shortcomings of existing NPUs by providing high accuracy and efficiency.

DEEPX’s innovation lies in creating NPUs that are not only small and cost-effective but also achieve accuracy comparable to, or even better than, GPUs. This success is attributed to DEEPX’s pioneering work in two core technologies: IQ8 (an INT8 model compression technology), and Smart Memory Access(minimizes D-RAM usage). DEEPX leads the market for low-power AI solutions, achieving the world’s highest power-to-performance ratio through proprietary advancements in hardware and software optimization

## Advancing Towards Mass Production

DEEPX has successfully demonstrated its original technology with sample units and is in the final stages of preparing its mass-production chip. Scheduled for market release in late 2024, the widespread adoption of products featuring DEEPX’s chips could establish the company as a technology leader in the on-device AI market by 2025.

The essence of on-device AI lies in low power consumption and the seamless integration of hardware and software. Since AI must operate on small devices, minimizing power consumption is crucial while maximizing AI performance within limited computing power.

Kim, drawing from his experience at Apple, emphasizes the importance of designing hardware and software with equal priority from the outset. Unlike Apple, which develops its own devices and services, DEEPX engaged with approximately 700 customers



DEEPX CEO Lokwon Kim Photo: DEEPX

during product development to understand their needs and find the optimal development point. This customer-centric approach has been pivotal in refining DEEPX’s products.

Four companies from Israel and the U.S., including DEEPX, are vying for dominance in the emerging on-device AI market. As the market begins to flourish beyond the server market, this competition is crucial for setting future momentum.

Kim is confident in DEEPX’s strategy. He notes that while some competitors prioritize rapid product releases, DEEPX focuses on price and optimization first. Given the rapid evolution of AI application services, Kim believes it is more important to align with market trends and offer high-performance features at a competitive cost rather than rushing products to market.

DEEPX’s innovative approach to developing versatile, efficient, and cost-effective NPUs positions the company as a formidable contender in the on-device AI market. By prioritizing customer needs and maintaining a balanced focus on hardware and software optimization, DEEPX is set to lead the industry with cutting-edge AI semiconductor solutions.

## DEEPX Company’s Philosophy: Valuing Technology Over Short-term Results

DEEPX is not focused on making AI semiconductors for autonomous cars or smartphones. Instead, DEEPX aims to advance the integration of AI into everyday life. The company’s products are designed to bring AI to areas such as CCTV and robots, pushing these technologies beyond outdated algorithms. DEEPX’s mission is to pioneer advancements that should have already been made in the AI semiconductor industry.

At the entrance of DEEPX, there is a note to employees that emphasizes the value of technology over monetary gains. It quotes Carl Sagan’s “pale blue dot,” reflecting the idea that life should be lived with value rather than chasing money and power. The message encourages the 70 employees to love their work and find it meaningful, aligning with the company’s ethos of creating technology for the greater good.

DEEPX believes in ‘technology that everyone shares.’ From a management perspective, AI is seen as one of humanity’s final inventions, marking the endpoint of human evolution. The goal is not to monopolize technology or chase profits but to lead the way in making AI accessible and beneficial for all. DEEPX’s AI semiconductors are the cornerstone of this vision, enabling widespread adoption of AI.

## Future Plan: Global Expansion and Innovation

DEEPX’s visionary CEO, Kim, has outlined a strategic plan for the future:

- Global Market Entry: Starting from the second half of this year, DEEPX plans to aggressively enter the global market with its first-generation product, consisting of four AI chips. This move is set to usher in the era of “AI Everywhere.”
- Technological Innovation: DEEPX aims to develop new technologies that enable super-scale AI services with power consumption of less than 5W. This innovation will make advanced AI technologies more accessible and practical for widespread use.
- Leadership in AI: DEEPX is committed to becoming a leading comprehensive AI chip company globally. By focusing on power and cost efficiency, the company seeks to provide core technologies that transition giant AI advancements from the realm of science to everyday applications.

DEEPX’s mission is to integrate AI into everyday life, advancing technology in meaningful ways. By valuing innovation over profit and aiming for global leadership, DEEPX is set to play a pivotal role in the future of AI. The company’s commitment to shared technology and global expansion highlights its dedication to making AI accessible and beneficial for humanity.

Join DEEPX’s Early Engagement Customer Program (EECP), and don’t miss out on their innovative products designed to enhance your AI capabilities. Discover how over 100 global companies, including Hyundai Kia Motors Robotics Lab, POSCO DX, Supermicro, and Dell, leverage their hardware and software to power their next-generation AI products.

For more information, you can follow DEEPX on social media or visit their official website.



InnoVEX 2023 Pitch Contest  
Startup Terrace Awards Winner



## Experience the Future of AI with DEEPX at Computex 2024

Taipei Nangang Exhibition Center  
4 ~ 7 June | Hall 1, 4F **#M0401A**



### Scan and Join Our Early Engagement Customer Program (EECP)

Discover how over 100 global companies, such as Hyundai Kia Motors Robotics Lab, POSCO DX, Supermicro, Dell, Kaytus, Inventec, Advantech, IEL Integration Corp., Innodisk, Aetina, Lex System, etc., leverage our hardware and software to power their next-generation AI products.



# Nordic semiconductor demonstrates integrated short-range, Wi-Fi and cellular wireless communication innovations at Computex 2024

## News highlights

The savvy smart location tags, asset trackers, connected health sensing devices, smart home appliances, advanced vehicles, and smart electric vehicle systems show that the popularity of Internet of Things (IoT) devices has penetrated nearly every aspect of our lives. The wireless communication technologies, coupled with powerful cloud computing and Artificial Intelligence (AI) technologies, play a significant role in driving technological advancement.

Nordic Semiconductor, a leading global provider of wireless IoT solutions, provides an exclusive product and technology showcase during Computex Taipei 2024. Strongly demonstrating the Nordic’s product series with short-range, medium, and long-distance wireless communication technologies, helps customers to develop modern devices and notable applications while continuously improving by delivering low-power multi-protocol wireless technologies.

In this interview with Mr. Richard Chen, regional sales manager of Nordic Semiconductor, he indicates that product development engineering teams are consistently struggling to implement the well-balanced features between connectivity, security, power saving, and other fancy functions while consumers’ endless demand for intelligent, sophisticated and instant-response IoT devices. These difficult trade-offs are achieving longer battery life and lower energy consumption in devices.

To help customers solve these tough problems, Nordic displays a series of wireless communication products and application demonstrations at the Taipei W Hotel in Xinyi District between June 4 to 7 2024. It covers several important product highlights below in this article. There are key product information and major demos on the following web pages of <https://www.nordicsemi.com/Events/2024/Computex-Taipei> for further reference.

### Short-ranged smart home applications for Matter and Bluetooth LE Audio + Auracast features

Nordic is set to showcase a range of smart applications based on Matter standard and Bluetooth wireless technology. The demos are highlighting Matter-over-Thread / Matter-over-Wi-Fi use cases. The booth shows an integration of Nordic’s development kits and devices including models of nRF54H20 DK, nRF54L15 DK, and nRF7002 EK.

Developed by Connectivity Standards Alliance (CSA), Matter is a smart-home connectivity standard that ensures connected devices interoperate reliably. Industry leaders such as Apple, Amazon, and Google have promoted Matter standards and are supported by their IoT ecosystems.

The recent released Matter 1.3 standard on May 8, 2024 brings support for more appliances like kitchen robotics.

More supported devices across ecosystems require bigger software to handle it. One of the strengthening features of Nordic solutions is the enlarged Flash memory size to satisfy customers to accommodate big software stacks and update firmware easily through Over-The-Air technology.

Another advantage of Nordic solution is the cyber security features like nRF9160 low power system-in-package (SiP) and nRF5340 system-on-chip (SoC) achieved Platform Security Architecture (PSA) Certified Level 2. The latest models such as nRF54H20 and nRF54L15 SoC have passed PSA Certified Level 3. These certifications provide Nordic customers with an assurance offering a secure platform to build IoT



Capable of supporting Bluetooth 5.4 and future Bluetooth specifications, plus LE Audio, Bluetooth mesh, Thread, Matter, and more, the nRF54H20 will be the foundation for a new wave of revolutionary IoT end-products. Its combination of advanced features will enable complex end-products that have previously been unfeasible.

products ensuring that all devices on the network are communicating securely and privately.

Meanwhile, the demos of Nordic’s Bluetooth Low Energy flagship SoCs are another spotlighting focus. Nordic’s award-winning Bluetooth LE series solutions will be displayed at the booth based on the nRF5340 Audio DK development kit delivering both low-energy and high-quality wireless audio as well as the broadcasting capabilities of Bluetooth LE Audio with Auracast.

Bluetooth LE Audio is the next generation of wireless audio streaming technology. The stand will also play host to a Bluetooth LE Audio and Auracast demonstration.

Bluetooth LE Audio pursues advanced sound quality with lower power consumption and reduces the delay between audio and visual experience to improve the immersive experience. Moreover, Auracast technology uses broadcast technology to share audio streaming services with multiple people.

In addition to hearing aid applications, Auracast is also aiming at some interesting use scenarios such as museums and art exhibitions navigation services to open up unlimited possibilities for broadcast applications in the future.

### Low powered Wi-Fi 6 and Cellular wireless communication applications

Nordic builds the mid-range product portfolio by introducing low-powered Wi-Fi 6 technology to explore new applications for high-speed data communication in the edge nodes. Taking edge computing use cases as an example, Chen described, the IoT sensors transfer sensing information wirelessly on the edge side to cloud servers for connecting Nordic’s Thingy:53 IoT machine learning platform for training a model in the cloud. The model can then be deployed straight back onto the SoC on the edge side for deployment and inferencing. The entire process is quick and requires no coding knowledge.

There are further examples in the real world. For example, the sensing audio data or abnormal vibrations

of industrial fans or motors in the factory can be captured for AI training in the cloud. After downloading the AI models to the edge SoC, the systems will have the prediction features to identify and predict the motor systems’ possible failures and malfunctions to ensure that the right motor maintenance strategy is implemented at the right time.

Talking about long distance wireless applications, Nordic focuses on cellular LPWAN technologies to connect devices through 4G/LTE, and 5G communication standards. The product series are ranging from nRF91 SiP product lines to newly released low-power cellular SoC chips to support LTE-M and NB-IoT protocol.

The product combination provides major benefits of low power consumption, cost-effectiveness, and size reduction due to the integrated multi-protocol cellular modem and transceiver. Giving battery-driven devices better prospects, Nordic cellular LPWAN solutions represent a leap forward in connecting remote meter-reading networks for water, gas, and electricity.

In the Nordic suite, the guests will witness the convergence of GNSS, Cell-based, and SSID-based Wi-Fi locationing in the demo stations. Using the nRF Cloud Location Services alongside DECT NR+ nodes, this showcase demonstrates asset tracking with precision and accuracy. Running predictive maintenance ML models by Edge Impulse’s ML software, there is a further demonstration station to show Thingy:53 IoT machine learning prototyping platform to allow guests to examine and feel the prediction functions.

In the demos of Human Interface Devices, Nordic uses Bluetooth LE gaming mouse device to emulate high-speed automatic packet transmission to a gaming PC. This gaming mouse delivers up to 8,000 data reports per second, ensuring high speed and minimal delay in intense



IoT shows Nordic’s goal to streamline cellular product development and support the entire product lifecycle. This is why we have integrated all the different parts of our cellular offering into a complete solution – a fully Nordic-owned and controlled offering that includes hardware, software, tools, cloud services and our world-class support.

and fast-paced gaming scenarios. This improved user experience that hasn’t been seen before.

As a pioneer of Bluetooth LE technology, Nordic has evolved into a full-spectrum wireless IoT company offering a range of solutions based on leading industry standards including LTE-M, NB-IoT, DECT NR+, Wi-Fi 6, Matter, and Thread, says Chen. Nordic’s diversified offering at Computex 2024 is showing greater market expansion not only to maintain a significant market share in Bluetooth Low Energy but also to expand their cellular IoT offerings. Nordic hopes to work with Taiwan’s electronics supply chains and ecosystem partners to embrace IoT business opportunities to spur growth in increasingly fast-paced markets.



Nordic is actively involved and one of the main contributors the development of Matter. Matter aims to make it easy for developers to create a secure and reliable solution. If you want your products to be interoperable with the major smart home ecosystems, Matter is the way to go.

The Synaptics logo, featuring a stylized blue leaf-like shape to the left of the word 'synaptics' in a white, lowercase, sans-serif font.

Three Synaptics chips are shown, each with the Synaptics logo and a stylized leaf icon. They are arranged in a slightly overlapping manner.

A white icon of a hand with a glowing blue point, representing sensing technology.

SENSE

A white icon of a microchip with a leaf inside, representing processing technology.

PROCESS

A white icon of a network node with three connections, representing connectivity technology.

CONNECT

A close-up of a hand interacting with a digital interface, with a glowing blue point of contact.

A detailed view of a microchip circuit, showing intricate patterns and a central processing unit.

A night cityscape with various digital overlays, including icons for Wi-Fi, cloud, and network, representing a smart city or IoT network.



# Fibocom builds AI-specialized “π” shape strategy upon 5G, AI, IoT to Empower Industry Transformation

## News highlights

### Remain Industry-focused and is confident to accelerate 5G Deployment

In the realm of 5G, this communication standard garnered significant market attention upon its release. The industry believes that 5G’s features such as high speed, low latency, and large-scale connectivity not only enhance user experiences in the consumer sector but also serve as a crucial driver for deepening communication technology applications across industries.

While the development of 5G has not been as fast as initially expected, communication giant Ericsson predicted in its 2023 report that the golden crossover period between 4G and 5G would be delayed until 2028. However, Tiger Ying pointed out that from the perspective of technical standards’ development speed, this delay is still within a normal cycle.

He mentioned that 5G’s current development is focused on Enhanced Mobile Broadband (eMBB), Massive Machine Type Communications (mMTC), and Ultra-Reliable Low Latency Communications (URLLC). However, large-scale commercialization has not been achieved. This trend also highlights the versatility of 5G IoT architecture and applications.

With the acceleration of 5G Advanced, it will truly replace 4G as the mainstream technology for IoT applications and successfully drive market development within five years.

Fibocom has responded to the trend of 5G development by formulating product strategies and solutions, actively expanding into different market sectors. Tiger Ying stated that the company’s primary task is to promote the large-scale commercialization process of eMBB.

They have already launched a diverse range of product portfolios. This includes the RedCap module, closely following the development of 5G technology and fully advancing the implementation of commercial plans.

In terms of application areas, Fibocom is focused on industrial intelligence and smart cities, primarily because these two major application scenarios have diverse demands for 5G technology. For instance, in machine vision, which integrates edge computing with 5G’s high bandwidth and low latency capabilities, it has been widely used in industrial quality inspection, promoting the development of industrial automation and intelligence.

In response, Fibocom provides a complete product portfolio, such as the 5G smart module SC171 with computing power up to 12TOPS, to meet specific demands in application scenarios. In addition, endpoint



Fibocom CEO Tiger Ying

devices with AI capabilities will also become a driving force in accelerating industrial intelligence processes.

Fibocom has invested significant research and development resources in these products, aiming to integrate 5G and on-device AI technologies, thus strengthening the company’s core competitiveness in industrial intelligence and smart city domains.

### Strengthening Edge AI Performance to Optimize Overall Cost-effectiveness

AI has become a recent global industrial focus, and Fibocom’s AI strategy centers on edge computing and the AI endpoint device solutions mentioned above. Since AI computing power can significantly increase the cost of endpoint devices, leading to reluctance from enterprises with implementation needs, cost control is crucial.

Tiger Ying pointed out that the solution to this problem varies depending on the AI architecture. Devices using an offline AI architecture require precise control of edge computing power to balance cost and effectiveness, while devices using an online AI architecture require strong communication capabilities to avoid affecting user experience due to communication delays.

Tiger Ying further stated that Fibocom has integrated the above two solutions into one through long-term research and development. The company has invested in smart module development, leveraging the computational efficiency and heterogeneous algorithm capabilities of modules to assist clients in rapidly deploying AI-capable terminal devices and optimizing costs.

Its product portfolio includes 5G data modems, SoC 5G solutions based on Linux, and Android, and built-in AI computing capabilities. He mentioned that while AI may slightly increase device costs, its excellent performance leads to overall cost optimization considering macroscopic aspects such as user experience and value creation.

**The wave of intelligence is revolutionizing industries globally, with 5G, AI, and IoT playing key roles in this trend.**  
**Fibocom, as China’s first stock-listed wireless communication modules and solutions provider (stock code: 300638), has continuously employed innovative thinking and strategic market positioning in recent years. CEO Tiger Ying pointed out that, regardless of the technological or application aspects, in facing the transformative effects brought by these three major technologies, the company is well-prepared and ready to assist clients in seizing vast smart business opportunities.**

He also mentioned that Fibocom has constructed heterogeneous computing on the network, enabling the scheduling of CPU, GPU, NPU, and DSP processor performance as needed. In addition, Fibocom has developed its toolchain and integrated it with heterogeneous algorithms, RTK, high-precision positioning, and other functions into firmware.

Its products based on Linux, Android, and Windows architectures can meet the needs of different industry customers. Looking ahead, Fibocom will focus on areas such as robotic lawnmowers, low-speed autonomous vehicles, PCs, and robotic development platforms, providing deeper vertical domain solutions.

### Enhancing π-shaped Capabilities to Meet Three Major IoT Demands

Regarding IoT, Tiger Ying pointed out that although

the IoT architecture and concept have been around for over a decade and have become increasingly widespread in recent years, practical constraints mean that adoption varies across industries. To accelerate IoT adoption, several key factors are necessary.

First is the ability to apply AI capability to IoT devices, as AI significantly impacts data collection costs and processing quality. IoT leveraging AI will be more readily accepted by enterprises, thus speeding up adoption in specific fields.

Second is security design, which has become a major focus in IoT in recent years. Only IoT architectures that can ensure data integrity and privacy protection will be able to deploy in vertical markets.

Lastly, communication technology is crucial. Advanced technologies like 5G-Advanced (5.5G) and Non-terrestrial Networks (NTN) ensure that communication coverage is no longer a weak point, broadening the scope of IoT applications.

Fibocom has introduced corresponding products to address AI, data privacy, and satellite communications needs. In terms of AI, the company’s 5G smart module SC171 and SC151 series are suitable for various 5G smart terminal devices.

For security, blockchain technology was deployed on 4G networks and commercialized. On the satellite communications front, Fibocom released NTN-supported communication modules in 2023.

Moving forward, the company plans to integrate these technologies and products, working with ecosystem partners to focus on specific vertical solutions, providing customers with high-quality and high-performance IoT solutions.

Fibocom consistently prioritizes intelligence in product innovation. Tiger Ying stated that this will continue to be the foundation for future development, with a market-driven approach to introducing cross-domain AI solutions.

He emphasized that while the market previously demanded “T-shaped” capabilities combining vertical telecommunication expertise to serve the horizontal industries, the AI era requires an additional vertical capability—AI specialization. Through “π-shaped”capabilities, which integrate horizontal and dual vertical expertise, Fibocom aims to create high-performance IoT architectures tailored to customers’ needs.

The company is progressively enhancing these “π-shaped” capabilities. It will continue to collaborate with ecosystem partners, integrating 5G, AI, and IoT technologies to help clients seize smart business opportunities.

# Innodisk expands edge AI applications and intelligent solution at Computex 2024 with new brand strategy

## News highlights

Innodisk, a leading global AI solution provider, is showcasing its comprehensive new products at Computex 2024 with a focus on edge AI. Key highlights include the industry’s pioneering CXL 2.0 memory expansion, E3.S SSD, and innovative MIPI over Type-C camera technology for machine vision.

Live demonstrations of smart manufacturing PPE recognition and intelligent people tracking will emphasize Innodisk’s custom edge AI capabilities. This event also unveils Innodisk’s new brand vision, “Architect Intelligence,” bringing out its commitment to building intelligent solutions with global partners.

