Next Generation Memory Interfaces Final Capstone Report

Sinan Liu

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2015-95
http://www.eecs.berkeley.edu/Pubs/TechRpts/2015/EECS-2015-95.html

May 14, 2015

University of California, Berkeley College of Engineering

MASTER OF ENGINEERING - SPRING 2015

Department: EECS

Concentration: Integrated Circuits & Physical Electronics

Next Generation Memory Interfaces Final Capstone Report

Sinan Liu

This **Masters Project Paper** fulfills the Master of Engineering degree requirement.

Approved by:

1. Capstone Project Advisor:

Signature: _____ Date _____

Print Name/Department:

2. Faculty Committee Member #2:

Signature: _____ Date _____

Print Name/Department:

Table of Contents

Project Introduction	**3**
Industry, Market and Trends Analysis	**4**
Intellectual Property Analysis	**18**
Technical Contribution	**23**
Concluding Reflection	**35**
Bibliography	**36**

Project Introduction

Our project, *Next Generation Memory Interfaces*, aims to develop a physical interface for the latest memory standard i.e. DDR4 SDRAM (Double Data Rate 4^{th} generation, Synchronous Dynamic Random Access Memory). The DDR4 standard allows for memory to be accessed at twice the data rate of its preceding standard, DDR3 while simultaneously reducing the total power consumption and increasing the memory density. As memory devices eternally seek to be faster, denser and extremely low power consuming systems, our interface will get us another step further in this quest.

In September 2012, JEDEC Solid State Technology Association, the organization that defines standards for the semiconductor industry, released the JESD79-4 for DDR4 SDRAM. With the release of this standard, companies such as Samsung, Micron Technology, Agilent Technologies and several others are developing SDRAMs compliant with this standard. The intent of our project is not to compete with the industry, but to explore the research and development opportunity presented by the new standard. To achieve the higher speed and low power requirements of DDR4, several changes are required at the architecture and circuit level. We aim to implement the strategies required to design a functional interface that meets the specifications of the DDR4 standard.

SDRAM is one of the commonly used types of memory in computing systems. It is a volatile memory that requires periodic refreshing to store the data.

Owing to the speed and structural simplicity, DRAMs are often used as the main memory in personal computers and workstations. DDR is a class of memory that transfers data twice as fast as a SDR (Single Data Rate) memory since the transmission occurs on both positive and negative edge of the clock. DDR4 is the fourth generation of DDR SDRAM which operates at a supply voltage of 1.2 V with data rates up to 3.2 Mbps – twice that of the preceding DDR3 standard.

Industry, Market and Trends Analysis

Integrated circuit design is an important sector in the semiconductor industry. The semiconductor industry is known to be highly competitive in nature, and the trend has been increasing over the years (Ulama 2014:19). Product life cycles are short, as more technologically advanced products replace older ones. Adoption of products is majorly affected by performance and reliability. The major companies in the memory technology industry are Intel Corporation and Samsung Electronics with 7.6% and 5.3% of the global semiconductor and electronics parts manufacturing market, respectively (IBISWorld 2015: 27). Broadcom Corporation, Texas Instruments Inc., Advanced Micro Devices (AMD) Inc., Micron Technology Inc. are few of the other major companies that compete in this industry (Ulama 2014:19). Noticeably, the major companies in this industry are all fairly large and well established, and compete over products and technologies. High demand for products and extremely low pricing intensify the competition in the industry. This poses a significant barrier to entry for new and smaller companies leading to only several companies currently building

DDR4 memory chips and controllers.

As we strive to make a smaller, faster, and more efficient memory controller, we have to compete with the products and research and development efforts of competing companies. Our competitive landscape does not merely include semiconductor companies, but also technologies that have similar features and functions when compared against our project. Existing memory technologies, such as DDR3, 3D stacked (3DS) - DDR3, and GDDR4, compete with DDR4 on various parameters such as cost, speed, and use-cases. While DDR4 is faster than previous memory generations, the higher cost of the new chip technology would make the cheaper DDR3 technology a strong competitor.

Emphasis is placed on the significant performance improvements that DDR4 presents over DDR3 technology. The following table shows a brief comparison of the key features between two technologies.

Table1. Comparison between DDR3 and DDR4 [1]

	DDR3	DDR4
Power supply voltage	1.5V	1.2 V
Speed	1.6~2.1 Gbps	1.6~3.2 Gbps
Density	8GB(max)	16GB(max)
Price	$100 avg	$200 avg.

The first comparison is from the aspect of power efficiency, not only does DDR4 have a lower supply voltage, but it also implements a new algorithm to control its energy consumption by entering its “standby” mode more frequently and precisely than DDR3. The improvements lead to better performance in both the power consumption, and operating temperature.

Furthermore, the most essential feature, memory speed, has been improved significantly in DDR4. The analogy between the memory speed and highway traffic speed is very descriptive. The speed of the memory is the amount of data can be transferred in a certain period of time. There are two factors determining the speed, which are, interface width and frequency of the memory’s operation. Considering the analogy, the bandwidth is the quantity of lanes on a highway, and frequency is the travel speed of its vehicles. Within a fixed time period, having more lines and a faster speeds will allow for more vehicles to travel. Similarly, having an improved working frequency, along with an enlarged bandwidth, DDR4 achieves a data transmission speed which is approximately 1.5 times faster than DDR3, as table 1 indicates. The increase of the speed is benefited from the revolutionary bank-group management technology.

Another differentiating factor is the density, or say, the space of a single memory chip. The advancement the memory chip’s encapsulation gives DDR4 a 50% density increase, with regard to maximum space. With a larger storage space, DDR4 is able to process more information simultaneously. However, similar to every emerging new technology, the current price of DDR4 memory is 30% to 50% more expensive than

DDR3, which can achieve an akin functionality. With a large-scale adaptation for DDR4 memory, and hardware compatibility of its peripheral devices, the manufacture price would quickly become more affordable in the future.

The GDDR3 and GDDR5, which stands for Graphics Double Data Rate 3 , are a kind of memory specifically designed for image processing. Despite the similarity in terms of the name, the graphic memory is named one generation ahead regular memory. This means that the core technology of GDDR3 is essentially an upgraded version based on DDR2 technology, rather than DDR3. The graphic memory is designed to have lower energy consumption, and an optimized performance when dealing with graphical-data processing. Since the application area of these two kinds memory differentiates amongst each other, they do not compete directly. The Graphic DDR is typically developed based on the previous generation of DDR memory technology, with improvements on speed and application-specific functional modifications.

As for the market analysis of our product, the main markets include traditional memory devices and consumer electronics - and they are booming. Based on transparency market research, it states that the global next generation memory technologies market was worth $207.8 million in 2012, and is projected to be worth $2,837.0 million by 2019, growing at a 46.1% average growth rate from 2013 to 2019 (Transparency Market Research, 2014). The report divides the overall market for next generation memory technologies on the basis of certain parameters: interface type, application, and geography. On the basis of interface, the market for next generation

memory technologies can be categorized into SATA, SAS, DDR, and PCIe and I2C (Transparency Market Research, 2014). The main applications of next generation memory technologies include embedded MCU and smart card, mobile phones, mass storage, cache memory, enterprise storage, and automotive.

Geographically, the global next generation memory technologies' markets can be divided into North America, Europe, Asia-Pacific, and the rest of the world. This industry is always looking for ways to decrease power consumption, increase density, and develop clever architectures. The new generation memory technologies market has gained significant momentum in recent years due to growing demand for faster, highly scalable, and cost-effective memory solutions.

