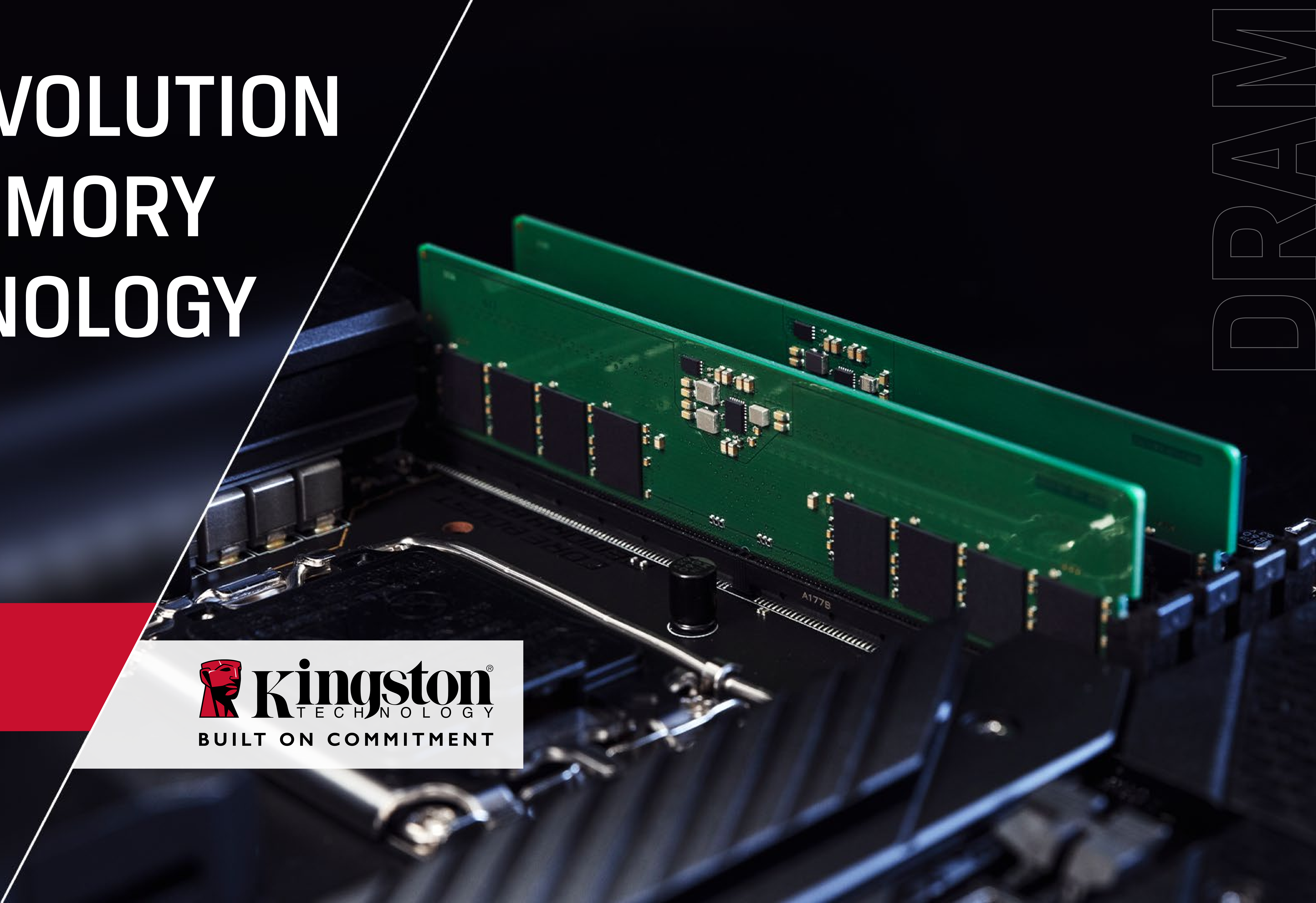
THE EVOLUTION OF MEMORY TECHNOLOGY



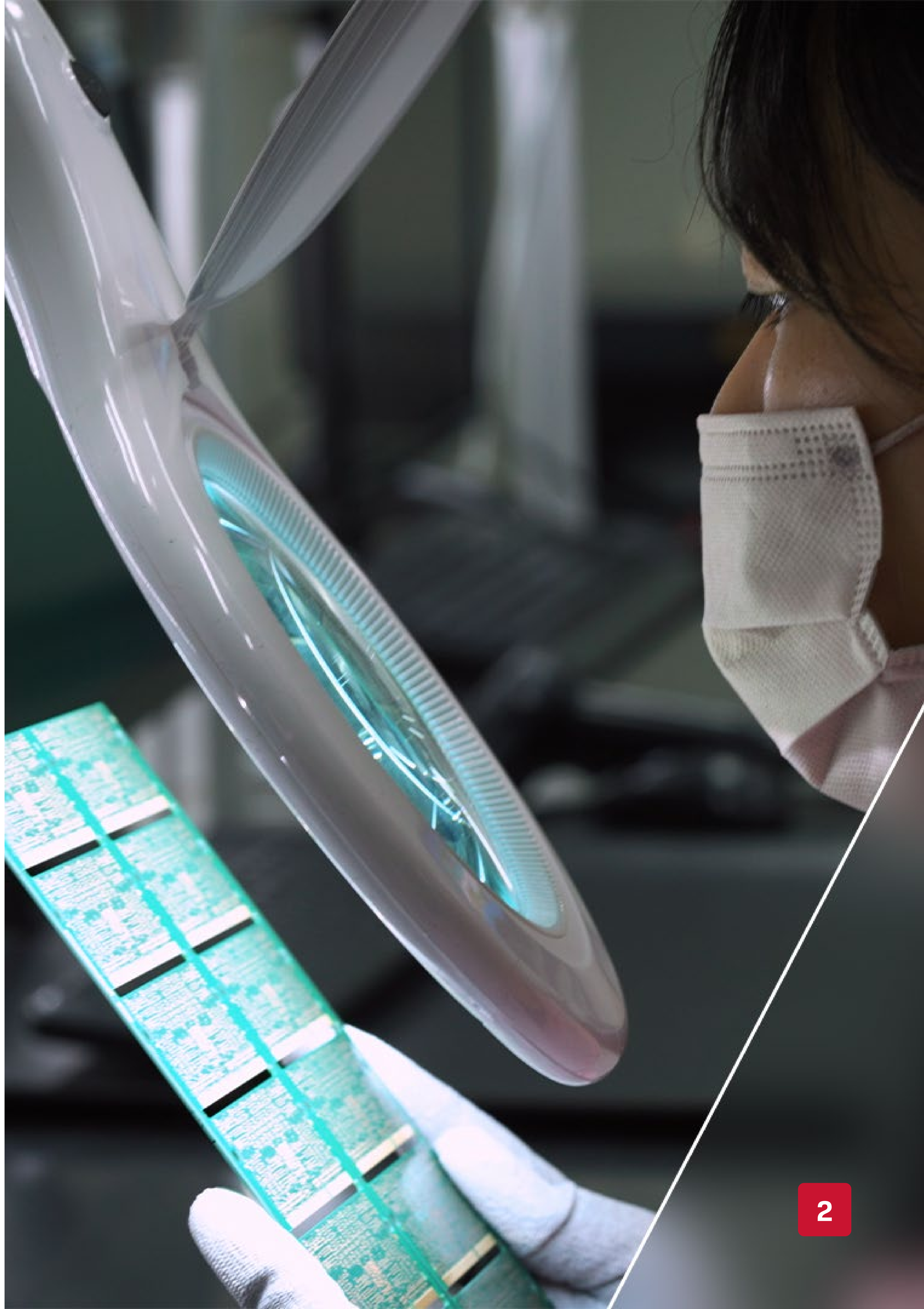
Kingston TECHNOLOGY

BUILT ON COMMITMENT

Foreword and Contents

The evolution of Dynamic Random Access Memory (DRAM) from Fast-Page Mode (FPM) to Synchronous (SDRAM) to Double Data Rate (DDR SDRAM), now in its 5th generation (DDR5), represents a significant progression in computing technology, driven by the need for higher performance, greater bandwidth, and improved energy efficiency. Amid the wave of AI applications, this demand continues to mushroom, with DDR5 and HBM (High-Bandwidth Memory DRAM), gaining significant traction in data centers and client systems. Considered the workhorse of semiconductor memory, DDR SDRAM holds a unique place in the industry thanks to its low power consumption and high performance, able to transfer data to the processor quickly.

