iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** VII **Month of publication:** July 2022

DOI: https://doi.org/10.22214/ijraset.2022.45912

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VII July 2022- Available at www.ijraset.com

A review on Buffer Chips and growing Memory Bandwidths

Manoj Y[1], Dr. B S Kariyappa[2]
[1, 2]*Department of ECE, R V College of Engineering, VTU, Belgaum*

Abstract: ***The rapid development of CPU technology and the increase in the number of functionalities involved in a chip demand improvement in the memory bandwidth to feed data to the growing pipeline technology. Slower memory bandwidth would result in reduced CPU utilization and make the processor stay idle. The rise in memory bandwidth comes with the challenge of maintaining the signal integrity and affects the number of DIMM (Dual Inline Memory Module) slots in a channel and reduces memory capacity. The challenge of maintaining the signal integrity with the increase in the frequency of operation is solved by introducing additional interfacing chips between the memory controller or CPU and memory module. Buffer chips are used in servers to maintain signal integrity and meet the timing parameters for the commands and addresses sent to the memory modules. Buffer chips help in increasing the frequency of operations as well as the number of memory modules connected. Register Clock Driver (RCD) is one buffer chip used between the memory controller and DRAM that supports the bandwidth and capacity needed in the next-generation data centres. This paper aims at reviewing importance of buffer chips and their application in modern day data centres.***

Keywords: RCD, Buffer chip, DDR5, DIMM.

I. INTRODUCTION

Increasing demand for high-speed operation comes with the challenge of improving the memory bandwidth to feed the faster circuits, with an increase in the frequency of operation maintaining the signal integrity becomes the primary need to prevent the underutilization of the device. Buffer Chips play an important role in maintaining signal integrity and allow a higher number of DIMM slots per channel as well as support low-power applications. Buffer chips are typically used in server memory modules as they improve the integrity of command sent and meet timing parameters.

A. DDR SDRAM

DDR stands for Double Data Rate memory. A pathbreaking invention in the memory technology [1]. This technology the data transmission on both posedge and negedge of the system clock. That in other words increases the data transmission by two times without increasing the frequency of operation. As the increased frequency brings up many problems such as contention in the bus due to the data on other bus and reduces the reliability of communication, sampling the data on both edges helps to keep the frequency at lower level but increased data transmission rates.

Synchronous Dynamic Random Access Memory (SDRAM) is a type of memory that can run at faster speed than the conventional DRAM memory and it is used as RAM in many computer and server memory modules. After SDRAM introduced, the DDR RAM 's has entered the market such as, DDR-1, DDR-2, DDR-3, DDR-4, DDDR-5 memory modules. The use of SDRAM was so effective that it took only about four years to overtake DRAM as the predominant form of computer memory due to its high operating speed since its introduction in July 1996. Today, SDRAM-based memory is the predominant type of dynamic RAM used throughout the computing spectrum, especially in computer random access memory.

B. Advantages and Disadvantages of SDRAM

It is always wise to know both the pros and cons of using technology before using it in new electronic circuit designs or projects. However, SDRAM has effectively cornered many computer memory situations, especially the RAM market. The

SDRAM has several important advantages.

1) Simple design, cheap and speed.
2) Do not requires complicated manufacturing process
3) The DDR version of SDRAM doubles the data rate of the base SDRAM by using both edges of the clock cycle.

For most applications, SDRAM has few drawbacks.

a) High power consumption
b) This is volatile memory. Data will be lost if power is turned off
c) Data needs to be updated
d) Slower than SRAM

When choosing all kinds of computer memory, such as RAM and other features, one should carefully weigh the pros and cons.

C. SDRAM types and DDR Versions

SDRAM technology has made great strides. As a result, several contiguous families of memory have been introduced, each with better performance than the previous generation. The most important development related to SDRAM was the introduction of a new approach called DDR. Currently, different versions of SDRAM are categorized according to DDR version. DDR2, DDR3, DDR4, DDR5, etc. This type of SDRAM doubles the data rate by transferring data across the clock, i.e., Rising edge and falling edge. This doubles the opportunity to transfer data synchronously. With the widespread use of SDRAM chips, the industry has organized chip behavior and interface standardization with the support of JEDEC, the Joint Electron Device Engineering Council.

1) SDR SDRAM

This is the basic type of SDRAM first introduced and released in 1993. This is called a single data rate SDRAM or simply SDRAM. SDR SDRAM was only able to sample data on single edge of clock.

