



DDR4 WHITE PAPER

Introduction

DDR3 has been with us for a long time, and Corsair has been there pushing the bleeding edge of performance, cooperating with Intel, AMD, and motherboard manufacturers to produce the fastest memory consumers can buy. Yet DDR3 is getting long in the tooth and modern processor architectures are becoming increasingly demanding. When today's hardware needs exotic, high-speed, carefully binned memory, it's time to look to a new technology.

Introduced in conjunction with Intel's new high end desktop platform based on their Haswell-E architecture and the X99 chipset, DDR4 is designed to meet the needs of present and future platforms. It offers higher performance, lower power consumption, higher density, increased reliability, and a remarkably forward-thinking design geared for heretofore unprecedented scaling. In short, DDR4 is the memory technology we need, now and for tomorrow.

The Demands of Modern Hardware

While Moore's Law has slowed somewhat for x86 processors, the steady march of progress in terms of both performance and hardware integration has resulted in a hardware ecosystem that is very slowly reaching the limits of what can be achieved with DDR3. Increasingly powerful x86 cores are now being married with substantially powerful graphics hardware on a single die, boosting the stress placed on the memory bus and raising the amount of memory bandwidth needed to keep these new integrated processors fed.

With Intel's Haswell and AMD's Kaveri architectures, we are seeing more and more situations where performance can be bottlenecked by a lack of memory bandwidth. Content creation (CAD, video editing, et al) continues to need as much memory capacity as possible. Meanwhile, mobile devices (notebooks, tablets) focus more and more on power efficiency. All of this, to say nothing of the ever escalating demands on enterprise and server hardware, hardware that often requires as much bandwidth and capacity as can be delivered while needing to reduce power consumption wherever possible.

The demands of current and future hardware and software architectures can be met with DDR4.

The History of DDR3

To understand DDR4, we need to have at least a working knowledge of the history of DDR3. With rare exception, introductions of new memory standards operate cyclically. When the original double data rate (DDR) memory was introduced, users were still working with SDRAM running at 133MHz, and the benefits of moving to this faster standard were limited. The outgoing technology will always be more mature and have a better price-to-performance ratio than its successor at time of launch; this was true of the transition from DDR to DDR2 and again from DDR2 to DDR3.

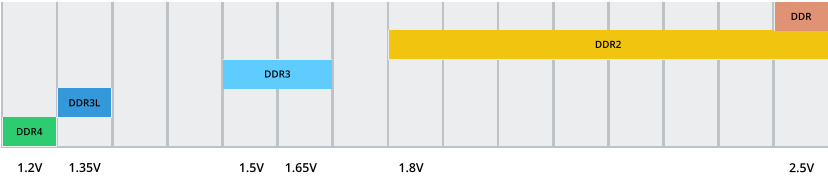


Figure 1: Operating voltage of DDR standards.

DDR3 was introduced in 2007, and at the time, DDR2’s effective mainstream speed had standardized at 800MHz with JEDEC’s peak spec at 1066MHz. DDR3’s introductory speeds were 800MHz and 1066MHz, but performance could actually be slightly lower in some instances due to higher latency. Yet DDR2 was a mature technology and at its limit; mainstream DIMM density topped out at 4GB and operating voltage was between 1.8V and 2.5V.

DDR3 was more forward thinking; specified voltage was 1.5V, speeds were designed to scale well past 1066MHz, and Intel (working with Corsair) created and incorporated the XMP specification that allowed end users to easily fine tune speeds and timings for high performance memory. So while DDR2 had a better value proposition when DDR3 launched, it was quickly eclipsed by the new specification’s headroom.

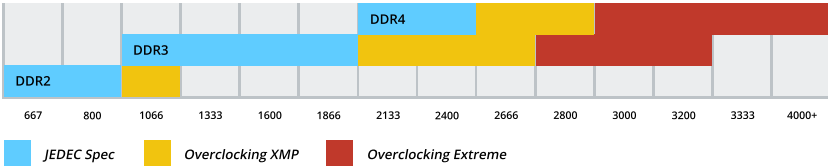


Figure 2: A history of DDR SDRAM memory speeds.