### Expanding Industrial Product Lines to Meet AI and Edge Serve Standards

Reliable and high-performance memory and storage are essential for AI and edge computing, particularly as edge servers evolve and demand high data throughput

and resilience in extreme environments. At Computex, Innodisk will unveil its first industrial-grade CXL 2.0 memory expansion, which is aligned with advanced industry standards and offers high scalability.

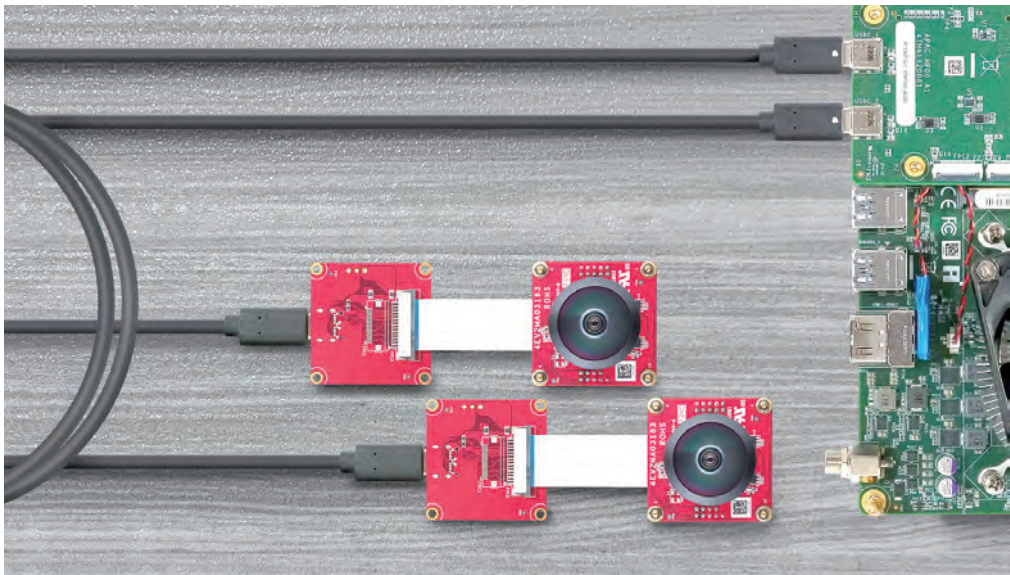
The showcase will feature award-winning E1.S Gen4 SSD series, alongside wide-temperature SSDs tailored for edge servers, such as the E1.S/E3.S, available in various form factors. The range also includes 16TB high-capacity SSDs designed to meet the evolving storage needs.

Further, Innodisk’s recent focus on embedded camera modules for AI machine vision development is crucial for advancing edge AI applications. Debuting at Computex, the innovative “MIPI over Type-C” camera technology converts signal through exclusive adapter board design and enhances platform compatibility, significantly extending the length of MIPI camera cables.

This unlocks applications in areas such as autonomous mobile robots (AMR) and shared mobility. Innodisk will also display a full range of camera module products for USB and MIPI interfaces.



Innodisk is showcasing its comprehensive new products at Computex 2024 with a focus on edge AI



Innodisk is showcasing the exclusive MIPI over Type-C technology at Computex 2024

### AI in Action: Smart Cities and Industrial Applications

Beyond industrial modules, Innodisk integrates AI computing technologies with software, hardware, and third-party technologies to create various edge AI solutions. At Computex, Innodisk will introduce “InnoTracking,” a smart people-tracking solution capable of accurately identifying and tracking specific individuals across multiple cameras and timeframes, enhancing public safety in cities, financial institutions, entertainment venues, and retail spaces.

Under the simultaneous trends of ESG and industrial automation, worker safety in hazardous environments becomes a crucial focus for enterprises. Innodisk’s PPE recognition solution uses an FPGA platform

to process real-time images from four cameras simultaneously, ensuring workers maintain adherence to safety protocols. In case of violations, the system promptly notifies supervisors with detailed information to initiate appropriate responses.

The “iCAP Air” air quality management solution addresses ESG requirements by combining patented air quality detection technology with sensors, edge servers, cloud management interfaces, and IoT ventilation controllers. This solution is suitable for smart factories and healthcare facilities, enabling real-time intelligent decision-making.

### “Architect Intelligence” – Innodisk’s Brand Strategy for the AI Era

With nearly two decades in the industry,

Innodisk has elevated its brand and become a global leader in product technology. The new “Architect Intelligence” concept aligns with Innodisk’s edge AI strategy, emphasizing the practical implementation of AI at the industrial edge.

Under this concept, Innodisk integrates various product lines into nine intelligent series, each represented by diverse colors, symbolizing the limitless potential of AI applications and Innodisk’s innovative spirit. Furthermore, these colored blocks represent the flexibility of solutions and collaborative effort to architect AI solutions with customers.

See the future of edge AI in actions! Visit Innodisk at Booth J0110, Hall 1, 1F, Nangang Exhibition Center from June 4 to June 7 to explore more.



# Skymizer launches ET2 IP solution, creating more possibilities for LLM with hardware and software platform

## News highlights

After announcing its foray into the LLM (Large Language Model) IP market, Skymizer recently unveiled a series of hardware and software solutions centered around LLM IP.

These offerings are designed to bring more imagination to the LLM application service market. Skymizer’s IP solution series is codenamed EdgeThought, with the first market-ready solution named ET2, capable of efficiently handling all current edge devices requiring LLM, including the recently released Llama3 with a parameter scale of up to 8 billion.

Before introducing the IP solution, Skymizer’s primary solutions were in the compiler space, bridging the gap between chips and software. This background

has endowed the company with extensive experience in overall system hardware and software integration and optimization. As the demand for LLM rises, Skymizer, leveraging its solid market foundation, has entered the IP market to cater to diverse vertical application needs.

**Launch of integrated hardware-software platform: ET2 features edge computing, LLM, and AI inference**

According to Skymizer Executive Vice President William Wei, ET2 encompasses three key elements: edge computing, LLM, and AI inference. Besides accommodating various LLMs in the market, ET2 can flexibly expand computing resources to meet client needs.

If the parameter scale of the LLM to be processed is

too large, expansion can achieve the required computing power, naturally increasing memory capacity and power consumption. In addition to the existing IP solutions, Skymizer also launched the SkyGenie SDK (Software Development Kit).

The SDK can address various categories of LLMs, including general, domain-specific, and private, assisting various industry applications. This enables software developers to create corresponding applications based on different LLM types, optimizing overall system performance. During COMPUTEX 2024, Skymizer will further demonstrate applications such as smart factories’ Autonomous Mobile Robots (AMRs), Drive-thru ordering smart assistants, and smart automotive scenarios using ET2 and other hardware-software solutions.

Wei emphasized that Skymizer’s comprehensive hardware-software system development experience allows the broad tech industry ecosystem to benefit from Skymizer’s complete platform solutions. He revealed that while the market currently sees edge AI GPUs performing at about 20 tokens per second, Skymizer’s ET2 tests show around 32 tokens per second, with implementation costs at just 1/100 of Edge AI GPUs. This high cost-performance ratio makes ET2 an ideal choice for cost-sensitive end applications.

**First Chip with ET2 solution to debut at CES 2025**

In the semiconductor domain, Wei candidly shared that Skymizer is open-minded and actively collaborating with domestic and international design services and IP firms. He also revealed that ET2 is highly expandable, from small IoT MCUs to high-performance Edge Servers.

When paired with higher bandwidth memory interfaces, it can function as a server-level inference engine for multi-user, multi-batch processing. The first chip adopting ET2 is expected to debut at CES 2025, positioning ET2 as a game-changer for edge device LLM inference.

**Conclusion**

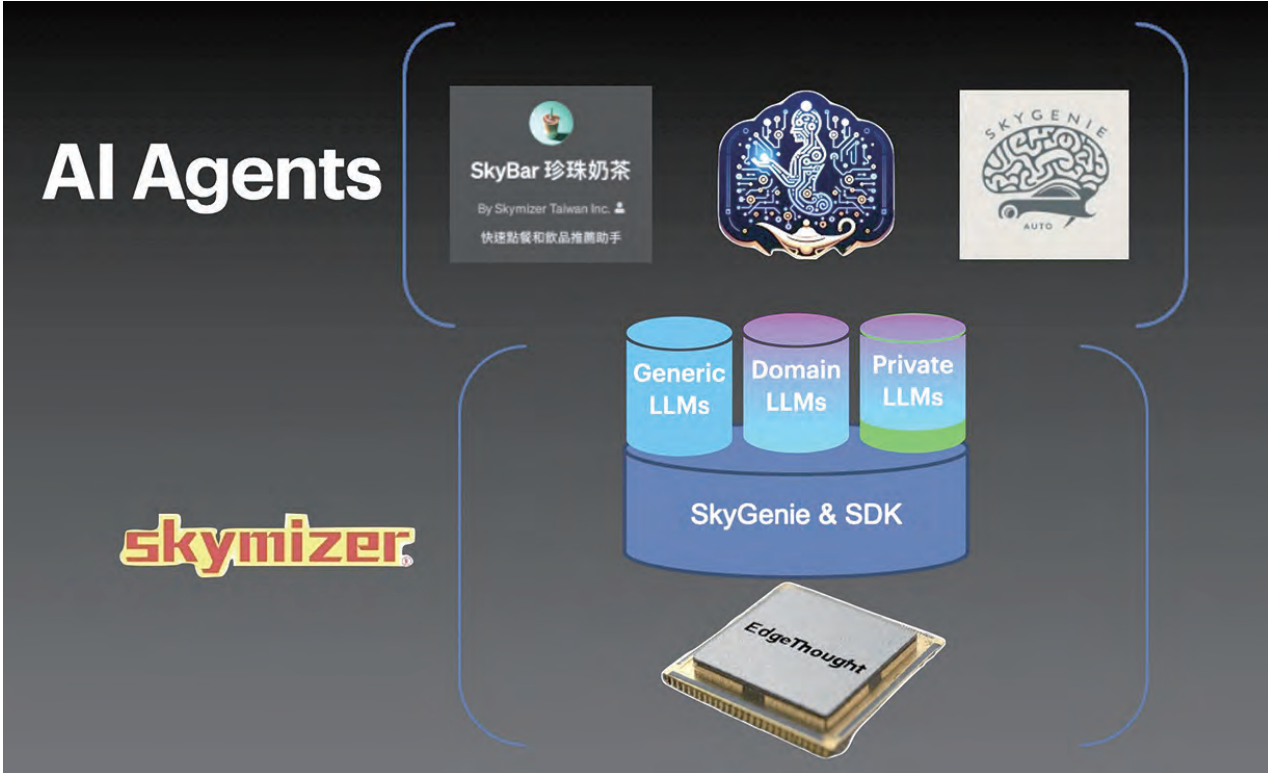
Skymizer’s launch of its first IP solution, ET2,



Skymizer Executive Vice President William Wei

marks a new milestone for the company. With rich hardware-software integration experience, Skymizer not only provides high-performance, cost-effective edge computing solutions but also equips developers with powerful tools through the SkyGenie SDK to meet diverse LLM application needs.

These innovations enhance the performance of different vertical market applications and create new possibilities for smart factories, autonomous mobile robots, smart automotive, and more. Skymizer’s comprehensive platform solutions are poised to be a significant driving force in the future LLM application market.



Skymizer not only provides LLM silicon intellectual property solutions but also offers the SkyGenie SDK, supporting various types of LLMs. This makes it easier and faster for AI application developers to create applications, enabling hardware chip partners to achieve higher integration and better meet market demands

## VIA Labs launches PortSense™: AI-ready features for enhanced USB hub management

## News highlights

At Computex 2024 in Taipei, Taiwan, VIA Labs announced PortSense™, a suite of manageability and intelligent connectivity features for USB Hubs that sets a new standard for docking station functionality in business and professional environments.

PortSense™ is an exclusive VIA Labs hardware and software solution embedded in the latest revisions of VIA Labs hub products, and it enables supported products to retrieve USB descriptor information from connected devices, even without a host system. USB descriptors contain vital details about the connected USB devices, such as the device class and capabilities, the product name and manufacturer information, serial number if present, and much more.

Typically, a host system uses this information to identify, configure, and interact with connected devices. However, with PortSense™, a managed docking station can collect usage data, perform tasks like pre-configuration and device inventory, and assist in implementing corporate policies.

PortSense™ is available in VIA Labs VL817 USB 5Gbps Hub and VL822 USB 10Gbps Hub. The VL832 USB4 Endpoint Device has an integrated USB 10Gbps hub and supports PortSense™.

While PortSense™ can function in autonomous mode, its true potential is unlocked when integrated into a connected platform where AI could be applied for analytics and policy control. Hubs with PortSense™ can communicate with an external controller using a standard I²C interface to share collected USB descriptor information and offer a

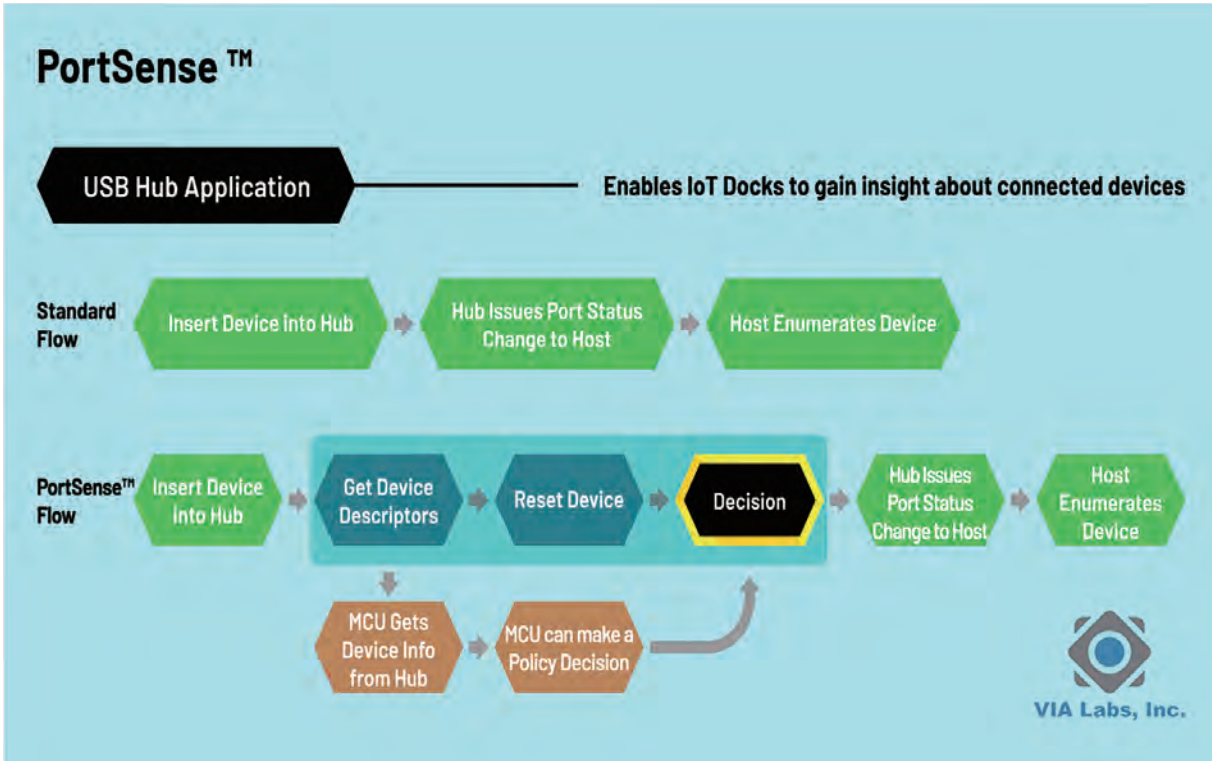
range of manageability controls, such as enabling or disabling ports, changing connection speeds, resetting devices, and toggling USB battery charging.

These controls can be applied on a per-port basis, including the upstream port, providing granular control over each connection. PortSense™ can be used to collect detailed user activity and enable analysis of usage patterns over time, making it perfect for hot-desking setups or as part of a comprehensive office management solution.

With PortSense™, VIA Labs is enhancing the modern workspace with more intelligent and connected solutions. The advanced capabilities of PortSense™ extend beyond basic data collection and control, providing valuable insights that help businesses manage their USB peripherals more effectively.

For instance, IT administrators can use PortSense™ to maintain an inventory of connected devices, ensuring that only authorized devices are in use. In addition, the AI-Ready features of PortSense™ support enforcing security policies, such as maintaining an allowlist or blocklist of devices or limiting the use of specific USB device classes in sensitive areas.

By acting as one layer of a comprehensive security strategy, PortSense™ helps reduce the risk of data breaches and unauthorized data transfers. This approach enhances operational efficiency and security, helping companies maintain a reliable and secure technology infrastructure. VIA Labs has recently released a white paper with more technical details about PortSense™, which can be found here: [https://www.via-labs.com/pressroom\\_show.php?id=98](https://www.via-labs.com/pressroom_show.php?id=98)



VIA Labs PortSense™: An exclusive suite of manageability and intelligent connectivity features for USB Hubs

## KIOXIA Taiwan to showcase advanced memory technologies and ecosystem applications at COMPUTEX

## News highlights

COMPUTEX Taipei 2024 will be held from June 4 to June 7 at Taipei Nangang Exhibition Center, Halls 1 and 2. KIOXIA Taiwan will showcase its full range of memory products and technologies under the theme of “Memory Maker – Making Flash Memory Solutions for Every Application” at booth X0001 on Floor 3 of Hall 2 in the Semiconductors & Hospitality Suites.

In addition to exhibiting KIOXIA’s memory products, the booth will feature joint displays with multiple ecosystem partners to highlight the critical role KIOXIA plays in the related ecosystem. This exhibition underscores KIOXIA’s indispensable position in the current technological landscape and the significant benefits its products and technologies provide.

**Business Memory Solution**

In the business memory solution area, KIOXIA will showcase cutting-edge flash memory technologies, including the 8th generation BiCS FLASH and CMOS directly Bonded to Array (CBA) architecture, alongside memory products for various applications.

**Business Solid State Drive Solution**

The area of business solid-state drive solutions (SSD) will feature a range of current and new products. The highlight will be the data center grade SSD - CD8P series, which utilizes the PCIe 5.0 interface and comes in both 2.5-inch and E3.S form factors.

Optimized for the performance, latency, power consumption, and thermal requirements of data center environments where power and cooling efficiency is critical. KIOXIA’s CD8P Series provides the predictability

and consistency needed for a seamless user experience.

In addition, several leading ecosystem partners will showcase their server systems compatible with KIOXIA products, providing visitors with a comprehensive understanding of the vital role and applications of KIOXIA’s business SSDs in nowadays’ large-scale computing demands. As one of the leading manufacturers in memory technology in business SSD, KIOXIA is focused on enhancing the performance, offering stable and reliable memory solutions.

**Personal Memory Solutions**

In the personal memory solution area, KIOXIA will exhibit a variety of SSDs, memory cards, and USB flash drives. A key highlight will be the debut of the EXCERIA with Heatsink SSD, scheduled for release in the second half of 2024.

This PCIe 4.0 SSD offers a sequential read speed of up to 6,200MB/s and is designed for PC systems with M.2 interfaces and PlayStation 5, providing a new option for gamers and general users. In addition, KIOXIA will showcase the newly launched EXCERIA PLUS G2 microSD card, offering up to 2TB of storage, 100MB/s read speed, and support for UHS-I and U3 (V30) speed classes, providing users a microSD with larger capacity and faster reading options.

**Highlights of KIOXIA Taiwan’s Booth**

Visiting the booth to experience the advanced memory technologies and understand KIOXIA’s pivotal role in the ecosystem through product displays and joint displays with ecosystem partners, KIOXIA Taiwan invites industry leaders and relevant professionals to join KIOXIA and Make Memories at COMPUTEX Taipei 2024.



