Development of 3D Integrated Circuits for HEP

R. Yarema [a]

[a] Fermilab, P. O. Box 500, Batavia, Illinois*

yarema@fnal.gov

Abstract

Three dimensional integrated circuits are well suited to improving circuit bandwidth and increasing effective circuit density. Recent advances in industry have made 3D integrated circuits an option for HEP. The 3D technology is discussed in this paper and several examples are shown. Design of a 3D demonstrator chip for the ILC is presented.

I. INTRODUCTION

Requirements for High Energy Physics front end electronics and detectors continue to push the limits for lower mass and power, and higher resolution. One example is pixel vertex detectors, where multiple scattering within the readout electronics and detectors limits track resolution. Low mass is required to limit scattering, which leads to low power designs. Higher resolution is needed for better tracking, which leads to higher function density within the pixel. Significant progress has been made in the last decade to address these issues by integrating sensors and front end electronics within the pixel cell. MAPS (Monolithic Active Pixel Sensors) began the move toward integrated sensors and readout electronics. More recently success has been demonstrated in integrating sensors and CMOS circuitry using SOI technology. Now, 3D integrated electronics is offering exciting new possibilities.

A. Monolithic Active Pixel Sensors

MAPS have generated much interest within the HEP community. To reduce mass, MAPS combine detector and front end electronics on the same substrate in a commercial CMOS process. Currently, numerous groups are working on MAPS development.

As impressive as MAPS are, there are several practical limitations. First, the size of the detector signal is generally small and dependent on the thickness of an epi layer which varies from process to process. Due to the nature of the process, only NMOS transistors can be easily used. Addition of PMOS devices in an N-well reduces the charge collection efficiency and affects the spacial resolution. Finally, the functionality that can be squeezed into a small pixel for higher resolution is limited.

B. SOI Integrated Detector Development

SOI detector wafers are formed by bonding together a top wafer with low resistivity and a bottom wafer with high resistivity, using a silicon oxide bond. A buried oxide is formed between the wafers. After bonding, the top wafer is thinned to just a few microns using one of several different techniques. Later, vias are etched through the buried oxide to implant diodes in the bottom wafer and CMOS circuitry is built on the top wafer.

SOI integrated detectors have several advantages over MAPS. One big advantage is that both NMOS and PMOS transistors can be easily accommodated in the design. The devices inherently can have larger detector signals since the thickness of the depleted detector substrate can be controlled. Finally, since the substrate can be fully depleted, less charge spreading and higher speed is possible compared to MAPS.

Early detector work was done in a 3 micron SOI technology [1]. Recent work has been done in the OKI 0.15 micron SOI process [2]. Fermilab has an arrangement to work with American Semiconductor Inc. and OKI on future SOI detector developments.

Although SOI has significant advantages over MAPS, the functionality that can be placed in a small pixel cell is still rather limited. 3D integrated circuit development reduces that limitation.

C. 3D Integrated Circuit

A 3D integrated circuit is generally referred to as a chip comprised of two or more layers of semiconductor devices that have been thinned, bonded, and interconnected to form a "monolithic" circuit. Often the layers (sometimes called tiers) are fabricated in different foundry processes.

Industry is moving toward 3D circuits which permit shorter interconnects to reduce R, L, and C and provide higher speed. 3D also provides higher functionality per unit area and can reduce interconnect power and crosstalk.

3D integrated circuits represent the physicist's dream of combining large functionality with processes optimized for high performance, in a very small area as shown in Figure 1.
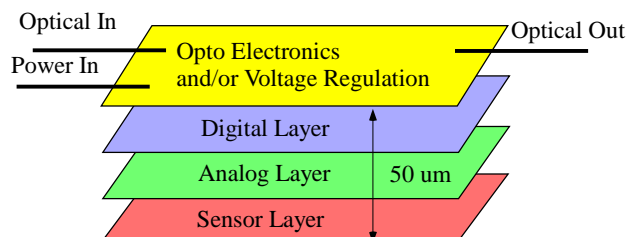

Figure 1: Thin 3D circuit with multiple tiers for HEP

Much work is being done on pixel arrays using 3D that could make this dream come true.

II. 3D CIRCUIT DEVELOPMENT

A. Advantages

A good way to see the advantages of 3D is to look at the layout of a conventional MAPS device and an equivalent 3D device as shown in Figure 2.

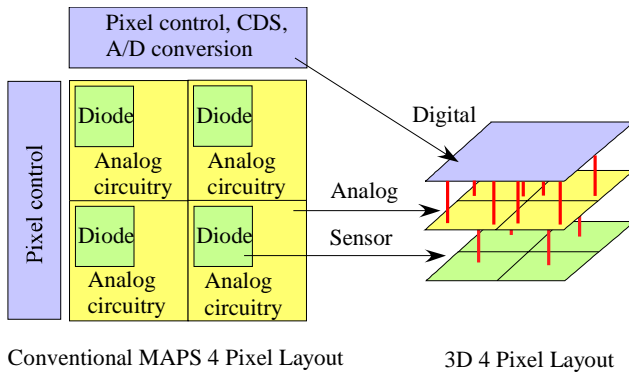Conventional MAPS 4 Pixel Layout   3D 4 Pixel Layout

Figure 2: Comparison of MAPS and 3D Pixel Layout

By placing circuitry on a separate layer from the sensor diodes, both NMOS and PMOS transistors can be used in the circuit design. Having a separate layer for digital functions, allows most of the circuitry usually placed around the periphery of the array to be placed above the array, providing a large amount of functionality in each pixel.