Seven years later, JEDEC has specified DDR3 speeds all the way up to 2133MHz, mainstream DIMM density is 8GB (with server DIMMs at 16GB), and an adjoining DDR3L standard has been adopted that operates at a reduced 1.35V. High performance DDR3 even exceeds 2133MHz; internal Corsair testing has seen Intel Haswell processors show continued performance benefits up to 2400MHz, while the GCN-architecture graphics cores powering AMD’s Kaveri processors are demonstrably bottlenecked even then. While the mainstream has settled on 1600MHz as the standard speed for DDR3 and Corsair memory continues to drive speeds well beyond that, it is possible to be performance limited by even exotic 2400MHz memory on current generation hardware. Beyond that speed price can increase dramatically, reflecting the careful binning of ICs that has to take place to produce those extremely high speed DIMMs. DDR3 has scaled well beyond the performance of DDR2, but is now approaching its own limits.

What is DDR4?

DDR4 is a new memory standard designed to eventually replace DDR3. While it's true that when running at the JEDEC specified speeds of 2133MHz and 2400MHz DDR4's higher latencies may produce slightly lower performance clock-for-clock than DDR3, DDR4 is designed to reliably run at much higher speeds that more than offset the increased latency. In virtually every way, DDR4 is superior to DDR3: it's capable of being much faster, more efficient, more scalable, and even more reliable. As for cost, much like the transitions to DDR, DDR2, and DDR3, DDR4 will become progressively cheaper as economies of scale take effect.

Physically, a DDR4 module, or DIMM, is very similar to a DDR3 DIMM. DDR4 can use a slightly taller printed circuit board and ups the pin count from DDR2 and DDR3's 240 pins to 288. The key notch (to ensure the DIMM is not improperly installed) is also in a different place, and the overall shape of the connector has a slight "V" contour to aid installation.

Architecturally, DDR4 is designed to operate at higher speeds and capacities with lower voltage and adds reliability features not present in DDR3.

Why Do We Need DDR4?

When DDR4 is introduced, the initial 2133MHz and 2400MHz speeds will be accompanied by another increase in latency, just as each previous memory technology transition has been. These speeds are essentially the top of the ladder for DDR3, though; while DDR3 kits can be obtained at speeds as high as 3200MHz, ICs capable of performing at those levels are extremely rare. Meanwhile, DDR4 is expected to scale well beyond 3200MHz.

What DDR4 offers is scalability for the future: individual DIMM densities start at 4GB and 8GB and are expected to scale to 16GB in 2015. Bandwidth is also capable of scaling up tremendously. 2666MHz DDR3 isn't especially common right now; it operates outside of JEDEC spec and requires carefully selected ICs, yet already situations exist that demonstrate a need for increased bandwidth beyond that speed. DDR4 comes out of the gate at 2400MHz, with 2666MHz, 2800MHz, and 3000MHz SKUs already planned.

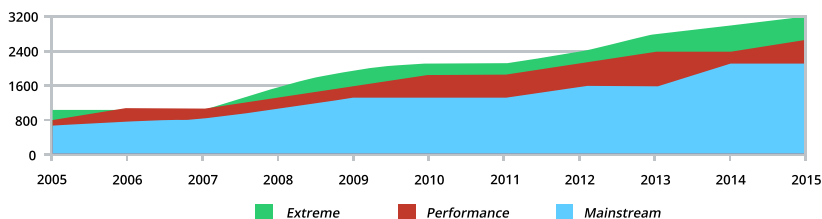


Figure 3: DDR bandwidth growth over time.

Finally, DDR4 operates at a nominal 1.2V and scales up to 1.35V, a reduction in operating power from DDR3's 1.5V standard and 1.65V mainstream high performance spec. Power efficiency has become increasingly important with each subsequent generation of CPU and GPU architecture from Intel, AMD, and NVIDIA, and DDR4 helps to enable that.

DDR4 in Detail

As previously mentioned, the primary benefits of DDR4 are increased bandwidth, a reduction in power consumption, increased density, and improved reliability.

Increased Bandwidth

While DDR3 kits are available all the way up to 3000MHz, JEDEC only specifies operating speeds up to 2133MHz. Scaling beyond that requires exceptionally high performance ICs, making higher speeds harder and harder to obtain. In memory bandwidth bound applications (such as Adobe Media Encoder and Handbrake), 4th Generation Intel Core i5 and i7 processors demonstrate performance scaling all the way up to 2400MHz. This is arguably a limitation of the DDR3 spec, which was never intended to scale this high.

