Hierarchical Multi-Scale Architectures with Spin Waves[1]

Mary M. Eshaghian-Wilner, Alex Khitun, Shiva Navab, and Kang Wang

Department of Electrical Engineering
University of California, Los Angeles
Los Angeles, CA 90095-1594
(310) 825-2647
maryew, ahit, shiva_n, wang @ee.ucla.edu

Abstract

In this paper, we present three hierarchical multi-scale architectures that are interconnected electro-optically at micro-scale level and via spin waves at nano-scale level. These architectures are derived from the Optical Reconfigurable Mesh (ORM), which is a MEMS architecture that supports several types of electrical and optical routings. The first multi-scale architecture presented here is a structure that consists of a set of nano-scale spin-wave crossbars that intercommunicate via the ORM's MEMS electro-optical routing. The second architecture consists of an array of nano-scale reconfigurable meshes that are interconnected with standard electrical reconfigurable switches at the micro-scale level. And, finally, the third architecture is a set of fully interconnected spin-wave clusters that can intercommunicate directly with one another using the ORM's MEMS free-space optical routing mechanism.

1. Introduction

The idea of fabricating tiny movable devices on chips was first conceived in the late 1960s, and strong research and development activity in this field started around 1980 [1]. Since then, many results have appeared that show theoretical modelings, new materials, fabrication processes, actuation mechanisms, and sensing methods. The impact that the MicroElectroMechanical Systems (MEMS) already have had in various applications, such as on sensing applications, has been very noteworthy. The art of integration via MEMS technology has led to the development of a huge array of integrated microsystems with rich and versatile functionality.

Optical MEMS is a relatively new and highly productive discipline within MEMS. The Optical MEMS conference started in 1996 and in just a few years has grown significantly. Several journals have been dedicated to Optical MEMS. For instance, an issue of Journal of Lightwave Technology is dedicated to Optical MEMS and Its Future Trends, and was edited by Lin, Wu, Sawada, and Mohr. It contains papers that discuss the integration of movable micromirrors in a chip. In [2], Ji and Kim show that they have designed and fabricated an addressable 4×4 array of micromirrors capable of providing up to 90 degrees of angular deflection. Each micromirror is comprised of a single crystalline silicon mirror plate supported by aluminum springs, which provides an extremely flat reflective surface and a compliant spring material that enables the integration of the device into a limited area without mitigating its performance. A mirror rotation angle of more than 80 degrees can be obtained by applying an external magnetic field. Furthermore, the authors state that this angle can be increased by the use of an electrostatic force. Each mirror plate and its associated springs occupy an area of 500x500 μm^2. A 10 μm thick layer of single-crystal silicon is used as the mirror plate for obtaining a flat surface. Their design allows for selective actuation of some of the micromirrors while the others remain clamped by electrostatic force.

The industry is significantly ahead of academia in designing micromirror arrays. The Digital Micromirror Device, also known as DMD, is built by DLP, which is a division of Texas Instruments. DMD is a micromirror system with approximately a million individually switchable micromirrors. Each mirror has a length of 13 μm and can be switched in 15 microseconds for maximum precision. The array of Micro-Opto-Electro-Mechanical System (MOEMS) mirrors built by Lucent Technologies is an array of 100 million switchable micromirrors. This essentially can operate as a peta bit switch that works for 1,296 ports, each containing 40 separate signals, and each of the signals can carry 40 gigabits per second.

The Optical Model of Computation (OMC) is an abstraction of computing chips with optical interconnects

---

[1] The authors are listed in alphabetical order.

[3,4] that was introduced in the late 1980s. Similar to the VLSI model of computation, which was proposed by Thompson in the late 1970s, the OMC model can be used to understand the limits on computational efficiency in using optical technology. Unique qualities of the optical medium are its abilities to be directed for propagation in free space and to have two optical channels cross in space with out interaction. These properties allow optical interconnects to utilize all three dimensions of space.