KIOXIA Taiwan invites industry leaders and relevant professionals to its booth at COMPUTEX Taipei 2024



# Bahwan CyberTek bets big on Taiwan’s tech ecosystem

News highlights

Bahwan CyberTek (BCT), a leading provider of digital transformation services, was established in 1999 with a simple, yet profound purpose - to prove that companies from the fringes of the world economy can become global players. 25 years later, BCT has more than 4,000 associates serving clients across 20 countries and is actively expanding in Taiwan.

In an exclusive interview with DIGITIMES Asia, Mr. Dharanendra Murthy (Dan), VP – Business, Southeast Asia, Bahwan CyberTek shares insights into the company’s plan for Taiwan and the evolving role of technology in shaping the business landscape of the future.

**Q:** *Congratulations on your 25th anniversary! Can you tell us a bit about Bahwan CyberTek, your offerings, and how you’ve stayed relevant for 25 years?*

A: Thank you! It is indeed a proud moment for all of us. The purpose behind Bahwan CyberTek (BCT) was simple, yet ambitious - to prove that companies from the fringes of the world economy can become global players. Today, 25 years later, BCT has over 4,000 associates working across 20 countries, serving over 2,000 customers, including Fortune 500 companies.

What’s helped us stay relevant and cut through the clutter is our unique combination of IP + Services + outcome-based business models. Equipped with deep-domain expertise in BFSI, Oil and Gas, Retail, Energy, and Telecom sectors, we’ve helped clients navigate complex digital transformation journeys with ease and helped them convert abstract goals into tangible results.

We lean on our IP expertise to help clients punch above their weight. Our award-winning products retina360, rt360, dropthought, FuelTrans, and CueTrans have helped brands enhance customer experience, improve operational efficiencies, and overall business performance. Built on a solid foundation of AI, IoT, and predictive analytics, our products not only help enterprises keep up with the pace of transformation but also stay ahead of the curve.

At BCT, we view talent as our intellectual wealth. Our knowledge capital, founded on mutual respect and trust, is a significant factor that sets us apart from our competitors. Our employees are the driving force behind our success and the core of our business decisions. This culture of belonging has been a major driver of our growth for 25 years.

**Q:** *What brings you to Computex this year? Can you highlight some of the flagship products and services that you’re showcasing at Computex?*

A: Computex 2024 and its theme “Connecting AI” promises to be exciting, and we’re all set to see AI in action. From software to hardware, from advanced AI to



Dharanendra Murthy, Vice President – Business (South East Asia), Bahwan CyberTek

intelligent cyber security, the event gives us a sneak peek into the future of AI and how enterprises leverage AI to revolutionize operations, turbocharge innovation, and ultimately create better customer experiences.

For BCT, the event brings us closer to clients in this region and allows us to demonstrate how our products, services, and solutions can deliver value. We will be showcasing retina360 and DropThought and sharing insights about our latest global projects for clients.

About our products, retina360 is an award-winning, USPTO-patented advanced predictive analytics platform, that leverages AI, ML, and predictive analytics, to offer real-time insights for decision-makers. retina360 has been proven to significantly reduce asset downtime and was awarded the Best R&D and Product Development award by the Federation of Indian Chambers of Commerce & Industry (FICCI) Tamil Nadu State Council.

The other product we will showcase is DropThought, an innovative experience measurement platform. DropThought is designed to empower businesses and organizations with actionable customer feedback. Utilizing analytics and AI, DropThought gathers insights through surveys, touchpoints, and Voice of the Customer (VoC) programs. Businesses have been using DropThought to enhance customer experience.

**Q:** *As a Middle East company with operations in India and the US, what brought you to Taiwan?*

A: Taiwan is not just at the center of the fourth industrial revolution but actively shaping the future of Industry 4.0. A high level of internet adoption, a

talent pool with deep knowledge of advanced digital technologies, and strong government intent to position Taiwan as the AI capital of the world make it the destination of choice for us.

Enterprises in Taiwan are not simply keeping up with the pace of transformation but are also looking at ways to stay ahead of it. We believe our products and services can perfectly complement this appetite for rapid growth. With 25 years of digital transformation experience across verticals, deep-domain expertise, and knowledge of the APAC market, we can help these enterprises take the digital leap, deliver enhanced customer experiences, and contribute to overall economic development.

Additionally, our shared commitment to sustainable development with maximum impact makes it an ideal match.

**Q:** *Can you tell us about your journey so far in Taiwan? What has contributed to your success in the region?*

A: Several factors drive our success in the regions. First, it is our deep understanding of customer needs and challenges. Second, being the right-size growth partner helps us scale and pivot swiftly as per customer needs, making us the ideal growth partner.

Our local team consists of highly skilled professionals with deep domain expertise who have worked on international projects, consistently maintaining a high-quality standard. Taiwan is recognized as a Center of Excellence for digital apps, and we have extensive experience in developing mobile applications, web applications, cloud solutions, and emerging technologies such as AI and AR/VR.

We have strong relationships with local clients and partners, allowing us to understand the unique needs and challenges businesses in Taiwan face. This enables us to offer tailored solutions that drive significant value and transformation.

**Q:** *Bahwan CyberTek has formed various strategic partnerships over the years. Can you discuss the importance of these partnerships and any recent collaborations you’re particularly excited about?*

A: Collaboration is a core value at BCT that has strengthened our global market presence. Customer-centric enterprises, especially in the age of AI, need strategic transformation allies to usher in the latest technological advances. Our long-term successful collaborations with partners like TIBCO, Oracle, and IBM and more recently our strategic partnership with Temenos have reinforced that speed, agility and efficiency can be achieved more quickly when we work with mutually driven partners.

**Q:** *What are your plans for further expansion in this region?*

A: Taiwan is at the heart of our global strategy, and we plan to expand our team here. We have successful cases internationally, and our expertise in industries such as BFSI, Oil and Gas, Retail, Energy, and Manufacturing can help local businesses achieve significant digital transformation and operational efficiency. Taiwan is an IT powerhouse with excellent hardware, software, and integration capabilities. We aim to find partners in Taiwan to collaborate and expand into markets in South Asia, the Middle East, North America, and other regions.

**Q:** *As an IP-focused company, what are your plans for products on emerging tech (AI, IoT, ML)*

A: Today, every business is a technology business. B2B or B2C, brands are obsessed with customer experience in a post-pandemic hyper-competitive market, and rightly so! Emerging technologies are making operations smarter, and digital transformation is helping brands extract maximum value from these technologies.

BCT is better positioned than ever to lead in this technology-first crowded market. Our innate focus on IP, innovation, and talent is already a differentiator. Huge market acceptance for our products and continued patronage for our future-fit services are opening new revenue opportunities and strengthening our existing partnerships.

Our focus on IP is based on a simple architectural concept - to sense; to analyze; to predict; and the capacity to drive business decisions. Our products, across verticals, are designed with this foundation and they continue to gain market share year on year.

We continue to evolve and meet market needs. Recently, we launched midas360 (Metrorail Intelligent Decision Analytics System), in India. midas360 is a cutting-edge platform developed by BCT and employs advanced data analytics and IoT predictive technologies to revolutionize how metro networks operate in cities. We also launched radar360, our advanced SaaS platform designed to enhance Employee Benefit Plan (EBP) audits.

**Q:** *Finally, is there anything you would like to say to your clients and partners as celebrate your 25th anniversary?*

A: First and foremost, I want to express my heartfelt gratitude to our clients and partners for their unwavering support and trust in our capabilities. It is their belief in us that has fuelled our growth and success.

I also want to acknowledge the incredible dedication and hard work of our employees, whose passion and commitment have been instrumental in shaping our journey. Together, we will continue to create innovative solutions, surpass expectations, and build a brighter future for Bahwan CyberTek, here in Taiwan, and across the globe.

## AI application transformation drives significant increase in storage technology demand, says Wallace Kou, President and CEO of Silicon Motion

News highlights

The rapid development of AI has had a massive impact on global industries. The emergence of innovative technologies such as generative AI and machine learning has not only enhanced the performance of devices like smartphones, PCs, data centers, automotive systems, and industrial applications but has also spurred the demand for high-performance storage solutions in the market. Whether used for AI training or inference, the most crucial aspect is data. Swiftly acquiring the necessary data, securely storing data generated through AI training or inference, and properly protecting it is a vital issue within the entire AI ecosystem. In response to this trend, Silicon Motion, the world’s largest NAND Flash controller supplier, has a comprehensive product layout and market strategy in place to help customers seize this immense business opportunity.

Recently, storage technology has gained significant market attention. Wallace Kou, President and CEO of Silicon Motion, identified five major driving factors. Firstly, the rapid development of AI technology, as mentioned earlier, has led to a surge in storage demand. Secondly, the expansion of AI application scenarios necessitates storage systems that provide high-speed data access, large storage capacity, and a stable and reliable operating environment. Thirdly, with the widespread adoption of AI applications, users place greater emphasis on protecting sensitive data, requiring storage systems with more sophisticated data encryption and access control mechanisms to ensure data security. Fourthly, storage technology continues to advance, including innovations such as QLC NAND, NVMe protocol, PCIe Gen5, providing new solutions to AI storage challenges. Finally, the continuous increase in storage performance and capacity demands in areas such as data centers and cloud services, driven by applications like AI and big data, further expand the market space for high-performance and large-capacity enterprise storage solutions.

### Rise of AI Edge Devices Demands Progressive Functionality in Storage Components

Wallace Kou further elaborated on the current market trends and changes in demand for storage components. He mentioned that current AI applications still heavily rely on the high computational capabilities of data centers and cloud servers to process large language models, leading to a significant surge in demand for High Bandwidth

Memory (HBM). However, with considerations for cost reduction, system latency reduction, and enhanced privacy protection, AI applications have started transitioning from cloud platforms to personal computers and AI smartphones, where miniature language models run on edge devices.

In this context, edge devices must improve performance, optimize human-machine interfaces, ensure seamless operation, and prioritize storage device performance, reliability, and data security. Moreover, to accommodate the vast amount of data generated by AI applications, edge devices need to expand their capacity while also managing cost increases, underscoring the importance of QLC NAND. It’s anticipated that QLC applications will expand from the current PC market to mobile phones and enterprise applications.

Another emerging trend is performance optimization. He highlighted the profound impact of data placement technology on AI performance, noting that major NAND manufacturers have recently introduced Zoned UFS, grouping similar types of data within the same storage block. For enterprise applications, ZNS and FDP technologies have adopted similar concepts, significantly enhancing read performance and extending device lifespans.

To meet the aforementioned demands, Wallace Kou believes that a comprehensive approach is needed, combining product strategy, design services, and organizational transformation to satisfy customer needs. On the product front, comprehensive customer support is essential, along with actively adopting new technologies to address market demands. Regarding design services, different application areas have varying requirements, and customers have a high demand for customized designs. Storage component suppliers must integrate advanced technology with robust firmware design capabilities to provide intelligent, efficient, and reliable storage solutions across various sectors. Lastly, enterprise organizations need to segment based on different customer applications, enhancing service quality and efficiency, focusing on delivering specialized domain products and customized services.

### Silicon Motion Unleashes Three Strategies to Meet Diverse Customer Needs

Responding to the aforementioned demands, Silicon Motion has implemented a comprehensive strategy. Wallace Kou highlighted that all product lines of the company support QLC, catering to customers’ requirements for high-capacity storage applications

while achieving cost reduction and rapid shipping objectives. Beyond product lines, the strategic level also encompasses technical support in storage and firmware. For instance, in terms of edge devices, Silicon Motion’s PCIe Gen5 SSD controller and UFS4.0 controller can fulfill the needs of AI PCs and AI smartphones with ultra-high performance and low power consumption. In enterprise storage, the MonTitan enterprise SSD development platform embraces key enterprise storage technologies like ZNS and FDP, optimizing data placement within systems. Moreover, Silicon Motion integrates its PerformaShape and NANDCommand technologies to guarantee maximum bandwidth performance and reliability, meeting the rigorous requirements of AI applications for high performance and reliability.

Regarding design services, Silicon Motion maintains long-term collaborations with major smartphone manufacturers, PC OEMs, hyperscalers, car makers, industrial product manufacturers, and other partners to deeply grasp customer pain points and requirements. This enables them to design products that align better with real-world application scenarios. Wallace Kou noted that firmware design capability stands as one of Silicon Motion’s significant competitive advantages in the NAND flash controller field. Over time, they have merged advanced technology with robust firmware design capabilities to deliver intelligent, efficient, and dependable storage solutions for various sectors. Silicon Motion’s firmware solutions not only boast highly optimized and stable features but also enable customized services, tailoring optimal solutions based on customer application scenarios, performance requirements, and specific needs.

### Fine-tuning Organizational Structure to Enhance Firmware Customization Service Advantages

As part of its organizational restructuring, Silicon Motion has recently established two major business groups: Client & Automotive Storage (CAS) and Enterprise Storage & Display Interface Solution (ESDI) to assist customers across different industries in overcoming challenges in the AI era.

The CAS group focuses on providing solutions such as client SSD controllers, mobile controllers, Ferri embedded storage, and expandable storage controllers. Its applications primarily cover areas like PCs, smartphones, automotive, gaming consoles, and industrial applications. On the other hand, the ESDI business group concentrates



Wallace C. Kou, General Manager of Silicon Motion, stated that the company’s long-term collaboration with partners in various fields enables them to deeply understand customer needs and provide storage solution products that fit actual application scenarios

on enterprise SSD controllers and display interface products, targeting markets such as data centers, servers, workstations, enterprise applications, as well as USB display and embedded GPU applications. Wallace Kou stated that a specialized organizational division of labor can help Silicon Motion efficiently meet the diverse needs of different customers, accelerate innovation, and enhance overall competitiveness. Whether in consumer, enterprise, automotive, or industrial markets, Silicon Motion can provide more professional and tailored solutions closer to the demands.

In the mobile sector, Silicon Motion has established strategic partnerships with several mobile phone manufacturers and mobile chip manufacturers to jointly develop high-performance, low-power, high-capacity storage solutions. Through firmware design capabilities and customized services, they meet specific requirements of phone manufacturers, optimizing storage performance and power consumption for AI smartphone development. In the data center domain, Silicon Motion has partnered with several major enterprise storage providers and hyperscalers in North America and China, securing commercial cooperation and providing enterprise storage solutions that meet performance and reliability demands.

Looking ahead, Wallace Kou believes that with the continuous evolution of technologies, the demand for storage in the market will continue to expand and Silicon Motion will increase its R&D investment to leverage its competitive advantages.



# JMicon Technology Corp. teams up with KaiKuTek Inc. for a sensational debut at COMPUTEX 2024 - Innovation Ignites Technology Evolution

News highlights

JMicon Technology Corp. teamed up with its wholly owned subsidiary, KaiKuTek Inc., to exhibit a variety of high-speed interface bridge controller solutions at the Sea Hall of the Taipei Marriott Hotel in Nangang during COMPUTEX 2024.

They also showcased various application solutions integrating artificial intelligence and millimeter-wave radar technology, including air gesture control and object sensing. JMicon Technology Corp., known for its high-speed interfaces bridge controller chips, unveiled its latest generation of products.

These chips represent a significant leap forward in transmission speed, stability, and efficiency, garnering attention from industry insiders and attendees alike. They also expanded the application fields of external storage to various smartphone platforms, simplifying cross-platform data exchange and positioning themselves as an industry focal point for future growth.

Tony Lin, VP of Marketing & Sales Center at JMicon Technology Corp., emphasized the company’s commitment to extensive collaboration on next-generation bridge control chips with its primary customer base. This initiative aims to lead trends in high-speed interfaces and storage applications, creating new market opportunities in collaboration with customers.

KaiKuTek Inc. showcased groundbreaking products, leveraging its leading position in 60GHz millimeter-wave radar technology. The company focuses on radar sensing, AI/ML technology, antenna design, and gesture recognition, enhancing various smart applications.

New products include TWS earbuds and intelligent eye massager with air gesture control for intuitive user experiences. In addition, a smart fan solution incorporates air gesture control and precise positioning through millimeter-wave radar, achieving automatic fan direction adjustment.

Contactless products are increasingly

integrated into daily life, with applications developed using millimeter-wave radar sensing technology and artificial intelligence projected for significant growth. The collaboration between JMicon Technology Corp. and KaiKuTek Inc. injects vitality into the industry, exploring broader horizons for technological innovation and enhancing user experiences.

Mike Wang, General Manager of KaiKuTek Inc., anticipates applying accurate tracking and positioning and air gesture recognition technology in various fields, including industrial, automotive, personal consumer electronics, and IoT. Through system-level AI radar sensing solutions, KaiKuTek Inc. aims to realize non-contact long-distance object detection and operation, revolutionizing human-machine interaction interfaces.

The success of COMPUTEX 2024 underscores the bright future for JMicon Technology Corp. and KaiKuTek Inc. Together, they will continue innovating, driving technological progress, and shaping the industry landscape.



JMicon teams up with KaiKuTek for a sensational debut at Computex 2024

## InWin showcases one-stop service: Incorporating AI ecosystem to create infinite possibilities

News highlights

Imagine yourself aboard a spacecraft named, “Infinite” showcased by InWin (6117-TW) at Computex 2024 in Taipei, Taiwan.

Every facility within the cabin is built with AI as the core concept, providing human-centric and reliable services. This journey travels to a boundless universe where you control your final destination.

InWin’s focal theme of 2024 is ‘AI hardware ecosystem,’ outlining product lines and future development directions. The product line includes AI personal computers, workstations, servers, and other related solutions, catering to various application needs. Customers can choose suitable models based on their needs and benefit from InWin’s decades of professional experience, knowledge, and technical services for customized production and assembly testing.

The surge of generative AI in 2023 increased the demand for AI training and implementation for many businesses. Issues such as system setup, secure training data management, and stable high-speed transmission have become top priorities for enterprises.

AI training often requires high-end GPUs, professional cooling modules, and stable power supplies to meet long-term, high-energy computational demands. Standard chassis specifications struggle to accommodate the expanding needs of AI system deployment.

Any hardware changes may render the original chassis inadequate, leading to the need for additional purchases or complete system replacements. This will increase the amount of hardware investment pressure for businesses.

To assist enterprises in planning their system configurations and retaining flexibility for adjustments, InWin has introduced hybrid storage servers. These servers support PCIe Gen5 high-speed transmission and data computation.

With internally integrated backplanes, Retimer Cards, and SSDs, these servers offer a complete PCIe 5.0 high-speed transmission and storage solution. They support efficient transmission rates for NVMe SSDs and provide flexible customization options ranging from 1U to 4U with liquid cooling modules to meet AI computing needs across various industries.

In addition to hybrid storage servers, InWin offers water-cooled servers, an essential aspect of cooling high-powered computing such as big data processing during the AI era. The water-cooling circulation module significantly enhances heat dissipation efficiency, ensuring continuous and stable system operations. Its compact design is especially suitable for edge computing.

At this exhibition, InWin unveiled its cutting-edge

high-U AI server products and workstation system solutions developed for brand clients, demonstrating its expertise in high-end system assembly. In addition, InWin collaborated with Phison Electronics to display AI server systems tailored for mutual clients.

By leveraging the aiDAPTIV+ software, these systems offer easy deployment, cost-effectiveness, and privacy and security advantages. This enables small and medium-sized enterprises to conduct mainstream LLM computations without concerns about expensive hardware setups.

With years of proven excellence, InWin has become a trusted global partner for computer and motherboard brand clients, providing comprehensive services, including institutional design, production, system assembly, and burn-in testing. Its reliable quality and comprehensive services are pivotal in driving the widespread adoption of AI technology.

In addition to enterprise-level products, InWin has unveiled the “Infinite” concept case, which embodies the idea of “exceeding infinite possibilities.” It marks the industry’s first one-piece, 180-degree curved glass case, breaking through limitations in glass processing technology concerning maximum area and curvature.

Moreover, it boasts a mechanically operated internal structure enhancing user convenience during system assembly and unique exhibition opportunities. With support for back-connect motherboards, it offers a cleaner panoramic view.

The “infinity” transparent glass reflects ARGB lighting, further complementing the spaceship-like hatch design, providing viewers with an immersive experience akin to embarking on a journey into space. This signature chassis will be a limited edition worldwide.