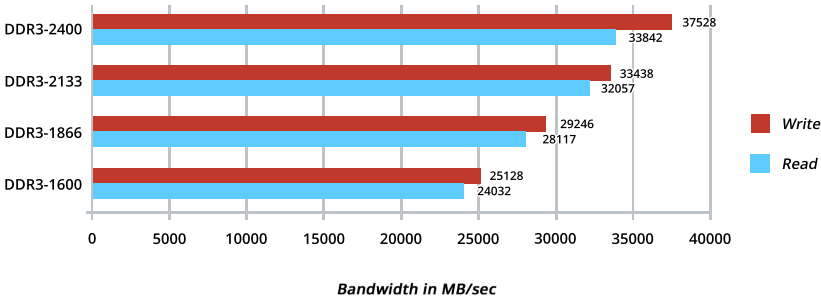


Figure 4: Intel Core i7-4770K, dual-channel DDR3 bandwidth measured in AIDA64.

DDR4 is being introduced to the mainstream at 2133MHz and 2400MHz, but the specification allows scaling well beyond that. Launch speeds are as high as 3000MHz, and keep in mind that this is only in the first year. While DDR4 inherits DDR3's timing methods and will feature another bump in timing latencies similar to the one DDR3 received from DDR2, substantial internal architectural changes and tighter yield specifications allow it to achieve an overall increase in potential bandwidth with room to continue scaling. Mainstream users benefit from these higher speeds, while overclockers and enthusiasts that have been bumping their heads against the limits of DDR3 now have an entirely new frontier to play in.

Reduced Power Consumption

Power consumption is improved in multiple ways. First, there's a baseline 20% reduction in the voltage required to operate DDR4. While DDR3 requires 1.5V and DDR3L reduced this requirement to 1.35V, DDR4 starts at 1.2V and is expected to receive a lower specification in the future that further decreases operating power.

Beyond that, DDR4 introduces additional power-saving technologies. DDR3 operates at a single universal voltage specification that then has to be scaled internally for some operations. Increasing voltage internally generates heat, draws more power, and is generally less efficient. DDR4 features a secondary voltage specification, V_{pp} , operating at 2.5V. Instead of losing efficiency by having to scale voltage, DDR4 has the needed voltage available from the start.

DDR4 also adds multiple new refresh methods designed to reduce operating and idle power consumption, and timings can be adjusted depending on DIMM temperature to improve efficiency and stability. This is more of a means of improving stability than efficiency, and while it won't be implemented on standard UDIMMs, we can expect to see this on RDIMMs and SODIMMs.

The DDR4 specification allows for more finely grained control over power to the memory itself. DDR3's method of VDDQ termination, dubbed center tapped termination or CTT, resulted in unused areas of memory still receiving power. DDR4 changes that by switching to PODs, or Pseudo Open Drains, which effectively allows power to simply drain through unused memory instead of being wasted.

The net result of these changes and additions to the specification is that DDR4 can be as much as 30% more efficient than DDR3 when operating at the same performance levels.

Increased Density

DDR4's architecture was specified to allow for considerably higher IC and DIMM densities than DDR3's. Mainstream DDR3 DIMMs currently scale to 8GB, but a modified DDR3 specification (Load Reduction DIMM) allows for an increase to 16GB or 32GB DIMMs. This is unfortunately only visible chiefly on the enterprise side; while AMD's processors can handle 16GB DIMMs, current generation mainstream Intel processors can't support individual DDR3 DIMMs larger than 8GB.

The limitations of capacity in DDR3 are addressed from several angles in the DDR4 specification, with each contributing to an overall massive improvement in potential density.

First, while a DDR3 IC is specified for up to eight internal banks, DDR4 doubles this number to 16 while also organizing them into four addressable bank groups (improving overall performance and efficiency). DDR4 also increases the potential density of the IC itself, all the way up to 16Gb.

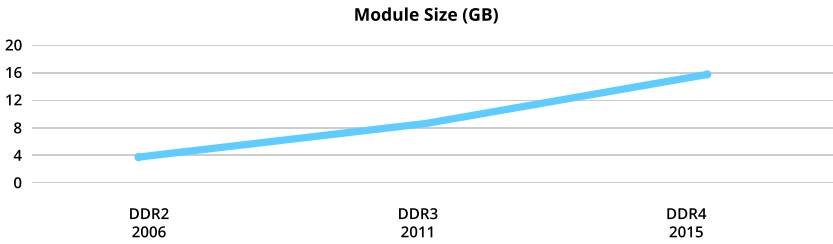


Figure 5: Maximum mainstream module size.