Due to the mentioned advancements in technology, we are now able to implement the OMC-based models with reconfigurable mirrors, such as the Optical Mesh or the ORM architecture [3,4]. The NanoElectroMechanical Systems (NEMS) implementation of these types of architectures will open yet another gate of possibilities towards the efficient implementation of VLSI architectures with reconfigurable mirrors. Recently, an architecture was described that integrates the MEMS implementations of ORM with Quantum Cellular Automata (QCA) [5,6]. The implementation of this architecture faces several challenges such as the low temperature requirement. In addition, if two bistable devices are connected together in series, some isolation is required between the input and output so that the input drives the output and not the reverse.

In this paper, we present three hierarchical multi-scale architectures that are interconnected electro-optically at micro-scale level and via spin waves at nano-scale level. These architectures are derived from the Optical Reconfigurable Mesh (ORM) which is a MEMS architecture that supports several types of electrical and optical routings. The first multi-scale architecture presented here is a structure that consists of a set of nano-scale spin-wave crossbars that intercommunicate via the ORM's MEMS electro-optical routing. The second architecture consists of an array of nano-scale reconfigurable meshes that are interconnected with standard electrical reconfigurable switches at the micro-scale level. And, finally, the third architecture is a set of fully interconnected spin wave clusters that can intercommunicate directly with one another using the ORM's MEMS free space optical routing mechanism.

The rest of the paper is organized as follows: In section 2, we present a brief introduction to spin waves as well as the ORM architecture. We then describe our proposed hierarchical architecture in section 3, which is followed by our conclusion and future work in section 4.

2. Preliminaries

In this section, we first provide a brief introduction to spin waves, and then we present a brief description of the optical reconfigurable mesh (ORM) model and the different types of routings that are used in that model.

2.1. Preliminaries on Spin-Waves

The use of spin waves for computation is an entirely new idea. Spin waves are used for both information transmission and information processing. We employ the classical type of computing as opposed to quantum, and the architecture presented can operate at room temperature. The following is a brief description of how spin waves can be used for computation and communication as presented in [7]. In the ideal case, spin waves can be used to provide an "LC" coupling of devices, without dissipative resistance. With the spin-wave concept, the spin rotates as a propagating wave and there is no particle (electron/hole) transport.

2.1.1 Our Prototype and Experimental Results

In Figure 1, we have schematically shown the prototype of the logic device structure. The core of the structure consists of a 100nm-thick CoFe film deposited on a silicon substrate. There are two asymmetric coplanar strip (ACPS) transmission lines on the top of the structure. The distance between the lines is 8μm. The lines and the ferromagnetic layer are isolated by a 300nm silicon oxide layer. The dimensions of the ACPS lines are adjusted to match 50Ω of the external coaxial cable.

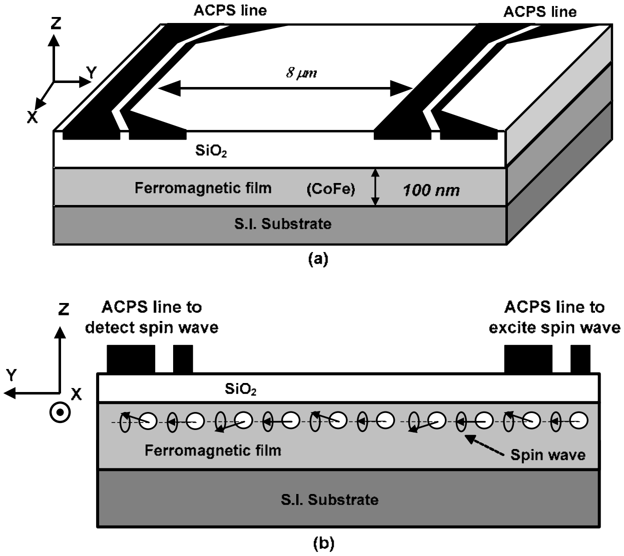


Figure 1 – Prototype of the Logic Device Structure

One of the ACPS lines (shown on the right) is used for spin-wave excitation. Hereafter, we will refer to this ACPS line as the "excitation" line. A voltage pulse applied to the excitation line produces a magnetic field and excites a spin wave (spin-wave packet) in the ferromagnetic layer. Because it is excited, the spin-wave propagates through the ferromagnetic film. The other ACPS line (shown on the left) is used to detect the inductive voltage signal produced by the propagating spin wave. Hereafter, we will refer to this line as the "detection" line.