Understanding the necessity of our effort follows from understanding the industry dynamic, which our product tempts to enter. After understanding the landscape within which we stand, we remain to have reason to believe that our project is valuable to our stakeholders. We remain to reason that our stakeholders should be more interested in receiving a completed deliverable from us over any other, equally qualified, external competitor.

Our first differentiating quality is that we offer to provide "non-contracted" work. Contracted work is any work commissioned by one party to be executed by another party. To begin such work, both parties must agree on the terms defined within the agreement document prior to the work's commencement. The agreement is realized through means of a binding contract that both parties agree to enact. Once the contract is created, it typically cannot be altered or modified, unless the consent of all parties is

evident. This could place the requesting party into a stiff situation if it discovers that its priorities have changed mid-way through a contract.

Upon the project's completion, the completed work is commonly handed off "as-is." This means that no additional support is to be provided in the future (unless explicitly negotiated upon within the original contract). Any additional requested support or modification requires for a new contract to be written up. Not only is this financially inconvenient, but it can also be logistically inconvenient for the recipient. Without support, the deliverable is handed off with a decreased utility. The recipient of the deliverable is stuck with using the deliverable solely within its original scope.

Our stakeholder, BWRC, benefits from ownership over the development process. A common clause added to most contractual work instills a limit on interim design modification requests. This clause exists to prevent the requesting party from overexerting the contracted party without compensation. Internal control over the developmental process allows for precise design-source malleability during development, and full exposure of the design files. Design-source malleability allows for the BWRC professors to more closely guide our direction through the project's development. It allows for them to change the path that we follow if new interests arise. There is no contractual overhead to worry about in this scenario.

BWRC benefits from retaining access, and owning, the source code and designs. The design-source exposure enables BWRC to question every aspect of the implementation until they understand it completely. With contracted work, this information is typically unavailable to the requesting party due to trade secrets being

used in a design. Owning the source enables BWRC to have permanent design-source access. Long-term source access enables cost-effective and effort-effective technology adaptation into any future BWRC projects. Along with adaptation, owning the source creates the opportunity for growing in-house expertise at BWRC through education.

The benefits mentioned above align very closely with our stakeholder's interests. The stakeholder, being BWRC, is interested in three main attributes from the project. First, BWRC wants a fully customizable deliverable due to unpredictable future demands. Second, BWRC wants the freedom to optimize the design for unique implementations that would require the modification of the source on a per-use basis. Third, BWRC wants to avoid the financial, temporal, and contractual overheads associated with third-party work. Our project delivers on all three attributes. By choosing to complete this project through our team, rather than a team of contractors, BWRC satisfies its internal interests.

Our team anticipates BWRC's decision to work with us as opposed to larger suppliers. The current semiconductor marketplace is saturated with both customers and suppliers. As Ulama describes (Ulama 2014:28), "Established operators in this industry have been able to develop solid relationships with customers, and it can be extremely difficult for new companies to gain contracts with customers when existing semiconductor manufacturing operators have built reputations over a long period." To exemplify the significance and the weight carried by the previous statement, note that the Semiconductor and Circuit Manufacturing industry is one of the largest exporting industries in the United States (Ulama 2014:5). It indirectly provides jobs to 250,000

Americans, is currently valued at $79.5 billion, and has grown at an annual rate of 4.8% (Ulama 2014:5).

The current players, both customers and producers, are very well established, and very tightly connected. Penetrating into the customer base that the massive producers currently support is near impossible for a small team like ours due to lack of reputation. Aside from penetrating, the customers in this segment of the market are a significantly strong force due to two reasons: 1. The intrinsic competitiveness of the current suppliers, and 2. "The electronics marketplace is continually under pressure to improve product functionality, decrease size, increase speed, and decrease cost." (IBISWorld Global Semiconductor & Electronic Parts 2015:33)

Our team has set our target in a completely different direction. Instead of focusing on the massive customers, who are already served very competitively, we direct our focus at an interestingly under-served segment in this market space. In part, our choice of direction is due to the methods through which our Capstone project was decided upon. The decision process confined the scope of the project to target academic goals and provide solutions for academic institutions. Thus, our customer space currently only encompasses the Berkeley Wireless Research Center, but is functionally able to serve any academic or small-scale organizations.

As we currently stand, with one effective customer in our sights, we are subjecting ourselves to a very strong customer market force. This is an undesirable outcome due to the limited size of the space, which we choose to attack, but success in this space will send positive signals at other research institutions. We would be able

to expand to encompass more academic institutions because they would prefer to acquire the product through us. Our effective results are comparable to their current methods of operation, but with the benefit of reduced fixed-cost expenditures – which arise when placing orders with large design and manufacture firms.

A majority of the market belongs to other companies; most all companies are well establishing large corporations including Micron (IBISWorld 2015: 27) and Texas Instruments (IBISWorld 2015: 30). The barriers that cause this include "access to latest technology and intellectual property, the level of investment…, access to skilled employees, and the dominance of existing players" (IBISWorld 2015: 25). In the memory industry, the companies compete over a very specific set of criteria including price, performance, features and power consumption, all of which are highly measurable and quantifiable metrics (IBISWorld 2015: 24).

If the dimensions of competition between companies in a given industry converge, then the companies are left to compete solely on price (Porter 2008: 12). In the integrated circuit market, the industry has converged heavily on these metrics of performance, features and power consumption, which has resulted in fierce price competition. Because "economies of scale can be significant in this industry" (IBISWorld 2015: 25), new entrants must manufacture large volumes to stand a chance against the bigger companies. This requires up-front capital that many smaller new entrants do not have available. Entering the market attempting to compete on these highly competitive dimensions would result in "zero sum competition" (Porter 2008: 13), and and would not be a viable business strategy.

When instead of converging on the same dimensions, companies target different segments of the customer base, the result can be "positive sum" competition: competition that increases the profitability of all companies (Porter 2008: 13). We plan to employ this strategy with our DDR4 memory controller. We've learned from BWRC that their needs are different than the typical semiconductor customer. BWRC fabricates chips in low volumes, so price is not a significant factor. Also, they require only a subset of the industry-standard feature set for DDR4 memory controllers, enabling us to reduce the size of the design. Finally, they need very specific portions of the controller designed, not the entire IP block that most competitors would offer.

Although the memory technology industry is highly competitive, growing, and difficult to penetrate, the market is growing fast due to this a demand for consumer electronics, an industry which is expected to grow 5.3% annually to nearly $300 billion dollars by 2019 (IBISWorld 2014: 4). This high demand and new market bring some space for new companies to enter and grow. These new entrants usually emerge during the transition between the technological revolutions and each one has its own specialty.

From the year of 2002 to 2013, DDR memory industry has undergone 4 significant technological transitions, all of which are aiming at improving in three performance aspects and achieving a denser data processing capability. As Darryle stated in the article, the product with "high levels of performance, reliability, quality and low levels of power consumption" (Bach,2014:6) can gain an apparent advance in the competition of memory design industry. Being the three largest manufactures of

memory chip and developer of DDR memory technology, Samsung, Crucial (Micron) and Hynix have already invested millions of dollars in their R&D sector to develop the new generation DDR4 memory interface in order to reinforce their dominating market share.

Given such a giant market, other major memory designers such as Kingston keep fastening their pace to catch up the memory controller design for the recent DDR3 to DDR4 transition. Besides the companies who are already in the market, there are significant numbers of new companies or say, new entrants, trying to seize this opportunity. According to the statement made by Darryle, "the latest Census data indicates that 64.1% of operators in this industry have fewer than 20 employees" (Ulama 2014:25). The development strategy of those new entrants is highly focused on certain features, and "specializing in a small number of product lines to serve niche markets" (Ulama 2014:25) in order to avoid a direct competition with large companies.