According to industry experts, there's no end in sight for DRAM as the core memory technology. But what exactly makes this memory type so adept at meeting the needs of today's businesses, in terms of performance and architecture? Are some types better suited for servers versus desktops? How is the technology evolving and what are the typical challenges and use cases? This eBook will address these questions and explore what the future holds for DRAM, with the help of some of Kingston's technical experts.



Contributors

This eBook has been created by two Kingston experts.

Mike Mohney | Kingston Technology

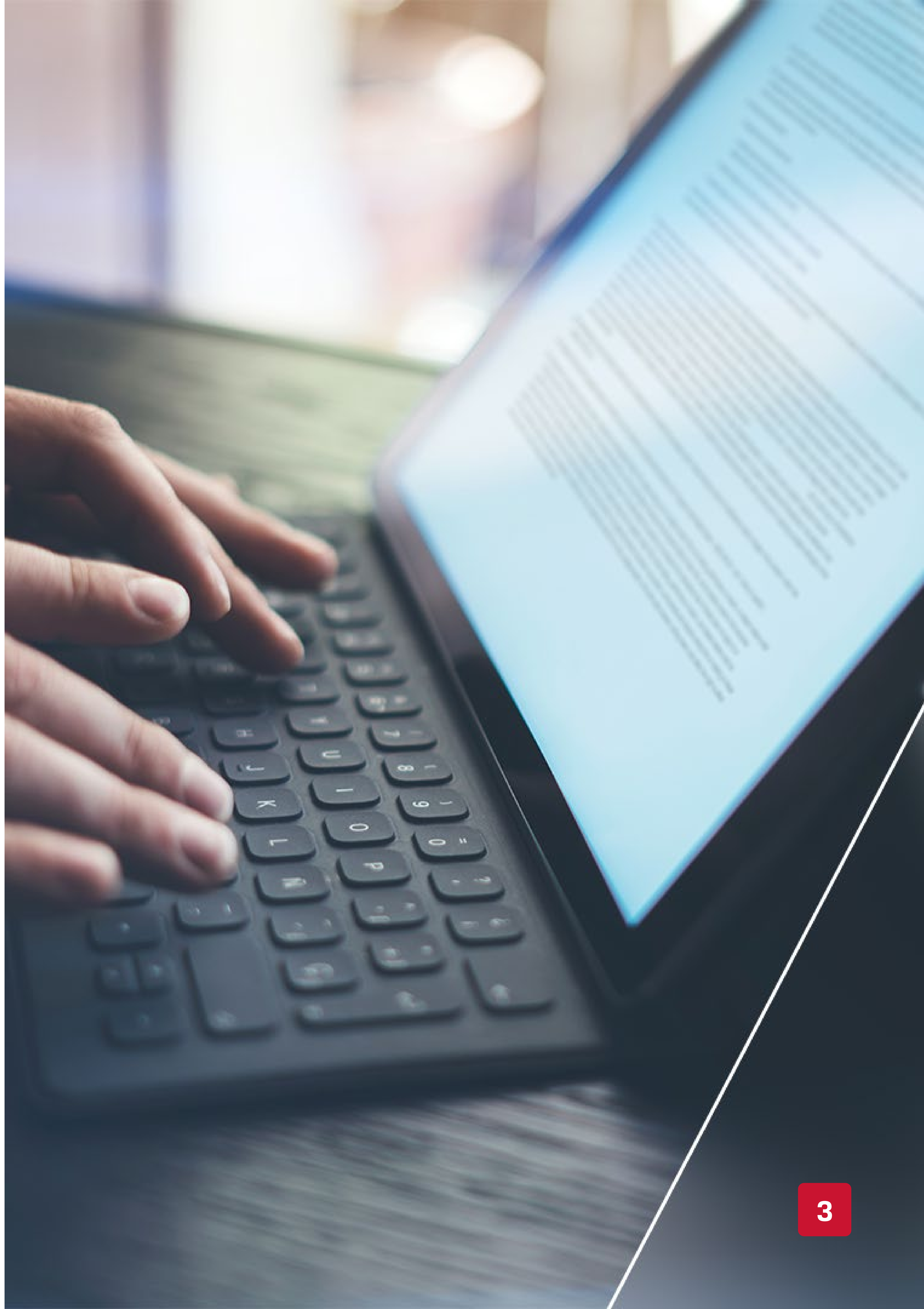
Mike Mohney is a Senior Technology Manager at Kingston Technology, based in Fountain Valley, California. He has been with Kingston since 1996, bringing over 28 years of experience to the company.