In Figure 2, we present experimental data on spin-wave detection by the time-resolved inductive voltage measurement technique [8]. The dashed line depicts the voltage pulse applied to the excitation line. The pulse characteristics are as follows: pulse amplitude 24.5V; rising time 1.2ns; and pulse length 20ns. The solid line depicts the inductive voltage signal detected by the detection line. One can see the inductive voltage oscillation at the detection line caused by the inductive coupling via the spin waves. The output voltage signal has maximum pulse amplitude 26mV, and the period of oscillation is 9ns.

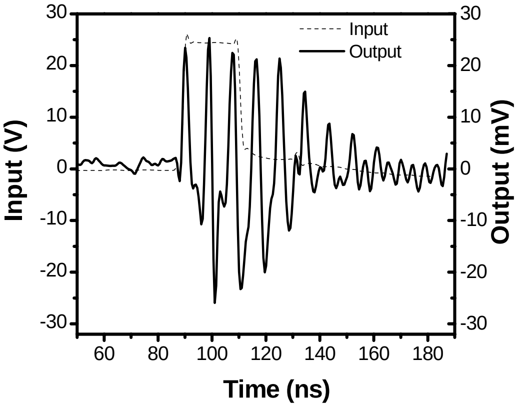


Figure 2 – Experimental Data on Spin-Wave Detection

These experimental data illustrate the possibility of signal transmission by spin waves over micrometer range distances. The attenuation time is about 20ns, and the signal-to-noise ratio is satisfactory (at least for 8μm propagation distance). We would like to stress that the utilization of spin waves is prominent for short-range in-chip communication. There are two important issues that require additional consideration: (i) power dissipation in spin-wave bus, and (ii) signal gain. In principle, the energy per spin wave can be scaled down to several kT, to be just above the thermal noise level. On the other hand, in order to compensate the damping of spin waves, one needs to include an amplification mechanism (gain). One of the possible gain mechanisms may be parametric spin-wave amplification, which has been experimentally demonstrated [9].

2.1.2. The Arithmetic and Logic Functionality of Spin Waves

The utilization of spin waves provides an opportunity to perform different logic functions in one device by controlling the initial phases of spin waves. The set of logic gates, i.e., the two-bit gates, AND, OR, and the one-bit NOT, can be realized on the prototype device mentioned in the previous section. The input information is coded into the phase of the spin wave or, in other words, in polarity of the voltage pulse applied to the edge ASPC lines (for example, V_{input} = +1V corresponds to the logic state 1, and V_{input} = -1V corresponds to the logic state 0). In order to detect the output signal Vind, we use the time-resolved inductive voltage measurement. Two spin-wave packets coming in phase enhance the amplitude of the produced inductive voltage, and otherwise cancel each other when coming out of phase.

When mapping an algorithm to this architecture, one should take into consideration the fact that two different types of data detections are possible at the nodes. Once the spin waves are detected by the receiver ACPS lines, the transmitted data can be digitized, or they can be left analog. In analog detection mode, the ACPS line detects the inductive voltage produced by the superposition multiple waves. For example, if ten waves are sending a "1," then their analog sum through their cumulative amplitude is computed instantly as 10. Also, this property can be used to compute logical functions as described previously. In digital detection mode, this value is digitized to just a "1," and then the computations are continued digitally.

It is possible to realize different logic gates AND, OR, and NOT controlling the relative phase of the spin waves. The voltage measured in the output port is compared to a reference voltage to determine logic state 1 or 0. This measurement is performed at the moment of spin-wave packet arrival to the detecting ACPS line area tm. The exact choice of tm depends on the logic function one needs to realize. For more information, refer to related publications [10].

2.2. Preliminaries on Optical Reconfigurable Mesh

In this section, we describe the ORM architecture including its processing and the deflection layer, and the three different types of routing that are used in this architecture.

A 4 × 4 optical reconfigurable mesh (ORM) is shown in Figure 3. There are two layers in the ORM: the deflection layer and the processing layer. The deflection layer consists of N^2 deflecting units while the processing layer consists of N^2 processing units. The processors on the processing layer are interconnected as a reconfigurable mesh and can also intercommunicate optically using the deflection layer [11].