Table2. Representative new entrants in DDR4 memory development

Company Name	Specialized market/feature
Century Micro INC.	Small physical size & low energy consumption
Montage Technology	Fast operating speed & low energy consumption
G.SKILL	Enhanced gaming performance

Three unique, representative companies are provided to conduct the analysis

of the new entrant. The table 1 above shows a brief comparison of three distinguished new entrants' key product features. It indicates each new entrant is trying to gain its market share by specializing its product from the three technical aspects mentioned in the previous paragraph.

The Japanese based company Century has just halved the physical size of DDR4 memory in their most recent product at the year of 2014. The China-based Montage Tech is more focusing on developing fast speed and lower power rate DDR4 memory for large scale server use. "Less power draw means less heat and longer battery life", which indicates "the servers are expected to be the biggest beneficiaries of the jump to DDR4" (Andy, 2014:6). Meanwhile, G.SKILL put majority of its resources into developing DDR4 memory controller with improved gaming performance. These companies are increasingly securing their niche markets by making breakthroughs in design of the memory controller while the major developers are still dominating the memory chip manufacturing area.

Big companies enjoy economies of scale, making it difficult to compete with them in manufacturing the integrated circuits (ICs). Based on the analysis of the new entrants, in order to build immunity for our design, we plan to segment the market to research institutes like BWRC. Their needs are different from most, and provide an opportunity for us to develop a product that satisfies these needs better than the competition. Since the design of our project is specifically for BWRC internal research use, there will be no direct competition and obvious threat from these new entrants either.

The threat from other technologies is weaker, as our DDR4 interface is more advanced than existing DDR3/GDDR5 interfaces. Therefore, we focus on developing the intellectual property and targeting the specific needs of the academic communities. This specific category of consumers require more customizable, and open, circuit designs at a lower volume, a need that is unmet by the larger companies that package their circuits in black boxes, manufacture in high volume, and allow little to no customization. By segmenting the market based on unmet needs, and our abilities to satisfy them, we hope to entrench our position as a profitable part of the semiconductor industry.

From the perspective of semiconductor circuit design, it is a complicated process to design a controller and integrate it with the memory chip. Therefore, our technology suppliers include both software side and hardware suppliers. Software suppliers are those who provide coding languages, design platforms, and simulation tools. Hardware suppliers are those who provide electrical specifications, datasheets, and other fabrication characteristics relating to memory chips.

Software suppliers mainly provide programming language support. Verilog and SystemVerilog are the two main programming languages we are using. They are hardware description languages used to model electronic systems. They are most commonly used in design and verification of digital circuits. Cadence, a company that provides electronic design automation software, covers many language design platforms, including Verilog and SystemVerilog. As an all-in-one suite, Cadence is our main software supplier.

Hardware suppliers provide descriptive information about the memory chip technology. Our controller is on a software level, but it will be integrated with the next generation memory chip technology, the DDR4 technology. Each generation of memory chips has new fabrication breakthrough. Thus, during our controller design, the latest information about memory chips is critical, such as voltage supply of the chips and the memory bank structure. Our hardware suppliers, such as Micron Technology, Intel Corp., and Samsung, are big semiconductor companies in this industry. In Semiconductor & Circuit Manufacturing in the US Industry Report, Intel Corp. and Samsung have 18% and 13.8% market share in 2014 (Ulama 2014:4). Although they seem like our competitors from the sales end, they also have the best research departments and technical experts in the chip fabrication domain. Samsung competes in the Semiconductor and Circuit Manufacturing industry via its fabrication and research and development facilities in the United States (Ulama 2014:4). They will release the paper and datasheet of their latest research results about DDR4 memory chip. According to the information provided by these large semiconductor-manufacturing companies, we are able to define the interface and design our memory controller.

Powerful suppliers capture more of the value for themselves by charging higher prices, limiting quality or services, or shifting costs to industry participants. As mentioned above, Intel Corp. and Samsung are both suppliers and competitors for us. If they limit our access to their latest technology about DDR4 memory chip, it will be hard for us to compete with them. However, the good news is that the DDR4 memory

specification is becoming a standard, so we will be less dependent on them.

There are certain aspects that we can focus on to succeed in this capital-intensive, and research-intensive, memory design industry. New companies are trying to explore the market by boosting their expertise in faster-speed designs, smaller dimension layouts, and highly customized application-specific designs. With increasing maturity of the DDR4 technology, the competition is becoming fiercer. This increased competition will largely benefit the semiconductor industry's evolution speed, as well as provide customers with cheaper and higher efficiency devices. Our project will not only encourage further development from competing companies and research groups, but also benefit BWRC's exploration of the utilization of DDR4's capabilities.

Intellectual Property Strategy

The PHY interface provides us a good scope for creating a patentable Intellectual Property (IP). The physical layer has been split into 5 major parts - serializer, transmitter, receiver, deserializer and timing circuits. Each of these allow for novel implementations and innovations in circuits design. As we are working at the cutting edge of technology, we would have to adopt ingenious techniques to meet the specifications for high data rates of DDR4. One or more of these implementations can provide us a patentable IP. This paper will discusses why this technology may be patentable, the advantages and disadvantages of seeking a patent, the current state of the semiconductor IP space, and the risks associated with not seeking a patent.

In the context of IP, creative designs and creative solutions fall cleanly under the category of patentable assets. In essence, the purpose for securing IP is to declare discernible ownership over a design or utility (USPTO, 2013). As an independent entity, we can draw benefits from securing patents and owning patents. The benefits we pose to secure range from monetary compensation to strategic industrial presence.

From a monetary perspective, owning patents allows our team to claim ownership to a recognizable asset. After incorporating our team as a legal entity, a patent opens us to the opportunity of being acquired. The proceeds from an acquisition could be used to finance additional ventures, which our team currently does not have the financial freedom to pursue.

A secondary monetization strategy that patent ownership affords us, is the option to license our technology to independent entities who wish to avoid committing R&D expenses for the purpose of developing said technology independently. Aside from the legal expense that we would need to undertake, the licensing option is financially robust.

The third and final benefit is an unquantifiable benefit. The third benefit arises from establishing a reputation as an entity. Acquiring a patent will demonstrate that we, as a team, know how to drive concepts into patentable ideas, and patentable ideas into awarded patents. Successfully acquiring a patent will demonstrate to that we are capable as a team, and will instill external confidence into our capabilities. This reputation will position us to open new leads amongst skeptical and risk averse customers.

The disadvantage of applying patent is obvious: it burns money. Filing a patent is not as simple as people imagine. Normally attorney fee becomes a big piece of the cost. Determined by the type of invention, the attorney fees are range from $5000 to more than $15,000(Quinn, 2011). Adding the government filing fee and all kinds of application fees, the total cost of preparing and filing a patent may exceed ten or twenty thousand dollars. In addition, the maintenance fees would be another big part of the cost. Depends on how many years the owner wants to keep, the maintenance fees float from $490 for small entities and $980 for large entities due at 3.5 years to $2055 for small entities and $4110 for large entities due at 11.5 years (Stim, 2012).

Considering that this IP would be used only for research or instructional purpose, it would be non-profitable. Therefore, applying for a patent brings financial burden to the owner. In other words, it would not be worth for individuals to applying patent for this IP. However, if the owner switches from individuals to college or Berkeley Wireless Research Center (BWRC), the conclusion would be different. First of all, the college or lab has budget to cover the cost. Furthermore, the patent would bring them reputation, which is far more important than profit for them. Thus, it would be worth to apply a patent for this IP for BWRC.