2) DDR SDRAM

DDR significantly speeds up previous generation SDRAM technology at the time of deployment [2]. DDR SDRAM / DDR1 SDRAM was the first adaptation of the DDR technology, achieving increased speed by transferring data twice per cycle, i.e., At both the rising and falling edges of the clock signal [3]. The increased speed of DDR1 SDRAM resulted in its quick adaptation and mode SDR SDRAM become extinct. Table 1.1 shows the DDR SDRAM data rates and clock speeds. DDR SDRAM memory has multiple banks in memory. DDR SDRAM was able to provide multi-interleaved memory access by using multiple banks in memory. Each memory bank is equal to an array of memory. Multi- interleaved memory access increased the overall memory bandwidth.

TABLE 1.1
DDR SDRAM DATA RATE AND CLOCK SPEEDS

DDR SDRAM type	Data Rate (MB/s/Pin)	Memory clock speed (MHz)
DDR-266	266	133
DDR-333	333	166
DDR-400	400	200

3) DDR2 SDRAM

DDR2 memory is more complex than previous memory. Memory cells are designed to work with external buses [4]. Similar to DDR, DDR2 was also quipped to transfer data on bith rising and falling edges of cloak and operating frequency was doubled compared to DDR, which makes DDR2 a fastest memory than DDR1. Table 1.2 gives the data rate and clock speed of DDR2 memory [5]. The use of prefetch buffers and off-chip drivers significantly contributed to the significant increase in the clock speed. The problem with DDR2 is that the buffer introduces twice the latency of DDR. This requires doubling the bus speed to counteract the latency.

TABLE 1.2
DDR2 SDRAM DATA RATE AND CLOCK SPEEDS

DDR2 SDRAM type	Data rate (MB/s/pin)	Clock speed (MHz)
DDR2-400	400	200
DDR2-533	533	266
DDR2-667	667	333
DDR2-800	800	400
DDR2-1066	1066	533

4) DDR3 SDRAM

DDR3-SDRAM has significantly improved performance and has been widely used [6]. These further speeds up the processor and computer and improves the performance of the other dependent elements. Table 1.3 shows the DDR3 SDRAM data rate and clock speed.

High data rate: DDR3 SDRAM provides a data rate of 800 Mbit / s per pin at a clock rate of 400 MHz [7].

Lower supply voltage: The supply voltage is reduced by 0.3 V to 1.5 V compared to the DDR2.

Low Power Consumption: Simply lowering the supply voltage reduces the power consumption of an equivalent chip (if any) by a factor of 0.69.

Memory Capacity: The memory size of the DDR3 SDRAM chip started at 512MB and increased by 8GB

TABLE 1.3
DDR3 SDRAM DATA RATE AND CLOCK SPEED

DDR3 SDRAM type	Data rate (MB/s/pin)	Clock speed(MHz_)
DDR3-800	800	400
DDR3-1066	1066	533
DDR3-1333	1333	667
DDR3-1600	1600	800
DDR3-1866	1866	933
DDR3-2133	2133	1066

5) DDR4 SDRAM

DDR4 SDRAM was the 4th generation SDRAM that made many advances to enable faster operation [8]. DDR4 SDRAM was developed because of the growing industry need for higher performance memory, in this case SDRAM [9]. It was first introduced in late 2016. Some of the key features of DDR4 SDRA include,

High Data Rate: When DDR4 was introduced, DDR3 was expected to peak at a data rate of 1.6 giga transfer per second per pin. Therefore, this was set as the entry point for DDR4 SDRAM. This transmission rate is expected to increase to twice this level. i.e., 3.2 giga transfers per second, potential increase [10].

Internal Banks: The internal banks increase to 16 (4 bank selection bits), up to 8 ranks per DIMM.

Operating Voltage: The DDR4 chip uses a 1.2V power supply with a 2.5V auxiliary power supply for word line boost called VPP. This is compared to the standard 1.5V for DDR3 chips and a 1.05V low voltage variant will be available after initial deployment.

DQ Bus: One of the other features included in the DDR4 SDRAM standard is the pseudo-open drain interface for the DQ bus.