In his role, Mike has been instrumental in managing and advancing Kingston's technology initiatives, particularly in the DRAM and memory solutions sector. His expertise and leadership have contributed significantly to Kingston's position as a leading third-party manufacturer of DRAM solutions.

Geoffrey Petit | Kingston Technology Europe

Geoffrey Petit is the Technical Resources Group Team Leader at Kingston Technology Europe. He joined Kingston in 2016 as a Technical Support Engineer, providing post-sales technical support to customers based in EMEA, delivering technical training to colleagues and new employees, and testing new products.

Geoffrey and his teams' responsibilities are to provide technical support for customers and handle pre-sales queries from business managers, Marketing, the Inside Sales team, customers, and key personnel.



> *Higher densities were now possible, meaning more memory capacity could be achieved on a single module, which was critical for servers and high-performance computing.*
>
> **Geoffrey Petit | Kingston Technology Europe**

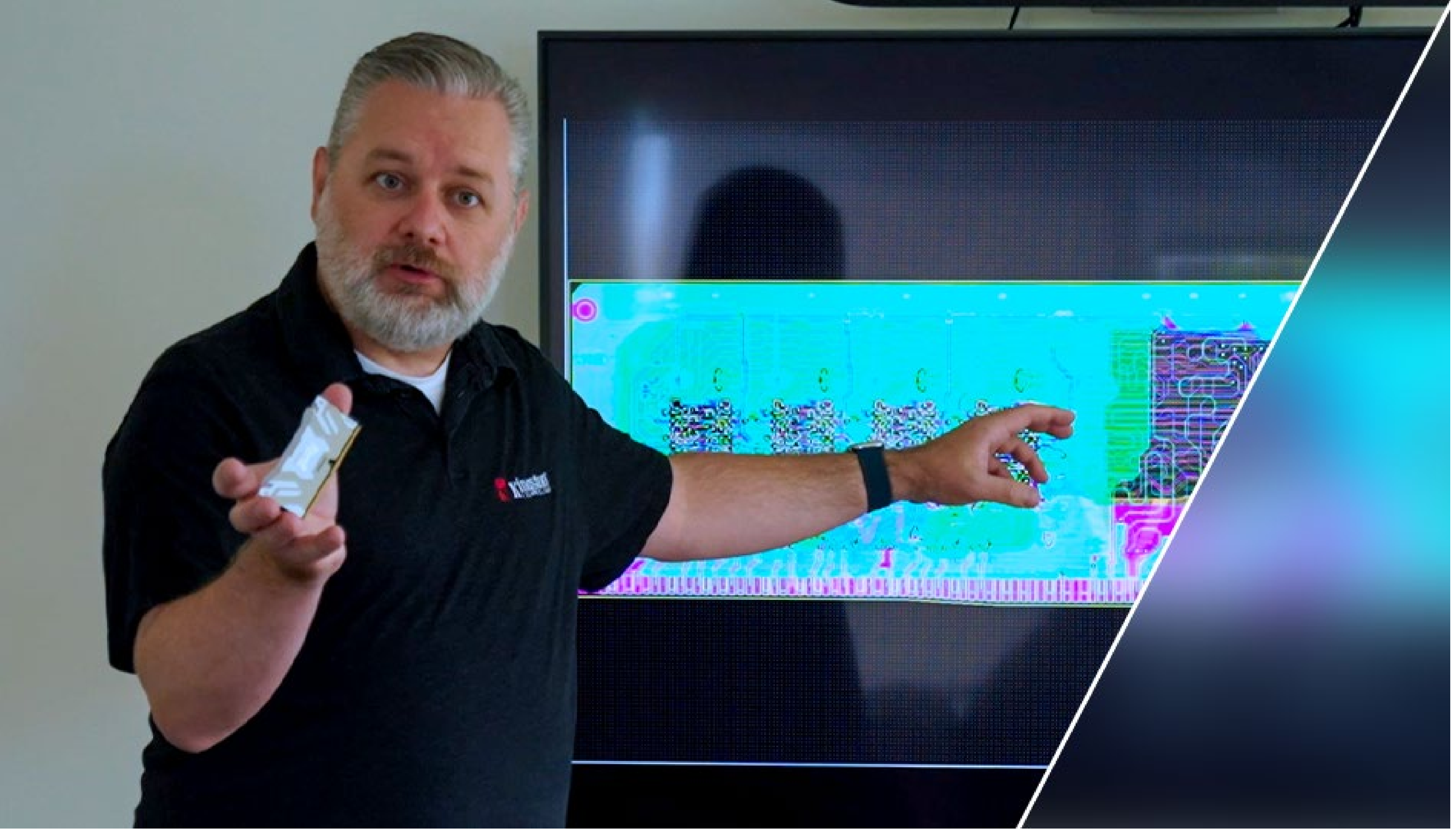
The Evolution of DRAM: From FPM to DDR5

By the mid-1980s, the PC revolution was in full swing with the introduction of the 80486 processor. Fast Page Mode (FPM) DRAM on SIMMs (Single In-Line Memory Modules) was the main memory technology. The need to increase performance led to the development of EDO (Extended Data Out) DRAM in the early 1990s, followed quickly by SDRAM and the DIMM (Dual In-line Memory Module), which operated more efficiently by aligning with the CPU clock and at a single data rate. In 2000, the first DDR (Double Data Rate) SDRAM was launched, and doubled the data rate by transferring data on both the rising and falling edges of the clock signal. It was also more power efficient than its predecessor, dropping to 2.5V per module compared to 3.3V. DDR SDRAM continued to evolve under the careful planning of the industry standards body (JEDEC), with the 2nd generation DDR (DDR2) launching in 2003. This was followed in 2007 with DDR3, then in 2014 with DDR4. Each generation increased memory speeds and capacities, and lowered the operating voltage, capitalizing on improvements to semiconductor wafer lithography and shrinking memory cells.