Unfortunately, the semiconductor IP market can be difficult for smaller entrants like ourselves. The rate of patent enforcement by larger corporations has not increased over the past few decades (Hall 2007, 5). However, in attempts to increase market share and presence, they have increased the number of patents they file. In the 1980's, the median number of patents filed by an employee was less than one, whereas during the

turn of the century it was near eight (Hall 2007, 10). While larger corporations have a broad and ever expanding portfolio, smaller firms focus on particular market segments in attempts to perfect and own this portion of the total revenue stream. Unfortunately for these smaller firms, this means that if and when larger corporations expand into their territory, they have no choice to defend what little they have. It is for this reason that smaller firms tend to more be more aggressive in enforcing their patents (Hall 2007, 3). Thus, it can be expected that we would have to actively enforce our patent. If our patent (or patents) focused solely on DDR4 memory control and interfacing, then we would have no choice but to defend the few eggs in our basket.

The risks associated with not patenting the design are significant. Since the integrated circuit design is based on following certain physical requirement and universal specifications, hundreds of similar design and product can be invented in the short time of period based on a same standard. In DDR4 memory design particularly, JEDEC standard is the critical specifications that everyone need to comply. There is high possibility that other individuals or companies will come up with very similar or even the same design. As Gene indicated in his article, engineers who are working on solving a certain problem "are likely to find solutions that are similar" (Gene 2009:8). If a similar design is first patented by other entities, the potential financial loss is irreparable and a great amount design effort would be wasted. Furthermore, without patent the design appropriately, competitors and free-riders can easily take advantage of the design or embedded our inventions into their products without any recognition of our work and having any consequence. Besides these two factors, without right

patenting, it is almost impossible to conduct technology transferring or licensing. And this would greatly impede the process of commercialization of the invention or designs.

Therefore, there are a great number of critical risks involving in not patenting the design and our memory controller design should be patented when its major functions and specification are met. The management of the patent can be done via creating a patent portfolio. By using management software or having regularly review, updates, categorization and balancing of the patent portfolio, the management with quality could be achieved.

Trade secret is one kind of intellectual property with unlimited time of protection. It can be one method to protect our technology, but it is not the optimal. The DDR technology evolves every three years averagely, an unlimited protection time is unnecessary. Giving the fact that circuit design industry is highly standardized and reverse engineering of circuit is quite mature, it would be difficult to protect the design with only trade secret but not patent since the trade secret sufferers from commercial espionage and high cost of protection. Moreover, the trade secret cannot prevent the similar or same product from being designed. Due the nature of the circuit design industry, trade secret won't play an excellent role in limiting other similar designs. As Shane said, patent is able to "protects your rights regardless of what anyone subsequently develops" (Shane 2007:8). Therefore, for technological inventions such as circuit design, patenting would be the optimal method to protect its originality.

Ultimately, deciding whether to seek or not to seek a patent for our design depends on the novelty of the final product. If we discover and implement a new

physical layer architecture that provides performance, costs, and/or feature improvements over the competition, then the patent's value overcomes the cost associated with filing it. If the final outcome is unique, but provides only marginal benefits compared to the competition, then there will be no benefit in filing the patent.

Technical Contribution

Overview

Our main task is designing the physical layer interface (PHY) for the latest memory standard, Double Data Rate 4th generation (DDR4). DDR4 memory is the next generation memory chips that would be used in computers and servers.

Before getting hand on design of interface, everyone was assigned task of literature research on DDR4 structure. DDR4 is a new memory standard with many latest functionalities. It is necessary to get familiar with DDR4 structure before we getting start our design work. We divided into organization, states, transition, operation modes, and PHY. These five sections cover all fields of memory structure, controller, and physical interface. I took charge of research on organization & addressing. The insight is introduced in literature review section.

After knowing the DDR4 standard, we moved to our main design task, PHY. The interface consists of five blocks: timing, serializer, transmitter, receiver, and deserializer. These five parts enable the efficient and precise data transmission between controller and memory chip. Considering that these five parts are equally important, we distributed them to five members. Kyle Dillon took charge of serializer;

Chenyang Xu took charge of transmitter; Kalika Saxena took charge of receiver; Miron Veryanskiy took charge of serializer; I took charge of timing. My part would generate a two-phase 1.6 GHz clock signal and provides necessary clock delay for other blocks in order to capture the data correctly.

Literature Review

Bank Group

Bank group is one of the biggest features used in DDR4 organization. Based on *DDR4 Bank Groups in Embedded Applications* published by Synopsys, "bank group used in DDR4 was borrowed from the GDDR5 graphics memories. In order to understand the need for bank groups, the concept of prefetch (burst length) must be understood. Prefetch (burst length) is the term describing how many words of data are fetched every time a column command is performed with DDR memories" (Allan, 2012). It will be helpful to have a quick look at the evolution from DDR1 to DDR4 in order to better understand bank group. Burst length of DDR1 is two. It means two data bits are sequentially output from the memory for one read operation. The number of global Input/output (GIO) channel is calculated by burst length multiplies the number of output data (DQ), as shown in Fig.1. Burst length of DDR2 was doubled to four. It reads four data bits at the same time, so DDR2 needs four GIO channels for each DQ. DDR3 did the same thing: burst length was doubled to eight; clock speed was doubled; the number of GIO channels was doubled as well.

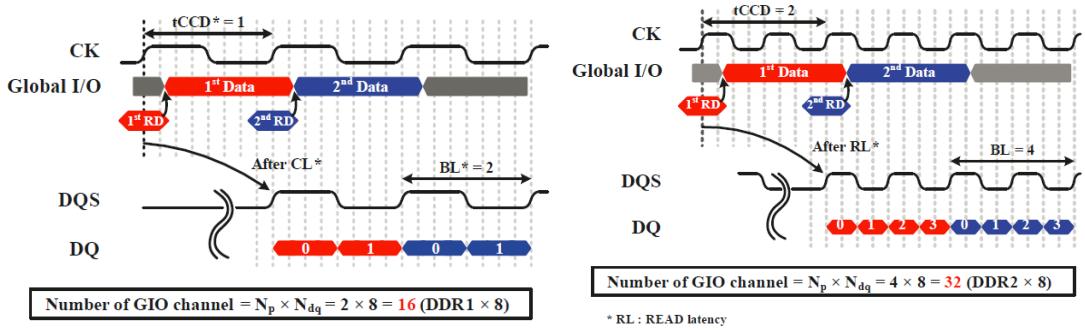

Fig.1 Timing diagram of read operation for DDR1 & DDR2

Continuing the trend with DDR4 would have required DDR4 to adopt a burst length of 16 (Allan, 2012). However, with increased burst length, the number of GIO increased as well, which will make the size of memories much bigger and expensive. In other words, the size was increased with burst length. Bulky size and high cost becomes the tradeoff of quick data speed, so designers saved cost and space by not going to double burst length any more. "DDR4 solves this problem by introducing the bank group. With bank groups, a prefetch of eight is executed in one bank group, and another prefetch of eight can be executed in another independent bank group. The bank groups are separate entities, as shown in Fig.2, such that they allow a column cycle to complete within a bank group, but that column cycle does not impact what is happening in another bank group", written by Graham Allan.