Some quick math illustrates the increase in capacity: Sixteen of DDR3's common 4Gb ICs running in dual rank operation equals 64Gb (16 ICs * 4Gb IC), resulting in DDR3's signature 8GB DIMM (64Gb / 8). However, DDR4 allows for 8Gb ICs and larger; sixteen of DDR4's 8Gb ICs running in dual rank operation equals 128Gb (16 ICs * 8Gb IC), resulting in a DDR4 DIMM density of a healthy 16GB (128Gb / 8). DIMMs at this density are not expected to reach the market until 2015, but the specification allows for them out of the gate, while DDR3 has yet to hit this density in the mainstream.

Possibly most impressive is the forward-thinking nature of DDR4's specification, though. DDR4 is the first memory technology specified to handle three-dimensional silicon stacking, or 3DS, allowing up to eight slave DRAMs to be stacked on top of a single master DRAM and addressed through that master DRAM.

Finally, the increased density DDR4 offers allows for improvements in cost. An 8GB DDR3 DIMM requires sixteen 4Gb ICs running in dual rank, while an 8GB DDR4 DIMM, once 8Gb ICs become available, would require only half that. This ignores the DDR4 specification allowing up to 16Gb ICs.

Improved Reliability

One benefit of DDR4 not often mentioned but worth discussing is the overall improved reliability and stability. While reliability and stability are appreciated on the consumer side of the market, they're vital on the enterprise side. The additions to the DDR4 specification that address this benefit everyone.

First, DDR4 implements a CRC, or cyclic redundancy check. As operating frequencies scale, the potential for write errors to memory increases. Employing the CRC allows for real-time write error detection and can correct intermittent errors, as well as enabling more robust error reporting to the system itself.

Command and address parity error detection and recovery have also been added to DDR4, and changes to the DDR4 register allow it to block commands upon detection of a parity error. While DDR3 would pass those commands on to the DRAM itself, DDR4 essentially stops them at the gate.

Operating temperature is a well known inhibitor with most computing technologies, as high temperatures can reduce signal clarity and produce errors. DDR4's thermals can be monitored and its timings adjusted to account for changes in temperature, though this is only expected to be implemented in SODIMMs where heat becomes a more serious issue.

Conclusion

Much like each memory technology transition that preceded it, DDR4's value upon introduction may seem nebulous compared to the incumbent DDR3. DDR2 operating at the same frequency as mainstream DDR was slightly slower, DDR3 had the same issue with DDR2, and DDR4 will have the same issue with DDR3. This is expected; each transition meant slightly looser timings. Yet each time, the newer specification was implemented less to improve the present and more to meet the needs of the future. DDR4 is no exception.

Where DDR4 does deviate substantially from its predecessors is in how dramatic the changes in specification and architecture are. Each iteration of DDR memory brought with it the same three improvements: increased speed, increased density, and reduced power. The differences in scale, however, are massive, and DDR4 in particular is the most spectacular leap yet.

SPECIFICATION	DDR2	DDR3	DDR4
Operating Voltage	1.8V to 2.5V	1.35V to 1.65V	1.2V
V_{DDQ} Voltage	-	-	2.5V
Speeds	400, 533, 667, 800, 1066	800, 1066, 1333, 1600, 1866, 2133	2133, 2400, 2666, 2800, 3000, 3200+
IC Densities	512MB to 4GB	512MB to 8GB	2GB to 16GB
Internal Banks	8	8	16 (in banks of 4)
DIMM Densities	512MB to 4GB	512MB to 16GB	4GB to 32GB
Module Pins	240	240	288

Figure 6: DDR specification comparison.

Improvements in speed and voltage have been fairly steady, but potential capacity sees an absolutely staggering increase from DDR3. Even ignoring the potential of 3DS, DDR4 DIMMs are expected to hit 16GB in 2015 and specified to go up to 32GB. None of this accounts for the improvements in reliability and stability baked into the DDR4 specification.

DDR4 is the most forward-thinking mainstream memory technology yet. Today it brings all of these improvements to Intel's 2014 Haswell-E architecture for enterprise and high end desktop processors. The future will see increases in density, reductions in operating voltage, and higher speeds, giving this new technology the potential to continue to meet the demands of its host system for years to come.