Fast forward to 2021 and DDR5 SDRAM was introduced, representing a major advancement in memory technology. DDR5 debuted at a speed of 4800MT/s, representing a 50% increase in bandwidth over the final speed of DDR4 at 3200MT/s. In addition to speed, DDR5 modules incorporated a Power Management IC (PMIC) which helped regulate the power required by the various components of the memory module, providing for better power distribution than previous generations, improving signal integrity and reducing noise. The trend to reduce power consumption continued, with DDR5 only requiring 1.1V to operate. Significant improvements to data integrity were also designed in, such as On-Die ECC (Error Correction Code) that could catch and correct bit errors within the DRAM component, reducing the likelihood of corrupted data.

> *In addition to the primary performance, power consumption, and density improvements, many other features have been designed in to each new generation. This includes enhanced error correction technology, improvements to signal integrity, mitigations added to prevent hardware hacking vulnerabilities, and new form factors.*
>
> **Mike Mohney | Kingston Technology**

Since launching, DDR5 has had four planned speed increases, supported by successive Intel and AMD platforms. Historically, memory speed increases happen once annually, following a predetermined cadence set by memory industry standards and enabled by new chipsets. DDR5's skipping over speed bins was in part due to competition from the chipset and processor manufacturers, and demand for high-performance memory to handle memory-bandwidth-intensive applications such as AI.



DRAM Module Types and Key Differences

The JEDEC industry standards body doesn't just define the specifications for DRAM memory, it also determines the form factors DRAM resides on to suit various compute platforms and environments.

Unbuffered modules, such as **Unbuffered DIMMs (UDIMMs)** and **unbuffered Small Outline DIMMs (SODIMMs)** are the most common type of memory modules used in consumer desktops and laptops.

The addition of DRAM components to support Error Correction Code (ECC) make for **ECC UDIMMs** and **ECC SODIMMs** in support of mainstream workstation-class systems. These provide critical data integrity support for memory-intensive applications.

For the single or multi-processor server, **ECC Registered DIMMs (RDIMMs)** feature a register component on the module, which buffers the data between the DRAM and the memory controller. This is critical in environments where large amounts of memory are required and where data reliability is key.

Load Reduced DIMMs (LRDIMMs) feature data buffers to reduce loads on the memory controller, which would otherwise toggle memory speeds down to compensate. LRDIMM technology enables large-capacity modules without sacrificing performance and was first introduced in 2012 for DDR3, then refined for DDR4 in 2014.

Low Power DDR (LPDDR) entered the market in 2006 as a solution for mobile devices to save on battery power. Although typically mounted directly to a system board, since 2024, LPDDR5 can also be used on the CAMM2 (Compression Attached Memory Module) form factor, providing a modular solution manufacturers can use in laptops or small-form-factor PCs.

Besides DDR SDRAM, the fastest-growing category of memory is **High Bandwidth Memory (HBM)**, developed by AMD in 2008 to address the increasing demand for high-performance and high-capacity memory to support GPUs with lower power requirements. HBM uses a high-speed interface to manage a 3D stack of SDRAM layers within a single chip package. This provides wide (128-bit+) addressable memory accesses and squarely targets graphics cards, on-package processor memory, and AI accelerator cards.

> *HBM evolved through successive generations over the last decade to increase support for higher memory capacities in more layers, wider data buses, and higher-performance throughput. That said, HBM memory is not currently used on memory modules and is not considered a viable alternative technology to DDR DRAM to scale in price per GB.*
>
> **Mike Mohney | Kingston Technology**

The Importance of Latency and Speed

Latency and speed are two key attributes defined by the memory industry standards body (JEDEC) that are used as performance metrics.