Another product that complements the theme of infinity is the ModFree ecosystem. Unlike traditional chassis, this product features a modular design where modules seamlessly interconnect with each other, offering unlimited expansion capabilities.

From mini-ITX to E-ATX systems to multi-node workstations, all is possible through InWin’s continuously updated ModFree modules and expansion kits. This product has also received the prestigious Red Dot Design Award from Germany.

In addition to the rich exhibition content, a large LED screen was set up on-site, allowing visitors an opportunity to experience an aerial guided tour. Interested industry professionals and vendors were invited to explore the celestial splendor found at the InWin station.

2024 COMPUTEX TAIPEI

Exhibition Time: June 4th to June 7th, 2024  
Location: Floor 1, Nangang International Exhibition Center, Hall 1, InWin Booth Number J0806.



InWin showcases its PCIe Gen5 high-speed transmission hybrid storage servers and a concept case named “Infinite” at Computex. These products demonstrate InWin’s capability in the AI hardware ecosystem, R&D design, innovative realization, and to Level 11 system assembly and testing services.

## Fueled by Apple, SDC’s and LGD’s tablet OLED shipments saw massive surge

Jessica Tsai, Taipei; Jack Wu, DIGITIMES Asia

With Apple’s launch of the new iPad Pro tablet that features OLED displays, South Korean providers Samsung Display (SDC) and LG Display (LGD) saw a significant increase in their tablet OLED panel shipments for March 2024.

Citing data from market research firm Omdia, Seoul Economic Daily reported that in March 2024, SDC and LGD’s tablet OLED panel shipment volume reached 980,000 units, representing a 350% monthly increase. Among them, SDC shipped 385,000 units, a 97.4% monthly increase, while LGD’s shipments reached 595,000 units, a staggering 2,486.96% monthly increase compared to 23,000 units in February 2024.

While Chinese company BOE’s panel shipments also rose from 65,000 to 100,000 units, it represents only a 53.8% growth rate, much lower than the growth rates reported by SDC and LGD.

Analyses believe that the strong growth in tablet OLED panel shipment momentum is primarily due to the influence

of Apple’s new iPad Pro. On May 7, Apple launched the iPad Pro with OLED displays. It comes in two sizes, 11-inch and 13-inch, with LGD supplying both the 11-inch and 13-inch OLED panels and SDC supplying the 11-inch OLED panels.

Because the OLED panels used in the iPad Pro deploy the two-stack tandem technology, BOE ultimately didn’t become a panel supplier for the OLED iPad Pro due to a technical gap and the lack of mass production capabilities.

Some South Korean perspectives hope that BOE’s inability to become an OLED supplier for the iPad will allow South Korean companies to expand their dominance in the OLED market.

Moreover, the performance of South Korean display companies is expected to rebound earlier than anticipated. Reportedly, the average supply price for smartphone OLED panels is around US\$50-\$60, while the price for iPad OLED panels (based on the 13-inch model) is around US\$380-\$390.



## India and Vietnam largest beneficiaries of China-plus-one, says Nomura

Jingyue Hsiao, DIGITIMES Asia, Taipei

As multinational companies recognize the importance of diversifying supply chains, India and Vietnam are emerging as key beneficiaries. The electronics and machinery sectors in India are likely to see rising exports in the coming years, says Nomura.

CNBC-TV18, Money Control, and the Economic Times citing Nomura, reported that India’s exports are expected to grow at a compound annual growth rate (CAGR) of 10%, reaching US\$835 billion by 2030, compared to US\$431

billion in 2023. Vietnam’s exports are projected to grow at a CAGR of 11.4% through 2030, with Mexico also seeing gains. China’s role in global value chains is evolving, leading companies to explore other options.

A significant portion of investments in India comes from the US and developed Asian countries. While strong enthusiasm exists for India’s potential, investors are advised to be patient in the short term. India is expected to see export-led investments in electronics, automobiles, capital goods, semiconductors, energy, and pharmaceuticals.

Nomura predicts that electronics will become the fastest-growing sector in India, with exports increasing at a compound annual growth rate of 24%, nearly tripling to US\$83 billion by 2030. Machinery exports are also expected to more than double, rising from \$28 billion in 2023 to \$61 billion by 2030.

“The shift in supply chains away from China has triggered what economist Kaname Akamatsu described as the wild-geese-flying pattern of economic growth. This pattern sees production move from the leading nation to developing countries,” explained Nomura.

Data from the Department for Promotion of Industry and Internal Trade of India shows that over the past years, computer software and hardware has become one of the largest sectors attracting foreign investments. The industry had an accumulated foreign equity inflow of US\$98.33 billion during 2000 and 2023, behind just the services sector.

FDI inflow in India, FY22-24 (US\$m)				
Sector	FY22	FY23	FY24	2000-2023
Services Sector	7,131	8,707	5,187	108,042
Computer Software & Hardware	14,461	9,394	3,417	98,329
Trading	4,538	4,792	2,661	42,192
Telecommunications	668	713	271	39,315
Automobile Industry	6,994	1,902	913	35,657
Infrastructure Activities	3,248	1,703	3,841	33,527
Townships, Housing, Built-Up Infrastructure And Construction-Development Projects	125	146	185	26,541
Drugs & Pharmaceuticals	1,414	2,058	913	22,377
Chemicals	966	1,850	770	22,072
Power	526	698	1,583	18,168

Source: DPIIT, compiled by DIGITIMES, May 2024



# Taiwan notebooks, 1Q 2024

Jim Hsiao, DIGITIMES Research, Taipei

## Introduction

Taiwan’s notebook shipments in the first quarter of 2024 had a performance stronger than DIGITIMES Research expected, dropping only 5.1% sequentially, smaller than the 8.4% on-quarter decline experienced in the same quarter a year ago and also better than the average 15% seasonal decrease during the years prior to the COVID-19 pandemic. (Note: Unless otherwise indicated, all figures and tables in this report refer to output from Taiwan makers.)

The volumes were up 3.9% from the same quarter a year ago in the first quarter.

Taiwan had a worse sequential shipment decline than the global average (-3.5%) in the first quarter as notebook demand was stronger in China than worldwide, resulting in more orders obtained by Lenovo’s in-house production lines and Chinese notebook maker Huaqin. This also led to a further decline in Taiwan’s share of global notebook shipments to 71.6%.

In the second quarter, Taiwan’s shipments are estimated to rise 3.2% sequentially, but will plunge 11.6% on year.

The shipment growth will be attributed mainly to enterprise models, Chromebooks, and Qualcomm-based Windows on Arm (WoA) models, which are primarily made by Taiwanese ODMs. Thanks to these products, Taiwan had a shipment growth stronger than the global average in the second quarter, while the share will also return to above 72%.

(Note: DIGITIMES Research treats detachable notebook products as tablets. Convertible notebooks with undetachable keyboards are considered notebooks.)

Global notebook shipments performed better than DIGITIMES Research originally expected in the first quarter of 2024, down only 3.5% from the earlier quarter, significantly lighter than that of the same quarter in 2023 and the average sequential decline of 15% for the quarter of the past years.

The first-quarter 2024 volumes

were up 12.2% from a year ago, the second consecutive quarter of seeing on-year growth for global shipments.

With notebook brands’ keen promotions to clear out excess inventory reaching an end in the early fourth quarter of 2023, overall shipments in the quarter only had a limited benefit from the seasonality and established a rather low comparison base for the first quarter of 2024.

The consumer sector of North America was heavily influenced by the slow season in the first quarter, but demand from Europe and China remained strong. Coupled with the fact that some channels had pulled in their orders early in anticipation of an upcoming cost hike, these factors had all contributed to a better-than-expected performance in the quarter.

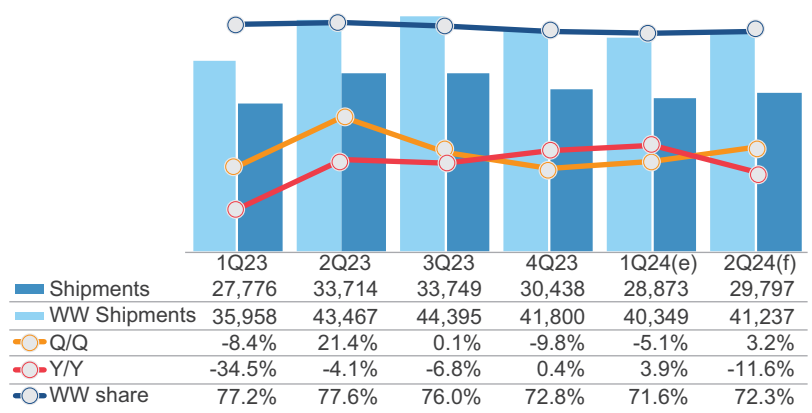
Global volumes are expected to pick up 2.2% sequentially in the second quarter, far weaker than second-quarter 2023’s over 20% increase and also inferior to the average 5-10% growth of the usual seasonal performance. On-yearly, the volumes are projected to decline by 5.1%.

For the consumer sector, compared to the shipment strategy used for the second quarter of 2023 to focus on digesting the inventory of models powered by the CPUs of the one-to-two-generation earlier platforms, a large portion of notebooks planned for the second quarter of 2024 are powered by the latest platform from Intel, such as Meteor Lake, AMD, or Qualcomm.

Because of the new platforms’ higher overall costs and rising memory quotes, the price tags of brand vendors’ new notebooks for the second quarter are expected to be high, which should hinder channel retailers’ order pull-ins for the quarter, leading to slipping shipments for consumer models.

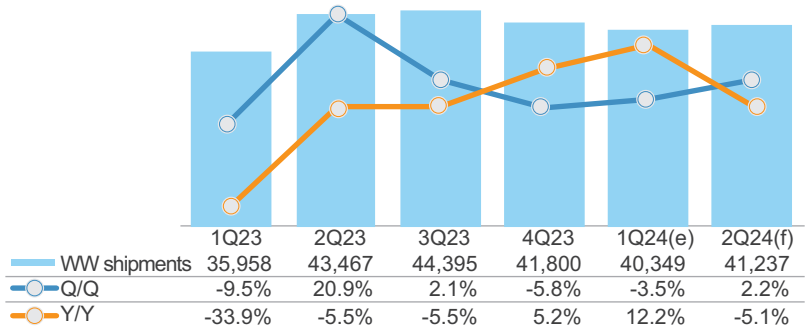
Although the second quarter was traditionally a strong season for the education procurement sector, the high shipment comparison base in the first quarter and Japan not yet starting its annual procurement project for 2024 will result in a weaker influence of seasonality this year than the previous ones.

## Notebook shipments, 1Q23-2Q24 (k units)



Source: DIGITIMES Research, April 2024

## Global notebook shipments, 1Q23-2Q24 (k units)



Source: DIGITIMES Research, April 2024

## Shipments breakdown

### Clients

Asustek’s orders with Taiwanese ODMs went up 14.6% sequentially in the first quarter as its orders given to Quanta surged 50% on quarter.

Lenovo was keen on meeting its notebook shipment target for fiscal 2023/2024, which ended on March 31, 2024, and had been placing aggressive orders for enterprise models with Taiwanese ODMs.

HP has placed robust orders for new consumer notebooks and education procurement products and its shipment share is expected

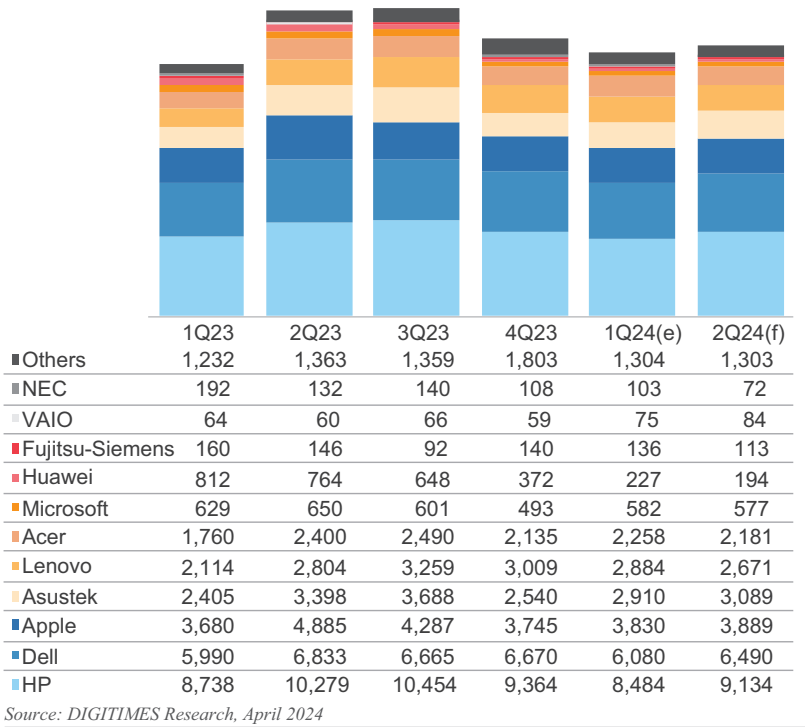
to rise back to above 30%.

Dell is expected to push for its enterprise and Qualcomm solutions-powered notebooks in the second quarter.

Asustek Computer’s shipments were up 12.7% sequentially in the first quarter of 2024 as the company had finished digesting its excess inventory ahead of other competitors, allowing it to establish a low comparison base in the fourth quarter of 2023.

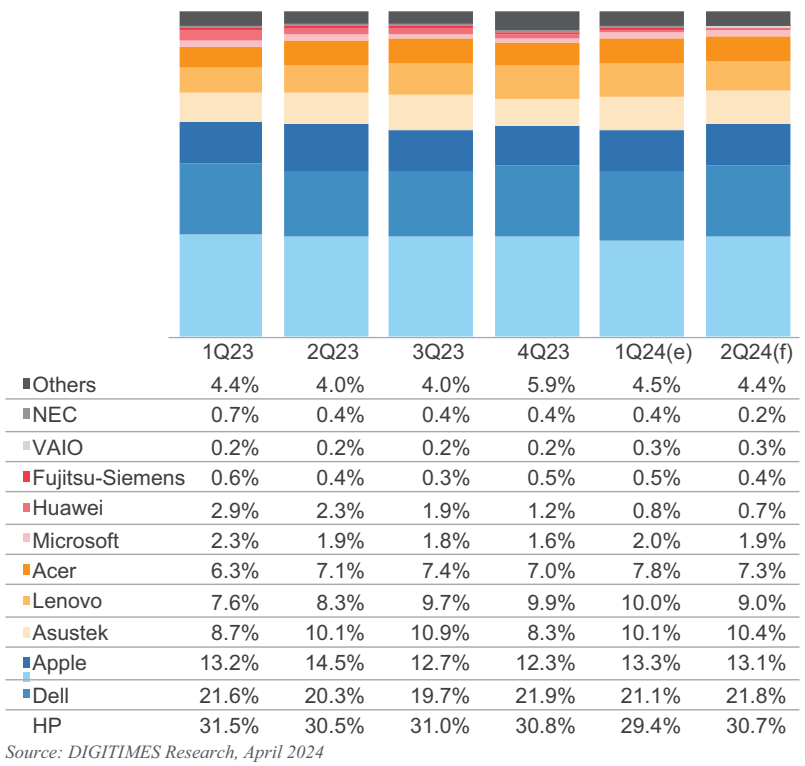
With Intel’s launch of the new CPU platform, Asustek was keen on pushing its shipments for related models, thus enjoying a good performance for the quarter.

## Shipments by major client, 1Q23-2Q24 (k units)



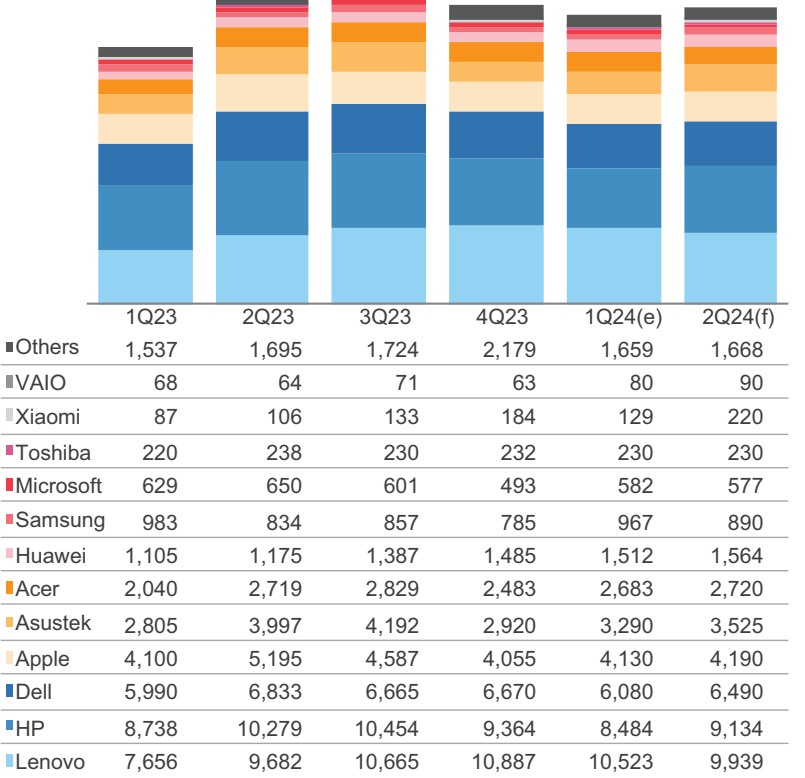
Source: DIGITIMES Research, April 2024

## Shipment share by major client, 1Q23-2Q24



Source: DIGITIMES Research, April 2024

## Global shipments by major vendor, 1Q23-2Q24 (k units)



\*Note: Shipments of NEC and Fujitsu-Siemens are included in Lenovo’s volumes.  
Source: DIGITIMES Research, April 2024

Acer enjoyed a sequential shipment increase of 8.1% in the first quarter thanks to its conservative approach over its shipments in the fourth quarter of 2023 as well as customers pulling in their education procurement orders and orders for Europe and the US retail channel early.

Apple’s shipments were up almost 2% on quarter in the first quarter thanks mainly to the update on its MacBook Air, which drives up demand. Samsung’s shipments picked up more than 20% sequentially as the first quarter is the peak season for South Korea’s consumer sector.

HP will see a sequential

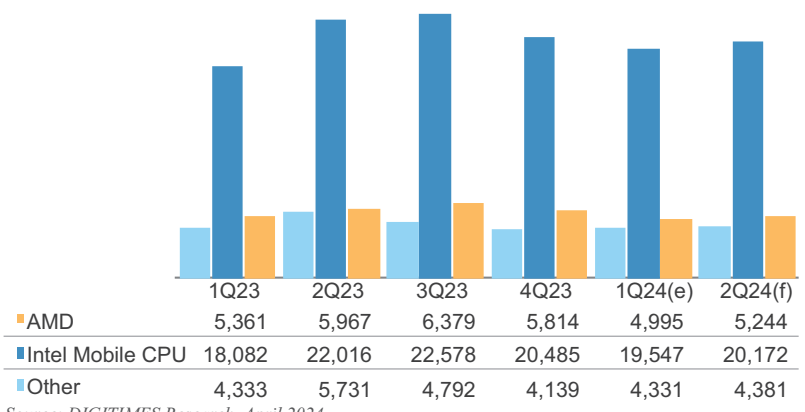
shipment recovery of 7.7% in the second quarter as the US-based brand will focus on pushing new consumer models with new CPUs and education procurement orders.

Dell’s shipments will uplift 6.7% from a quarter ago thanks to a demand recovery in the enterprise sector and the release of Qualcomm solutions-powered AI notebooks.

### CPUs

Taiwan’s shipments of notebooks powered by Arm-based processors (Other) rose 5% from a quarter ago in the first quarter of 2024 as shipments of Apple’s MacBooks, which are equipped with an in-house developed M series processor, were undermined

## Shipments by CPU, 1Q23-2Q24 (k units)



Source: DIGITIMES Research, April 2024

by overall weak sales of high-end notebooks in the fourth quarter of 2023 and an update to MacBook Air in the first quarter.