II. THE GROWING DATA AND MEMORY BANDWIDTH

The advent of newer technologies led to the growth of shear growth in the volume of data [11]. The research from IDC shows that only 2 zeta bytes of data was created and consumed throughout the world in the year of 2010. With 41 zeta bytes of data consumed in the year of 2019, the data consumption grows to an increasingly high amount of 64.2 zeta bytes. Figure 2.1 depicts the increase in the data consumption and creation over years. The jump in data is due to the employee's accessing data for work from home [12]. The data creation and replication are expected to grow beyond 180 zeta bytes by the end of 2025 [13]. The memory module must be equipped with the ability to manage large volume of data. The reason behind the growth of data includes

1) Increased data flow due to widespread advancements in Artificial Intelligence in data centers, 5G, Automotive and Internet of Things (IOT).
2) Increased demand for edge computing
3) Increase in the number of mobile users, and amount of data used per month.

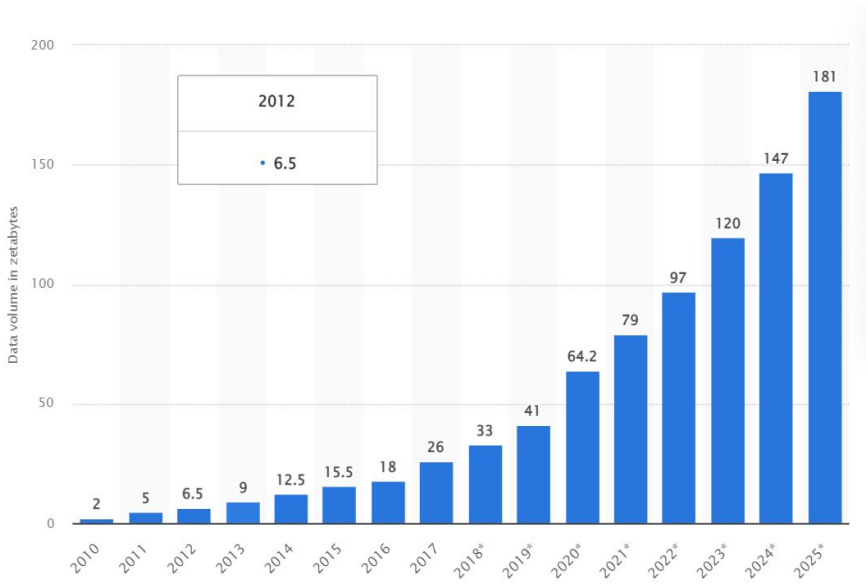


Fig. 2.1 Data consumption through years (source IDC) [14]

According to economic survey of India, wireless data usage has been increased by more than 7 times in the first quarter of 2022 compared to 2017-18 data usage. It is now 32397 peta bytes, also data usage per month of a person has increased from 1.24 GB to 14.1 GB in the same time [15]. The growing data puts a lot of loads on the data centers and increases demand for faster processing technology.

A. Growing Memory Bandwidth

The arrival of DDR5 memory module as opened new dimension of computation that can improve the performance of a system way beyond reachable by DDR4 memory module [16]. Table 2.1 shows a comparison of DDR5 memory specification with it predecessor DDR DRAM technology.

The DDR5 technology promises a higher bandwidth than DDR4. DDR4 stopped at 32 GB of memory size per stick, whereas DDR5 can go up to 128 GB per stick. DDR4 had a single 72-bit channel (64 bit + 8 bit of ECC) whereas DDR5 supports two 40-bit channels ((32-bit data + 8-bit ECC) *2). The additional ECC feature help in tackling any error during data transmission and improves efficiency and reduces data access latency. DDR5 comes with better power management system with its operating voltage scaled down by 0.1 v compared to DDR4. But the overclocked memory modules with strict timings can scale up the operating voltage. So far DDR5 has reached 1.35 volts on DDR5-6800.

TABLE 2.1
DDR TECHNOLOGY SPECIFICATION COMPARISON [17]

	DDR3	DDR4	DDR5
Max Unbuffered capacity	16 GB	32 GB	128 GB
Bandwidth	6400-17067 MB/s	12800 – 25600 MB/s	38400 – 5700 MB/s
Transfer rate	800 – 2133 MT/s	1600 – 3200 MT/s	4800 – 7200 MT/s
Base Frequency	400-1067 MHz	800 – 1600 MHz	2400 – 3600 MHz
Effective Frequency	800 – 2133MHz	1600 – 3200 MHz	4800 – 7200 MHz
Voltage	1.5 V	1.2 V	1.1 V
On die ECC	NO	YES	YES

The growing memory bandwidth not only comes with advantage of higher data rate and RAM capacity but brings problem in signal integrity management. The problem of maintaining signal integrity is solved by adding additional silicon in the form of buffer chips. The Buffer chips are usually used between CPU or memory controller and DRAM memory module.