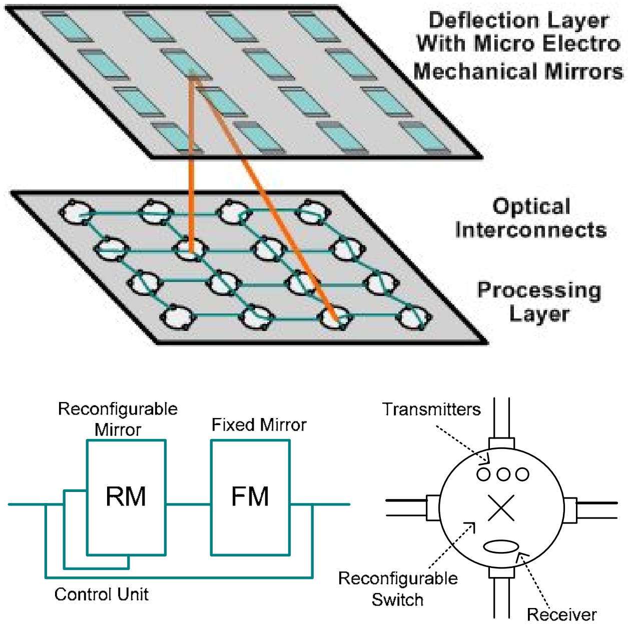


Figure 3 – The ORM Architecture

The reconfigurable mesh model used in the processing layer is standard. The reconfigurable mesh of size N^2 consists of an $N \times N$ array of processors connected to a grid-shaped reconfigurable broadcast bus, where each processor has a locally controllable bus switch. The switches allow the broadcast bus to be divided into sub-buses, providing smaller reconfigurable meshes or reconfigurable bus segments. The detailed structure of a processing unit in the processing layer and the detailed structure of a deflecting unit in the deflection layer are also shown in Figure 3. In the following subsections, we describe each of those components [3].

2.2.1. The Processing Layer

There are $N \times N$ processing units on the processing layer. There are three optical transmitters and one receiver residing in each processing unit. One of the transmitters, TR(1), is directed towards the control unit of the deflection unit. The second one, TR(2), is directed towards the reconfigurable mirror (RM) of the deflection unit, and the third one, TR(3), is directed towards the fixed mirror (FM) of the deflection unit. Each processing unit has a constant number of $\log N$ bit memory cells and simple computation capabilities. It is connected to other processing units in the mesh by the electrical reconfigurable buses. Each processing unit controls the internal reconfigurable switches and is responsible for sending and receiving data to and from the other processing units. We index the processing unit in the i^{th} row and the j^{th} column of the mesh on the processing layer as $P(i,j)$ in which $1 \leq i, j \leq N$.

2.2.2. The Deflection Layer

The deflecting layer contains $N \times N$ deflecting units. Each deflecting unit consists of two mirrors and an arithmetic control unit. One of the mirrors is a fixed mirror (FM), which transfers data from the processor under it to a fixed address whenever it is used. Another mirror is a reconfigurable mirror (RM). The control unit receives an address from the processor under it, translates the address, and controls the direction of the RM. Since the angle of the FM is fixed, the processor can send data directly from one dedicated transmitter to its destination without going through the control unit. We define each deflecting unit (a mirror and the related control unit) located directly above $P(i,j)$ as $M(i,j)$.

2.2.3. Data Movement in ORM

The data can be routed in three different ways in this architecture. In the first, *electrical routing*, the routing is done only through electrical buses. The second type, *optical routing*, uses free-space optics. The third type, *electro-optical routing*, uses electrical and optical free-space connections to allow a complete connection among N processors. Each of the movements is described below.

2.2.3.1. Electrical Routing

The electrical routing in ORM is similar to those for reconfigurable meshes. This type of routing is any routing from one node to another or a broadcast, which uses electrical buses in the reconfigurable mesh only. This type of communication is suitable for providing arbitrary configuration of the buses in the processing layer.