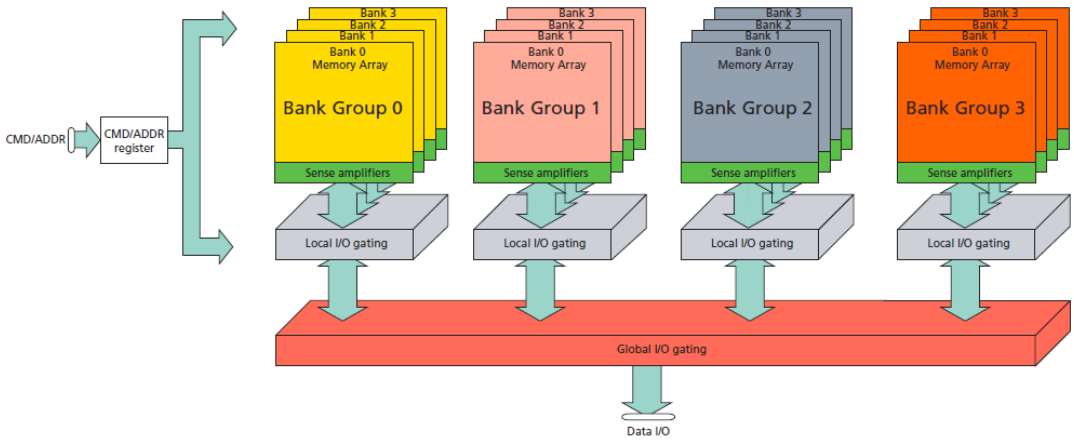

Fig. 2 Bank groups of DDR4 SDRAM (JEDEC)

There are two new timing specifications, tCCD_S and tCCD_L, coming from the introduction of bank groups. The "CCD" stands for "column to column delay". The "_S" means "short", and the "_L" means "long". In *DDR4 Bank Groups in Embedded Applications*, Graham Allan wrote "going from one bank group to another bank group does not have dependencies on these specifications. That is the case for using the tCCD_S specification, which, at four clock cycles, is unrestrictive in a manner similar to DDR3. However, going from command to command while staying within the same bank group, requires attention to DDR4's new tCCD_L specification, which is typically greater than four clock cycles. This case can impact the performance of design, which is particularly important in embedded applications." (Allan, 2012)

PVT

After research on the new memory structure and specification of DDR4, every member moves to individual block research. I took charge of timing block of the whole PHY, which is a critical part in the whole system. It is written in *DLL/PLL on a DRAM* on Rambus website "synchronous communication is used to achieve high data transmission speed to and from the memory. A clock signal always acts as a timing reference so that data can be transmitted and received with a known relationship to this reference. A difficulty in maintaining this relationship is that process, voltage, and temperature (PVT) variations can alter the timing relationship between the clock and data signal (DQ), resulting in reduced timing valid window" (Rambus, 2014). This

issue becomes more obvious in high frequency system, which limits the ability of systems to transmit data at high speeds. As shown in Fig. 3, in low frequency, the PVT variations cause data uncertainty area and shorten the data valid window. In high frequency, it may result in no data valid window.

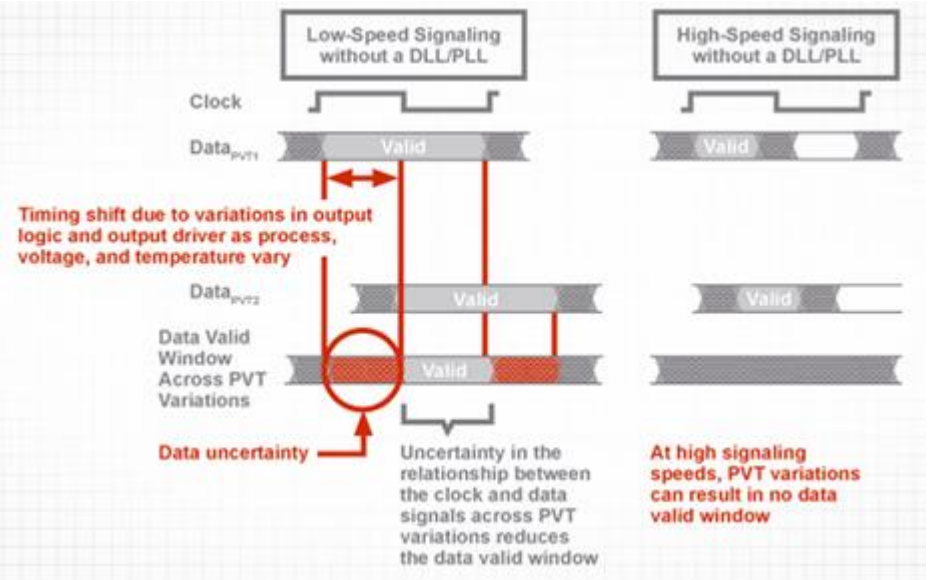

Fig. 3 Reduced valid window due to PVT variations

Receiver cannot estimate how much the data valid window changes caused by PVT variations. The data de-serializing and handing over to internal write command synchronized to the internal clock was another design challenge at high data rates due to large PVT variation (JSSCC, 2013). Building on these issues, my job is to generate delay on clock signal to compensate PVT variations and enable the receiver to capture the output data correctly.

Data strobe timing

In *Data Strobe Timing of DDR2 using a Statistical Random Sampling Technique*, the authors introduced "the DDR memory technologies utilize source-synchronous

double data rate techniques to achieve higher data bandwidth. A group of data bits is sent over parallel wires along with a strobe. The timing of the strobe with respect to the data is very critical and differs for read and write operations. During the write operation the edges of the strobe have to be center-aligned with the burst of data bits to maximize timing margin; during read cycles, edges of the strobe are edge-aligned with the data signal. Thus it leaves the responsibility of adjusting the phase of the strobe to capture the data correctly to the receiving logic. Conventional DDR memories use DLL or PLL to correct the phase of strobe relative to data signal and launch the data and strobe with balanced timing" (Bhatti, Denneau, Draper, 2007).

DLL & PLL

Delay Lock Loop (DLL) and Phase Lock Loop (PLL) is widely used in the field of integrated circuit. They have many similarities. As described by Rambus article, "both of them can be used to maintain a fixed timing relationship between signals in environments where PVT variations cause these relationships to change over time. DLL and PLL work by continuously comparing the relationship between two signals and providing feedback to adjust and maintain a fixed relationship between them" (Rambus, 2014).

DLL is used to detect and lock the timing difference between clock signal between output data signal. DLL mainly consists of phase detector, charge pump, loop filter, and voltage-control delay line (VCDL), as shown in Fig. 4. Phase detector is

one of the most significant parts of DLL, which detects phase differences between the clock and

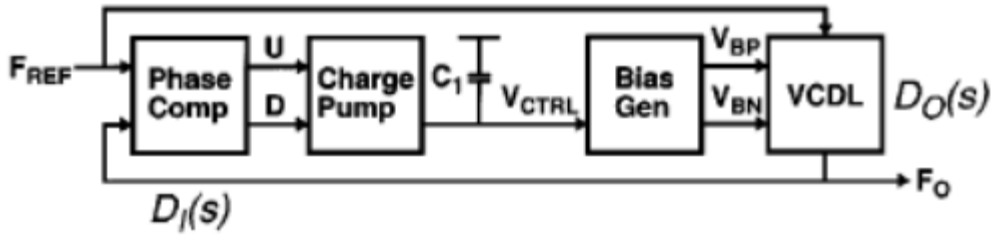

Fig. 4 Block diagram of Delay Locked Loop (DLL)

data signal. The phase detector detects this phase difference, and transfer the information through a low pass filter to a delay line that outputs the delayed clock signal to maintain the desired timing relationship (Rambus, 2014). It was said in Rambus article that "one of the difficulties of maintaining phase relationships between these two signals is that the loop which provides feedback to the phase detector must account for the timing characteristics of the output logic and output driver. This is important, as it estimates the phase differences between the clock and the data being driven by the output driver. In order to accomplish this, circuits that mimic the behavioral characteristics of the output logic and output driver are inserted into the feedback loop to model timing delays and changes in behavior as PVT variations. Maintaining the timing relationships between the clock and output data in this manner with DLL and PLL results in improved timing margins, as shown in Figure 4, and addresses an important limitation to increasing signaling speeds" (Rambus, 2014).