B. Advantages of DDR5 over DDR4 DIMM's

1) High memory bandwidth and efficiency.
2) Better power efficiency and scalability
3) Low latency and higher capacity of DIMMs.

III.BUFFER CHIPS

Buffer chips are used in high-speed memory applications such as DDR4 and DDR5. Buffer chips replaced the "stub bus" architecture used in DDR1 and DDR2 and replaced it with an hierarchic bus system where only p2p or p2mp connections are used [18]. A memory system with very large number of memory chips can be produced by cascading. Buffer chips are usually connected with more than one memory arrangement, and they parallelize the said data and command and send them to appropriate memory arrangements. The buffer chips consist of data interface that receives the data from a memory controller. The internal logic of the chip parallelizes the data, and the second data interface that write the parallelized data to the memory blocks. This property of buffer chips makes them possible to deliver a single load of clock and command/address signal to the entire DIMM.

A. Buffer Chips in Data Centers

Data centers stores and shares application and data [19]. These centers support the storing, sharing, and disseminating of an organization's data and applications. With 79 zeta bytes of data created and consumed throughout the world, it is expected to grow beyond 180 zeta bytes by the year 2025. The rapid increase in the amount of data consumed also results in the computational overhead on data centers [20]. Memory bandwidth and CPU speed miss-match have become an important reason behind the computational bottleneck observed in servers or data centers. Servers employ two types of DIMMS one known as RDIMM (Registered DIMM) and other being LRDIMM (Load-Reduced DIMM). Both memory modules are supported with additional silicon to maintain signal integrity. LRDIMM employs an additional Data Buffer which improves signal integrity on the data bus. The challenge of maintaining the signal integrity is resolved by using buffer chips in data center memory modules. Buffer chips also help in maintaining the signal integrity of commands sent to memory modules by a memory controller and meeting the timing parameters. They can also be used over the information sent when the memory module must support many DIMM slots per channel of a DRAM memory module.

B. Register Clock Driver

RCD is a buffer chips used between the memory controller and DRAM memory module [21]. With the growth of need for the faster data, the demand for in-memory computation is also increasing. The RCD buffer chip is one such mission-critical enabler of DDR5 DIMM memory to deliver the bandwidth and capacity needed for next-generation servers. DDR5 RCD buffer chip is an important component of both RDIMM and LRDIMM memory modules of DRAM [22]. Figure 3.1 shows how RCD is interfaced between the memory controller and DDR memory module of the type LRDIMM. RDC buffer chip supports both SDR (Single data rate) where, data is sampled only on the rising edge of the clock and DDR (Dual Data Rate), where the data is sampled on both rising and falling edge of the clock. RCD utilizes an even parity mechanism which supports error checking and correction mechanism.