2.2.3.2. Optical Routing

The optical routing in ORM is the routing through optical free-space interconnections only. The data transfer does not use any electrical bus in the system. All N^2 processors can communicate in unit time delay as long as there is only one read or write from or to each location. In the following, we describe how such an optical connection is established between two processors through the RM.

A connection phase consists of two cycles. In the first cycle, each processor sends the address of its desired destination processor to the arithmetic control unit of its associated mirror using its dedicated laser TR(1). The arithmetic control unit of the mirror computes a rotation degree such that both the origin and destination processors have equal angle with the line perpendicular to the surface of the mirror in the plane formed by the mirror, the source processor and the destination processor. Once the angle is computed, the mirror is rotated to point to the desired destination. In the second cycle, the connection is established by the laser beam, TR(2), carrying the data from the source to the mirror and then from the reflected mirror towards the destination. An example of an optical routing from processor P(2,2) to processor P(4,3) is shown in the first ORM figure above.

The read operation has two phases. In the first phase, the read requirement and the reader's address are sent to the processor, which stores the desired data. In the second phase, the data is sent back to the reader depending on the reader's address. Both phases use the two-cycle write routing method.

2.2.3.3. Electro-Optical Routing

This communication mechanism establishes an efficient full connectivity among only the N processors situated diagonally in the processing layer on the N^2 processors in the ORM (i.e., for processors P(j,j) where $1\leq j \leq N$). This routing technique uses electrical buses on the processing layer and fixed mirrors on the deflection layer.

The connection for electro-optical routing is implemented as follows. Each processor P(j,j) is associated with the j^{th} row of the deflection unit, where the row contains N fixed mirrors. The i^{th} fixed mirror in that row for $1 \leq i \leq N$ is directed to the processing unit P(i,i). There are two possible types of routing: Exclusive Read Exclusive Write (EREW) and Concurrent Read Concurrent Write (CRCW). In EREW, any PE P(i,i) sends data to P(k,k) in the following way: First, P(i,i) sends the data to P(i,k) through the electrical row bus; Then, P(i,k) sends data to P(k,k) through transmitter TR(3) and its deflector M(i,k). The variety of techniques available in this architecture makes ORM a very powerful computing model [12].

3. The Hierarchical Multi-Scale Architectures

Here we describe three multi-scale hierarchical architectures that embody the ORM architecture as well as the spin-wave technology described above. In these three architectures, in addition to the standard processing layer of ORM, at micro-scale level, there is another processing layer at nano-scale level. The nano-scale processing layer includes a set of spin-wave-based nano-scale computing modules. Similar to the ORM architecture, there is a deflection layer that is responsible for optical interconnectivity.

3.1. Hierarchical Multi-Scale Crossbar

This hierarchical architecture consists of a set of nano-scale spin-wave-based crossbars that use the electro-optical routing mechanism of ORM to communicate among modules, as shown in Figure 4. In the electro-optical routing, as mentioned in the previous section, no reconfiguration of mirrors is necessary, and only fixed mirrors are used. The data communication in each row is through electrical interconnections, while the fixed mirrors provide vertical paths among processors. Since only electrical switches and fixed mirrors are used, this architecture has a switching time of nano-seconds. In Figure 4, the electrical connection between P(4,1) and P(4,4) has been highlighted, and the electro-optical routing from processor P(1,1) to processor P(3,3) is shown. P(1,1) makes an electrical connection to P(1,3) while P(1,3) is connected to P(3,3) using M(1,3) fixed mirror.

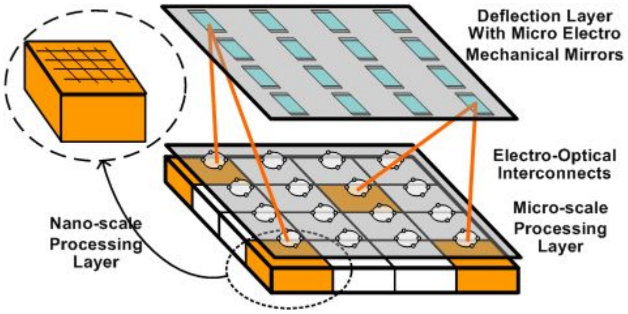


Figure 4 – The Hierarchical Architecture with Spin-Wave-Based Crossbars in the Computing Layer

Crossbars are attractive architectures because they can realize any permutations of N inputs to N outputs. However, the main shortcoming of regular VLSI crossbars is that N^2 switches are used to transmit only N pairs of data. The nano-scale crossbar described here, while requiring the same number of switches as standard crossbars, is capable of transmitting N^2 data elements. This is because each spin-wave bus is capable of carrying multiple waves at any given time using different frequencies. An example of the proposed spin-wave cross-bar architecture is shown in Figure 5. Note that a set of column spin-wave buses on the bottom and a set of row spin-wave buses on the top are connected via the vertical spin-wave switches [13].

