

Introducing Micron[®] DDR5 SDRAM: More Than a Generational Update

By Scott Schlachter and Brian Drake

Introduction

DDR5, the successor of DDR4, has been developed to deliver performance improvements at a time when system designers are feeling increasing pressure from continuous technological advancements—where current memory bandwidth is simply unable to keep up with newer processor models that have increasing core counts.

DDR5 is the fifth-generation double data rate (DDR) SDRAM, and the feature enhancements from DDR4 to DDR5 are the greatest yet. While previous generations focused on reducing power consumption and were driven by applications such as mobile and data center, DDR5's primary driver has been the need for more bandwidth.

Compared to DDR4 at an equivalent data rate of 3200 megatransfers per second (MT/s), a DDR5 system-level simulation example indicates an approximate performance increase of 1.36X effective bandwidth. At a higher data rate, DDR5-4800, the approximate performance increase becomes 1.87X—nearly double the bandwidth as compared to DDR4-3200.

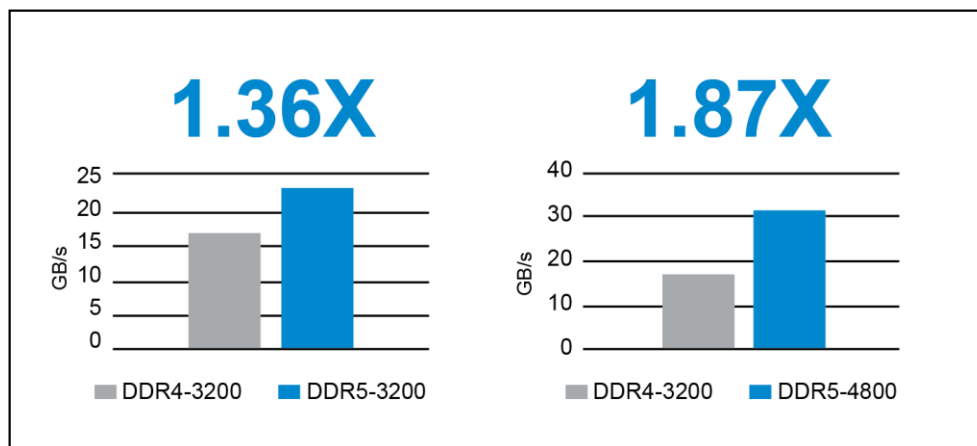


Figure 1: Effective Bandwidth: DDR4 vs. DDR5¹

Driven by data rates up to 6400 MT/s and key architectural improvements, Micron's DDR5 is pushing potential system bandwidth even higher. This white paper discusses some of the key architectural improvements of DDR5 and, specifically, how they enable significant bandwidth growth over DDR4.

1. Source: Micron. Bandwidth normalized to x64 interface, 64B random accesses, 66% reads, dual-rank x4 simulation, 16Gb. Best estimates; subject to change.

Meeting Next-Generation CPU Requirements

At the system-level, despite only modest clock rate improvements, the transition to multicore CPU architectures has enabled continuous year-over-year compute performance gains.

Signal integrity, power delivery, layout complexity, and other system-level challenges are limiting advances in CPU core frequency. Simultaneously, CPU core counts are continuously increasing, limiting the available memory bandwidth-per-core. New memory architectures beyond DDR4 SDRAM are needed to meet the next-generation bandwidth-per-core requirements.

The chart below illustrates the bandwidth trend in real systems and shows how bandwidth-per-core has been relatively constant and is even beginning to trend down (as illustrated in the lower right corner of the chart).