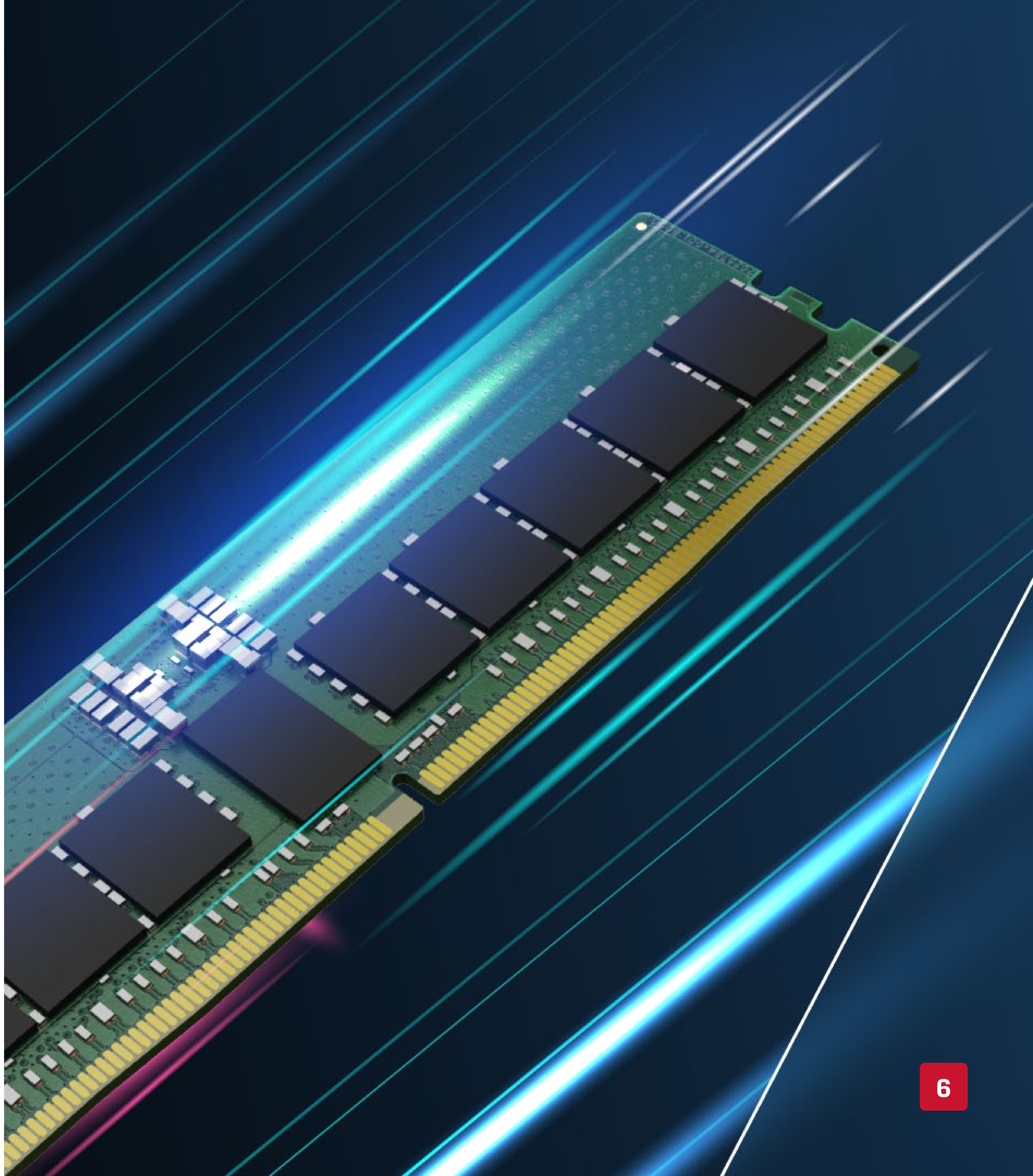
In computing, there are many different types of applications that may utilize one piece of hardware over another for its workload. Memory-bound applications will benefit from high-performance memory speeds and lower latencies, as opposed to those that are storage- or GPU-focused.

Mike Mohney | Kingston Technology

For every new memory technology, JEDEC specifies the standard speeds and timings, which are used by memory manufacturers, processor and chipset architects, and motherboard/system manufacturers to stay in alignment. As industry-standard memory speeds increase, latencies also increase. This is often a point of controversy and misunderstood by users who believe faster standard speeds are negated by increased CAS latency timings. Total latency however, which is a combination of speed and timings, is a more accurate way of measuring memory performance in nanoseconds. This refers to the time it takes for the processor to receive data from memory.

When it comes to the impact on computing tasks, unbuffered memory is ideal for desktops and workstations needing quick response times. Server-class memory, like Registered and Load Reduced DIMMs, excels in data centers where stability, error correction, and handling large datasets are more critical than latency.

Geoffrey Petit | Kingston Technology Europe



Typical Use Cases and Workload Impact

End-users and data center architects choose platforms based on their application and workload needs. In turn, application demand for memory capacity and performance dictates the type of modules chosen and configured.

In the memory industry, we differentiate between client (aka PC) and server classes of components and module form factors. Client-class systems encompass desktops and laptops using industry-standard non-ECC memory in the unbuffered DIMM (UDIMM/CUDIMM), SODIMM/CSODIMM, and CAMM2 form factors. Meanwhile, server-class systems, including desktop workstations and mobile workstations, use memory modules supporting ECC (Error Correction Code).

Consumer devices favor the simplicity and speed of unbuffered memory. Desktop and laptop computers aren't designed to operate 24 hours a day and are typically powered off when not in use. The types of applications and workloads on these systems also don't push the tolerances of memory components the way servers do, so there's no need for ECC support.

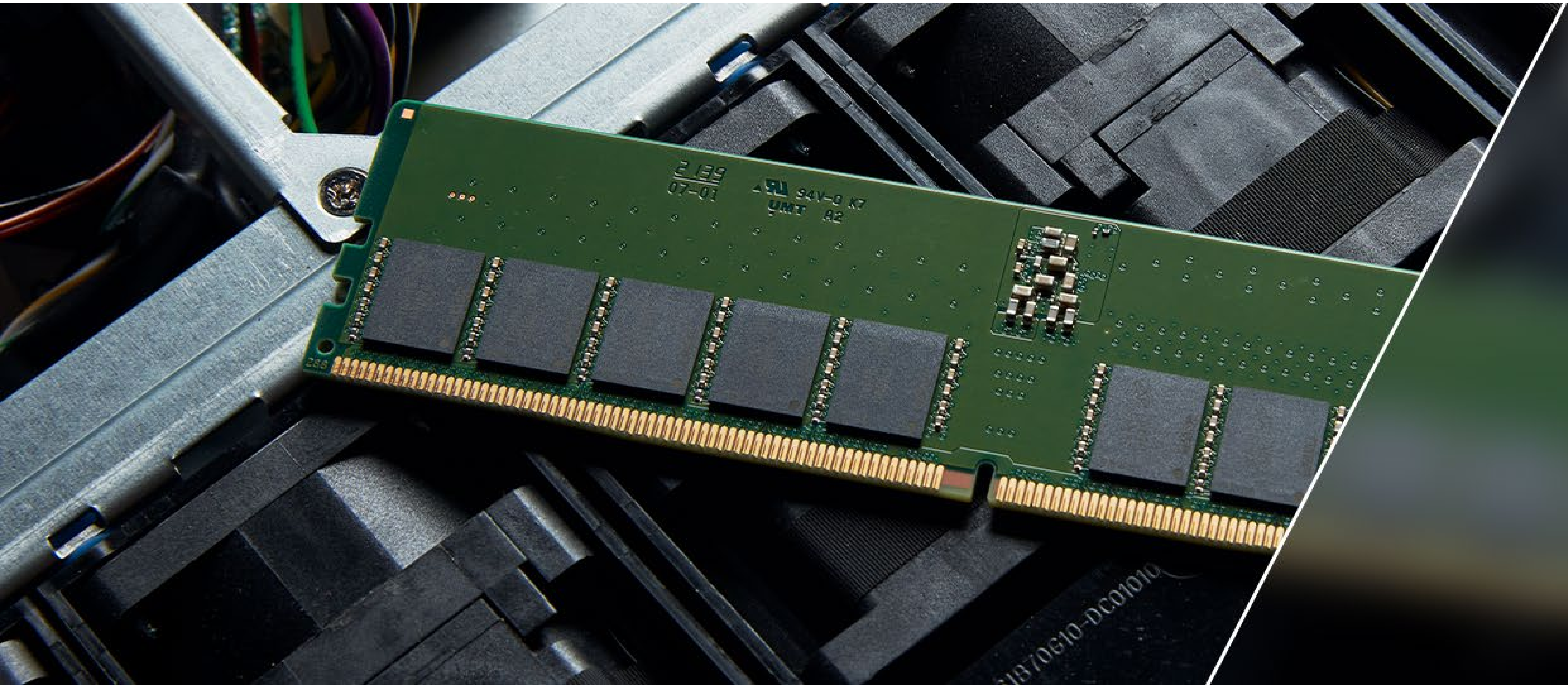
On the other hand, more complex systems, like servers and high-performance workstations designed to be always on, benefit from the added stability and reliability offered by ECC Registered (RDIMMs) and Load Reduced DIMMs (LRDIMMs). ECC-class modules support error correction for corrupted data, preventing the server from crashing or losing critical information. These modules also feature a higher grade of DRAM components, tested to higher tolerances and with lower failure rates.

