Wafer Scale Integration for Massively Parallel Memory-Based Reasoning*

Hiroaki Kitano and Moritoshi Yasunaga†
Center for Machine Translation
Carnegie Mellon University
Pittsburgh, PA 15213 U.S.A.
hiroaki@cs.cmu.edu, myasunag@cs.cmu.edu

Abstract

In this paper, we describe a design of wafer-scale integration for massively parallel memory-based reasoning (WSI-MBR). WSI-MBR attains about 2 million parallelism on a single 8 inch wafer using the state-of-the-art fabrication technologies. While WSI-MBR is specialized to memory-based reasoning, which is one of the mainstream approachs in massively parallel artificial intelligence research, the level of parallelism attained far surpasses any existing massively parallel hardware. Combination of memory array and analog weight computing circuits enable us to attain super high-density implementation with nanoseconds order inference time. Simulation results indicates that inherent robustness of the memory-based reasoning paradigm overcomes the possible precision degradation and fabrication defects in the wafer-scale integration. Also, the WSI-MBR provides a compact (desk-top size) massively parallel computing environment.

1. Introduction

The memory-based reasoning (MBR) is one of the mainstay approach in massively parallel artificial intelligence research [Kitano et. al., 1991] [Waltz, 1990]. The basic notion of MBR is to place memory as a foundation of intelligence. Instead of having abstract and piecewise rules to make inferences by chaining them, MBR simply stores large numbers of actual cases to carry out similarity-matching in order to make inferences from the most similar cases in the past. MBR has been succesfully applied to English pronounciation task [Stanfill and Waltz, 1986], census data classification[Creecy et. al., 1990], protein structure prediction[Zhang et. al., 1988], machine translation[Kitano and Higuchi, 1991][Sumita and Iida, 1991][Sato and Nagao, 1990], natural language understanding [Kitano and Higuchi, 1991] and other data-intensive domains.

MBR, however, requires massively parallel hardware such as the CM-2 connection machine [Thinking Machines Corporation, 1989]. The problems of current massively parallel machines are (1) economically expensive, (2) huge physical size, and (3) parallelism does not suffice for some large-scale applications. For example, the CM-2 is a multi-million dollar machine requires installment in the machine room and attains only 64K parallelism (VPR = 1) with 1-bit processing elements. Given the fact that many serious MBR applications (national statistics, crime investigation, tax records, virtual reality, etc) require more than a few million parallelism with reasonably low economic expense, current massively parallel machines do not suffice for further penetration of the MBR and other massively parallel AI approachs for real-world applications.

The solution which we will offer in this paper is to develop a wafer-scale integration for memory-based reasoning (WSI-MBR). The design and performance simulation indicates that we can attain about 2 million parallelism on a single 8 inch wafer with 0.3 micro fabrication technologies which would be available for commercial production around year 1995. WSI-MBR will provide nearly 200 million parallelism when the wafer stack cluster method was established, which is the state-of-the-art VLSI fabrication and assembly technologies[McDonald et. al., 1991]. The inference speed is on the order of nanoseconds by the use of a hybrid analog/digital computing circuits. The implication of this technology is significant. It means that we will be able to built desk-top or even lap-top massively parallel MBR systems.

The WSI-MBR is a run-time component. All data and weight are pre-computed on massively parallel machines such as the Connection Machine. Computed weights and data are loaded onto WSI-MBR to perform inferencing. In the large-scale memory-based reasoning system, the content of the memory-base is expected to be fairly stable so that up-date of weights and data will not be necessary for a certain duration of the operation. Separation of the main computing system and delivery (or run-time) system is justified even from a commercial point of view.

*This research was carried out as a program of the International Consortium for Massively Parallel Advanced Computing Technologies (IMPACT).

†This author is on leave from Hitachi Central Research Laboratory.

Since massively parallel machines would be quite large and expensive, development of run-time systems with extremely cheap and compact size would substantially push down cost-effectiveness trade-off point.

2. Memory-Based Reasoning Paradigm

Memory-Based Reasoning is a reasoning method based on a large set of examples. Stanfill and Waltz wrote that [Stanfill and Waltz, 1986]: *We consider the phenomenon of reasoning from memories of specific episodes, however, to be the foundation of an intelligent system, rather than an adjunct to some other reasoning method.* This approach counters the traditional AI paradigm which places rules as the foundation of intelligence. It is not the scope of this paper to discuss benefits and limitations of the MBR itself. However, it should be noted that there are some sucessful reports and some commercial systems that have already been deployed using the MBR paradigm (English word pronounciation [Stanfill and Waltz, 1986], Prediction of protein structure [Zhang et. al., 1988], the Census classification [Creecy et. al., 1990], and others.)

The MBR method requires a large set of cases. Each case consists of a set of features and a goal. Features represent problems need to be solved, and a goal represents a solution in the case. In the MBRtalk example, each case has seven fields for characters of the word, and goal fields which represents a correct pronounciation and stress (Table 1).

Statistical approach is used to determine significance of each feature in making correct reasoning. While relatively simple and homogenious operations are performed on each datum, the memory-based reasoning is particularly suitable for SIMD-type massively parallel machines such as the Connection Machine.

A similarity measure will determine the weight in which each feature affects the result of reasoning. In case of the MBRtalk task, the measure is computed based on the following equations:

$$w_f^g(D,\tau.f) = \sqrt{\sum_{v\in V_g}\left(\frac{|D[f=\tau.f][g=v]|}{|D[f=\tau.f]|}\right)^2}$$

$$d_f^g(D,\tau.f,\rho.f) = \sum_{v\in V_g}\left(\frac{|D[f=\tau.f][g=v]|}{|D[f=\tau.f]|} - \frac{|D[f=\rho.f][g=v]|}{|D[f=\rho.f]|}\right)^2$$

w_f^g is a weight of feature f on the field g. d_f^g is a value difference metric. Such metric differ in each task, but they are similar enough to be hardwired in VLSI chips. Please refer [Stanfill and Waltz, 1986] and other papers for details of the MBR approach.

Since benefits of the MBR over other reasoning paradigms, such as rule-based systems and neural networks, have been discussed in [Stanfill and Waltz, 1986], we simply point out merits which motivated us for the development of the WSI-MBR. One of the computational advantages of the MBR is that it allows data-parallel computing because the similarity of each record can be computed independently from other records. It is also a very attractive scheme for direct hardware implementation due to its simplicity and potential robustness against noise and faults.

3. Wafer-Scale Integration

Wafer-Scale Integration (WSI) is the state-of-the-art VLSI fabrication technology (See *Proc. of International Conference on Wafer Scale Integration* for recent progress in WSI. Also, [Cavill et. al., 1991] provides a good overview of the area.), and has been applied to various domains such as neural networks [Yasunaga et. al., 1991]. It fabricates one large VLSI-based system on a wafer as opposed to conventional VLSI production which fabricates over 100 chips from one wafer. The advantage of WSI is in its size (high integration level), performance, cost, and reliability:

Size: WSI is compact because nearly all circuits necessary for the system are fabricated on a single wafer.

Performance: WSI has substantial performance advantage because it minimizes wiring length.

Cost: WSI is cost effective because it minimize expensive assembly line.

Reliability: WSI is reliable because it eliminates the ponding process which is the major cause of circuit malfunctions.

However, there is one big problem in WSI fabrication: defects. In the conventional VLSI fabrication, one wafer consists of over 100 chips. Generally, we have certain percentages of defective chips. Traditionally, chips with defects have been simply discarded and the chips without defects have been used. To estimate the faults in the chip, we use the Seeds model[Seeds, 1967]:

$$Y = e^{-\sqrt{DA}}$$

where Y is a yield of the wafer, D is the fault density which is, in the fabrication process we are going to use, about 1 fault per cm^2, and A is the chip area. This is a reasonable rate for the current fabrication process. However, even this level of fault would cause fatal problems for such an attempt to build an entire IBM 370 on one wafer. Unless sophisticated defects control mechanisms and robust circuits are used, a single defect collaspes an entire operation. But, redundant circuits diminish the benefits of the WSI. This trade-off has not been solved.

In the WSI-MBR, we take a radically different approach. We accept a certain rate of defects. Rather

Fields	Input Fields							Output (Goal)	
Rec. No.	n-3	n-2	n-1	n	n+1	n+2	n+3	Pron.	Stress
00001	-	-	-	f	i	l	e	f	+
00002	-	-	f	i	l	e	-	A	1

Table 1: A part of the memory-base for MBRtalk task

than trying to control defects, we propose to use WSI for implementing a more robust computing mechanism so that a small amount of defects does not seriously affect the overall operation of the chip. We believe that MBR is ideal for this solution because it does not rely upon any single data unit. The essence of the MBR is a bulk data set which gives stable reasoning capability.

4. WSI-MBR

WSI-MBR is a digital/analog hybrid WSI specialized for memory-based reasoning. We decide to employ a digital/analog hybrid approach in order to increase parallelism and performance.

First, in the digital computing circuit, a floating point processor part takes up most of chip area. On the other hand, the analog circuit requires only a fraction of area for implementation of equivalent floating point operation circuits. The digital approach has an advantage in its flexibility since various sequences of floating point operations can be programmed. The analog approach in inflexible since it hardwires a computing sequence. However, in the WSI-MBR, the sequence of computation is already well-defined so that no re-programming is required (See [Mead, 1989] for one other use of analog VLSI for neural networks). Use of a less area-demanding analog approach provides two major advantages over the digital approach: (1) increased parallelism, and (2) speed up due to relaxed wiring constraints (critical paths and wire width).

Second, the analog circuits provide a drastic speed up over digital circuits. It is well known that analog circuits have a natural advantage in speed of computation. This is also true in our model as we will discuss it later.

Figure 1 shows a wafer floor layout of the WSI-MBR. Figure 2 shows a chip selection circuit. One wafer contains 56 memory chips, 7 data-bus chips (denoted as B), and a serial/parallel converter chips (denoted as C). This layout is for a five inch wafer. The chip selection circuits consists of signal lines which identifies and selects a chip on the wafer which the data should be send and accessed. In the figure, CL denotes the chip location signal (3 bits column, 3 bits row, 1 bit left/right selection) which identifies a chip on the wafer. L/R, CC, and RC are address of the chip to be selected.

Each memory chip consists of 4K MBR memory cells. One MBR memory cell loads one record and goal data, and carries out MBR operations. Figure 3 shows a

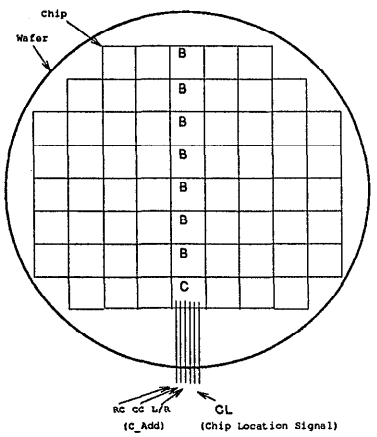

Figure 1: Wafer Floor Layout

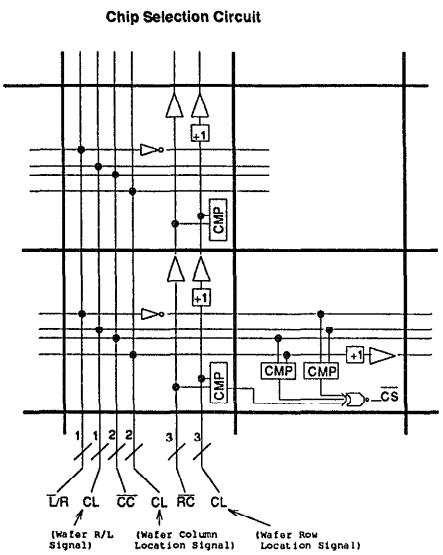

Figure 2: Chip Selection Circuit

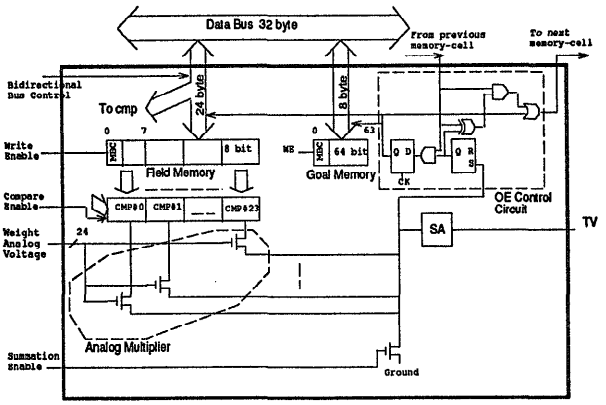

Figure 3: Schematic of the Memory Cell of WSI-MBR

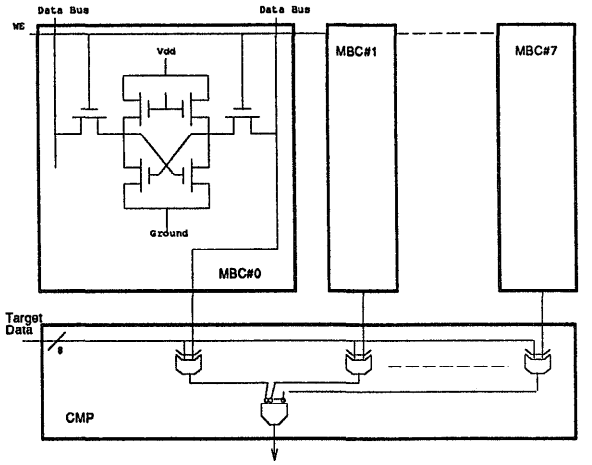

Figure 4: Schematic of Memory Bit Cells and Comparator

schematic of the MBR memory cell. In this design, the cell has 256 bits of digital memory (24 8-bits fields and a 64-bits goal memory), comparator, analog multipliers, a sense amplifier, and an Output Enable control circuits.

The input data will be broadcast through the data-bus to all MBR memory cells. In each MBR memory cell, each field value of the input data is compared with the value of the corresponding field. This comparison is done at each comparator CMP in parallel. Figure 4 shows a schematic of memory bit cells and a comparator. When the value is the same, CMP holds a high voltage output which will be multipled by a corresponding weight using one transistor multiplier. Results of each multiplication are added and produces a similarity value of the record (called the score voltage SV) on the MBR memory cell.

Now we want to take out the value of the goal memory of similar MBR memory cells. We use a threshold relaxing technique to retrieve the goal value in the order of the similarity rank. Initially a high voltage is given to the threshold voltage line (denoted as TV in Figure 3). This voltage of the TV line is compared with the score voltage using a sense amplifier (denoted as SA in Figure 3). The host computer, or the controller, controls the TV level through a digital-analog converter. The controller decreases the TV level until any of the MBR memory cell has higher score voltage. When the score voltage is higher than the threshold voltage, then a S/R latch is set which indicate that the goal memory of the MBR memory cell should be retrieved. The Output Enable control circuit (OE control circuit) runs serial throughout the single chip to allow asynchrous scanning. This means that a single fault in any part of the OE control circuit hampers an entire chip function (but, not an entire wafer). In order to improve the reliability, the OE Control circuit is built as triple redundancy circuit. Thus, unless all three OE control circuits in the same memory-cell cause faults (this would be an extremely low probability), the chip function would be maintained. Therefore, the WSI-MBR carries out asynchronous data fetching to attain fast data retrieval, but maintains high reliability.

One MBR memory cell uses about 4,000 transistors: 1,024 for memory bit cells, 1,920 for comparators, 16 for analog multipliers, and about 1,000 for the output enable controllers and other circuits such as the bus driver.

5. Performance Simulation

We have developed a simulator for WSI-MBR. The simulator is written in C language and run on UNIX-based workstation and on the CM-2 connection machine. The simulator can model factors critical to WSI designing such as (1) noise-levels with various noise distributions, (2) defects, and (3) run-time faults.

In this paper, we use Nettalk (MBRtalk in this case) task as an example of the performance data we have obtained. Although we have investigated the performance of WSI-MBR with several other tasks, we use MBRtalk for this paper because it is one of the most widely understood tasks. The MBRtalk system was implemented on the WSI-MBR simulator. Experiments were carried out for the MBRtalk system with the memory-base size of 908 (6506), 5908 (43166), 10908 (80394), and 19908 (146940) words (records). Also, we used a full set Nettalk data.

5.1. Parallelism

WSI-MBR attains a strikingly high level of parallelism. When the WSI-MBR design in the previous section is implemented on the 5-inch wafer using 0.5 μ CMOS fabrication technology, with the standard yield rate, the WSI-MBR provides 240K records each of which carry out completely data-parallel MBR operation. This is attained by the high density fabrication technology which can implement 16M transistors in the memory cell area, and by the compact design of the MBR memory cell which requires only 4,000 transistors per a cell. Tech-

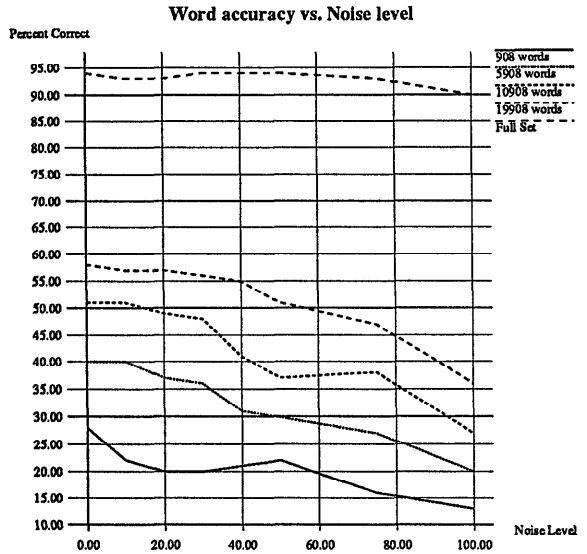

Figure 5: Word accuracy and noise

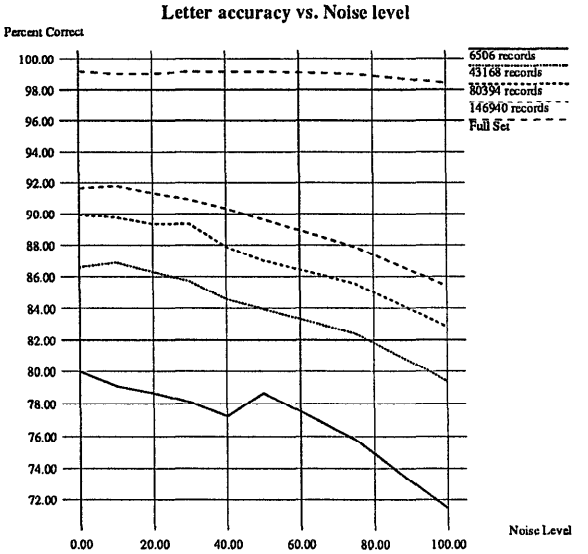

Figure 6: Letter accuracy and noise

nologies assumed in this design (5-inch wafer and 0.5 μ CMOS fabrication) are already available in selected chip producers. Within 4 to 5 years, the fabrication technologies will provide for us 8 inch wafers and 0.3 μm fabrication technologies[IEDM-91, 1991]. The WSI-MBR will be able to offer about 2 million parallelism on a single wafer. Table 2 shows the number of memory-cells on a single wafer using various fabrication technologies. The 1995 technology is an estimation from current trends of VLSI technologies (See [Gelsinger et. al., 1989] and [Dally, 1991] for brief estimations of VLSI technologies in future. [IEDM-91, 1991] is one other good source of information.).

5.2. Precision: Robustness against noise

One of the problems of the analog VLSI is the precision in computing. Inherently, some noise undermines computing precision. We have simulated the noise factor of the analog part of our circuit. We have confirmed that the anticipated noise level will not affect the accuracy of reasoning in the memory-based reasoning system. There are several noise sources such as shot noise, thermal noise, 1/f noise, burst noise, and avalanche noise. Also, there are some device-level deviation at production process of the wafer. Empirically, the total level of noise is known to be at the level of $\pm 30\%$. In our experiments, we varied noise levels starting from $\pm 0\%$ to $\pm 100\%$ which essentially deviates similarity weight matrices. The distribution of the noise is uniform within the given range. Results are shown in Figure 5 and in Figure 6.

We have discovered that the WSI-MBR is extremely robust against noise. There is only a minor degradation even with the $\pm 100\%$ noise with the full-data set MBRtalk. The MBR seems to be fairly practical on analog VLSI since expected maximum level of noise is about $\pm 30\%$ which does not cause any significant degradation in accuracy even with the small memory-base. It should be noted that the WSI-MBR is more robust against noise than NETtalk since NETtalk's letter accuracy degrades to under 80% with 100% noise[Sejnowski and Rosenberg, 1987] whereas MBRtalk on WSI-MBR maintains over 80% of letter accuracy with the memory-base larger than 45,000 records.

5.3. Fault Tolerance: Robustness against defects

WSI inherently involves faults. As we have discussed previously, there is virtually no means to eliminate or control the defects. While we recognize this problem, we have carried out a set of simulations to identify the robustness of MBR against device level defects. The result is shown in Figure 7. The accuracy degrades almost linearly. Since the expected run-time faults would be far less than 1.0%, we see no problem on the accuracy degradation of MBR due to the run-time faults. By the same token, the result can be applied to the production defects. Since the expected defects density in our fabrication process is about 0.5 defect/cm^2, the total expected number of defects on one wafer is about 50 (assuming 100 chips/wafer and 1 cm^2 chips). This is only 0.02% of the 240K memory-cells. This implies that we don't even have to test each memory-cell to ensure the expected level of accuracy.

5.4. Speed: Beyond Tera FLOPS

Computing time of the WSI-MBR is mostly a sum of the maximum transmission delay in the wafer, memory bit cell access time, and gate delay for multiplier and other

Year for Production	Design rule (μm)	MBR Memory Cell per Chip	Chips per Wafer 5 inch	8 inch	MBR Memory Cell per Wafer
1989	0.8	1K	60	—	60K
1992	0.5	4K	60	120	240K - 480K
1995	0.3	16K		120	1920K

Table 2: Number of MBR Memory Cells on a Single Wafer

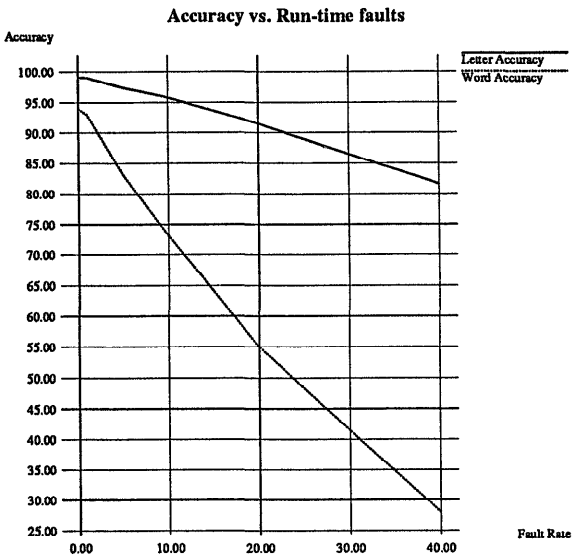

Figure 7: Fault rate and accuracy

circuits. Table 3 shows an estimated computing time of the WSI-MBR using various fabrication technologies. We use a data pipe-line in each data-bus chip to mitigate transmission delay.

Using the 1995 technology, we can attain 66 nanoseconds MBR operation and 2 million parallelism at the record level. Considering that each record contains 24 fields each of which requires weight computing, total arithematic operation for one MBR operation would be equivalent to 70 Tera FLOPS in digital computers. In the worst case, the data fetch may create a bottleneck slowing down the entire cycle to 1 milli-second. However, even in this case, the total computing power on a single WSI-MBR attains 0.48 Tera FLOPS. The knowledgable readers may notice that the system clock for the 1995-technology is 15 MHz, far less than 50 MHz which is the expected cycle for the VLSI processors. This is because we do not use intra-chip pipeline and memory banks, which are major methods for attaining higher system clock cycle, because these methods would complicate circuits, thus decrease parallelism. In our design simulations, the higher-level of parallelism was preferred over the introduction of pipeline and memory-banks. We can use higher system clock cycle such as 50 MHz, when the design decision was made to do so.

6. Conclusion

In this paper, we presented the design of the wafer scale integration for massively parallel memory-based reasoning. The unique feature of the WSI-MBR is an integration of broad range of the state-of-the-art technologies such as wafer scale integration, high precision analog circuit, massive parallelism, and memory-based reasoning paradigm. Our experiments indicate several significant advantages of the WSI-MBR:

- WSI-MBR attains an extremely high level of parallelism. Due to the compact design attained by the digital/analog hybrid approach, and by the state-of-the-art fabrication technologies, a single wafer WSI-MBR can host 240K MBR memory cells. By year 1995, the WSI-MBR can be fabricated with 0.3 μm technology which enables 2 million MBR memory cells to be hosted on a single 8 inch wafer. Each MBR memory cell has its own processing capabilities.
- WSI-MBR attains TFLOPS-order of aggregated computing power on a single wafer. Our estimation results demonstrate that WSI-MBR can attain 70 TFLOPS using the 1995 WSI technology. Even in the worst case where the data fetch creates a substantial bottleneck, the WSI-MBR attains 0.48 TFLOPS.
- WSI-MBR overcomes the problem of noise in analog VLSI due to robustness of the MBR approach. Our simulator experiments demonstrate that WSI-MBR does not cause significant accuracy degradation with the noise level anticipated in the actual implementation of the analog part of the VLSI.
- WSI-MBR is highly reliable because (1) it minimizes the ponding process which is the major cause of processing and run-time faults, and (2) it is highly fault-tolerant due to the distributed nature of the MBR approach. While WSI-MBR shows linear degradation as number of records are damaged, the expected run-time faults rate is extremely small, a fraction of a percent. Thus, our experiments indicate that the WSI-MBR can overcome the problem of the run-time faults.
- WSI-MBR can be built as a compact plug-in module or as an independent desk-top massively parallel system. This is due to the compact circuit design and by the use of WSI technology.

In summary, the WSI-MBR offers an extremely high level of parallelism, over TFLOPS of aggregated computing power, and the approach (combination of WSI technology and MBR paradigm) effectively overcomes the

Year for Production	Design rule (μm)	Max. Trans. Delay on Wafer (Two-way)	Memory Bit Cell Access Time	Gate Delay for Multiplier, etc.	Total Time	System Clock Cycle (MHz)
1989	0.8	20	60	16	96	10
1992	0.5	20	40	10	70	14
1995	0.3	32	30	6	66	15

Table 3: Estimated computing time of WSI-MBR at each technology (nanoseconds)

noise problems of the analog VLSI and the faults problem of WSI. In addition, the WSI-MBR is expected to be cheap (few thousand dollars) and compact (TFLOPS MBR machine in the lap-top computer size). Although, the high performance was attained due to the specialized architecture, the basic paradigm, MBR, is known to be useful for many areas of application. The computing power offered by the WSI-MBR would certainly explore new and realistic application areas. The utility of the WSI-MBR would be explored with data-intensive domains such as human genome sequencing, corporate MIS system, etc.

We believe the implication of our work is significant. We have shown that a compact and inexpensive massively parallel MBR system can be built even now. As far as the MBR is concerned, we are no longer constrained by the memory-space, the processor power, and the cost of the massively parallel machines. Millions processing elements are there, and TFLOPS are in our hand. The next step is to build applications to make use of the state-of-the-art WSI-MBR technology.

References

[Cavill et. al., 1991] Cavill, P. J., Stapleton, P. E., and Wilkinson, J. M., "Wafer Scale Integration: A technology whose time has come," *Proc. of International Conference on Wafer Scale Integration,* IEEE Computer Society Press, 1991.

[Creecy et. al., 1990] Creecy, R., Masand, B, Smith, S., and Waltz, D., *Trading MIPS and Memory for Knowledge Engineering: Automatic Classification of Census Returns on a Massively Parallel Supercomputer,* Thinking Machines Corporation, 1990.

[Dally, 1991] Dally, W., "Fine-Grain Concurrent Computing," *Research Directions in Computer Science: An MIT Perspective,* Meyer, A., (Ed.), MIT Press, 1991.

[Gelsinger et. al., 1989] Gelsinger, P., Gargini, P., Parker, G., and Yu, A., "Microprocessors circa 2000," *IEEE Spectrum,* Vol. 26, No. 10, 1989.

[IEDM-91, 1991] IEEE *Proc. of International Electron Device Meeting,* 1991.

[Kitano and Higuchi, 1991] Kitano, H. and Higuchi, T., "High Performance Memory-Based Translation on IXM2 Massively Parallel Associative Memory Processor," *Proc. of AAAI,* 1991.

[Kitano and Higuchi, 1991] Kitano, H. and Higuchi, T., "Massively Parallel Memory-Based Parsing," *Proc. of IJCAI-91,* 1991.

[Kitano et. al., 1991] Kitano, H., Hendler, J., Higuchi, T., Moldovan, D., and Waltz, D., "Massively Parallel Artificial Intelligence," *Proc. of IJCAI-91,* 1991.

[McDonald et. al., 1991] McDonald, J. F., Donlan, B. J., Russinovich, M. E., Philhower, R., Nah, K. S. and Greub, H., "A fast router and placement algorithm for wafer scale integration and wafer scale hybrid packing," *Proc. of International Conference on Wafer Scale Integration,* IEEE Computer Society Press, 1991.

[Mead, 1989] Mead, C., *Analog VLSI and Neural Systems,* Addison Wesley, 1989.

[Sato and Nagao, 1990] Sato, S. and Nagao, M., "Toward Memory-based Translation," *Proc. of the International Conference on Computational Linguistics (COLING-90),* 1990.

[Seeds, 1967] Seeds, R.B., "Yield and Cost Analysis of Bipolar LSI," *Proc. of IEEE International Electron Devices Meeting,* Oct., 1967.

[Sejnowski and Rosenberg, 1987] Sejnowski, T. and Rosenberg, C., "Parallel Netwroks that Learn to Pronounce English Text," *Complex Systems,* 1, 145-168, 1987.

[Stanfill and Waltz, 1986] Stanfill, C. and Waltz, D., "Toward Memory-Based Reasoning," *Communications of the ACM,* Vol. 29, No. 12, 1986.

[Sumita and Iida, 1991] Sumita, E. and Iida, H., "Experiments and Prospects of Example-Based Machine Translation," *Proc. of the Annual Meeting of the Association for Computational Linguistics (ACL-91),* 1991.

[Thinking Machines Corporation, 1989] Thinking Machines Corporation, *Model CM-2 Technical Summary,* Technical Report TR-89-1, 1989.

[Waltz, 1990] Waltz, D., "Massively Parallel AI," *Proc. of AAAI-90,* 1990.

[Yasunaga et. al., 1991] Yasunaga, M., et. al. "A Self-Learning Neural Network Composed of 1152 Digital Neurons in Wafer-Scale LSIs," *Proc. of the International Joint Conference on Neural Networks at Singapore (IJCNN-91),* 1991.

[Zhang et. al., 1988] Zhang, X., Waltz, D., and Mesirov, J., *Protein Structure Prediction by Memory-Based Reasoning,* Thinking Machines Corporation, 1988.