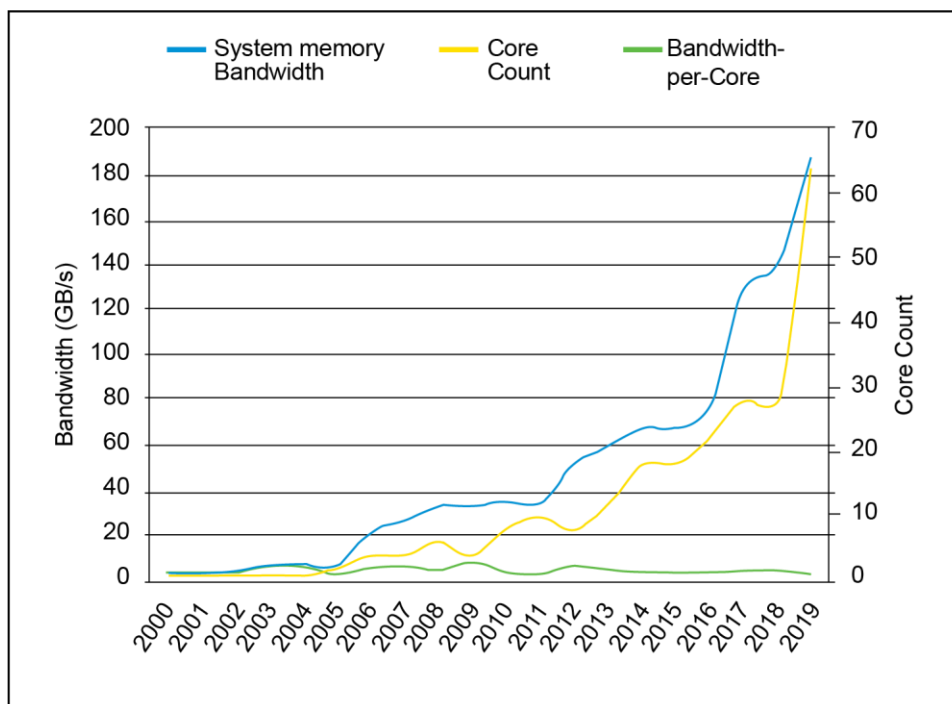


Figure 2: Theoretical DRAM Bandwidth vs. Core Count Trend

Source: Micron. Assumes 100% data bus utilization (theoretical bandwidth); for illustration purposes.

DDR5 Features

The transition from DDR4 to DDR5 represents far more than a typical DDR SDRAM generational change. DDR5 demonstrates a major step forward that has completely overhauled the overall DDR architecture with one primary goal: increasing bandwidth.

Increased Data Rates

A number of key feature additions and improvements enable DDR5's bandwidth increase. Primary among these is a dramatic increase in device data rates. While DDR4 spanned data rates from 1600 MT/s to 3200 MT/s, DDR5 is currently defined with data rates ranging from 3200 MT/s up to 6400 MT/s. This data rate increase alone will not only allow the existing bandwidth-per-core to remain equal as core-per-CPU counts increase (shown with the red arrow below), but it will also allow for far higher bandwidths.

The figure below includes data bus efficiencies (not shown) from a simulated workload to calculate potential *effective* bandwidth across different DDR4 and DDR5 data rates (this is different than the theoretical bandwidths shown in Figure 2).

Total Bandwidth GB/s (8Ch/1 DPC)		Core Count					
		24	32	40	48	56	64
DDR5 6400	298.2	12.42	9.32	7.45	6.21	5.32	4.66
DDR5 6000	288.1	12.00	9.00	7.20	6.00	5.14	4.50
DDR5 5600	276.8	11.54	8.65	6.92	5.77	4.94	4.33
DDR5 5200	264.5	11.02	8.26	6.61	5.51	4.72	4.13
DDR5 4800	250.9	10.46	7.84	6.27	5.23	4.48	3.92
DDR5 4400	235.3	9.80	7.35	5.88	4.90	4.20	3.68
DDR5 4000	218.6	9.11	6.83	5.47	4.55	3.90	3.42
DDR5 3600	201.0	8.38	6.28	5.03	4.19	3.59	3.14
DDR5 3200	182.5	7.60	5.70	4.56	3.80	3.26	2.85
DDR4 3200	134.3	5.60	4.20	3.36	2.80	2.40	2.10

*Data bus utilization applied

Figure 3: DDR5 Maintains Bandwidth with Increased Core Count

Source: Micron

Impressively, this dramatic increase in the I/O switching rate (data rate) is achieved without the need for differential signaling at the DQ pins; the DQ bus remains single-ended, pseudo-open drain (POD). However, there are critical new features that enable these higher data rates to be achieved. One of these is the addition of equalization in the form of a multi-tap decision feedback equalizer (DFE) in the DQ receivers. The DFE mitigates the effects of inter-symbol interference (ISI) at the higher rates by opening up the data eyes inside the device.

Other new features directly enabling data rate increases include:

- **Duty cycle adjuster (DCA) circuit** capable of adjusting both the DQ and DQS duty cycles for the read path internally. This helps to correct the small duty cycle distortions that occur naturally as those signals pass through the devices and PCB, ultimately optimizing the duty cycles for the DQ and DQS signals received by the controller.