DRAM Compatibility Issues and Upgrading Considerations

Along with use-case, application, and workload considerations, the importance of brand reputation cannot be overstated when assessing memory options. Generally, memory companies have fewer compatibility issues if they invest in testing infrastructure to validate their memory designs with the chipset architects (like Intel and AMD), and to work with motherboard and system manufacturers to perform memory qualifications. A strong ecosystem of checks and balances exists between Intel, AMD, motherboard manufacturers, major system brands, and memory manufacturers. But not all memory module vendors participate.

Installing incompatible DRAM can prevent a system from booting. When upgrading or replacing memory, always consult the motherboard manual or manufacturer's website before purchasing new memory and consider:

1. **Motherboard support:** Verify what specific memory technology and module type the motherboard supports (e.g., DDR4, DDR5, RDIMM vs. UDIMM).
2. **Speed:** Match or exceed the current DRAM's speed to avoid performance issues. Within a DDR generation, speeds are generally backward compatible. So, buying a standard 3200MT/s part will safely clock down to work on systems that require 2666MT/s.
3. **Capacity:** Choose modules to install in identical pairs or groups, matching the motherboard architecture, and always try to overprovision capacity to account for future memory needs.
4. **Mixing DRAM modules:** Mixing different DRAM types (width, density, brand) within pairs or groups can lead to instability. Installing in identical pairs or groups of modules according to the memory architecture of the motherboard reduces the chance of there being issues.
5. **Error correction:** If installing ECC Unbuffered modules into a client or mainstream workstation, be sure to verify the motherboard and processor model support the ECC function.



> *For example, DDR4 RDIMMs and LRDIMMs used the same module key (notch) as unbuffered DIMMs. When plugged into a desktop class system, RDIMMs and LRDIMMs will not work. DRAM component widths and densities can also affect compatibility, as some chipsets will not work with specific DRAM widths or high densities.*
>
> **Geoffrey Petit | Kingston Technology Europe**

DRAM Compatibility Issues and Upgrading Considerations

The memory industry is continuously in motion, designing for the demands of the next generation while keeping the needs of today's and yesterday's computers in mind. It's therefore critically important for memory module manufacturers to maintain an extensive archive of computer platforms going back several generations.

> *Testing new memory components on older systems, also referred to as regression testing, is one very important step skipped by some memory module manufacturers to reduce costs. This is an area that frequently exposes compatibility issues.*
>
> **Mike Mohney | Kingston Technology**

Maintaining an extensive database of system compatibility is also key to preventing issues. As one of the only memory module vendors in the world to actively maintain a database archive of over 40,000 computer systems, Kingston's engineers can more accurately communicate what memory upgrade options are compatible with the thousands of current and legacy computer models in the global market. Nuances between chipsets and generations of processors from Intel and AMD go frequently, and, in some instances purposely, undisclosed to users. Kingston's goal is to share the technical insights needed for users to understand how to choose the best, most compatible option for their computer.

> *Compatibility issues can occur when components haven't been vetted or optimized with a chipset or BIOS. Common compatibility issues also include using unsupported DRAM configurations or module types in a system.*
>
> **Geoffrey Petit | Kingston Technology Europe**

Overcoming DRAM Manufacturing Challenges

While compatibility can be an issue if not properly addressed, it's not the only challenge. Design complexity, precision in assembly, and quality control all contribute to primary DRAM manufacturing challenges. And as a leading third-party DRAM module manufacturer, these challenges are what define Kingston. However, each challenge has its own solution:

So, let's start with how we design our memory solutions.

- **Design complexity:** Each type of DRAM module we design has unique features, whether it's DDR4 vs. DDR5, Unbuffered vs. Registered, and many others—this adds complexity to the design. This requires advanced engineering and precise integration to ensure reliability and performance.
- **Solution:** To address this, we employ specialized design software and rigorous testing protocols to ensure that each memory type functions as intended without errors.