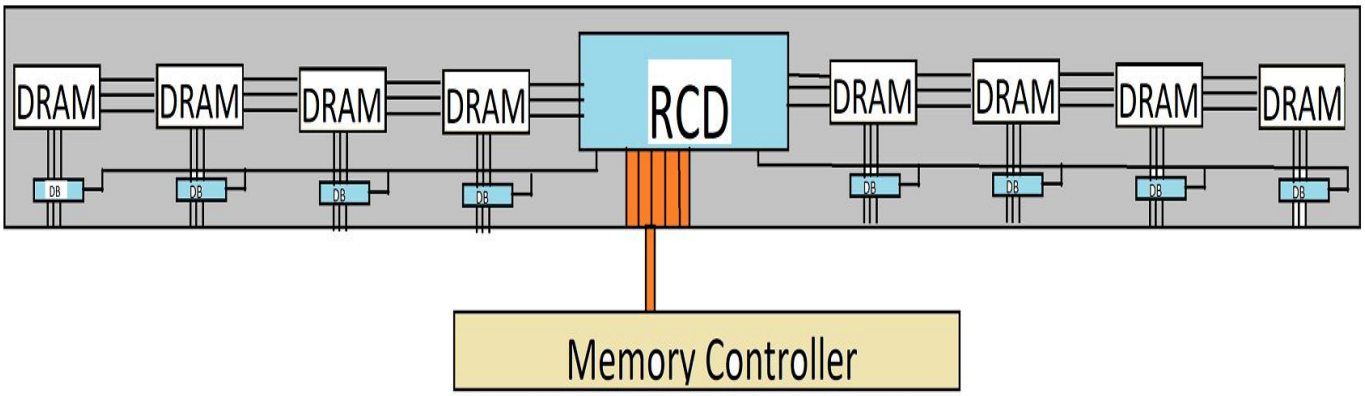


Figure 3.1 RCD Buffer Chip connected to Memory controller8

Table 3.1 shows the interfaces of an RCD buffer chip. Host interfaces is the input of RCD which is connected to the memory controller or a CPU. Since DDR5 has two channels each with 32 bit of data and 8 bit of error correcting code. The RCD also have channel-based system. There are two distinct sets of input interface that receive the data/command from the host controller.

TABLE 3.1
RCD INTERFACES

Interface	Pin Name	Description
Host Interface	DCA [6:0]	Command/Address Bus
	DCS [1:0]	Chip Select
	DPAR	Parity Signal
	DCK	Input Clock
	DRST	Asynchronous Reset
	ERROR_IN	Error input
DRAM Interface	QCA [13:0]	Command/Address Bus
	QCS [1:0]	Chip Select
	QCK	Clock at DRAM Side
	QRST	Reset
	ALERT_n	Alert
Data Buffer Interface	BCOM [2:0]	Command bus to DB
	BCS	Chip Select to DB
	BRST	Reset Signal to DB
	BCK	Clock to DB

Each channel of a DDR5 DIMM has 2 ranks, and RCD parallelizes the received data and drive to the ranks of DIMM module. There by reduces load on the memory controller. The data buffer interface is used to buffer command to the Data Buffers of a LRDIMM DIMM. The use of data buffers in LRDIMM memory modules advantages the DRAM.

Use of data buffer is advantageous in,

1) Machine learning, electric vehicle etc., have emerged as the latest technologies but the challenges with these applications are advanced signal equalization, and increased capacity and bandwidth.
2) Data buffering with its unique techniques helps in overcoming these challenges. Hence interfacing chips like RCD become important in data centers and autonomous vehicle and for machine learning applications.
3) Data buffers increase application performance by allowing synchronous operations such as file reads or writes to complete quickly instead of blocking while waiting for hardware interrupts to access a physical disk subsystem; instead, an operating system can immediately return a successful result from an API call, allowing an application to continue processing while the kernel completes the disk operation in the background.
4) The benefits of a data buffer are particularly evident in long-distance applications, when operating at higher data rates or in systems with a heavily loaded PCI bus.