- **DQS interval oscillator circuit** that allows the controller to monitor changes in the DQS clock tree delays caused by shifts in voltage and temperature. This enables controller designs to actively decide if, and when, retraining may be beneficial or necessary to keep the write timing optimized.
- **New and improved training modes**, including a new read preamble training mode, command and address training mode, chip select training mode, and a write leveling training mode. Write leveling provides the same capability as DDR4 that allows the system to compensate for timing differences on a module between the CK path to each DRAM device (which varies according to the fly-by path across the module) and DQ and DQS paths (which are short). Additionally, DDR5 has new functionality to compensate for the unmatched DQ-DQS receiver architecture, further enabling the faster data rates.
- **Read training patterns with dedicated mode registers.** The associated data patterns include the default programmable serial pattern, a simple clock pattern, and a linear feedback shift register (LFSR)-generated pattern, which ultimately provide more robust timing margin for the high data rates.
- **Internal reference voltages** for the command and address pins (V_{REFCA}) and chip select pin (V_{REFCS}). In addition to the internal reference voltage for the DQ pins (V_{REFDQ}), which improves voltage margin on the DQ receivers, these new internal reference voltages for the command/address and chip select pins improve the voltage margin on their respective receivers, further enabling the device to achieve higher data rates.

Protocol Features for Performance

In addition to higher data rates and improvements to the I/O circuitry, DDR5 introduces other new protocol features unrelated to data rate that are integral to increasing bandwidth and performance. For example, DDR5 DIMMs feature two 40-bit (32 bits plus ECC) independent channels. When combined with a new default burst length of 16 (BL16) in the DDR5 component, this allows a single burst to access 64B of data (the typical CPU cache line size) using only one of the independent channels, or only half of the DIMM. Providing this ability to interleave accesses from these two independent channels enables tremendous improvements to concurrency, essentially turning an 8-channel system as we know it today into a 16-channel system.

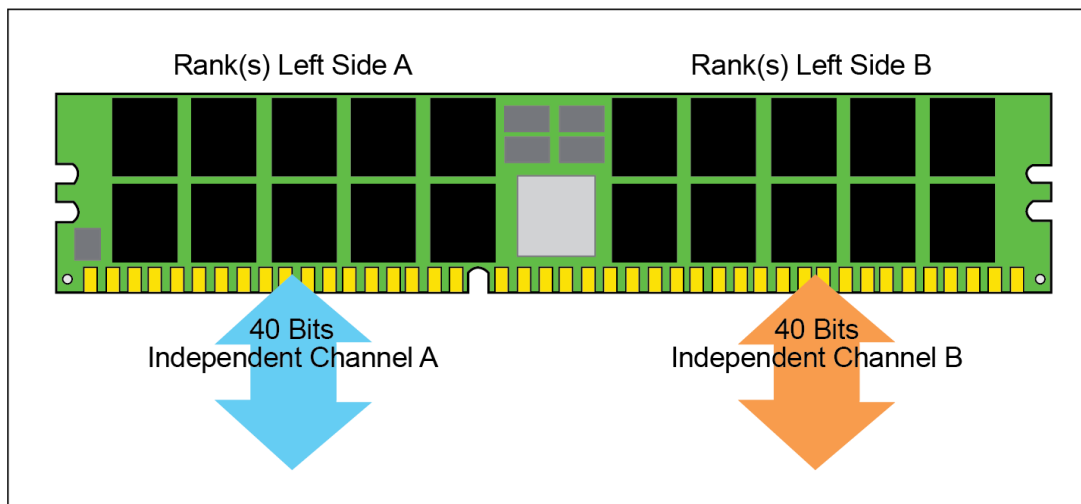


Figure 4: DDR5 DIMM

In the DRAM array, the number of bank groups (BGs) is doubling in DDR5 as compared to DDR4, keeping the number of banks-per-BG the same, which effectively doubles the number of banks in the device. This enables controllers to avoid the performance degradations associated with sequential memory accesses within the same bank (for example, causing t_{CCD_S} to be the sequential access restriction, instead of the much longer t_{CCD_L}). The addition of same-bank refreshes and improvements to the pre/postambles on the command bus

(by introducing an interamble) help to mitigate the traditional performance bottlenecks commonly observed in DDR4, improving the overall effective bandwidth of the memory interface.

A Massive Overhaul: DDR4 vs. DDR5

In addition to the features mentioned above, other features have been added or improved in DDR5 to dramatically improve device architecture and performance.