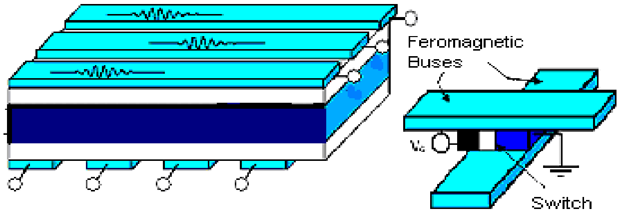


Figure 5 – The Nano-Crossbar with Spin Waves

3.2. Hierarchical Multi-Scale Reconfigurable Mesh

This hierarchical architecture consists of nano-scale spin-wave-based reconfigurable meshes, which at the module level are interconnected via electrical routing. Figure 6 shows this organization.

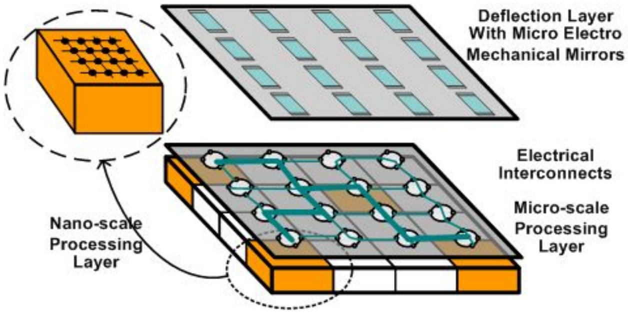
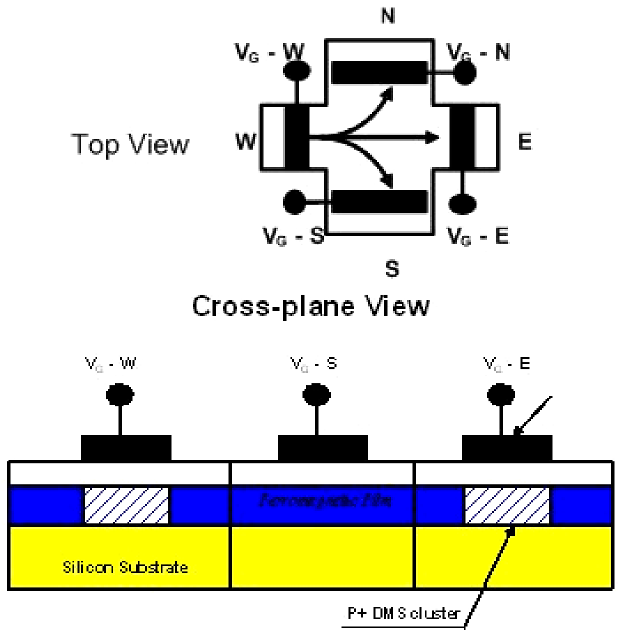


Figure 6 – The Hierarchical Architecture with Reconfigurable Meshes in the Computing Layer

Each of the nano-scale reconfigurable meshes used here consists of an *N x N* array of nodes connected to a reconfigurable spin-wave bus grid, where each node has a locally controllable bus switch. In this architecture, similar to the crossbar, a set of column spin-wave buses at the bottom and a set of row spin-wave buses on the top are connected via the spin-wave switches. Each switch is placed at the grid point of the mesh. Basically, except for the spin-wave buses, the nano-scale spin-wave-based reconfigurable mesh is similar to the standard reconfigurable mesh with a switching speed of a few nanoseconds. However, in the spin-wave-based version, at any given instance, O(N) messages, as opposed to one message, can be sent over each bus using different frequencies. The structure of the spin-wave reconfigurable switches is shown in Figure 7 [14].