PLL is similar to DLL, but have different elements. PLL does not have VCDL, but a voltage-controlled oscillator (VCO). PLL can also be used to divide-down or

multiply-up an external system clock frequency for use in other parts of a chip (Rambus, 2014). PLL is always used to provide a slower clock frequency to the core of a dynamic memory, while the interface operates at a higher clock frequency (Rambus, 2014).

Methodology and Materials

DLL vs PLL

Although DLL and PLL implements similar functions, DLL is preferred in our project. Jitter is the deviation from true periodicity of a presumed periodic signal (Wikipedia, 2015). The VCO of PLL accumulates jitter, but jitter does not accumulate as much in a DLL delay line. In DLL, the reference clock jitter and the noise induced by power supply noise or substrate noise disappear at the end of the delay line (Chung, Chen, Lee, 2006). The reason is that a jitter event simply gets transferred to the output of the delay line of DLL once and forgotten, unlike being re-circulated in a VCO. Furthermore, the order of the DLL is generally equal to the loop-filter order, which is often one. DLL stability and settling issues are more relaxed relative to PLL. Based on all these differences between DLL and PLL, DLL is more preferred in our project.

Methodology

The input clock signal of PHY is a one-phase clock signal coming from controller. In order to get two-phase clock signal, an inverter can be used to produce the reverse phase signal. However, the inverter would bring delay, which makes phase difference

between original phase and reverse phase. 2-3 splitter, as shown in Fig. 5, was used to solve this issue. It produced different delay on both lines, so finally I got a matched two-phase signal as shown in Fig. 6.

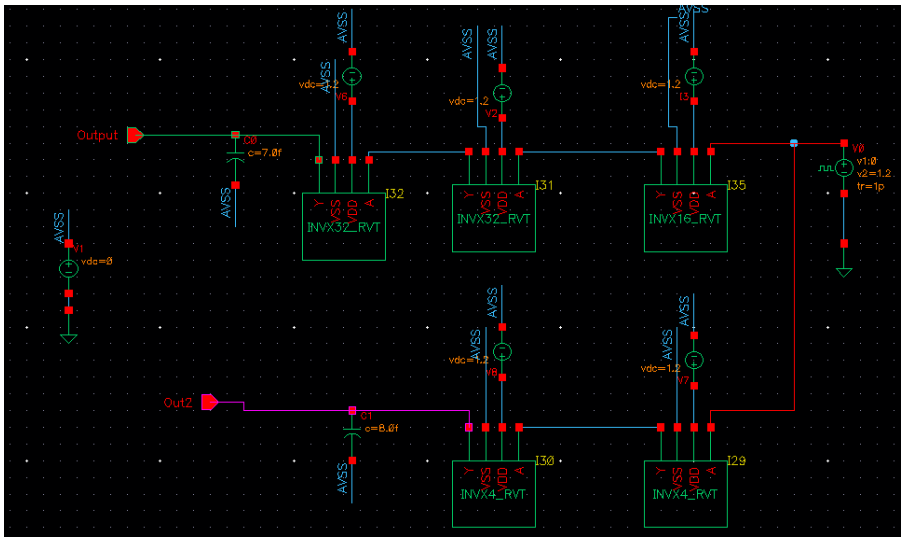

Fig. 5 Schematic of 2-3 splitter

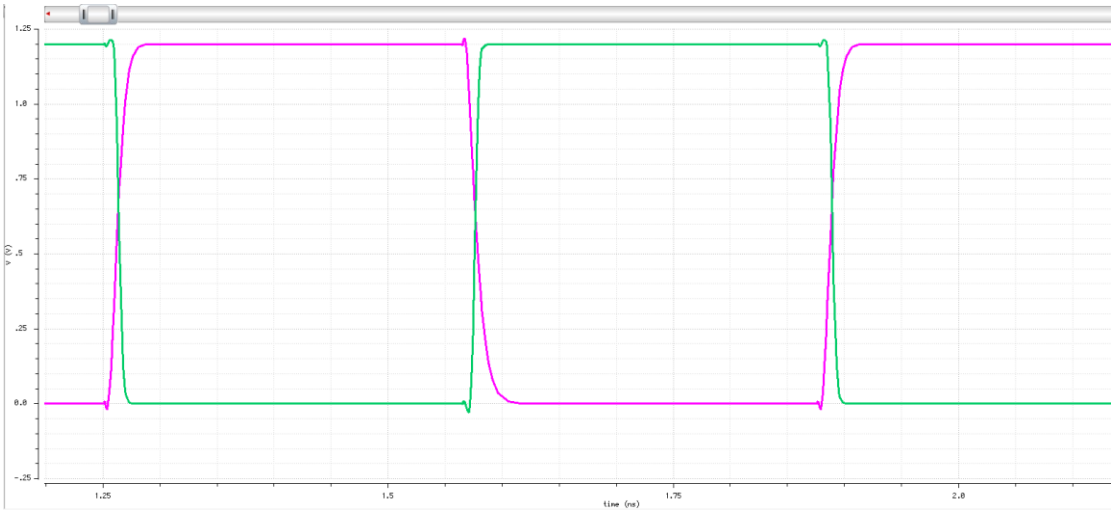

Fig. 6 Waveform of two-phase clock signal generated by 2-3 splitter

Delay line is the most important part in DLL. It consists of a group of inverters. Total delay of delay line should be a whole clock period, and each delay cell has identical delay. The block diagram is shown in Fig. 7. This delay line generated 8 different delayed clock signals. With select signals (S0, S1, and S2), multiplexer can

output one of these 8 delayed clock signal to other blocks.

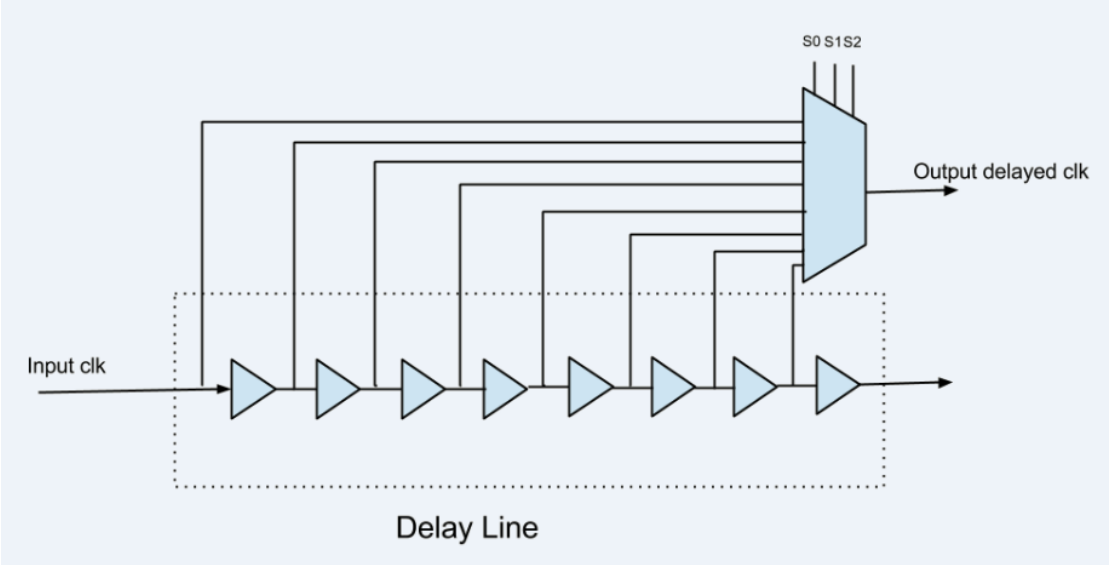

Fig. 7 Block diagram of delay line

In order to reach total one-period delay at the end of the delay line, the load should be carefully chosen. The delay of inverter chain is given by formula:

$$t_p = t_{inv}(\gamma + \frac{C_L}{C_{in}})$$

Result and Discussion

I designed the delay line using schematic in Cadence Virtuoso. The schematic of (part of) delay line is shown below in Fig. 8. In this figure, the wires annotated with "1eighth", "2eighth", "3eighth", and "half" stand for 1/8, 2/8, 3/8, and 1/2 of period delayed signals respectively.

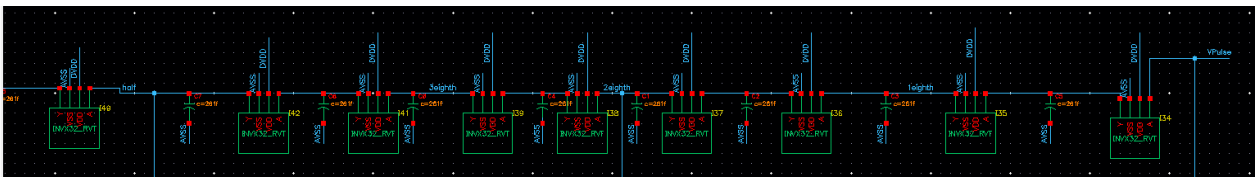

Fig. 8 Schematic of half of the delay line

The waveforms of 8 different delayed clock signals are shown in Fig. 9. For example, the red curve is reference clock signal, and white curve is one-period

delayed clock signal. From this result, it can be seen the middle point of rising edge of white curve overlapping with the middle point of rising edge of red curve. It implies that the total delay of delay line is exactly a whole clock period. Delay between adjacent curves is 1/8 of clock period. It suggests the design of delay line is successful. Finally, I combined the delay line with 2-3 splitter. With one-phase clock and select signals as inputs, the block outputs two-phase delayed clock signals for both transmitter and receiver side.

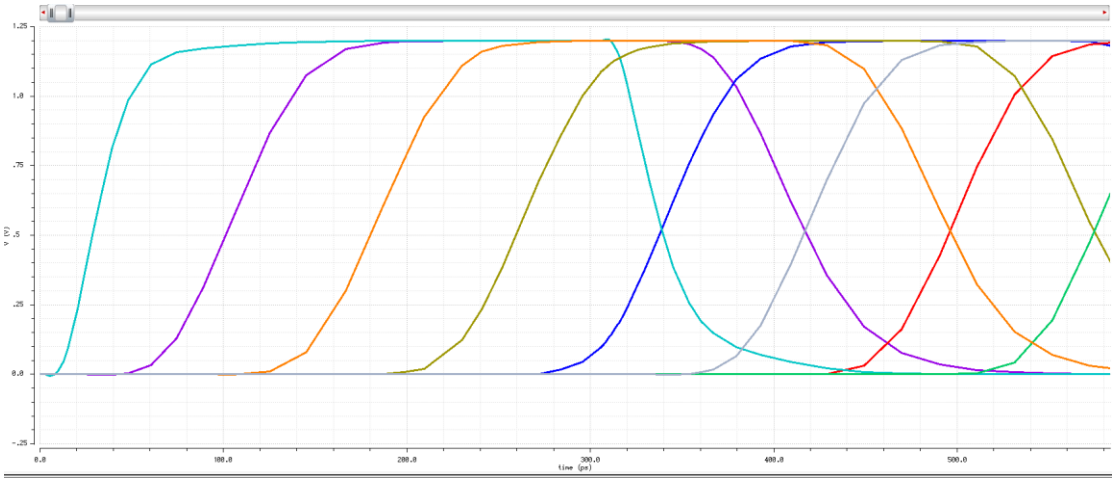

Fig. 9 Waveforms of all 8 delayed signals

In order to simplify the integration test, I only connected 4 signals (ref_clk, ¼ period delayed, ½ period delayed, and ¾ period delayed) to the multiplexer. The test bench schematic of the timing block is shown in Fig. 10. With input of select signals, it is able to transmit wanted two-phase delayed clock signals to transmitter and receiver side.

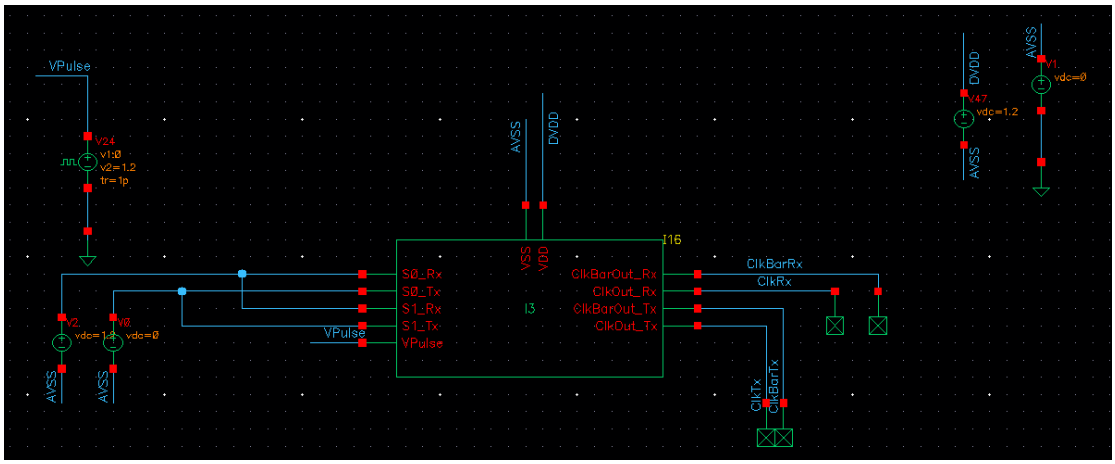

Fig. 10 Test bench schematic of timing block

For example, as test bench schematic showing, the input select signals are: S0_Rx = 1, S1_Rx =1, S0_Tx = 0, S1_Tx = 0. With select signal of "11", the output is 3/4 period delayed signal. With select signal of "00", the output is 0 delayed signal. The waveforms of output signals, ClkTx, ClkBarTx, ClkRx, and ClkBarRx, are shown in Fig. 11. In this figure, it can be clearly seen that ClkTx and ClkRx has 3/4 period phase difference.

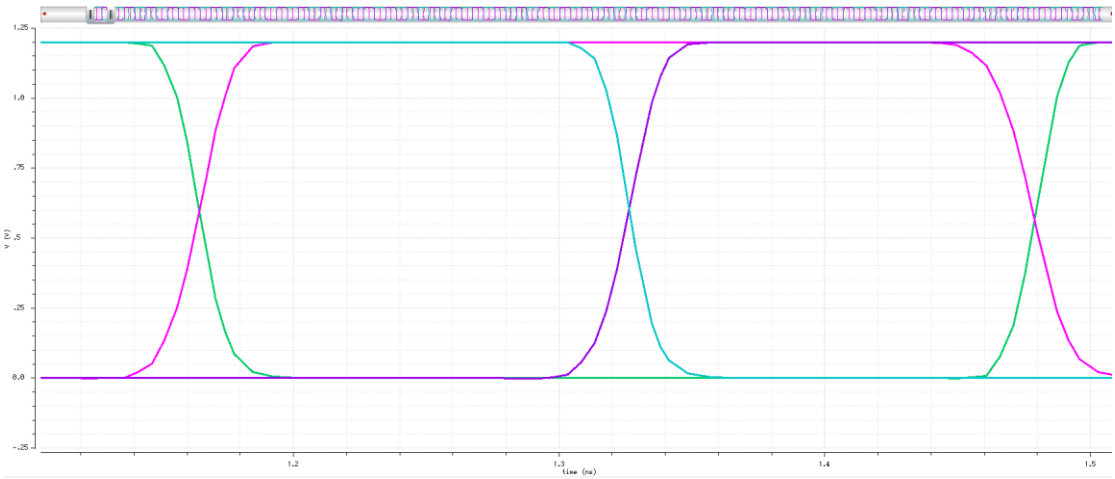

Fig. 11 Output waveform of test bench

The results imply the delay line is capable of generating expected delay signals. For PHY stage test, this delay line reaches expectation. However, if we want to integrate PHY with controller and memory in the future, the whole system of DLL

should be finished, including phase detector, up & down counter, and voltage-controlled delay line. The controller will send information about how much delay is needed for each block, and DLL can generate the required delayed signal.

Concluding Reflection

As described in above, the simple version of PHY integration test is successful. However, it is just a start point. Ideally, we should integrate controller, PHY and memory together. Controller will send information to make PHY read/write data from/to memory chip. In current stage, we didn't add controller and memory.

I got many insights from our whole year project. During our project, we kept changing our scope and direction of project. We wasted lots of time on broad literature review. For future research, I will recommend finish the general literature research and nail the direction as early as possible. Afterwards, we can choose a narrow field and focus on it to search more detailed reference materials. After that, we can start practical design with background knowledge in early period. It may give us longer period on design, test, and verification.

Looking back to our whole year project, we started with low efficiency, but finished with enthusiasm. We did lots of literature research in team and individually. Although final outcome is not any outstanding breakthrough, we finished the simple version of PHY with our own understanding. We tried to think big, but solve small problem.

Bibliography

Bach, Matt

Aug. 2014, Tech Primer: DDR4 RAM, http://www.pugetsystems.com/labs/articles/Tech-Primer-DDR4-RAM-589/ accessed March 1st ,2015.

Feng, J., et al.

2013, System level signal and power integrity analysis for 3200 Mbps DDR4 interface. Electronic Components and Technology Conference (ECTC), 63rd IEEE 1081:1086., accessed February 16, 2015.

Hulkower, Billy

August 2013 2013 Laptops and Desktop Computers - US. http://academic.mintel.com/display/637710/, accessed March 1st ,2015.

IBISWorld

2015 IBISWorld Industry Report Global Semiconductor & Electronic Parts. http://clients1.ibisworld.com/reports/gl/industry/default.aspx?entid=960, accessed March 1st ,2015.

Panek, Tracey

2015, Samsung Electronics Co., Ltd. http://subscriber.hoovers.com/H/company360/overview.html?companyId=55277000000000, accessed March 1st ,2015.

Patrizio, Andy

24 Jun. 2014,All about DDR4, the next-gen memory coming soon for PCs and mobile devices. PCWorld, http://www.pcworld.com/article/2365823/next-gen-memory-is-coming-fast-here-s-what-you-need-to-know-about-ddr4.html,, accessed March 1st ,2015.

Porter, Michael E., and Victor E. Millar.

1985, How Information Gives You Competitive Advantage, accessed March 1st ,2015.

Porter, Michael

2008, The Five Competitive Forces That Shape Strategy, accessed February 16th, 2015.

Sohn, K., et al.

2013, A 1.2 V 30 nm 3.2 Gb/s/pin 4 Gb DDR4 SDRAM with dual-error detection and PVT-tolerant data-fetch scheme. IEEE Journal of Solid State Circuits 48(1):168-177. , accessed February 16, 2015.

Transparency Market Research

May 2014, Next Generation Memory Technologies Market (Non-volatile and Volatile) - Global Industry Analysis, Market Size, Growth, Trends and Forecast 2013 - 2019, accessed October 14th, 2014

Ulama, Darryle

2014 IBISWorld Industry Report 33441a: Semiconductor & Circuit Manufacturing in the US. http://www.ibis.com, accessed February 16, 2015.

Hall, Bronwyn

January 2007 An Empirical Analysis of Patent Litigation in the Semiconductor Industry, accessed March 1st, 2015.

Quinn, Gene

October 30 2009 The Risk of Not Immediately Filing a Patent Application, accessed March 1st, 2015.

Quinn, Gene

January 2011, The cost of obtaining a patent in the US, accessed February 28th, 2015

Shane, Olafson

April 2007 Patents vs. Trade Secrets-Giving Your Business the Competitive Edge, accessed March 1st, 2015.

Stim, Rich

March 2011, The usual costs of filing a patent application, accessed February 28th, 2015

U.S. Patent and Trademark Office

October 2013 Description of Patent Types, http://www.uspto.gov/web/offices/ac/ido/oeip/taf/patdesc.htm, accessed March 1st, 2015

World Intellectual Property Organization

May 26 1989 Treaty on Intellectual Property in Respect of Integrated Circuits, http://www.wipo.int/treaties/en/text.jsp?file_id=295136, accessed March 1st, 2015

Dillon, Kyle; Liu, Sinan; Saxena, Kalika; Veryanskiy, Miron; Xu, Chenyang

Feb. 2015, Next Generation Memory Controller Capstone Strategy Project Paper, accessed March 16th, 2015

Allan, Graham

Sep. 2012, DDR4 Bank Groups in Embedded Applications

http://www.synopsys.com/Company/Publications/DWTB/Pages/dwtb-ddr4-bank-groups-2013Q2.aspx, accessed March 14th, 2015.

JEDEC Solid State Technology Association

Sep. 2012, JEDEC Standard: DDR4 SDRAM (JESD79-4), accessed March 14th, 2015.

JSSCC

A 1.2 V 30nm 3.2 Gb/s/pin 4 Gb DDR4 SDRAM With Dual-Error Detection and PVT-Tolerant Data-Fetch Scheme, accessed March 15th, 2015

Rambus

Feb. 2014, DLL/PLL on a DRAM

http://www.rambus.com/documentation/inventions/item/759-dll-pll-on-a-dram accessed March 14th, 2015.

Nagaraju, Manohar

Design of process invariant Delay Lock Loop (DLL) ECE 6770 –Final Report, http://www.kdstevens.com/stevens/6770/reports/08-dll-hardening.pdf

accessed March 14th, 2015

Chung, Ching-Che; Chen, Pao-Lung; Lee, Chen-Yi

An All-Digital Delay-Locked Loop for DDR SDRAM Controller, accessed March 15th, 2015

Bhatti, Rashed; Denneau, Monty; Draper, Jeff

Data Strobe Timing of DDR2 using a Statistical Random Sampling Technique, accessed on March 14th, 2015.