Thanks to the two factors, the shipment share of Arm-based notebooks also increased 1.4pp from a quarter ago in the first quarter.

Intel postponing the launch of its latest Meteor Lake platform to December 2023 led brand vendors to primarily allocate their Meteor Lake-based notebook shipments in the first quarter, resulting in a small rise in the shipment share of Intel-based notebooks.

Notebooks featuring AMD’s Hawk Point processor are set to begin volume shipments in the second quarter and should boost shipments of AMD-powered notebooks by almost 5% sequentially in the quarter.

AI notebooks powered by Qualcomm’s Arm solutions will start shipments in the second quarter. However, brand vendors are expected to stay cautious with

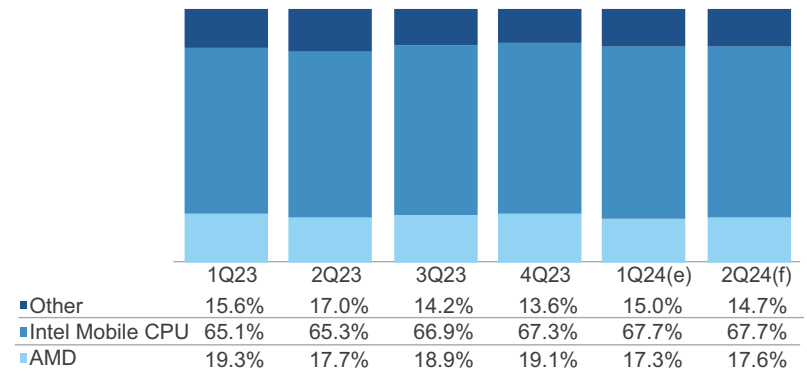
the shipments of related products initially due to their high price tags and the fact that they will not enter volume production until June.

The sub-12-inch notebooks were the only sector with a sequential shipment increase in the first quarter of 2024 as North America’s large-scale education procurement projects had started early in March 2024 and significantly drove up shipments of small-size notebooks (mostly 11.6-inch) to more than double from a quarter ago.

The second quarter is the traditional peak season for the education procurement sector and shipments of education models are expected to uplift more than 30% sequentially, relatively driving up shipments of sub-12-inch notebooks.

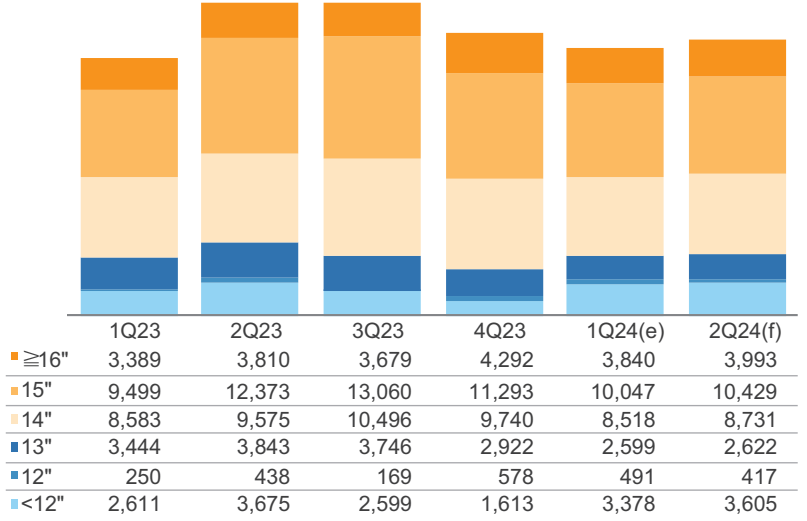
In the second quarter, most new mainstream x86 consumer and enterprise notebooks have adopted 15- or 16-inch panels which should slightly boost the two sector’s share.

## Shipment share by CPU, 1Q23-2Q24



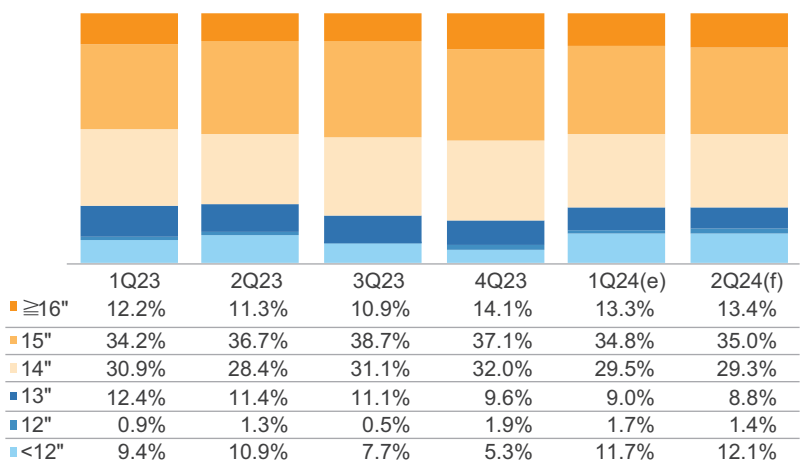
Source: DIGITIMES Research, April 2024

## Shipments by screen size, 1Q23-2Q24 (k units)



Source: DIGITIMES Research, April 2024

## Shipment share by screen size, 1Q23-2Q24



Source: DIGITIMES Research, April 2024

### Makers

The shipment share of the top-2 ODMs – Quanta and Compal was up 1.4pp from a quarter ago in the first quarter, better than DIGITIMES Research originally anticipated.

US-based brand vendors’ education procurement Chromebook orders were primarily given to Quanta and Compal as projects in North America started

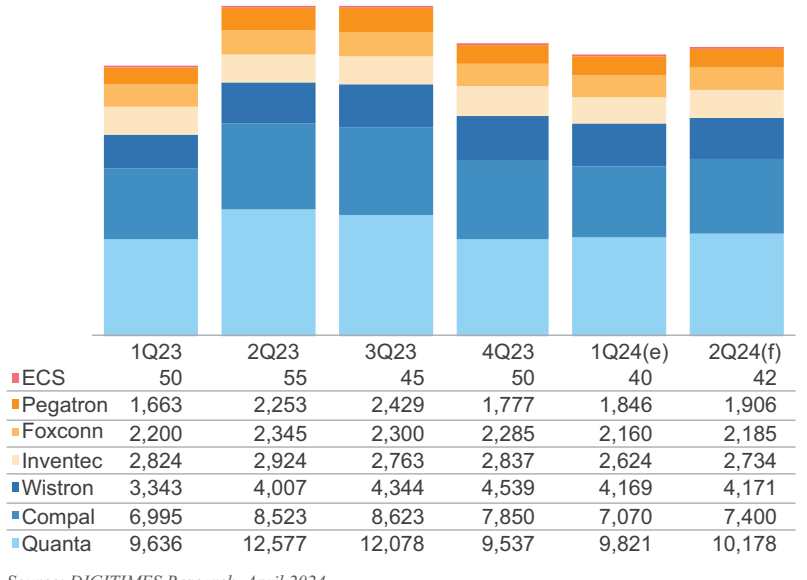
early.

The top-2 ODMs will continue to benefit from brand vendors’ Chromebook orders in the second quarter, which is the traditional peak season for the education procurement sector.

Brand vendors’ notebooks using Qualcomm’s Snapdragon X Elite processor are mainly outsourced to Compal for manufacturing.

Continued on page 9...

## Top maker shipments, 1Q23-2Q24 (k units)



Source: DIGITIMES Research, April 2024



# Global server market, 1Q 2024

Jim Hsiao, DIGITIMES Research, Taipei

DIGITIMES Research estimates that global server shipments in the first quarter of 2024 declined by 1.6% from the previous quarter – a result weaker than previously predicted.

### Introduction

DIGITIMES Research estimates that global server shipments in the first quarter of 2024 declined by 1.6% from the previous quarter – a result weaker than previously predicted.

The momentum of upgrading general-purpose servers by major US cloud service providers (CSPs) was weaker than expected, and although AI server shipments to the major US CSPs still saw strong growth, their overall server shipments increased only slightly quarter-on-quarter.

Brand vendors experienced higher-than-expected declines in the first quarter of 2024, as compared to the high base of the previous quarter amid significant off-season effects. In the second quarter of 2024, global server shipments are expected to increase by 1.9% sequentially, mainly contributed by brand vendors.

Brand vendors’ new general-purpose server shipments will increase compared with the previous quarter. The supply of high-end server GPUs will continue to improve, which will help boost brand vendors’ AI server shipments significantly. The overall shipments for large US CSPs are also expected to increase slightly in the second quarter.

The recovery of the overall

server market in the first quarter of 2024 was still limited. Large-scale US CSPs such as Google and Microsoft slowed down the deployments of general-purpose servers. The industry ended up growing only 1.3%.

American brand vendors experienced greater growth in the fourth quarter of 2023, increasing 9.5% quarterly, and the off-season effects in the first quarter were also more pronounced, with shipments falling by 4% quarter-on-quarter.

The performance of Chinese companies was better than expected, mainly because some datacenter companies began to update general-purpose servers.

Looking forward to the second quarter of 2024, US brands such as Dell and HPE have seen some customers start accelerating their replacement demand, while the supply of high-end GPUs such as H100 and MI300 has improved, which will enable the delivery of AI server shipments that have been previously delayed.

Among North American CSPs, Google and Microsoft have slowed down the purchases of general-purpose servers, but their shipments will still grow mildly in the second quarter. The shipment growth of brand vendors and CSPs in the second quarter will be more beneficial to Foxconn (Hon Hai) and Quanta among Taiwanese ODMs.

(Note: Shipments in server tracker reports refer to the number of motherboards shipped. A server motherboard supporting multiple CPUs is still considered one unit. A rack server that is equipped with multiple motherboards is counted based on the number of

its motherboards. Shipments of other networking devices such as switches, routers and storages, are not included in the figures.)

### Key factors affecting the global server market

On the demand side, there are more favorable factors than unfavorable ones for the global server shipments in the second quarter of 2024.

The conflicts in the Middle East and Europe have not abated, and inflation has not improved as much as expected. These are still restraining corporate and personal spending. The lack of significant recovery in demand for traditional cloud services is still constraining the shipments of general-purpose servers.

The rebound large US CSPs saw in the first quarter was not as strong as previously expected with only a slight quarterly growth of 1.3%. This was because some companies were updating their server platforms more slowly than expected.

In the second quarter, the recovery of demand from the cloud services market will continue to be slow, with overall purchasing

volume growing slightly by 1.7%.

Due to the low comparison base formed in the previous quarter and improvements in AI server shipments in the second quarter, US brand vendors’ overall shipments will increase by nearly 5% quarter-on-quarter, which is better than that for large CSPs.

On the supply side, there are also slightly more favorable factors.

Corporate customers who significantly postponed replacements in 2023 have begun to accelerate platform updates, driving brand vendors’ shipments for new models.

Major suppliers are still strategically controlling memory supply, and the earthquake that hit Taiwan in early April has also affected part of Micron’s DDR5 production. As a result, DDR and SSD prices have increased by about 10% and 20%, respectively in the second quarter.

Global server shipments in the first quarter of 2024 were lower than expected, declining by 1.6% from the previous quarter; compared with the same period in 2023, they declined by 6.7%.

The decline mainly resulted

from slower-than-expected updates to new processor platforms by large US CSPs. The purchase volume was originally expected to increase by 7% quarterly. However, the new round of general-purpose server deployments by Google, Microsoft and others was slower than expected. At the same time, Amazon’s shipments lapsed into declines in the first quarter. The large US CSPs only saw a 1.3% growth in the first quarter.

Against a higher growth for US brand vendors in the fourth quarter of 2023, (quarterly growth of 9.5%), the off-season effects became more significant in the first quarter of 2024, with shipments falling by 4% quarter-on-quarter.

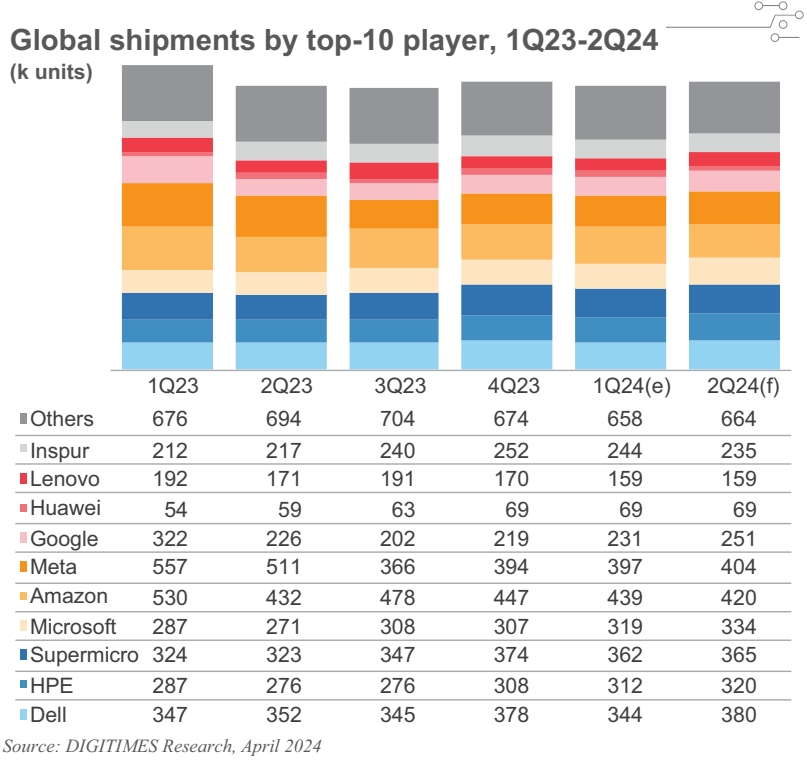
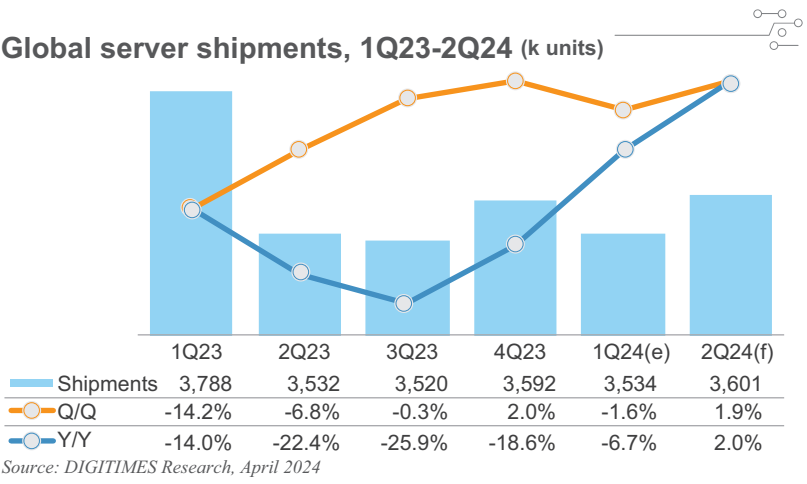
Chinese cloud companies’ shipments in the first quarter increased by 0.1% quarter-on-quarter, which was better than the previously expected quarter-on-

quarter decrease of 3.8%, thanks to the fact that some datacenter companies such as ByteDance and Tencent began to update general-purpose servers.

Global server shipments are expected to grow by 1.9% year on year in the second quarter of 2024, with the quarterly growth rate also expected to be significantly better than that for the same period in 2023 (quarterly decrease of 6.8%). It will be the first on-year growth since the third quarter of 2022.

Purchases by large US CSPs will increase by 1.7% quarter-on-quarter in the second quarter. Among them, Google and Microsoft will continue to accelerate the updates of general-purpose servers, and shipments of Nvidia high-end AI servers to Meta and Amazon will enter a high-volume stage.

Continued on page 10...



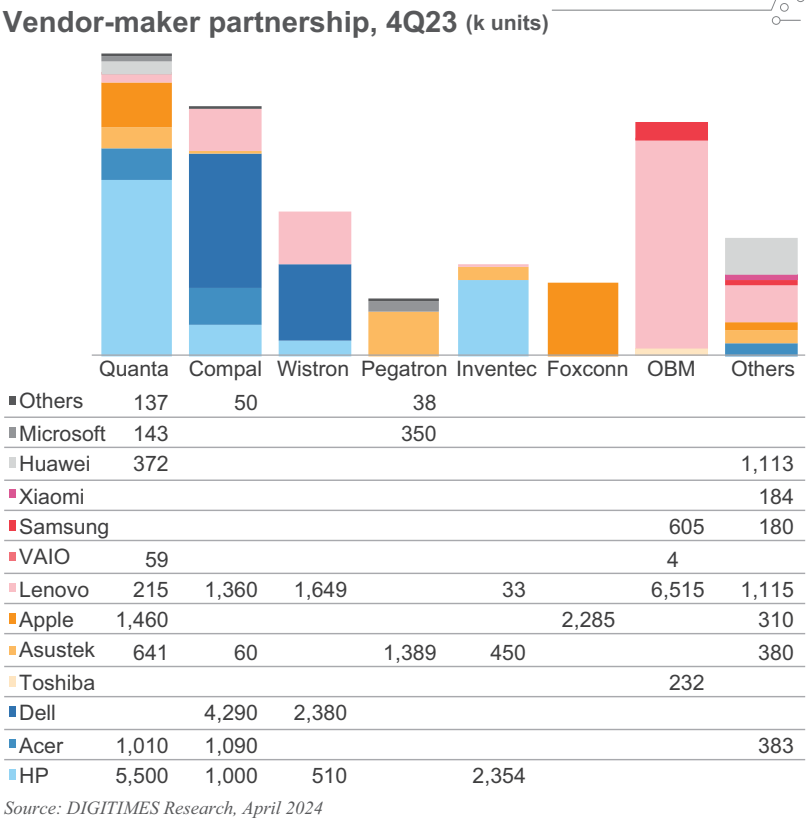
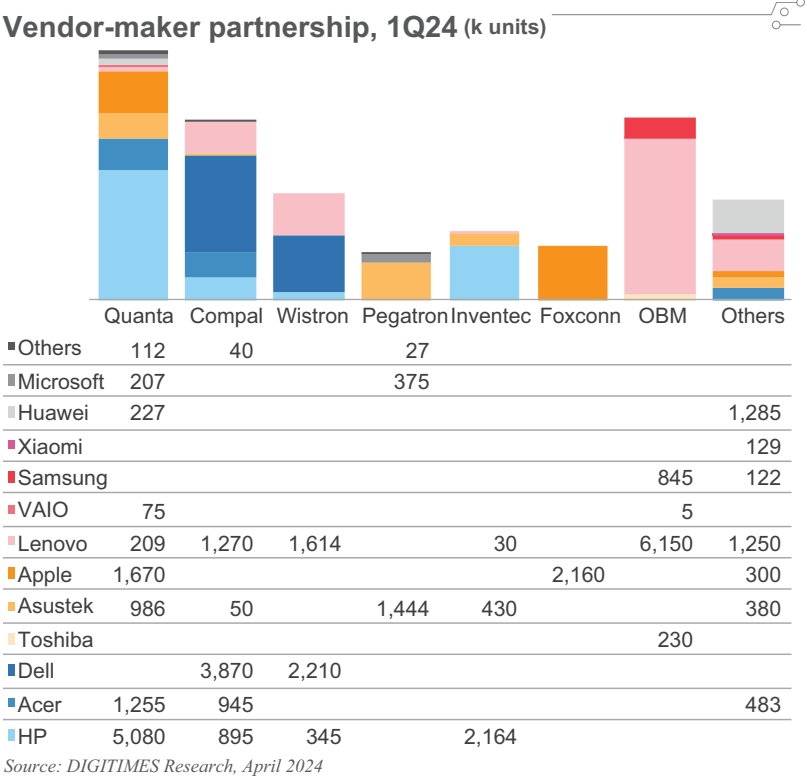
...Continued from page 8

Among the global top-5 ODMs, Quanta was the only maker with sequential shipment growth in the first quarter of 2024. The company’s shipments rose almost 3% from a quarter ago, which translated into a 3pp increase in shipment share to 34%.

The growth was contributed mainly by Apple’s increased

orders with Quanta also seeing its orders for consumer models from Taiwanese brands pick up from a quarter ago.

In the second quarter, Compal will see its shipments increase 4.7% sequentially, driven by increasing orders for Qualcomm WoA models and Chromebooks from clients in the US and Taiwan.



Inventec’s orders from its largest client HP will slump sequentially in the second quarter, but increasing orders for gaming models from Asustek will help offset the decline.

### Chromebooks

Global Chromebook shipments soared 45% sequentially and 36.1% on year to surpass four million units in the first quarter of 2024.

Lenovo was the largest Chromebook brand in the first quarter with shipments of over 1.2 million units. HP ranked second with nearly one million unit shipments.

Acer was the third largest brand in the first quarter, delivering over 800,000 units, followed by Dell at the fourth with over 600,000 unit shipments. Asustek surpassed Samsung to take fifth place with shipments of over 300,000 units in the quarter.

As the strong shipments in the first quarter had set a high comparison base for the second

quarter, the volumes are only expected to rise 32.3%, an increase that is inferior to those of the same quarter in the past years.

The growth in the second quarter will be driven mainly by education procurement projects in North America as projects in emerging markets will be as strong.

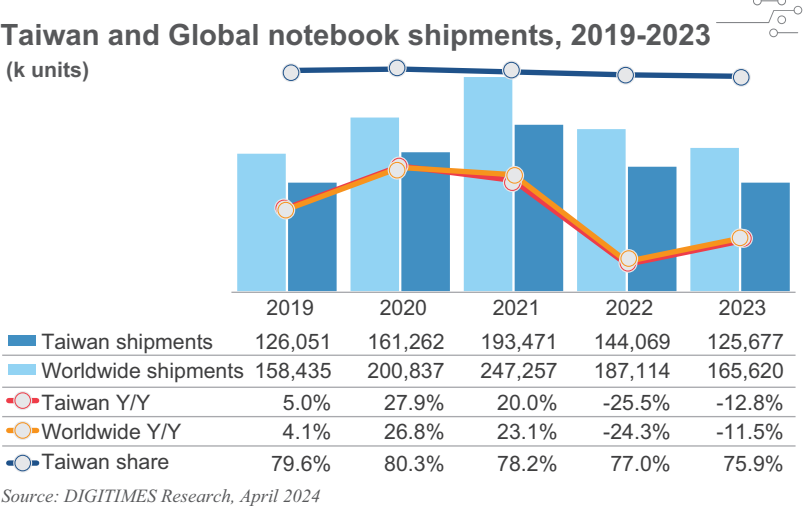
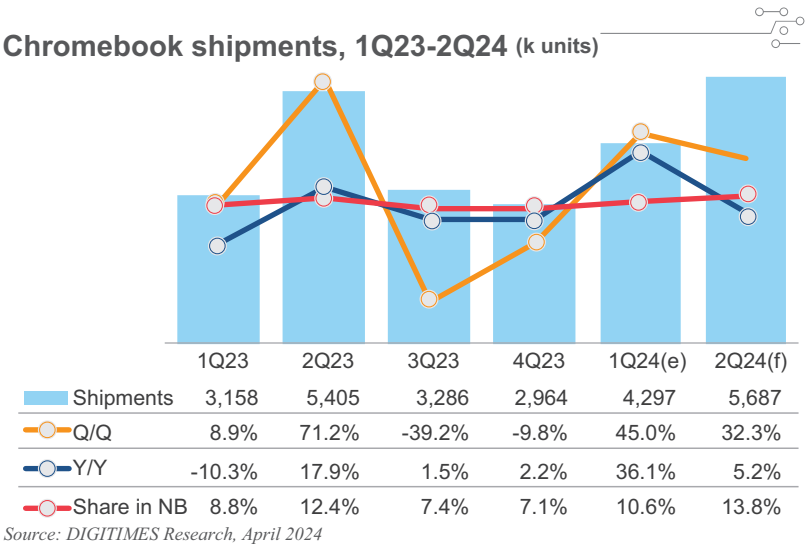
### Important factors

#### AI PC and memory

Microsoft is set to release a large multimodal model with related applications for offline use on WoA notebooks in May 2024 and the functionality is expected to stimulate demand for AI PCs.

The multimodal is expected to provide an offline voice/image assistance function, personalized fuzzy search function and smart voice control for Office.

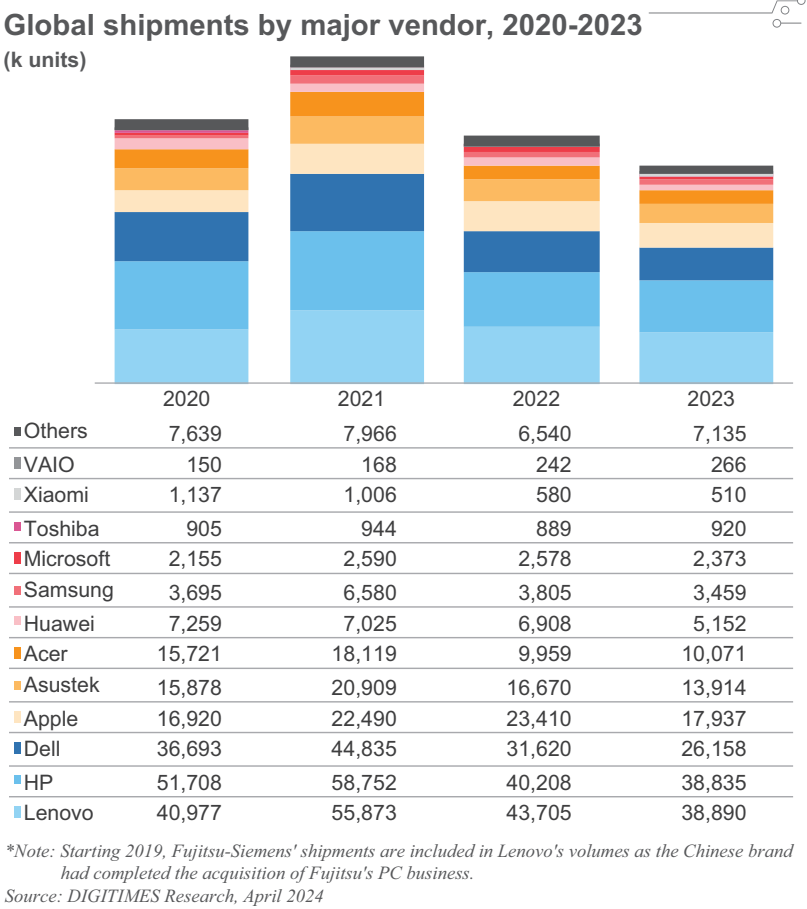
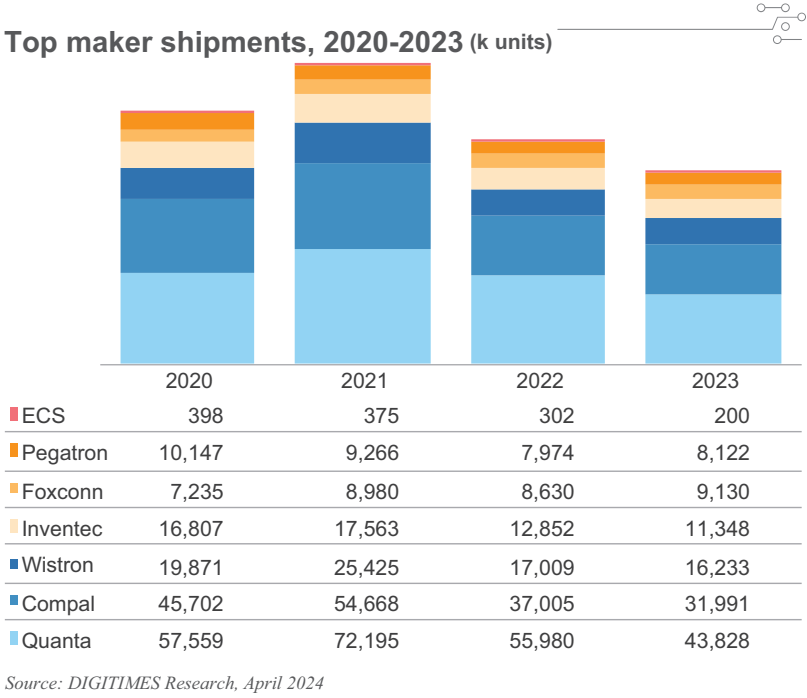
Notebooks featuring Qualcomm Snapdragon X Elite (Hamo) will be launched in May. The notebooks will feature high bandwidth



memory (HBM) and AI computing capability with the CPU being the only one so far recognized by Microsoft for operating its offline AI model.

Microsoft will also launch a Surface Pro and a Surface Laptop powered by the Snapdragon X Elite in the middle of May at an event.

### Annual shipments





...Continued from page 9

Brand vendors’ shipments in the second quarter are expected to grow by 4.2%, outperforming large CSPs. Seasonal demand will improve. Companies that were inactive in updating their machines in the past two years have begun to replace them. While supply of high-end GPUs such as H100 and MI300 has improved, brand vendors’ AI server shipments are expected to increase significantly.

Chinese CSPs are expected to experience a quarter-on-quarter decrease of 1.4% in the second quarter, mainly because of the high comparison base in the previous quarter. Some companies’ AI server orders are expected to increase in the second quarter, but traditional server shipments are expected to decline compared with the previous quarter.

Shipments breakdown

The global server brand and customer rankings changed in the first quarter of 2024. Among them, Supermicro surpassed Dell in shipments and rose to third place, while Microsoft leapfrogged HPE to fifth.

Supermicro surpassed Dell in the ranking as expected, mainly because Dell had more impact from the off-season effects and the GPU shortages in the first quarter. Supermicro’s shipments were better thanks to the stable supply of GPUs.

In addition to still having the highest purchase volume of high-end AI servers among all companies in the first quarter, Microsoft also updated some general-purpose server platforms, allowing its

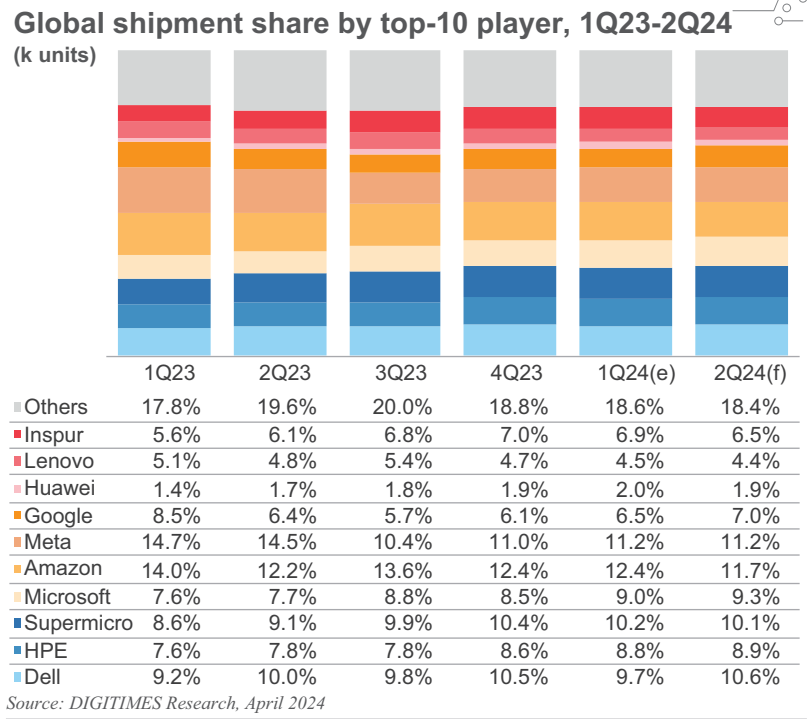
purchase volume to exceed that of HPE, which entered the off-season for shipments.

The global ranking is expected to change again in the second quarter of 2024. Dell’s shipments are expected to overtake Supermicro’s and its ranking will rise to third place. Google’s shipments are expected to surpass Inspur’s and its ranking will rise to seventh.

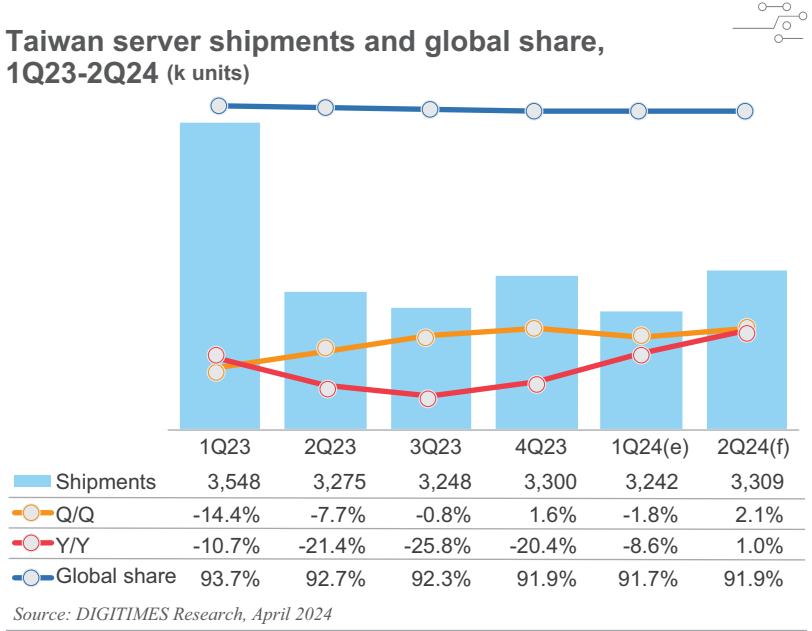
Dell will have a better supply of H100 GPUs in the second quarter, and AI server shipments will increase significantly compared with the first quarter. In addition, enterprise customers will increase their replacement demand for Eagle Stream and Genoa processor platforms, which will help Dell’s share of shipments exceed 10%.

Although Google’s shipment performance in the first quarter was not as substantial as expected, it remains on a quarterly growth track. In the second quarter, its shipments will increase in both traditional general-purpose servers and fifth-generation TPU AI servers, driving up its share to 7%.

According to DIGITIMES Research’s sources, Amazon suddenly lowered its shipment forecasts for ODMs and upstream components in April 2024. In addition to the impacts of the Middle East conflicts on its Israeli chip development team, Amazon probably has turned conservative about the outlook of the datacenter and cloud service markets. Amazon’s shipments are expected to decrease rather than increase in the second quarter, and its share of shipments will also drop to less than 12%.



Shipments from Taiwan makers



The server shipments of Taiwanese manufacturers in the first quarter of 2024 were lower than previously expected, with a quarter-on-quarter decrease of 1.8%. The performance was not as good as that of the global average, causing Taiwanese manufacturers’ global share to drop by 0.1pp to 91.8%.

The quarterly growth of shipments from Taiwanese manufacturers was slightly lower than that of the global average, mainly due to Chinese manufacturers Inspur and Lenovo lowering the proportion of orders released to Taiwanese manufacturers.

Among Taiwanese ODMs, Wiwynn and Quanta experienced quarterly growth, while the others saw quarterly declines that were mostly higher than the global

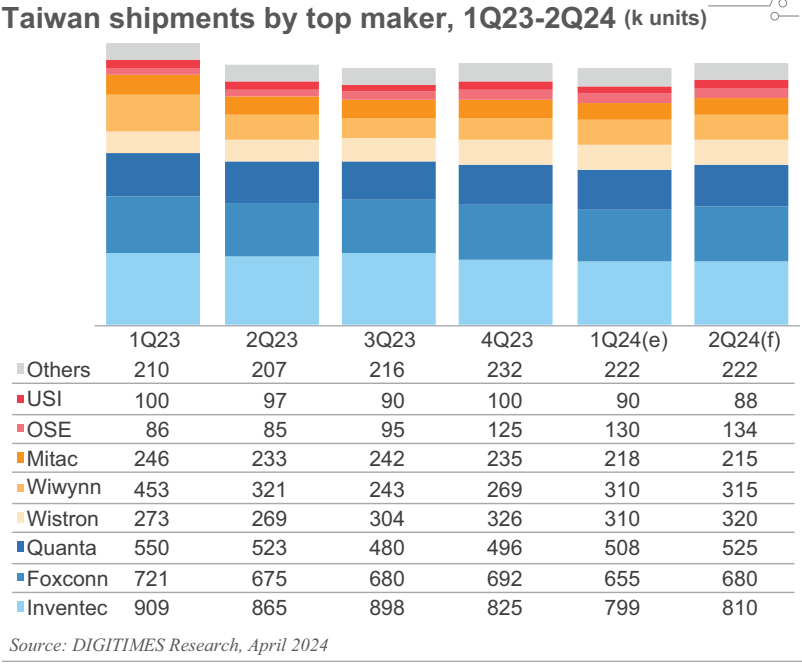
average.

In the second quarter of 2024, server shipments from Taiwanese manufacturers are expected to increase by 2.1% quarter-on-quarter, slightly outperforming the global average. This will help Taiwanese manufacturers’ share increase by 0.1pp to 91.9%.

This is mainly due to the fact that the shipment growth of American brand vendors, which mainly rely on Taiwanese ODMs, will be better than the global average.

Taiwanese manufacturers Foxconn, Wistron and Quanta will benefit from higher growth of orders from brand vendors and some CSPs, with better growths than fellow ODMs in the second quarter.

In the first quarter of 2024, Wiwynn and Quanta performed

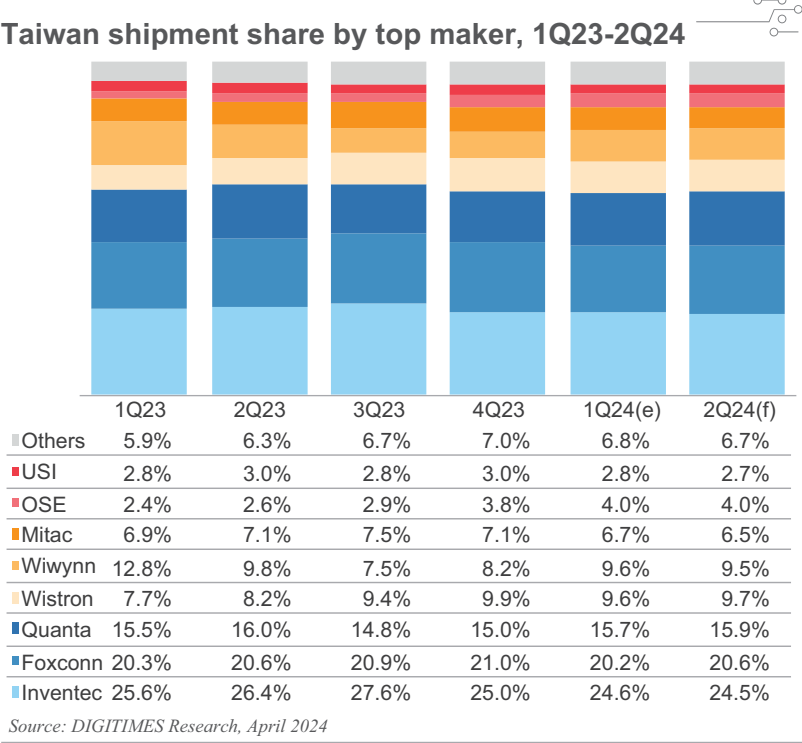


better in server shipments.