$V_G < 0$ - *Spin-Wave comes through the DMS cluster*
$V_G > 0$ - *Spin-Wave total reflection on the boundary*

Figure 7 – The Spin-Wave Reconfiguration Switch

Reconfigurable meshes are suitable architectures for image-processing tasks because images can be mapped onto them in a straightforward fashion.

3.3. Hierarchical Multi-Scale Fully Interconnected Clusters

In this hierarchical architecture, a set of nano-scale spin-wave-based fully interconnected clusters at the nano-scale level are fully interconnected through the free-space optical routing of the ORM using the reconfigurable MEMS mirrors at micro-scale level. Because of the reconfigurable mirrors used, the switching speed at the micro scale level is significantly slower than the switching speed in the nano-scale level.

In the fully interconnected spin-wave-based clusters, each node can simultaneously broadcast to all other nodes, and can concurrently receive and process multiple data. This architecture consists of N computing nodes placed around a circle on a magnetic film. Each node is an ACPS line, which can be used as a sender or receiver at each point of time. Figure 8 shows the top and cross-view of the layout of this architecture on a semiconductor chip. This nano design allows nodes to exchange $O(N^2)$ messages at any given time, using different frequencies, while having an $O(N^2)$ area [15].

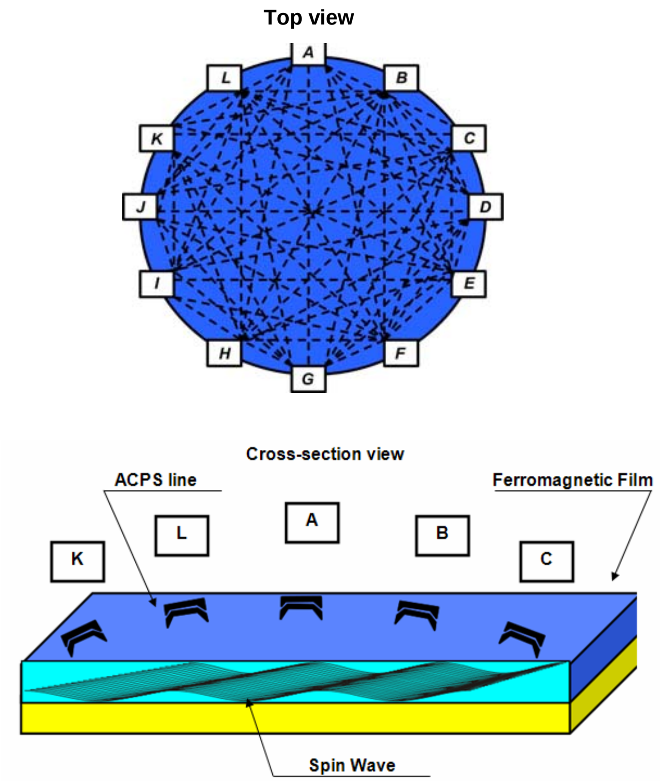


Figure 8 – The Top and Cross-Section View of the Architecture with Full Spin-Wave Interconnectivity

This architecture realizes full interconnectivity at each level for any type of random pattern, and therefore it is a desirable architecture for implementing types of applications that require high and random interconnectivity. Biologically inspired computations and neural networks are a few examples. Figure 9 shows this hierarchical structure.

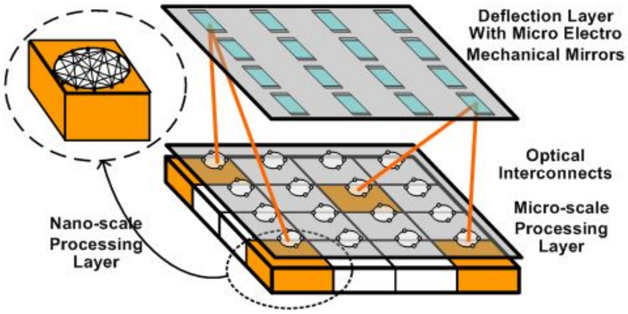


Figure 9 – The Hierarchical Architecture with Fully Interconnected Clusters in the Computing Layer