IV. CONCLUSIONS

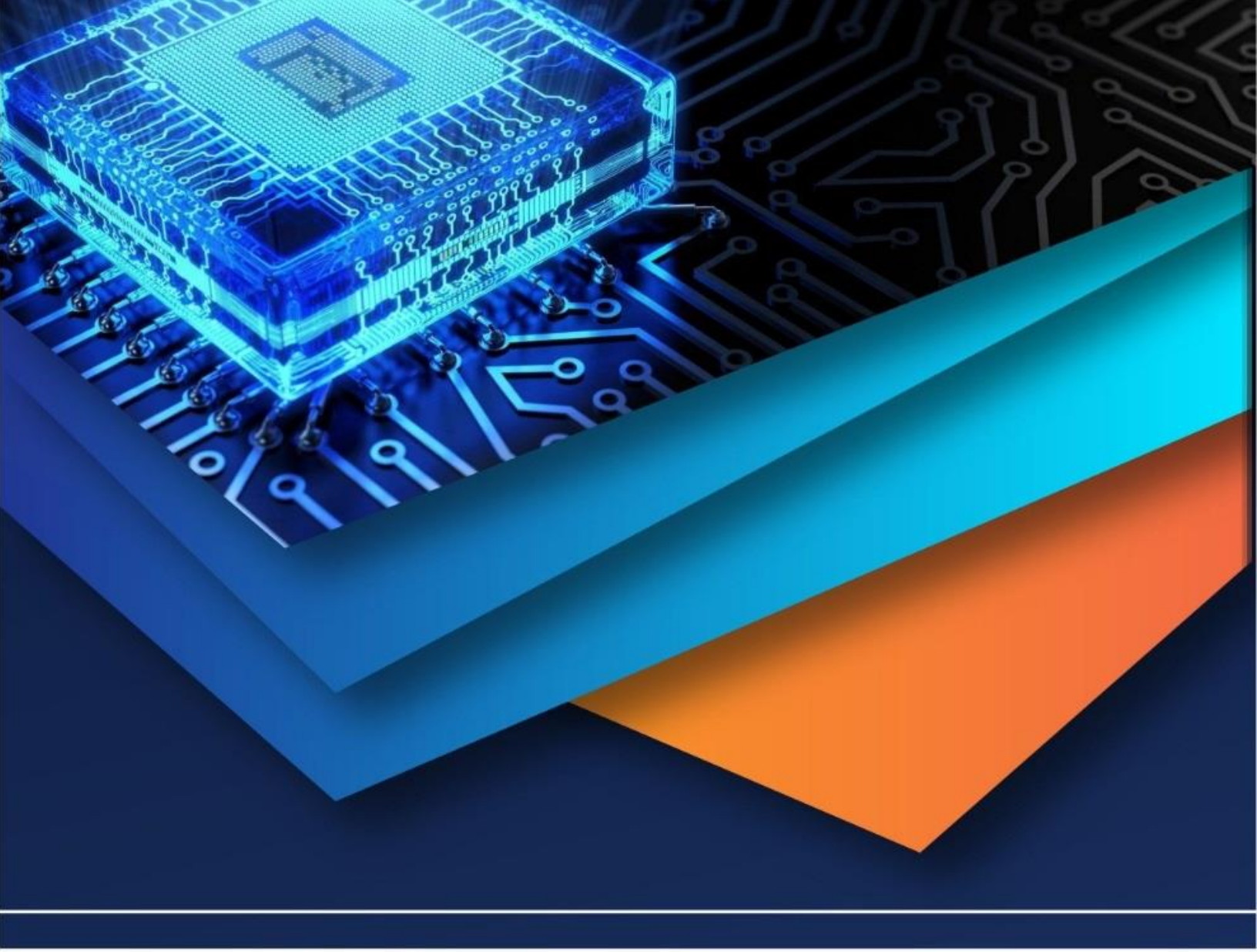
With the rapid evolution of CPUs and the integration of new features, memory system architects are now improving their memory bandwidth to feed the ever-growing computing pipeline. Memory bus data rates have increased to meet increasing bandwidth demands, but the increasing challenge of maintaining signal integrity impacts the number of DIMM slots per channel, and ultimately the processor, also affect the amount of memory available. Limiting memory bandwidth or capacity risks reducing processor utilization. Buffer chips are typically used in server memory systems to improve the signal integrity and timing of commands and addresses sent to memory modules. The benefits of memory module buffer chips extend to the system level. Additional silicon improves the data rate while allowing more modules to be connected to the memory bus. This allows the processor to run larger workloads and process them faster.

1) Memory buffer chips continue to evolve, as evidenced by the number of technological advances seen from one generation of silicon to the next.
2) The importance of server DIMM chipsets will continue to grow as the industry supports Moore ‘s Law and seeks to maximize the limitations of the von Neumann architecture in evolving data centers.
3) Buffer chips, GPUs, and FPGAs all help advance server and data center architectures, opening the door to new paradigms that can eliminate bottlenecks in current and future workloads.
4) Data centers are steadily moving away from the traditional approach of ’one size fits all’, which may continue and be strengthened in the coming years.