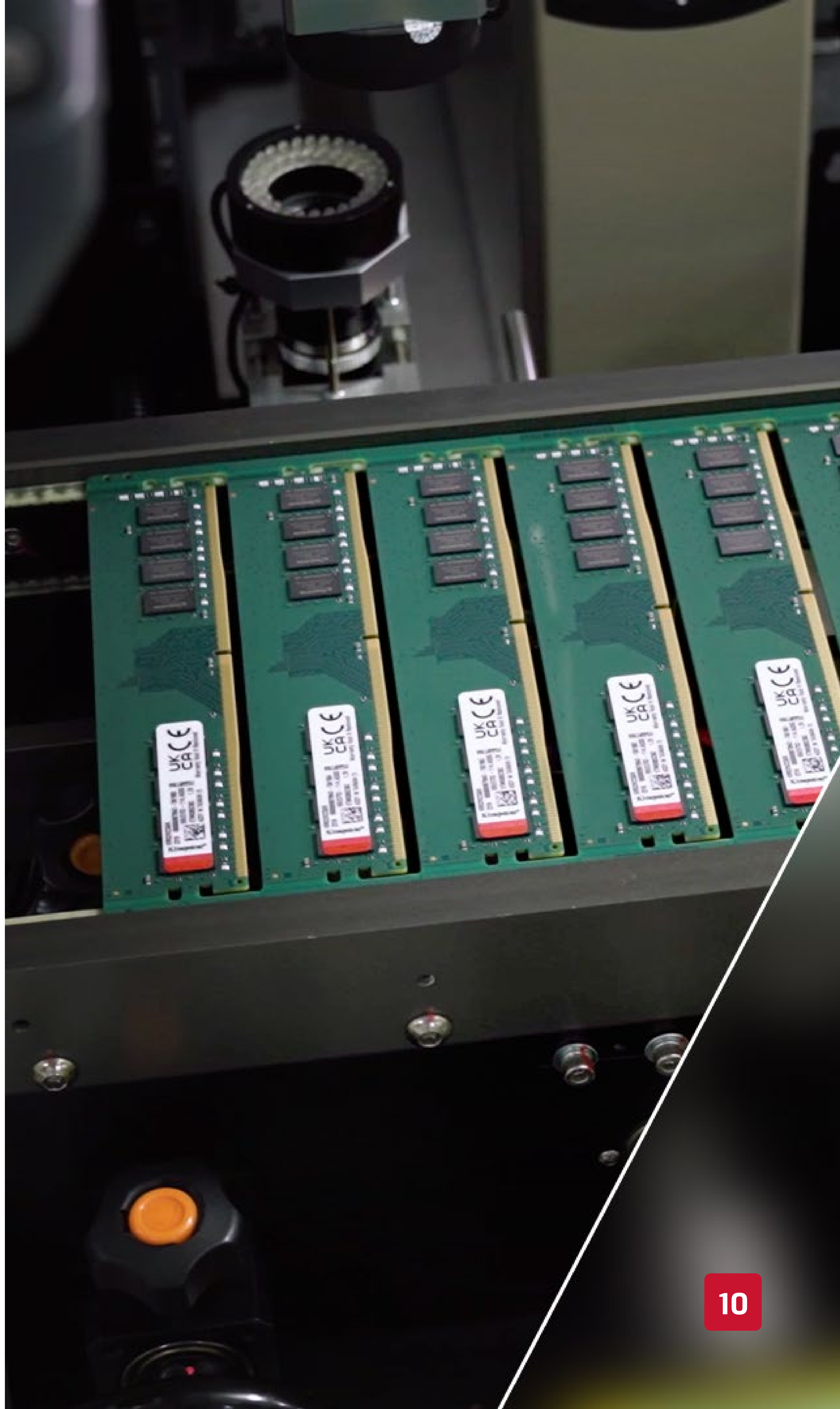
We then work with leading DRAM semiconductors.

- **Fabrication precision:** DRAM chips are produced by semiconductors on nanoscale processes, where even minute variations can lead to defects, affecting yield and performance.
- **Solution:** Kingston exclusively works with semiconductors that can guarantee high levels of performance and reliability. These semiconductors use cutting-edge lithography techniques and cleanroom environments to minimize defects, ensuring high precision and consistency during production. If they cannot meet these requirements, then we do not work with them.

Next is how we build and test our DRAM modules.

- **Quality control:** Once assembled, all our DRAM form factors must meet stringent performance and reliability standards.
- **Solution:** Extensive testing under various conditions, including temperature and stress tests, helps identify and eliminate faulty units, ensuring that only reliable memory reaches the market.

Through advanced technology and stringent quality assurance, Kingston produces high-performance, reliable DRAM solutions suitable for a variety of computing needs. We take this a step further by working closely with Intel and AMD to receive reference platforms that help us develop new memory technology, as well as prepare for advances needed in our production testing capabilities. Both hardware and software upgrades are constant for our production environments to support new memory speeds, new capacities, and to make improvements in the quality of the modules produced.



DRAM Development: The Influence of Market Trends

When it comes to shaping the development and adoption of different types of DRAM memory, market trends driven by the evolving demands of technology and consumer behavior play a significant role. Performance, efficiency, and scalability are all key factors that influence both development and adoption.

When looking back, the past few decades of computing and workload demands have had an impact on the types of memory developed. The mid-2000s saw the memory industry pivot to begin offering memory technologies that could save on overall power consumption, both in the mobile space and in the data center. In the mid-2010s, virtualization drove the demand for higher-capacity modules. At that time, performance loss with high-capacity modules due to chipset limitations eventually led to the development of Load Reduced DIMMs for DDR3 and DDR4.

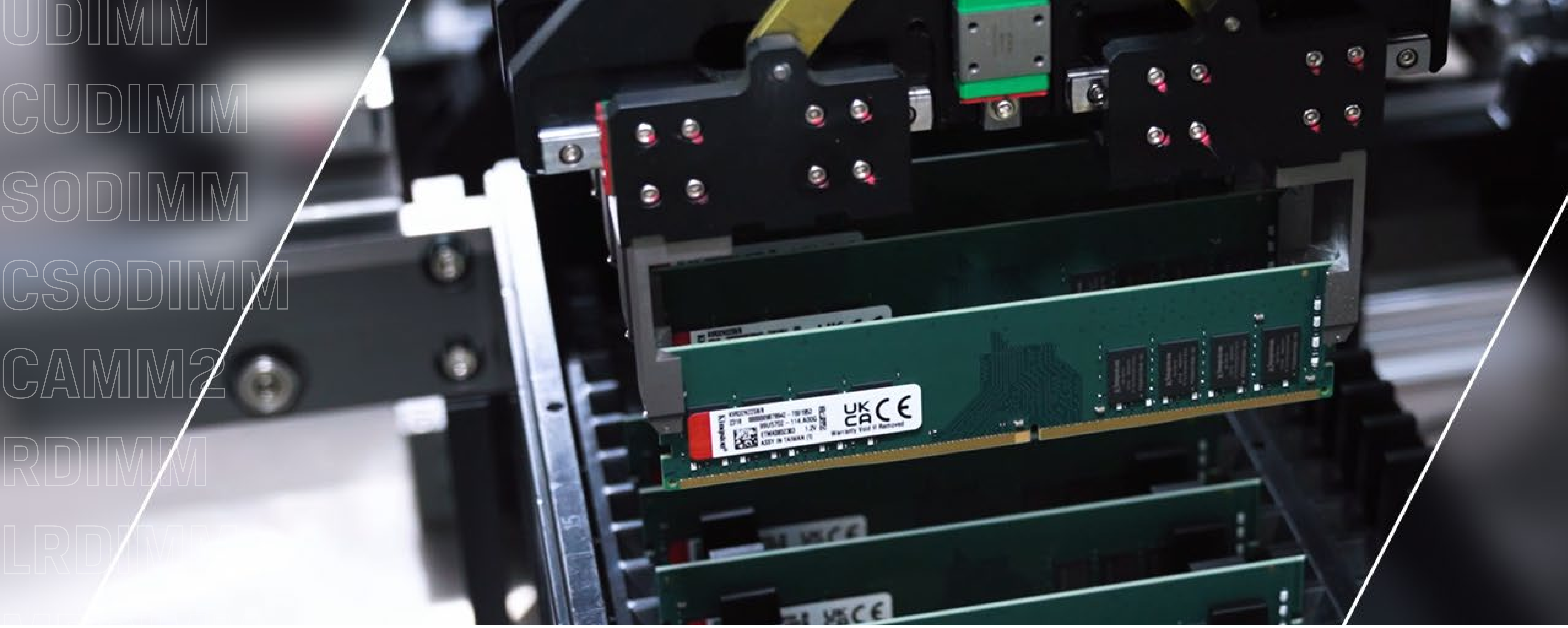
Today, industries like AI, gaming, and big data analytics continue to grow and are increasingly demanding high-speed, high-capacity memory. This drives the development of advanced DRAM module types like the Multiplexed-Rank DIMM (MRDIMM), which cater to these performance needs. The push for thinner, lighter devices has also influenced the adoption of compact and efficient memory solutions like the CAMM2, which offer manufacturers cost-effective modular solutions to replace DRAM-down, or multiple SODIMMs which would not physically fit into an Ultrabook-class laptop or tablet.

The ability to expand memory capacity beyond the traditional DIMM socket is another area in rapid development.

> *The performance demands of AI are another key driver fueling the creation of scalable, high-capacity, and high-performance memory like MRDIMMs, which specifically address the bottleneck of high-capacity memory performance.*
>
> **Geoffrey Petit | Kingston Technology Europe**

These are just a few examples of how the memory industry adapts to market trends, demonstrating how our ecosystem and standards body are always ready to tackle the challenges of tomorrow's memory needs.

> *Investing in infrastructure and scaling to support next-generation memory technologies is ongoing. Memory speeds increase annually, so having next-generation platforms well before launch is critical to ramping up production to support global demand when new systems hit the market.*
>
> **Mike Mohney | Kingston Technology**



> *These advancements, including CAMM2, CUDIMM, and CXL, indicate a future where memory technologies will drive faster, more efficient, and more flexible computing systems across various applications and industries.*
>
> **Geoffrey Petit | Kingston Technology Europe**

The future of DRAM technology

Looking ahead, the memory industry will continue to adapt and plan for the needs of the market. Future developments in DRAM technology are focused on increasing speed, reducing power consumption, and increasing density to meet the demands of advanced applications like AI, big data, and cloud computing. In addition, challenges with today's memory technologies and form factors are already influencing the next-generation DDR specification, DDR6, in development with JEDEC. Due to be finalized by 2027, DDR6 will likely focus on higher performance, with significant linear increases in data rates compared to DDR5, and a wider data bus.

Until then, DDR5 will continue to increase in speed and be utilized in new form factors. This includes the CAMM2, which has been projected to become the dominant module solution for mobile and small-form-factor systems in the next couple of years. The thin-profile CAMM2 can effectively replace two SODIMMs in traditional laptops, saving manufacturers significant costs by using a modular memory solution over discrete DRAM components mounted directly to a motherboard. Some motherboard manufacturers have even demonstrated the CAMM2 can be used in traditional desktop PCs. As an approved memory supplier to Dell for their original CAMM design, Kingston is strategically positioned to support the CAMM2 revolution, with the infrastructure and investment to manufacture and test this new form factor already in place. Stay tuned on the Kingston website for our CAMM2 solutions, set to launch in the first half of 2025.

CUDIMM is another new type of DRAM module that integrates a clock driver on Unbuffered DIMMs beginning at 6400MT/s DDR5. This component redrives the clock signal from the processor on the module, enhancing signal integrity and reducing incidence of error due to noise and jitter, which become problematic at higher speeds.

Then there is Compute Express Link, or CXL for short, another new DRAM category in its early stages. CXL is an open-standard protocol that operates on the PCI Express bus, much like NVMe for storage. The first focus for CXL products is memory expanders, which utilize DRAM (DDR4, DDR5, HBM) on various form factors, to increase the memory capacity and expand the usable memory pool for servers.

Summary

With the rise of AI, memory designers are racing to keep up. As the backbone of semiconductor memory, the evolution of DDR SDRAM with its large capacity and rapid data delivery to processors is continuing to advance. By tackling key compatibility and manufacturing challenges with investment and strict quality control, manufacturers can produce reliable, high-performance memory suited for diverse computing needs. But to meet the specific requirements of your environment, Kingston's experts are here to support you, helping you navigate the complexities of evolving chipsets, processor generations, and optimized memory configurations.

Built on Commitment

From big data to IoT devices, including laptops, PCs, and wearable technology, Kingston Technology is dedicated to delivering top-tier product solutions, service, and support. Trusted by leading PC manufacturers and global cloud providers, we value our long-term partnerships that help us evolve and innovate. We ensure every solution meets the highest standards by prioritizing quality and customer care. At every step, we listen, learn, and engage with our customers and partners to deliver solutions that make a lasting impact.