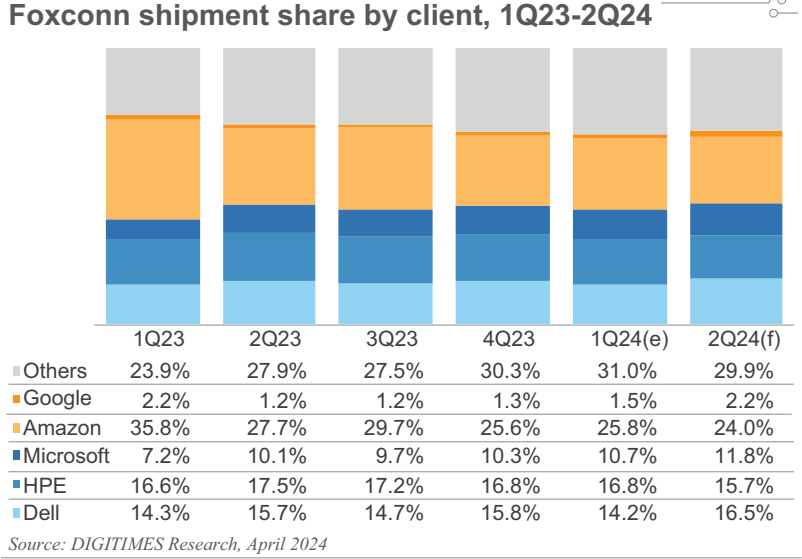
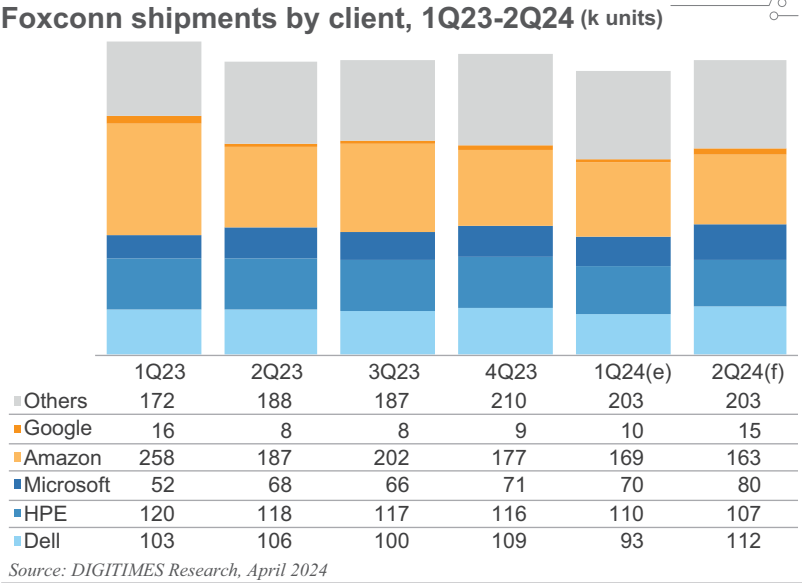
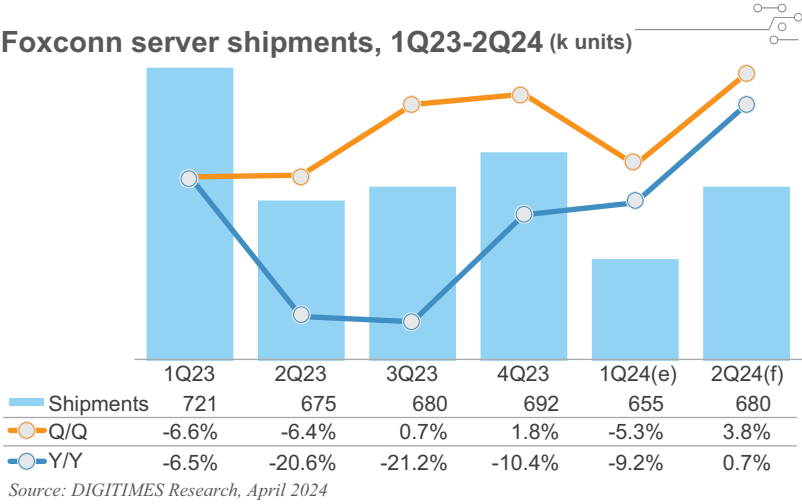
Wiwynn saw purchases from major CSP customers including Meta, Microsoft and Amazon increase significantly sequentially, and its share rose by 1.3pp compared with the previous quarter.

The growth in Quanta’s share was less than previously expected, but its share in the first quarter still increased by 1.1pp. Purchases from Google, its second-largest customer, and Microsoft, its third-largest customer, both increased by more than 10% compared with the previous quarter.

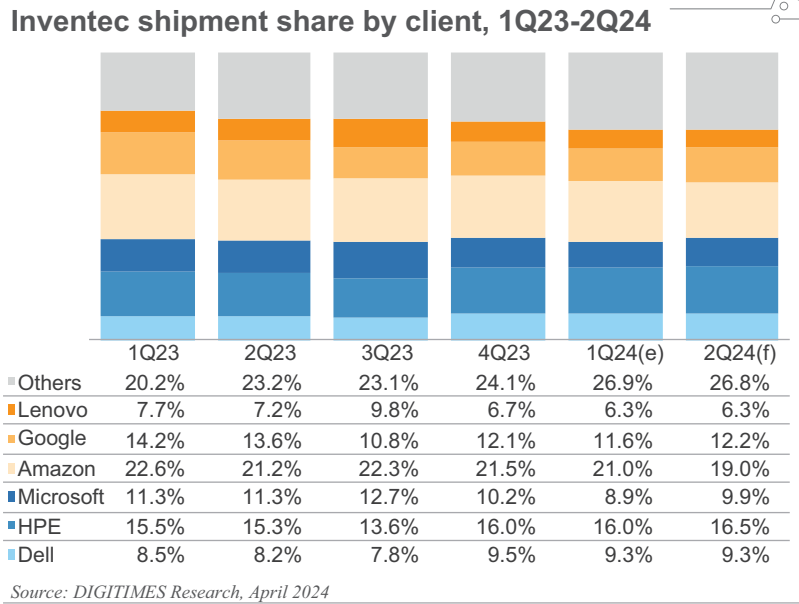
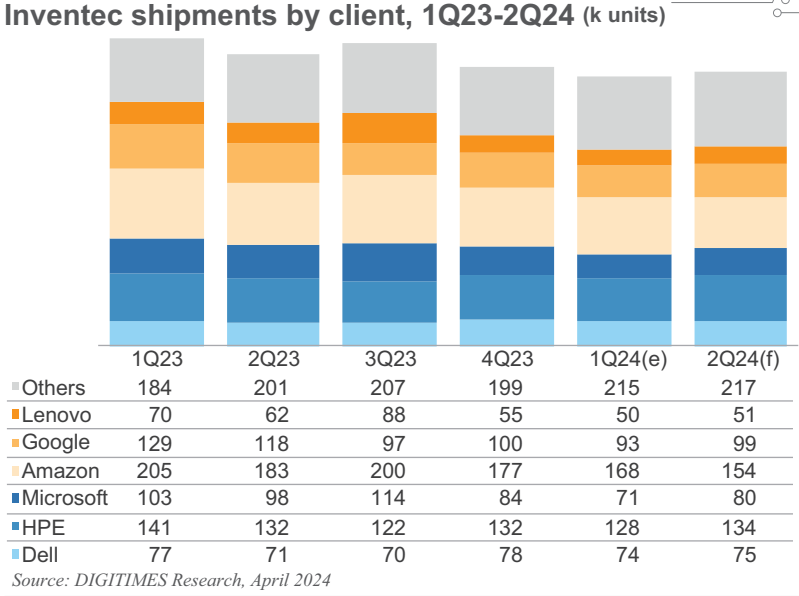
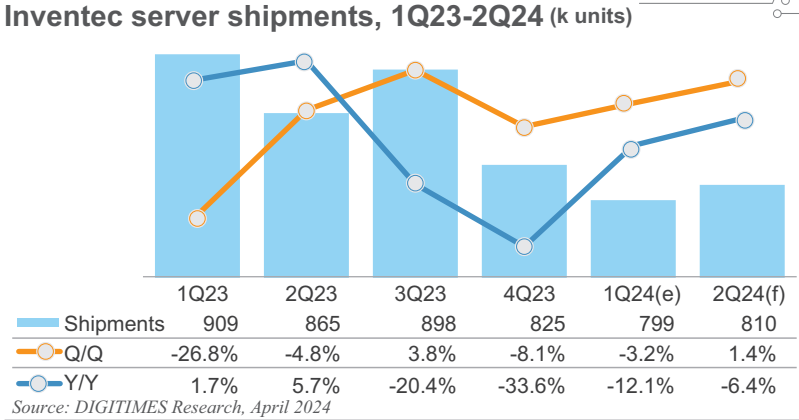
Foxconn, Quanta and Wistron



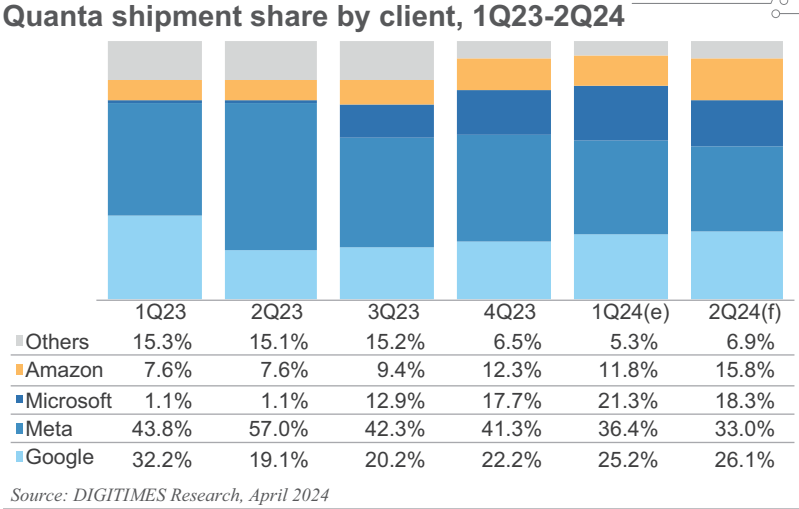
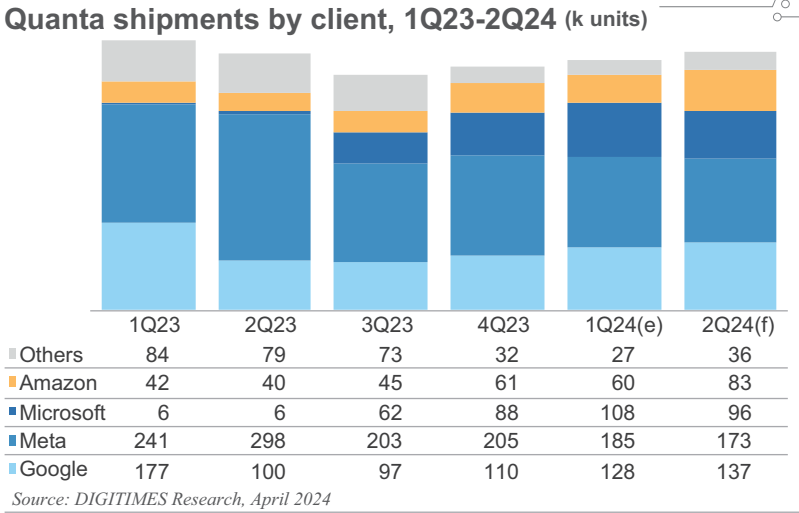
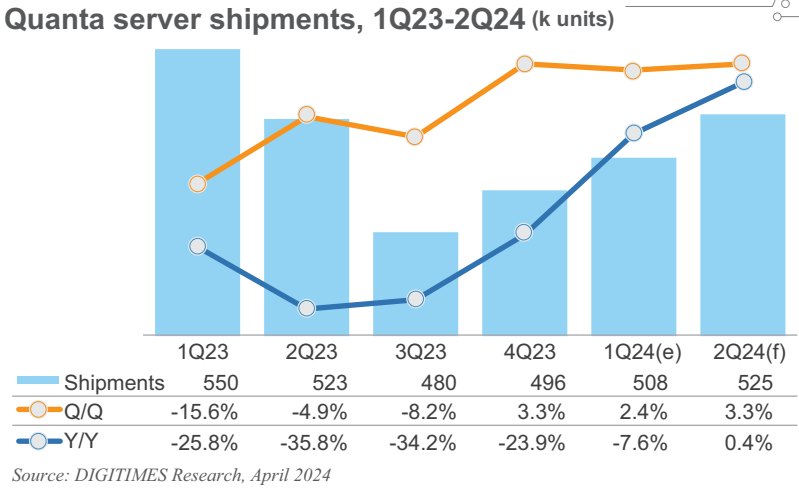
Foxconn



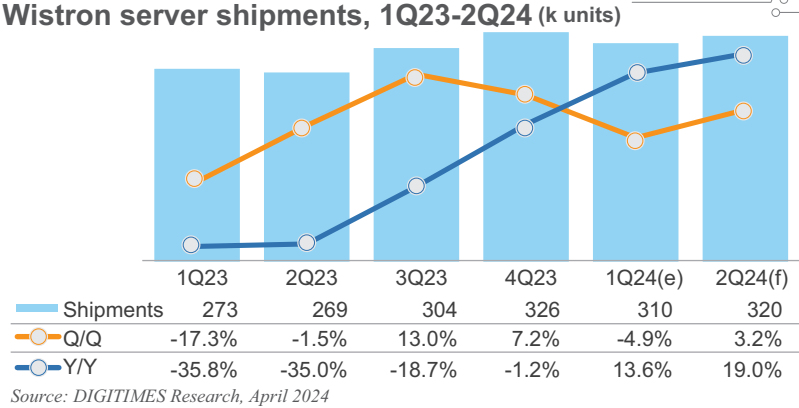
Inventec



Quanta



Wistron

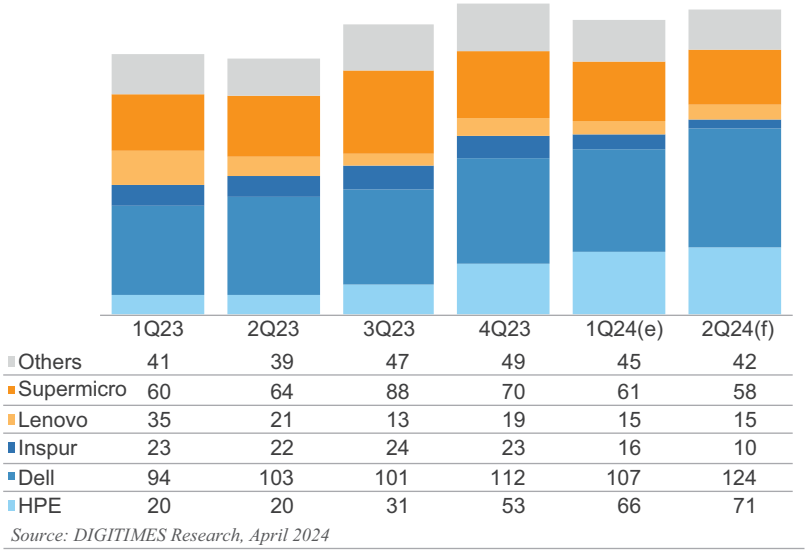


Continued on page 11...

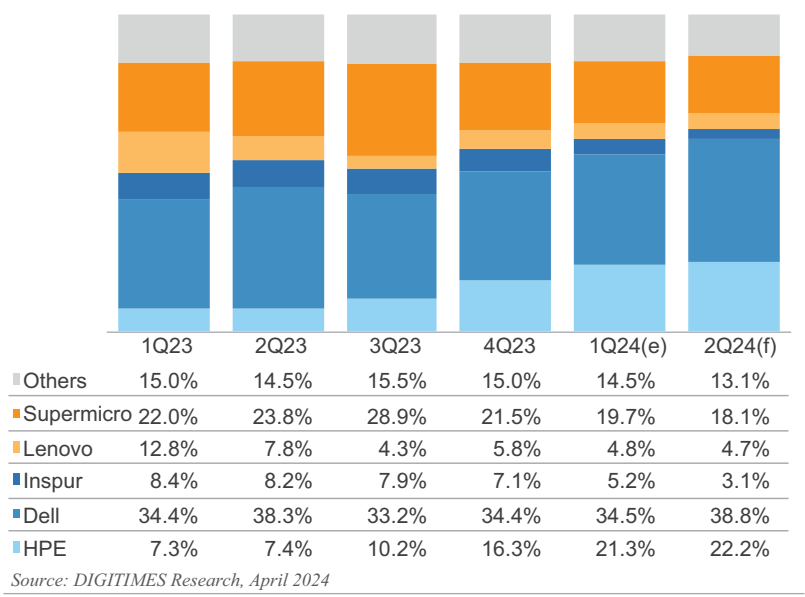


...Continued from page 10

Wistron shipments by client, 1Q23-2Q24 (k units)

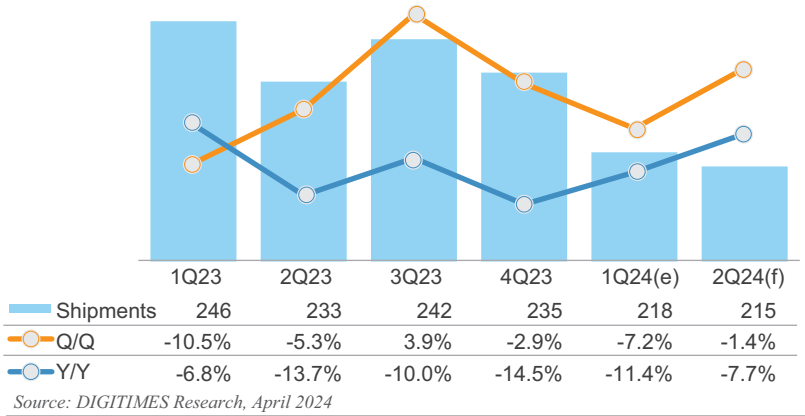


Wistron shipment share by client, 1Q23-2Q24

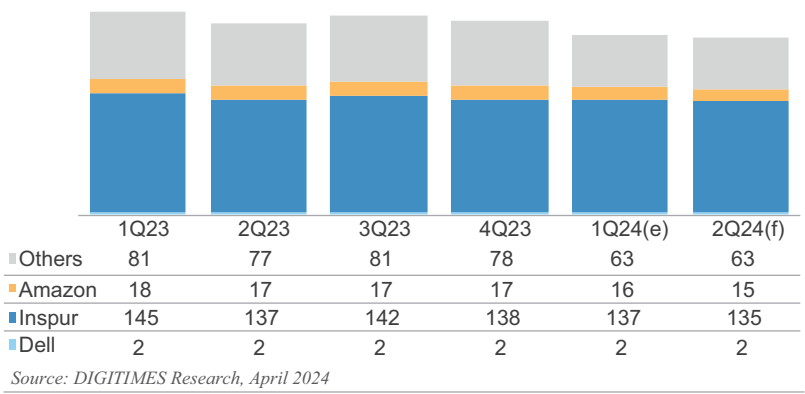


Mitac

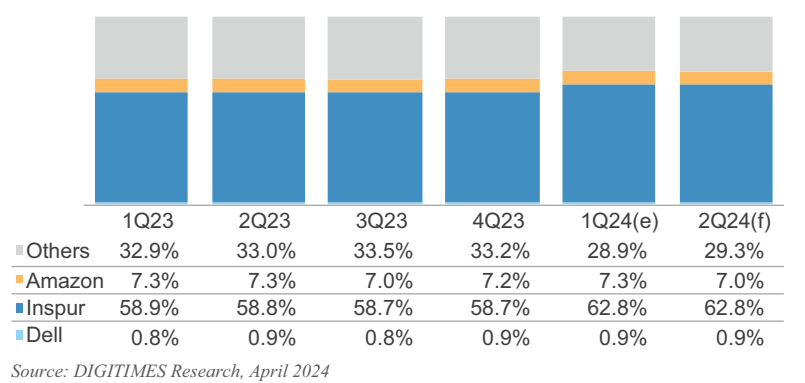
Mitac server shipments, 1Q23-2Q24 (k units)



Mitac shipments by client, 1Q23-2Q24 (k units)