REFERENCES

[1] “Understanding DDR” Truechip, the verification IP specialist, https://www.truechip.net/articles-details/understanding-ddr/440552344.
[2] H. Fujisawa et al., "1.8-V 800-Mb/s/pin DDR2 and 2.5-V 400-Mb/s/pin DDR1 compatibly designed 1Gb SDRAM with dual-clock input-latch scheme and hybrid multi-oxide output buffer," in IEEE Journal of Solid-State Circuits, vol. 40, no. 4, pp. 862-869, April 2005, doi: 10.1109/JSSC.2005.845555.
[3] E. Chan, Huabo Chen and Chee Yee Chung, "High speed DDR performance in 4 vs 6-layer FCBGA package design," 2004 Proceedings. 54th Electronic Components and Technology Conference (IEEE Cat. No.04CH37546), 2004, pp. 314-319 Vol.1, doi: 10.1109/ECTC.2004.1319357.
[4] N. P. Shettar, "Design of Arbiter for DDR2 memory controller and interfacing frontend with the memory through backend," 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), 2016, pp. 2905-2909, doi: 10.1109/ICEEOT.2016.7755230.
[5] N. P. Shettar, "Design of Arbiter for DDR2 memory controller and interfacing frontend with the memory through backend," 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), 2016, pp. 2905-2909, doi: 10.1109/ICEEOT.2016.7755230.
[6] P. Guoteng, L. Li, O. Guodong, D. Qiang and X. Lunguo, "Design and Implementation of a DDR3-based Memory Controller," 2013 Third International Conference on Intelligent System Design and Engineering Applications, 2013, pp. 540-543, doi: 10.1109/ISDEA.2012.132.
[7] K. Park, S. Baeg, S. Wen, and R. Wong, "Active-precharge hammering on a row induced failure in DDR3 SDRAMs under 3× nm technology," 2014 IEEE International Integrated Reliability Workshop Final Report (IIRW), 2014, pp. 82-85, doi: 10.1109/IIRW.2014.7049516.
[8] “DDR5 vs DDR4: Is It Time To Upgrade Your RAM” tom’s HARDWARE, https://www.tomshardware.com/features/ddr5-vs-ddr4-is-it-time-to-upgrade-your-ram
[9] J. Zheng, K. Yan, Y. Zhang and Z. Chen, "Design and Implementation of DDR4 SDRAM Controller Based on FPGA," 2018 2nd IEEE Advanced Information Management,Communicates,Electronic and Automation Control Conference (IMCEC), 2018, pp. 421-424, doi: 10.1109/IMCEC.2018.8469378.
[10] T. Choi and S. Kim, "Shielding can design for a DDR4 connector system to reduce RFI," 2017 IEEE Electrical Design of Advanced Packaging and Systems Symposium (EDAPS), 2017, pp. 1-3, doi: 10.1109/EDAPS.2017.8277040.
[11] “What’s the real story behind the explosive growth of data” , redgate, https://www.red-gate.com/blog/database-development/whats-the-real-story-behind-the-explosive-growth-of-dat
[12] H. Glaser and H. Halpin, "The Linked Data Strategy for Global Identity," in IEEE Internet Computing, vol. 16, no. 2, pp. 68-71, March-April 2012, doi: 10.1109/MIC.2012.39.

[13] A. Rauber, O. Witvoet, A. Aschenbrenner and R. Bruckner, "Putting the World Wide Web into a data warehouse: a DWH-based approach to Web analysis," Proceedings. 13th International Workshop on Database and Expert Systems Applications, 2002, pp. 822-826, doi: 10.1109/DEXA.2002.1045999.
[14] N. Na and H. To, "Effectiveness of Equalization and Performance Potential in DDR5 Channels with RDIMM(s)," 2019 IEEE 69th Electronic Components and Technology Conference (ECTC), 2019, pp. 1208-1214, doi: 10.1109/ECTC.2019.00187.
[15] "Internet usage worldwide - statistics & facts" source: IDC. www.statista.com
[16] "Massive data usage growth in India: Economic survey" THE ECONOMIC TIMES, https://economictimes.indiatimes.com/industry/telecom/telecom-news/massive-data-usage-growth-in-india-economic-survey/articleshow/89245236.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst.
[17] "DDR5 SDRAM-product core data sheet" MICRON Technology Inc.
[18] "Buffer chips and method for controlling one of more memory arrangement", George Braun, Holzkirchen, Hermann Rucker Bauer. United states patent, March 2004
[19] "Data centers evolution: DDR5 DIMMs advance server performance" RAMBUS Inc, 2021
[20] A. Abdusalam, A. R. bin Ramli, N. K. Noordin and M. L. Ali, "Real time data acquisition and remote controlling using World Wide Web," Student Conference on Research and Development, 2002, pp. 456-459, doi: 10.1109/SCORED.2002.1033156.
[21] "DDR5 Registering Clock driver-enabling breakthrough memory bandwidth and capacity in servers" . Rambus Inc, www.RAMBUS.com
[22] "DDR5 Registering clock driver definition (DDR5RCD01), JEDEC standard document, JESD82-511, August, 2021.
[23] "Understanding DDR" Truechip ,the verification IP specialist, https://www.truechip.net/articles-details/understanding-ddr/440552344.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY