Journal of Multidisciplinary Engineering Science Studies (JMESS)
ISSN: 2458-925X
Vol. 3 Issue 9, September - 2017

Proposed Architecture For Comparison Of DNA Computing With Silicon Computation

Halim Ullah
Faculty of Computer Science
Preston University Kohat
Islamabad Campus, Pakistan

Dr. Malik Sikander Hayat Khiyal
Professor Faculty of Computer Science
Preston University Kohat
Islamabad Campus, Pakistan
E-mail: m.sikandarhayat@yahoo.com

***Abstract*—The impact of Deoxyribonucleic Acid computation is increasing on the researcher due to fabulous potentials it contains, the storage capacity is massive, and parallelism in processing is incredible. This research is aimed at investigating cost of DNA based computing in contrast to the traditional approach. Different attributes of both techniques are compared and analysed. Comparison of Process, Storage, Cost and Power Consumption are made and Results are given. A model is designed for comparison of the mentioned attributes in both techniques. Energy consumption and production Cost is almost negligible. DNA based computing is remarkable for solving such combinatorial problems which are impossible to solve with traditional computing.**

Keywords—DNA Based Computing, Comparison of Traditional and DNA Computing, Future of Computation, DNA vs Silicon.

I. INTRODUCTION

Looking at the potential applications of Deoxyribonucleic Acid (DNA) in various fields of computation researchers explore it to store data in it because of DNA storage capacity, density, scalability, ideal for long term stability and archival of data.

DNA computing retains number of advantages over silicon based computation. Some leading advantages of DNA computing are given in [1].

1st, a process with DNA computing can be performed energy efficient in contrast to traditional computing technique. Efficiency is roughly recorded as one million times.

2nd, it has potential to store an astonishing quantity of material for a certain space.

3rd, the element constituents are too much in quantity and easily can found.

4th, DNA computing tolerates for enormously parallel computation. Most researchers and experts consider that it has capability of performing huge parallel processing [1].

Since the early ages the computation and calculations have been performed by human [2]. Changes in computational techniques occurred, from manual way of computation to mechanical technique and in modern ages to electronic technique. By definition Computation is a technique of getting input, perform some sort of manipulations store it and produce output. The size of it tremendously decreased and processing power increased. Computer contains silicon built integrated circuits [2].

For digital computer the requirement is that it has ability of using logical functions for arithmetic operations [3]. Adders are the logical circuits that perform addition in the central processing unit (CPU). The half adder adds only two binary digits because it has only the ability to add two bits, but it is a key for creating full adder, which can add three bits and can be cascaded to yield serial adders for adding multi bit integers. Therefore, to construct half adder for newly technique of computation is very hard. But molecular half adder logic gate have been construction and tested [3].

II. LITERATURE REVIEW

The computational skill of living thing has captivated researchers from last three or four decades. They are focusing on implementation of living things in computational environment [4]. Until the experiment of Adelman was not introduced there was rare focus on implementation process of computational codes in biological things. The fortunate thought was that it is only theoretical study and suggestions were all theoretical.

DNA is appealing media for data storage due to the very huge amounts of data storing capability. They vastly exceed the storage capacities of any conventional electronic, magnetic or optical media. A gram of DNA contains about one thousand and twenty one (1021) DNA bases, or about one hundred and eight (108) Tera-bytes (TB). Hence, a few grams of DNA may have the potential of storing all the data stored in the world [5].

Silicon is the most common semiconductor material used in transistors; it is economical and easily handles to process, but it has its limitations. Reducing the size of transistor for increasing the speed lead to other problems, the problems are; increased consumption of energy and huge deviation in the properties of transistor. In this study they put the idea to use carbon nanotube instead of silicon; it will result's in faster and small transistor. This nanotube is hollow cylindrical shaped molecule. The diameter of it is about a nanometer (nm) and width is 1/50, 000 of human hair [6].

The author in [7] calculated that if F be the fraction of sequential calculations and (1-F) is possibly parallelized fraction, then the maximum achievable speed by using P number of processors is (1/F+ (1-F)/P).

In [7] the author calculated that if 90% processing or calculation is parallelized and 10% remains sequential while 5 processors are used, it would roughly give you 3.6 maximum speed-ups. If processors increased to 10 then it will speed up the processing up to 5.3 and if the numbers of processors are increased to 20 then it will speed up calculations or processing to 6.9 which proves that doubling the hardware increase the speed only 30%. If 1000 processors are used for calculations it will speed up 9.9 that mean a limit will come that the speed will start decreasing [7].

To put the idea of DNA base computing in practice, the author in [8] selected a problem for computation that is Hamiltonian path problem. The author assumed a map of cities connected to each other with nonstop flights, for example "it is possible to travel directly from Boston to Detroit but not vice versa. The goal is to determine whether a path exists that will initiate at the starting city (Atlanta), finish at ending city (Detroit) and pass through each of the remaining cities exactly one time". Cities are given named in a manner that first name followed by last name of the desired city. "Every C is replaced by G, every G by C, every A by T and every T by A, for this particular problem only one Hamiltonian path exists, and it passes through Atlanta, Boston, Chicago and Detroit in that order. In computation the path is represented by GCAGTCGGACTGGGCTATGTCCGA, a DNA sequence of length 24" [8].

In [8] while solving the problem, the algorithm was applied as follows:

- Interconnect all the variables.
- Eliminate all paths except those which initiate from the beginner.
- Eliminate all paths that do not finish at the final variable.
- Exclude all paths which have length equal to the number of variables or exceed.
- Eliminate all paths that miss any variable.
- Check is whether any path remained left, is so then return "Yes" otherwise return "No" [8].

Scientists have stored audio and text on fragments of DNA and then retrieved them, a technique that may provide an approach to handle the vast data of the digital age. They encoded Martin Luther King speech "I HAVE A DREAM", a photograph and Shakespeare's 154 sonnets. Later they were able to retrieve them with 99.99 % accuracy [9].

Companies, governments and universities face a challenge of storing the flood of digitized information, videos, books, movies and songs over the internet, some experts explored the answer in biology. They have found ways to encode data in cells and bacteria but these biological elements eventually die. DNA is not alive, it could sit passively in a storage device for thousands of years, and it is dense, stable and durable. Researchers at Harvard University reported the encoding of an entire 54,000 word book in strand of DNA [9].

III. DESIGN AND ANALYSIS

To ensure that DNA based computing is suitable and efficient than the conventional computing according to cost, Speed, parallelism and density of storage. Hybrid research methodology means Quantitative and Qualitative approach will be used.

- Compare cost of DNA based computing with silicon based computing.
- Analyse parallelism in processing of both computational techniques.
- Mathematically prove density of storage capacity in both type of computation.
- Review the approach of both type of computation to hard problems.
- Educate students about the DNA based computing.
- Present the limitations of silicon based computing and their solutions given by DNA base computing.
- Review on the state of art in DNA based computing.

The limitation of unawareness about DNA based computing will be overcome by properly presenting the newly come-way of computation to the users of related field.

Mathematically comparison of cost, storage capacity, parallelism in processing and survey of achievements in DNA based computing will all help to properly analyse both computation.

The architecture shows that a hard combinatorial problem will pass through two different techniques of computation. One called Silicon Based Computing and the other is DNA Based Computing. Different attributes of them will compare and figure out the resultant data. Attributes like Process, Storage, Cost and Power Consumption are focused in this research see figure 1. The architecture is designed in thesis of authors [10]

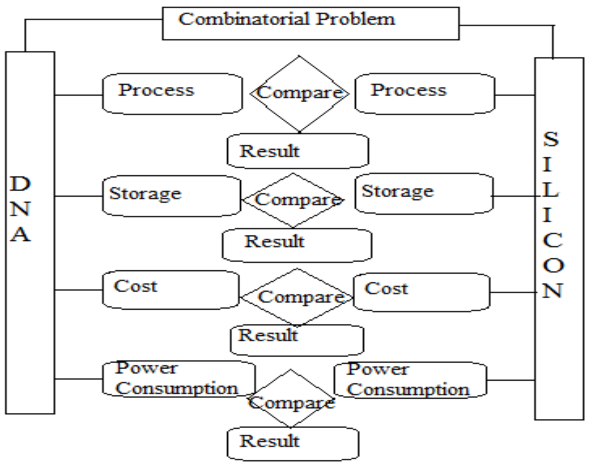

Fig 1: Architecture of Analyzing Hard Problem

A. A Process of Both Computing Techniques

For comparison of process, the process data is taken from the previous work and compare for both computational techniques. Their performances were analyzed carefully and results were made as given in [10].

DNA: In a second 10^{14} DNA flights number were simultaneously concatenated in only 1/50 of teaspoon of solution [8].The new computational technique tolerates exceptionally parallel processing. The philosophers made a conclusion that this newly technique has potential to perform at a level of ten power seventeen (10^{17}) operations per second (sec) or more, in other words a level that silicon-based computers will never ever be able to match [11].The author in [12] discusses that 330 trillion operations per second could perform easily with the DNA.

Silicon: The author in [7] calculated that if 90% processing or calculation is parallelized and 10% remains sequential while 5 processor are used, it increase the speed 3.6 times and the delay of 1.4 mean 28% occurs. If processors increased to 10 then its speed up the processing up to 5.3 and the delay of 47% notified, and if the numbers of processors are increased to 20 then it will speed up calculations or processing to 6.9 which proves that doubling the hardware increase the speed only 30%. If 1000 processors are used for calculations it will speed up 9.9 and the delay will also increase to 990.1 that is 99.1% delay. In conclusion; a limit will come that the speed will start decreasing.

B. Storing Data in Both Techniques

In [10] the designed diagram is for comparing the storage capabilities of both techniques, the comparison and analysis of the calculated data will be performed and results will be produced.

DNA: The Storage medium is Nucleic Acid and the potential of storage capacity in DNA is an attractive feature. Experiments have been made and successfully retrieved fifty seven to ninety nine (57-99) base pairs of data that are not property of bacterium. It is consider that a milliliter (ml) of liquid can have 10^9 bacteria. If assumption is made that an average of eighty (80) base pairs of data for each bacterium, using four (4) base pairs for each byte that can approximately give nineteen (19) Gigabytes (GB) of data for size of one (1)ml [1].

Above ten (10) trillion particles of DNA can adjust in space of less than one (1) cubic centimeter zero point zero six (0.06) cubic inches having density of one (1) bit per cubic nanometer. DNA computer can hold ten (10) terabytes data with the mentioned amount of DNA [2].

Data storing capability of DNA is its impressive potential; it has been noticed appealing media due to immense capability of storage. Ten power twenty one (10^{21}) DNA bases are found in a gram of DNA, or about ten power eight (10^8) Tera-bytes (TB). Hence, a few grams of DNA may have the potential of storing all the data stored in the world [5].

Silicon: The Storage medium is semiconductor and required ten power twelve ($10^{12)}$ cubic nanometers to store 1 bit [13].The current areal density notified for magnetic based storage stands at ten (10) GB per square inch (sq. inch). If it does consider that the thickness of hard disk platter is one (1) centimeter (cm) that gives one point five (1.5) GB/ml.

C. Cost comparison of Both Techniques

While comparing the cost; cost required for both techniques were taken and analyzed. The production cost and availing the facility, material availability is also focused; results were made and presented.

DNA: It was declared that DNA particles are required for newly born computational technique, its materials are available in every living thing, so until and unless there is cellular living thing it is available. This material is available everywhere and in enormous amount, it can obtain easily.

Cost is significant attribute for synthesis and sequencing of DNA. Synthesis cost is 0.05$ per base

[14]. With the increase in the amount of data to be stored in DNA there is negligible change in cost used to store unit MB associated with DNA based storage. Cost of DNA computing grows with the increasing of file size it using. The observation results that for long term archival using DNA as storage medium becomes moderate and economical.

Silicon: Non- Recurring and High Cost Investment over every new IC chip design will increase. As every new chip is more complex from the preceding one, therefore reducing size will not only add complexity to the design but it will also increase the overall cost per design from one chip to another [11].

Cost of silicon computing increases due to adaptation of newer versions of devices and also software. The upcoming device or software is presented in such a way that it is going to solve each problem, but in real it won't. Need of market for better, efficient devices increases with usage and interference of computing in daily life. The potential of silicon based computing for solving hard problems and storing large amount of data is not getting enhanced due to its limitations.

D. Energy Consumption of Both Techniques

To calculate power consumption of both techniques; attributes of power are gathered and compared in comparison module, and then results were produced. The power need for processing and storing data is analyzed carefully.

DNA: It was observed during the research that DNA computing can operates adjacent to thermal equilibrium, really reducing energy dissipation. It can perform about 2 x 10^{19} (irreversible) operations per joule [11].

Silicon: Supercomputers of current era are not very energy efficient, they executes power of 10^9 operations per joule [11]. An exclusive hurdle for advancement in silicon-based chips includes consumption power and heat dissipation. Power density has been grown with the rate of rate $S^{0.7}$ since couple of years for all generations. Size of transistors are becoming pretty small and that small transistors consume small amount of power, but Integrated Circuit (IC) chip becoming denser and denser due to the number of transistors are being placed on it, hence it consumes enormous voltage to run all transistors and generates massive heat.

In year 1971, the world's first single chip microprocessor was introduced by Intel, that first Intel microprocessor contained twenty three hundreds (2,300) Transistors at ten (10) μm, which was using 1/10th of watt. Transistors were increased for speed up and one of modern processor, a 3.2 GHz Pentium IV extra edition consumes one hundred and thirty five (135) watts.

In dual core 2.8GHz Pentium D the numbers of transistors were increased to 167 Million and notified two hundred and forty four (244) watts power which it consumes [11].

IV. RESULTS:

The results are made on the bases of designed architecture. Comparison diagrams for Process, Storage, Cost and Power Consumption are designed and results are made after comparison of both computational techniques.

A. Process:

After comparison of different modules of both traditional and DNA based computing the analysis are made. Results are remarkably attractive. Today's super computer performs operations maximum up to thirty five point eight (35.8) teraflops, which really high speed of operations occurrence, but the comparative operations of DNA is found hundred (100) teraflops [10]. This is huge number, hundred (100) teraflops means hundred (100) trillions floating point operations are operated in a sec see figure 2.

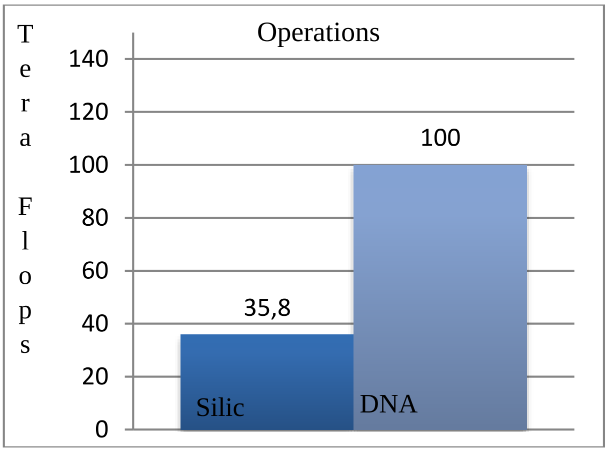

Fig 2: Operation Comparison Result

DNA based computing performs 0.1 PetaFlops operations per sec and in comparison the Silicon based computing only performs 0.0358PetaFlops operations per sec see figure 3. This is a huge difference in the performance of both techniques for computation.

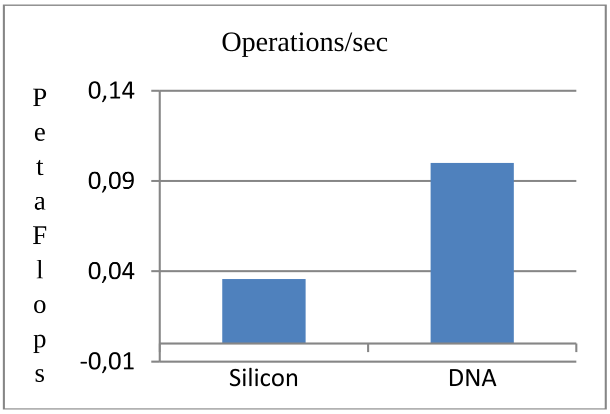

Fig 3: Performance Comparison

Performing combinatorial problem with traditional computer setups sequential search tree, generates each branch sequentially and save the shortest one. Performing for 20 cities by hundred (100) million instructions/sec and also requires forty five (45) million GB memory, mathematically it takes 2 years to generate the paths. While with DNA it looks possible due to 10^{15} is Nano mole of material. All operations are going to be performed in parallel.

Parallelism of DNA is really massive, one thousand and eighteen (1018) strands of DNA gives ten thousands time greater operations then today's super computer. The parallelism of silicon based computing were taken 90% and 10% process as sequential; the calculated speed is shown in figure 4.

The delay noticed in 90% parallel and 10% sequential processing is shown in figure 5. It proves that increasing number of processor on chip is not satisfactory any more.

Suppose if parallel processing increase to 95% and only 5% processing remains sequential then possible speed up with 5 to 2000 processors is shown in figure 6.

Only 5% processing remains sequential and 95% is parallel then the delay was noticed that a limit will reach where it does not increase the speed instead it will start decreasing the speed see figure 7.

The delay is seen increasing enormously, due to this increase in delay; the number of processors are not going to increase any more. In comparison the parallelism in DNA computing increases exponentially.

B. Storage:

In DNA of bacteria the replications occurs very fast that is equal to thousand (1000) bits/sec data storage rate, it looks handsome number but as compared to traditional computing it is very slow. This rate is significantly higher of silicon based computing, but after the 1st replication it increases exponentially in DNA that is 2^n. Number of strands increases with the replication and the bit/ sec data rate also increases, operations executes in parallel. After thirty (30) iterations it reaches one thousands (1000 Gbits/sec) see figure 8. In this situation the today's fast hard drive is beyond to DNA storage rate [1].

Today's video tapes acquire 10^{12} cubic nanometers space to store one bit data, but DNA requires only one cubic Nano meter for one bit data, that mean DNA stores data 10^{12} times more data in equal space. Data of trillion CD's adjust only in one cubic centimeter of DNA [14].

In [15] the author made comparison of different researcher's views about storing data. DNA has density of eighteen (18 Mbits/ inch), and traditional computing storage stores 1/100,000 times less information in same space.

C. Cost:

The cost of DNA and traditional computing is analyzed and the result is that increasing number of transistors on chip attracts users for hoping betterment but actually it costs them, nothing else.

The production cost of traditional computing goes significantly above the DNA cost because as far as there is availability of cellular organism there is DNA. No need to plant an industry and produce it, enormous availability makes it an attractive source. To achieve DNA level storage and parallel processes, exponential increase in cost will be required for silicon computing industry.

D. Power consumption:

DNA computing is one billion times energy efficient as compared to silicon computation [16]. The building block of DNA is chemical bonding and the bonding happens in absence of power source. In DNA $2*10^{19}$ operations performed per joule, while in silicon based computing 10^9 operations occur per joule shown in figure 9.

Heat and power consumption is seen a hurdle for the future of silicon based computing. $S^{.7}$ is the grown rate of power consumption in it, 2300 transistors upgrades to 167 million transistors and power consumption increased from 1/10th of watt to 244 watt in dual core 2.8 GHz processor [1].

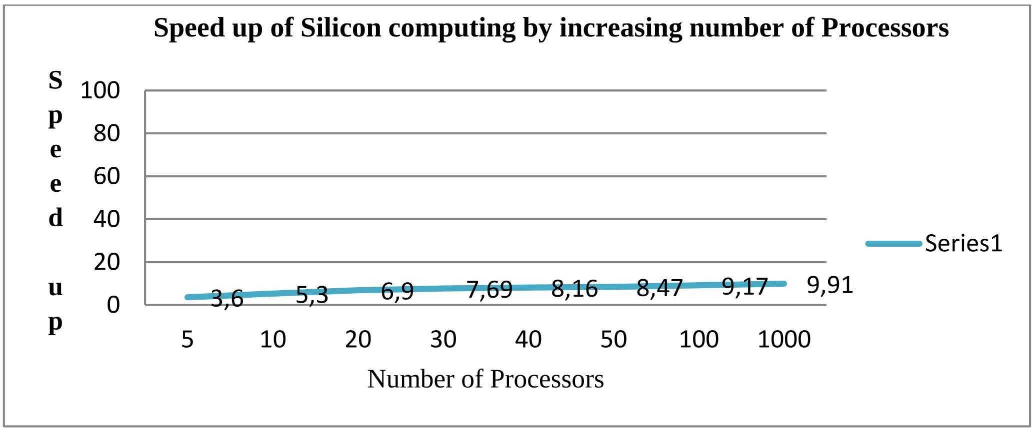

Fig 4: Speed Up

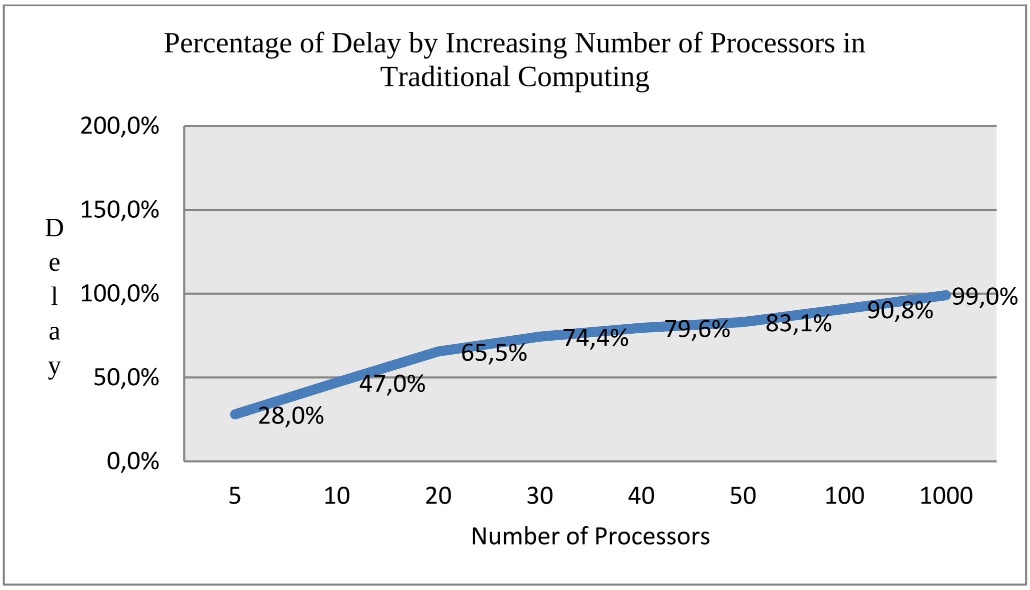

Fig 5: Delay

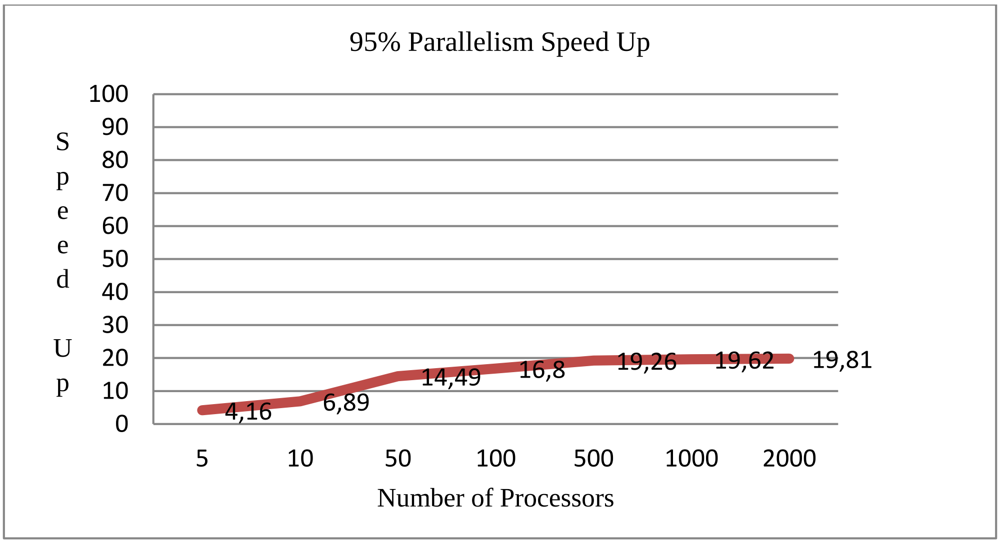

Fig 6: Speed Up with 95% Parallelism

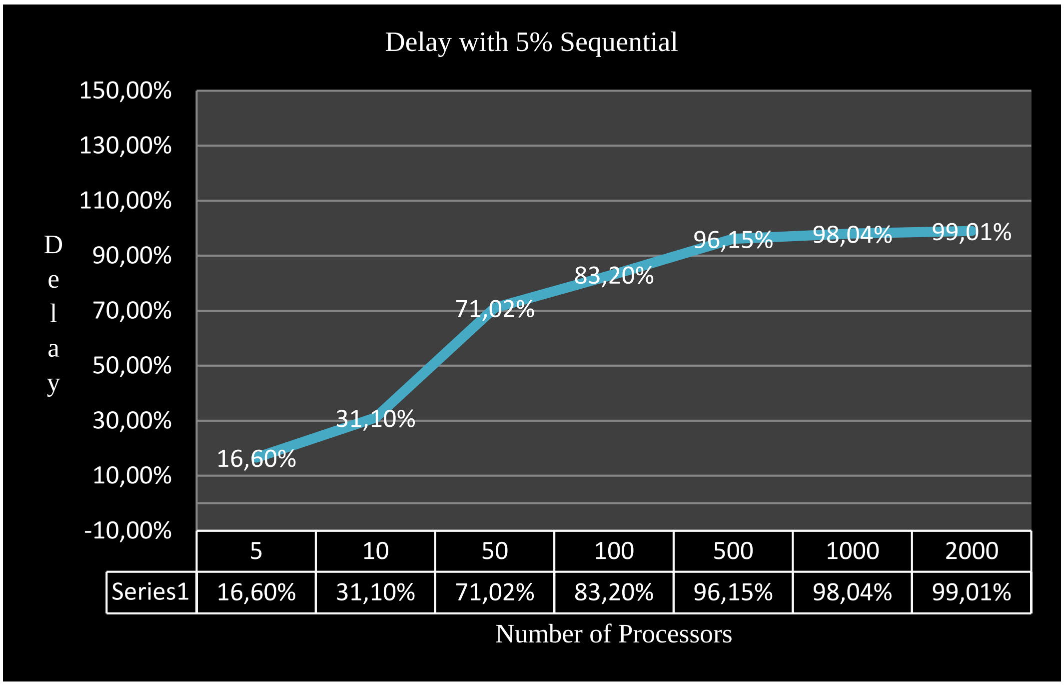

Fig 7: Delay with 5% Sequential Processing

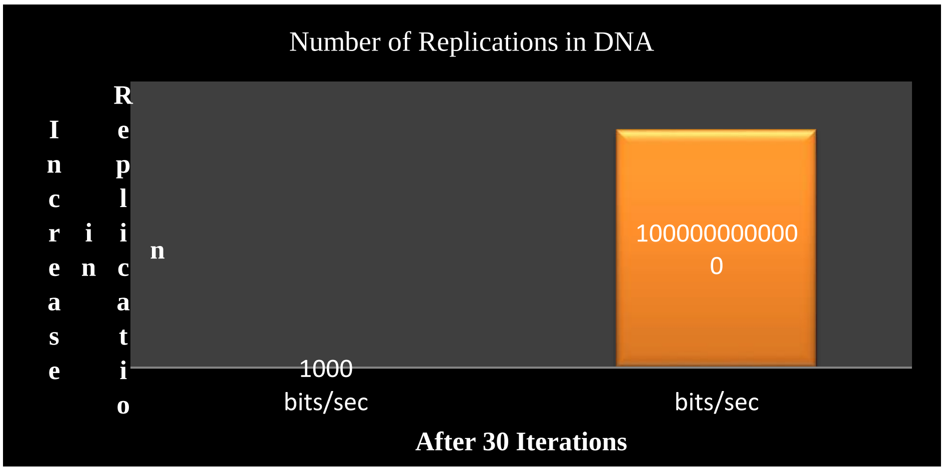

Fig 8: Exponential Increase in Replications

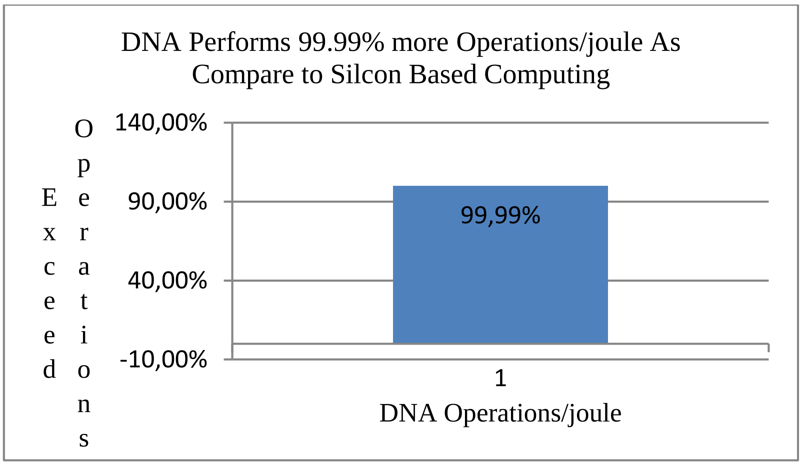

Fig 9: Number of Operations/Joule

CONCLUSION

It is concluded from the study that DNA is authentic medium for computation; it's getting famous round the world for the features it holds. The researcher works hard on it, the number of limitations in traditional computing is hopefully going to solve in

DNA computing. Especially the storage and processing are going to be enhanced to the level of today's global world need.

DNA based computing is suitable to use in all the cases compared in the study that is processing, storage, cost and energy consumption all are supporting this technique for future computation. There is still a lot to do for its maturity and research is newer stop process. There are certain limitations but will soon overcome and these limitations will be its advantages. This kind of computation is really beneficial for humanity in view of every aspect, either health care or social life. It helps human in each and every situation.

DNA based computing is the part of biological, chemical and computer study. It combines number of fields and researchers have attraction in it. It involves researchers of most dangerous diseases and they have solutions in it. Disease like cancer will hopefully diagnose and treated via this technique of computing.

The study is introducing the newly technique of computing in Pakistan to open doors of research for the student of Computer science, Information technology, Nano science and Biochemical. The study helps the organizations in selecting and acquiring of new technology according to their needs.

REFERENCES

[1] A. Maurya, A. Nair and S. Sanyal, "An Overview of the Evolutionary Trends in Molecular Computing using DNA", VeermataJijabai Technological Institute India, 2011.

[2] P. S. Sugathan, "DNA Computing", Department of Computer Science Cochin University of Science and Technology Cochin – 682022, 2010.

[3] M. N. Stojanovic and D. Stefanovic, "Deoxyribozyme-Based Half-Adder", Contribution from the Division of Clinical Pharmacology and Experimental Therapeutics, Department of Medicine, Columbia University, Box 84, 630 W 168th Street, New York, New York 10032, and Department of Computer Science, University of New Mexico, Albuquerque, New Mexico 87131, P: 6673-6676, 2003.

[4] R. Deaton, M. Garzon, J. Rose, D. R. F. chetti and S. E. Stevens, "DNA Computing: A Review", IOS Press, Fundamental Informatics 30, P: 23-41, 1997.

[5] A. Gehani, T. LaBean, and J. Reif, "DNA-Based Cryptography", Department of Computer Science, Duke University, Box 90129, Durham, NC 27708-0129, 2000.

[6] S. Kumar, G. Pant, V. Sharma and P. Bisht, "Nanotechnology in Computers", International Journal of Information & Computation Technology, ISSN 0974-2239 Volume 4, Number 15, P: 1597-1603, 2014.

[7] G. M. Amdahl, "Validity of the Single Processor Approach to Achieving Large Scale Computing Capabilities", AFIPS Conference Proceedings, P: 483–485, 1967.

[8] L. M. Adelman, "Molecular computation of solutions to combinatorial problems". Published by American Association for the advancement of Science, vol 266(11) P: 1021-1024, 1994.

[9] S. N Kumar, "A Proper Approach on DNA Based Computer". American Journal of Nano materials, 3(1), P: 1-14, 2015.

[10] M. S. H. Khiyal and H. Ullah, "Study of Processing, Storage and Cost of Deoxyribonucleic Acid Computing with Silicon Computation".LAP Lambert Academic Publishing, Omni Scriptum GmbH & co. KG Germany, 2nd Feb 2017.

[11] S. Hassan, Himaira and M. Asghar, "Limitation of Silicon Based Computation and Future Prospects", Second International Conference on Communication Software and Networks, 978-0-7695-3961-4/10, IEEE DOI 10.1109/ICCSN, 2010.

[12] S. Sharma, CSE-DNA-Computing-report, "Seminar Report on Working of Web Search Engine". Published in Studymafia.org, P: 1-25, Feb 2009 to May 2009.

[13] D. Nixon "GSEC Assignment Version 1.3 DNA and DNA Computing in Security Practices" Is the Future in Our Genes? GSEC Assignment Version 1.3 P: 1-14, 2002.

[14] V. Dhameliya, D. Limbachiya, M. Khakhar and M. K. Gupta, "On Optimal Family of Codes for DNA Storage", arxiv: 1501.07133v1 [cs.IT], 28 Jan, 2015.

[15] S. M. H. T. Yazdi, Y. Yuan, M. Jian, Z. Huimin and O. Milenkovic, "A Rewritable, Random-Access DNA Based Storage System", arxiv: 1505.02199 v1 [cs.IT], 8 May, 2015.

[16] A. Srivastava and J. Ramadas, "Analogy and Gesture for Mental Visualization of DNA Structure", Citation: Srivastava, Analogy and Gesture for Mental Visualization of DNA Structure. In Treagust, D.F. &Tsui, C.-Y. (Eds.), Multiple Representations in Biological Education, Dordrecht, the Netherlands: Springer. P: 311-329, 2013.