Table 1: Device Feature Comparison Highlights Between DDR4 and DDR5 SDRAM

Feature	DDR4	DDR5	Benefit/Improvement
Data rates	1600-3200 MT/s	3200-6400 MT/s	Increased performance and bandwidth.
$V_{DD}/V_{DDQ}/V_{PP}$	1.2/1.2/2.5	1.1/1.1/1.8	Lower power.
Internal V_{REF}	V_{REFDQ}	V_{REFDQ} , V_{REFCA} , V_{REFCS}	Internal $V_{REFCA/CS}$ rails significantly improve voltage margins for those pins, enabling higher data rates. This can save BOM costs by eliminating the need for an external reference voltage on the board.
Device densities	2Gb-16Gb	8Gb-64Gb	Larger monolithic devices.
Prefetch	8n	16n	Enables higher data rates while keeping the internal core clock range similar to DDR4.
DQ receiver equalization	None	Multi-tap DFE	Opens up the DQ data eye inside the DRAM, directly enabling high data rates.
Duty cycle adjustment (DCA)	None	DQS and DQ	Enables the controller to compensate for duty cycle distortion (DCD) on all DQS and DQ pins by adjusting the duty cycle inside the DRAM.
Internal DQS delay monitoring	None	DQS interval oscillator	Provides a method for the controller to decide if/when to re-train based on changes in DRAM delays caused by shifts in voltage and temperature; provides robustness against environmental changes.
On-die ECC	None	128b+8b SEC, error check and scrub	Strengthens on-chip RAS; reduces burden on controller.
CRC	Write	Read/Write	Strengthens system RAS by protecting read data.
Bank groups (BG)/banks	4 BG x 4 banks	8 BG x 2 banks (8Gb x4/x8) 4 BG x 2 banks (8Gb x16) 8 BG x 4 banks (16-64Gb x4/x8) 4 BG x 4 banks (16-64Gb x16)	Helps to avoid performance degradations from same-bank sequential memory accesses.
Command/address interface	ODT, CKE, ACT, RAS, CAS, WE, A<X:0>	CA<13:0>	Requires two cycles for some (but not all) commands, dramatically reducing the CA pin count.
ODT	DQ, DQS, DM/DBI	DQ, DQS, DM, CA bus	CA ODT provides improved signal integrity and saves BOM costs by eliminating the external termination resistor network for the CA bus.
Burst length	BL8 (and BL4)	BL16, BL32 (and BC8 OTF, BL32 OTF)	Combined with a 2-channel DIMM architecture, enables 64B cache line fetch using only half of a DIMM.

Feature	DDR4	DDR5	Benefit/Improvement
MIR (“mirror” pin)	None	Yes	Improves DIMM signaling by allowing shorter traces/stubs for clamshell modules and board designs.
Bus inversion	Data bus inversion (DBI)	Command/address inversion (CAI)	Reduces power and noise on the V _{DDQ} rail.
CA training, CS training, write leveling training modes	Write leveling training mode	CA training, CS training, and write leveling training modes	Improved timing margin on the CA and CS pins enables faster data rates. Write leveling training in DDR5 also compensates for the device’s unmatched DQ-DQS path, making it easier to support fast data rates with short write preambles and enabling shorter bus turnarounds.
Read training patterns	Possible with the MPR	Dedicated MRs for serial (user-defined), clock and LFSR-generated training patterns	Dedicated read training includes MRs for training pattern selection, including one that uses an LFSR to provide a PRBS pattern. This provides a more robust read timing margin, particularly at the higher data rates.
Mode registers	7 x 17 bits	Up to 256 x 8 bits (LPDDR type read/write)	Room to expand as needed for new feature support and improvements.
PRECHARGE commands	All bank and per bank	All bank, per bank, and same bank	SAME BANK PRECHARGE (PREsb) enables the precharging of a specific bank in each bank group, keeping the active state of all other banks unchanged.
REFRESH commands	All bank	All bank and same bank	SAME BANK REFRESH (REFsb) enables the refreshing of a specific bank in each bank group, keeping all other banks in the bank group free to access.
Loopback mode	None	Yes	Enables testing of the DQ and DQS signaling between the controller and the DRAM, isolating the actual memory array since read/write accesses are not needed.

Conclusion

The growing need for increased memory bandwidth due to scaling requirements and higher performance targets of next-generation computer systems creates an important challenge for today’s system architects. DDR5 SDRAM has been developed to address the need for higher bandwidth, offering massive improvements over previous generations of SDRAM.

With a robust list of new and enhanced features, Micron’s DDR5 SDRAM sets the bar higher than ever before when it comes to overall system performance—pushing the limits of high-speed signaling and directly addressing the memory bandwidth challenge.



Learn more at www.micron.com