4. Conclusion

In this paper, we presented three hierarchical multi-scale architectures that are interconnected electro-optically at micro-scale level and via spin waves at nano-scale level. These architectures are derived from the Optical Reconfigurable Mesh (ORM), which is a MEMS architecture that supports several types of electrical and optical routings. The first multi-scale architecture presented consists of a set of nano-scale spin-wave crossbars that intercommunicate via the ORM's MEMS electro-optical routing. The second architecture consists of an array of nano-scale reconfigurable meshes that are interconnected with standard electrical reconfigurable switches at the micro-scale level. And, finally, the third architecture is a set of fully interconnected spin-wave clusters that can intercommunicate directly with one another using the ORM's free-space optical routing mechanism. This paper presented a preliminary study on how various types of electrical, optical, and spin-wave-based technologies can be used for interconnecting nodes and modules at micro and nano-scale levels in a single chip. Each of these technologies has certain features and limitations. The architectures presented here combined them in a complementary fashion so that their resulting computational and communicational capabilities were maximized

5. References

[1] M. Mehta, “ISRC Future Technology Topic Brief,” Bauer College of Business Administration, University of Houston, Texas, 2003.

[2] C.H. Ji, and Y.K. Kim, “Electromagnetic Micromirror Array with Single-Crystal Silicon Mirror Plate and Aluminum Spring,” in Journal of Lightwave Technology, Vol. 21, No. 3, March 2003.

[3] M. M. Eshaghian-Wilner, L. Hai, "A Glance at VLSI Optical Interconnects: From the Abstract Modelings of the 1980s to Today's MEMS Implementations," book chapter "Handbook on Innovative Computing," 2006.

[4] M. M. Eshaghian-Wilner, “Parallel Algorithms for Image Processing on OMC,” IEEE Transactions on Computers, 40(7):827-833, July 1991.

[5] M. M. Eshaghian-Wilner, “The Architectural Design of a Nano-Scale Computing Model,” Journal of Systemics, Cybernetics, and Informatics, vol. 2, no.4 2004.

[6] M. M. Eshaghian-Wilner., A. H. Flood, A. Khitun, J. F. Stoddart, and K. L. Wang, “Molecular and Nano-Scale Computing and Technology,” book chapter “Handbook of Innovative Computing,” 2006.

[7] A. Khitun, and K. L. Wang, "Nano Scale Computational Architectures With Spin-wave Bus,” Superlattices & Microstructures. 38(3),184-200, September 2005.

[8] A. Khitun,, et al, submitted for publication, 2006.

[9] B. A. Kalinikos, et al, “Parametric Frequency Conversion with Amplification of a Weak Spin-Wave in a Ferrite Film,” in IEEE Transactions on Magnetics, vol. 34, no.4, pt.1, July 1998, pp.1393-5.

[10] A. Khitun, and Wang K. L., “Nano-Scale Computational Architectures With Spin-Wave Bus,” Superlattices & Microstructures. 38(3),184-200, 2005.

[11] M. M. Eshaghian-Wilner, L. Hai, “An Optical Interconnected Reconfigurable Mesh,” Journal of Parallel and Distributed Computing, vol. 61, no. 6, June 2001, pp. 737-747.

[12] L. Hai, “Efficient Parallel Computing with Optical Interconnects,” Ph.D. Thesis, New Jersey Institute of Technology, 1997.

[13] M. M. Eshaghian-Wilner., A. Khitun, S. Navab, and K. L. Wang, “A Nano-Scale Architecture with Spin-Wave Crossbar,” The 6th IEEE Conference on Nanotechnology, Ohio, USA, July 2006.

[14] M. M. Eshaghian-Wilner., A. Khitun, S. Navab, and K.L. Wang, “A Nano-Scale Reconfigurable Mesh with Spin Waves,” The ACM International Conference on Computing Frontiers, Italy, May 2006.

[15] M. M. Eshaghian-Wilner., A. Khitun, S. Navab, and K.L. Wang, “Nano-Scale Modules with Spin-Wave Intercommunications for Integrated Circuits,” The NSTI Nanotech 2006, Boston, MA, May 2006.