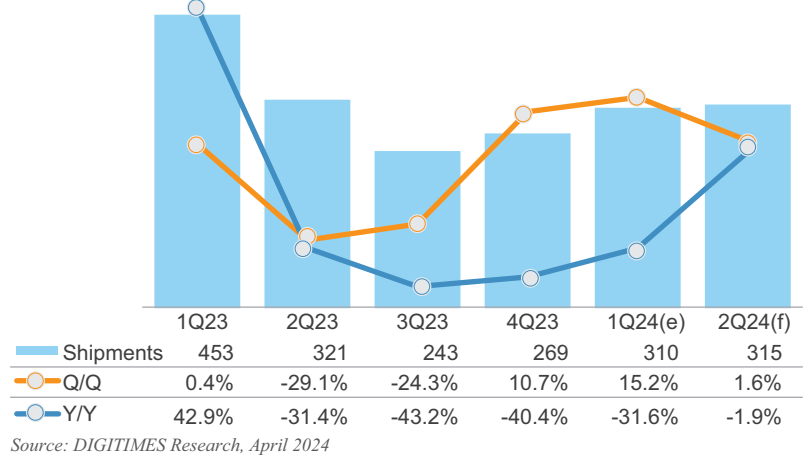


Mitac shipment share by client, 1Q23-2Q24

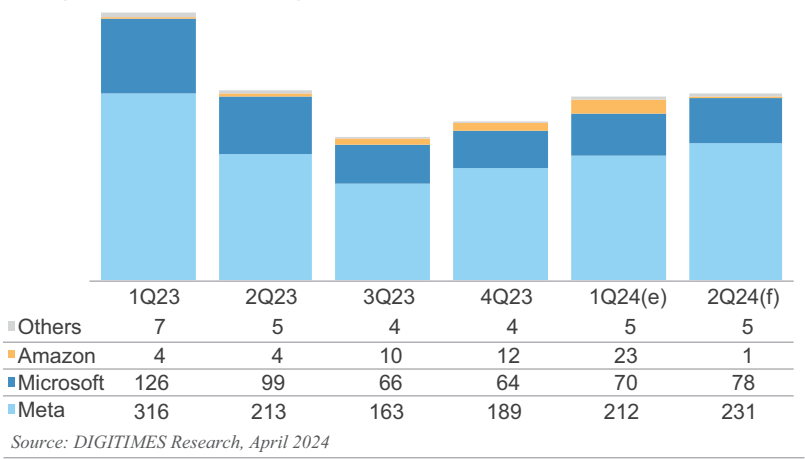


Wiwynn

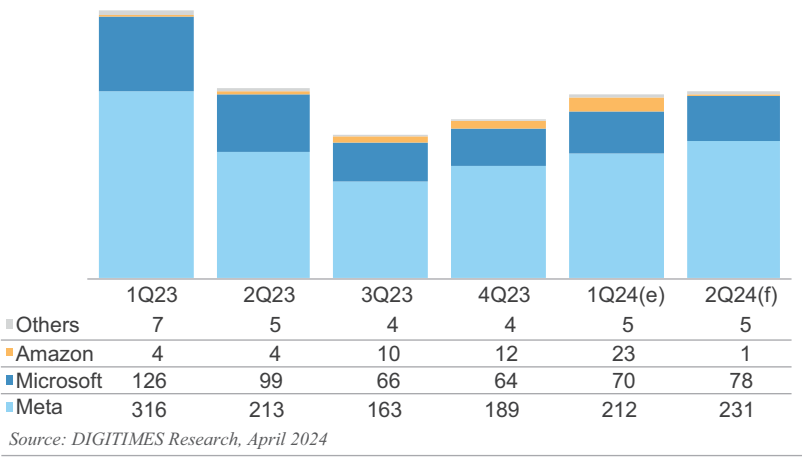
Wiwynn server shipments, 1Q23-2Q24 (k units)



Wiwynn shipments by client, 1Q23-2Q24 (k units)

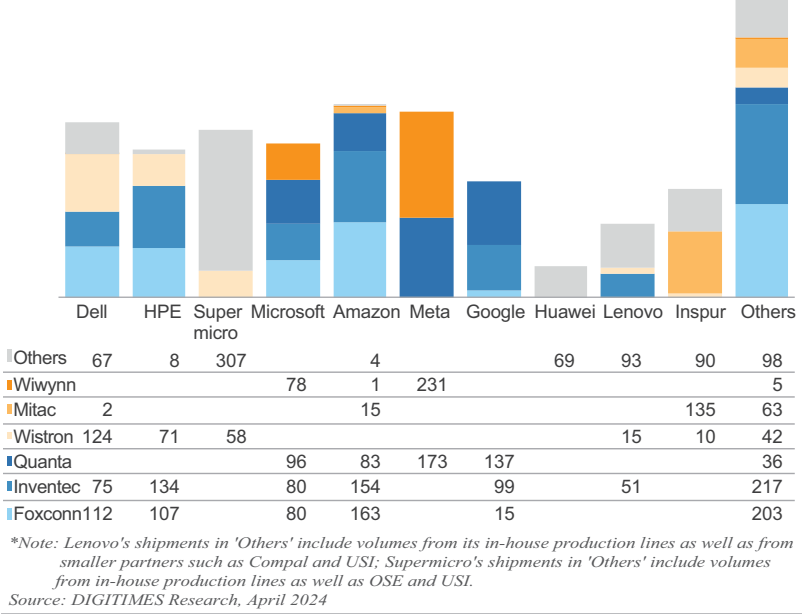


Wiwynn shipments by client, 1Q23-2Q24 (k units)



Client-maker server shipment matrix

Client-maker partnership matrix, 2Q24 (k units)



Amazon, Dell and Microsoft will see significant changes in the proportions of their outsourcing to manufacturing partners in the second quarter of 2024.

Amazon will increase the proportion of outsourcing to Quanta by 6.1pp to 19.8%.

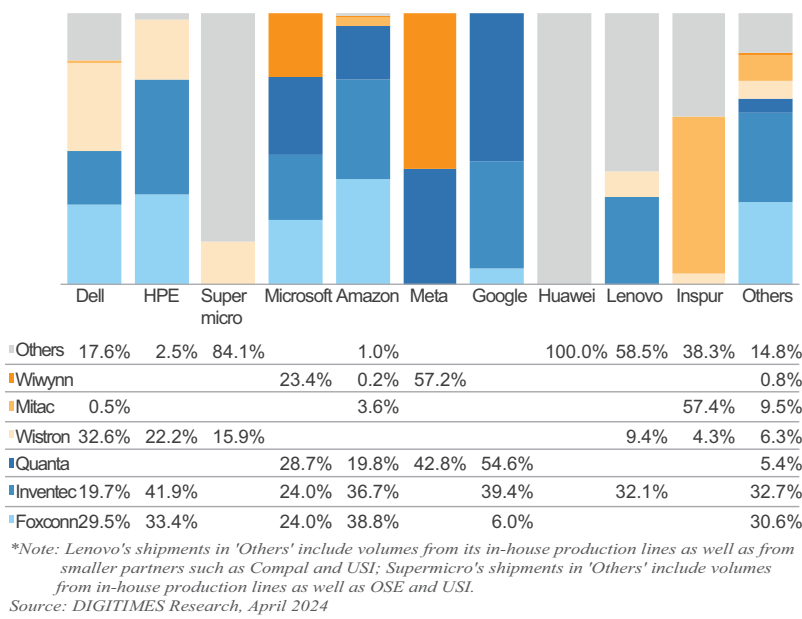
Due to design problems in Amazon's AI servers running on its in-house developed ASICs, it has also reduced the purchase of its own Graviton CPU. The main server ODM partners Wiwynn and Inventec will also be affected. Instead, Amazon is expected to increase purchases of Intel servers from Quanta, which will see its share of Amazon's orders rise.

Dell's outsourcing proportions for Foxconn and Wistron in the second quarter will increase by 2.3 and 1.7pp respectively to 29.4% and 32.7%. This is mainly because Dell's new models launched in the second quarter will be mainly manufactured by Foxconn and Wistron.

In the second quarter, Microsoft's outsourcing proportions for Foxconn, Inventec and Wiwynn will increase by 2.4, 1.7 and 1.2pp respectively to 24.1%, 23.9% and 23.2%.

In the second quarter, Microsoft will ramp up its purchase of new AMD models, with Foxconn being the main supplier.

Client shipment share by maker, 2Q24



In the second quarter of 2024, Wiwynn, Wistron and Quanta will see significant changes to their main server customer structures.

Wiwynn's order dependence on Meta and Microsoft will increase by 5pp and 2pp respectively to 73.3% and 24.6%. Amazon's AI ASIC has technical issues and has lowered its server orders for Wiwynn, while Meta has been more active in purchasing general-purpose servers from Wiwynn in the second quarter.

Wistron's dependence on Dell and HPE for orders in the second quarter will increase by 4.5pp and 1pp respectively to 38.9% and 22.3%.

Dell will rush to ship new Intel products and AI servers in the second quarter at the same time, with Wistron as its main ODM.

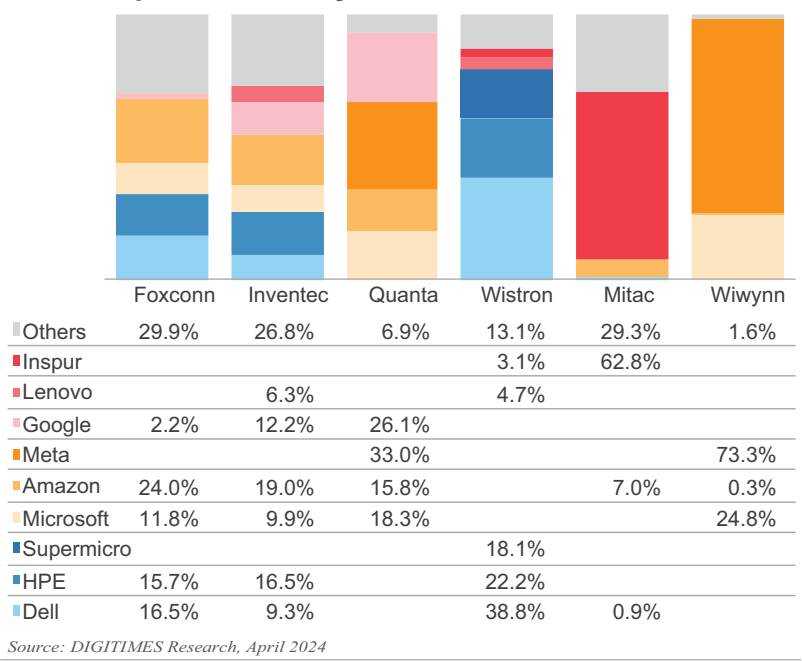
HPE's AI server collaboration with Gigabyte is coming to an end, and AI server orders in 2024 will mainly be handled by Wistron.

Quanta's dependence on Amazon and Google for orders in the second quarter will increase by 4pp and 1pp respectively to 15.9% and 26.1%.

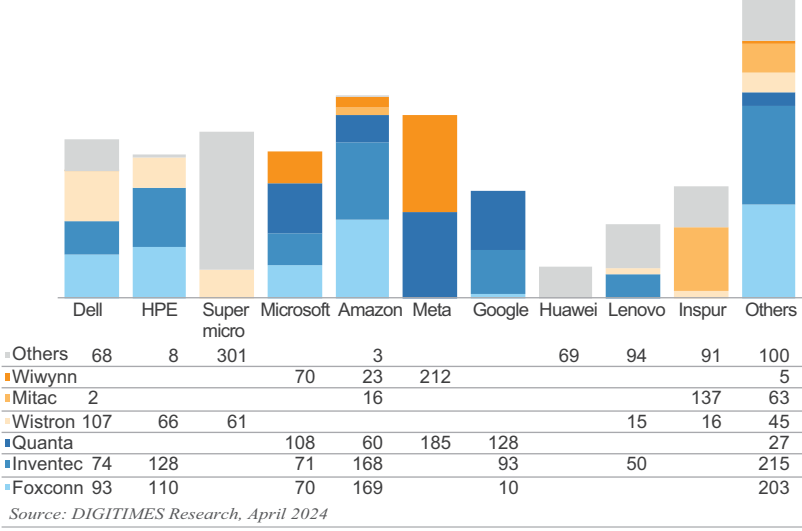
Amazon is expected to increase its purchases of Intel servers from Quanta in the second quarter.

Google will continue ramping up purchases of new general-

Maker shipment share by client, 2Q24



Client-maker partnership matrix, 1Q24 (k units)



purpose and AI servers – both from Quanta – in the second quarter, which will see Quanta become more dependent on Google.

In the first quarter of 2024, Microsoft, Meta and Google changed significantly the proportions of orders for their outsourcing partners.

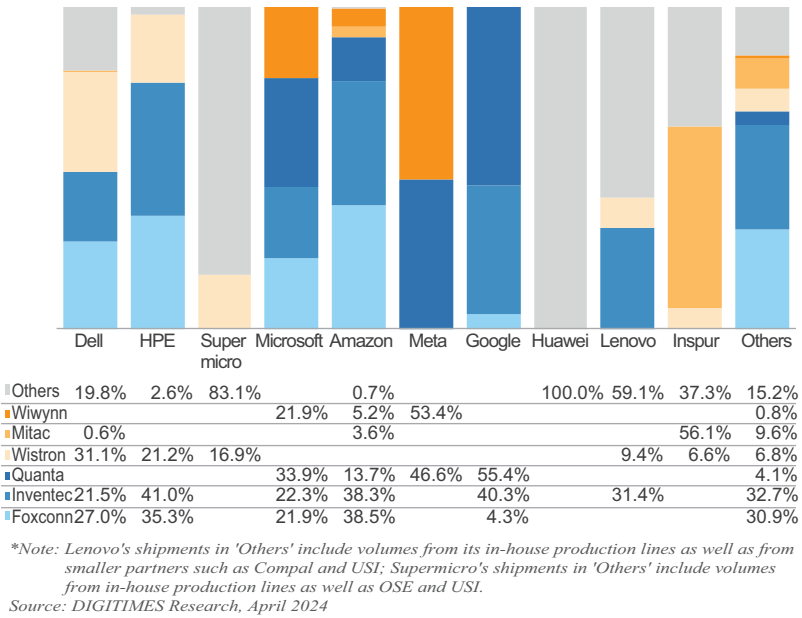
Microsoft's purchases of Intel general-purpose servers from Quanta continued to increase in the first quarter, and shipments of high-end AI servers also increased compared with the previous quarter. Microsoft increased the proportion of orders for Quanta by 5.2pp

to 34%. Microsoft's purchases of general-purpose servers from Wiwynn also rose slightly sequentially.

In the first quarter, Meta released most of its orders for new general-purpose server models to Wiwynn, increasing the proportion of outsourcing for Wiwynn by 5.5pp to 53.4%.

Google's purchases of new-generation Intel servers relied mainly on Quanta. Coupled with the increase in purchases of TPU and GPU AI servers, its proportion of orders released to Quanta increased by 5.1pp to 55.3%.

Client shipment share by maker, 1Q24



In the first quarter of 2024, Quanta, Wistron and Wiwynn saw significant changes to their server customer structures.

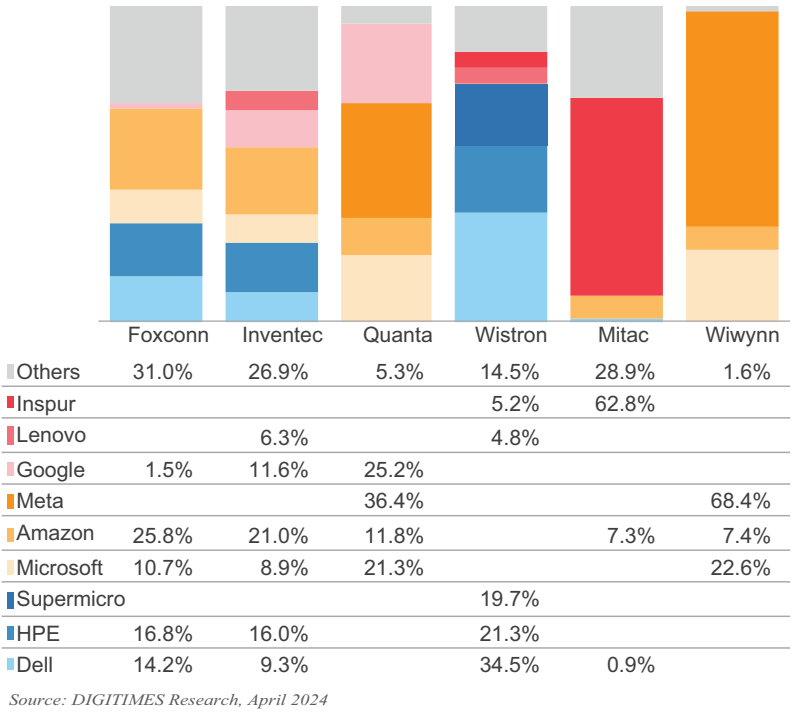
Quanta's dependence on orders from Microsoft and Google increased by 3.5pp and 3pp respectively to 21.3% and 25.1%.

Quanta's shipments of new Intel models and high-end AI servers to Microsoft continued to grow in the first quarter.

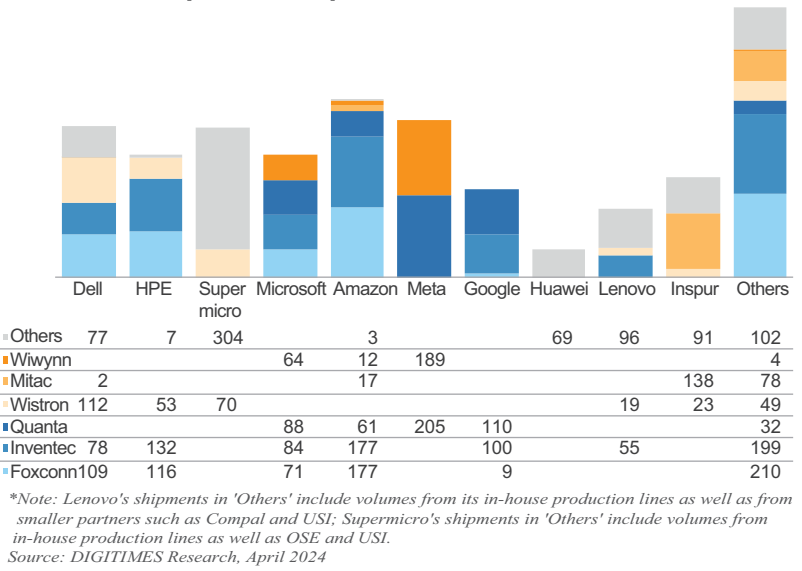
Wistron's dependence on HPE's orders increased by 5pp to 21.3%. HPE relied mainly on Wistron for manufacturing new platform models in the first quarter.

Wiwynn's dependence on Amazon increased by 3pp to 7.4% in the first quarter. Amazon depended on Wiwynn as its main manufacturer for AI servers using its own ASIC, and shipments in the first quarter continued to increase.

Maker shipment share by client, 1Q24



Client-maker partnership matrix, 4Q23 (k units)



To get the latest figures and information from DIGITIMES Research's studies such as notebooks, servers, AI applications, foundries and IC designs, please visit <https://www.digitimes.com/reports/>







Global Leader in  
AI-Powered Cybersecurity



# AI-Powered Cybersecurity

2024.06.04 ~ 06.07



## Securing AI Workforce

Explore AI-Misuse &  
Safeguard AI PC &  
Usage



## Safeguarding Enterprise GenAI

Explore Risks & Best  
Practices for Securing  
Private LLMs



## Securing NexGen AI Data Center

Essential Strategies for  
Protecting Tomorrow's  
Data center



## AI-Powered Cybersecurity

Empowering CISOs &  
Enterprise Resiliency

# Visit Trend Micro booth and get a nice coffee for free!

📍 AI Security Cafe  
R103/R104, TaiNEX 1  
(next to 1F Area J Entrance)