B. Key Technologies

Dozens of companies and organizations are working on 3D integrated circuits in the United States, Europe, and Asia. Companies such as IBM, Intel, Phillips, STM, Toshiba, and Samsung, and organizations such as MIT Lincoln Labs and the Fraunhofer Institute IZM-Munich are very active in 3D. Each of them has their own approach to working on the four key technologies needed for 3D: 1) bonding between layers using oxide to oxide fusion, copper tin eutectic bonding, or polymer bonding, 2) wafer thinning using a combination of grinding, lapping, etching and CMP, 3) through wafer vias using different processes (some of which need to include hole passivation), and 4) high precision alignment of parts before bonding [3].

C. Approaches for HEP

There are two different approaches that are useful for HEP as shown in Figure 3. The first approach uses a die to wafer or die to die bonding technique. Dice are tested and sorted before being bonded to a target wafer, resulting in a higher overall yield. This approach is conducive to using devices from different processes. Thus one can bond SOI to CMOS, CCDs to CMOS, or DEPFETs to CMOS, etc.
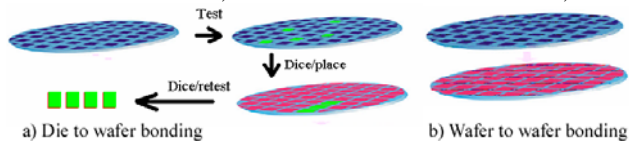

Figure 3: a) Die to wafer bond b) wafer to wafer bond

In the second approach, wafers are bonded to other wafers. For efficient use of silicon, the dice on the different wafers should be of the same size. Tight alignment tolerances are required over the entire wafer area and yields will be lower. The wafers could be of different technologies but there are significant advantages if the wafers are all made in an SOI process. The SOI process allows the wafer to be easily thinned to the buried oxide layer, resulting in active circuit layers that are less than 10 microns thick. Furthermore, the via formation in SOI is easier than in CMOS processes.

1) Die to wafer bonding:

With the die to wafer approach, the parts may be bonded face to back (circuit side to substrate side) or face to face. Figure 4 shows one way in which the devices are bonded and the vias are formed. It is necessary in the design to leave space for the inter-device vias. If vias pass through a CMOS layer, the vias must be insulated from the substrate.
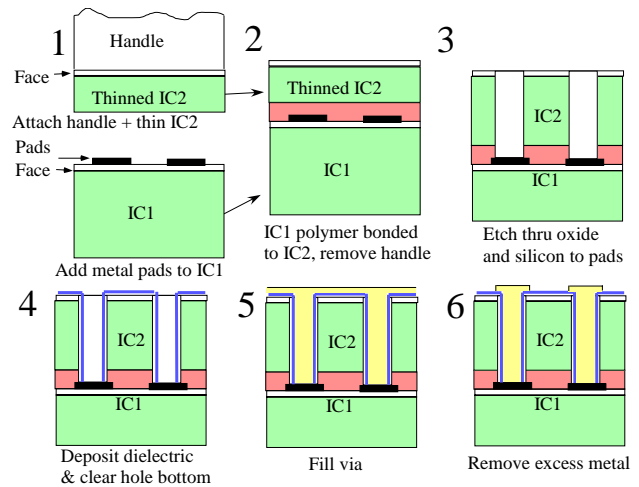

Figure 4: Steps showing polymer bonding and via formation for die to wafer process.

The polymer bond technique was used by RTI International to build a 3D infrared pixel array [4]. The array is a stack of 3 layers fabricated in three different processes: 1) the sensor is HgCdTe, 2) the analog circuitry (30 um thick) is made in 0.25 um CMOS, and 3) the digital part is made in a 0.18 um CMOS process. The array is 256 x 256 with 30 micron square pixels. Vias (4 μm dia.) were etched using the Bosch process.   A chip cross section is shown in figure 5.
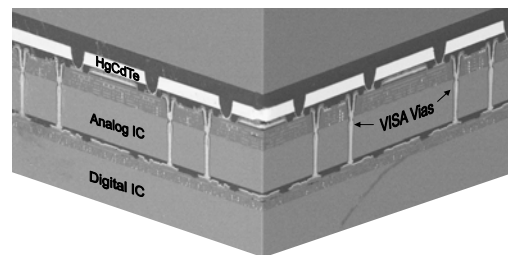

Figure 5: Cross section of 3D infrared pixel array

Face to face bonds can be made using a Cu_3Sn eutectic bond for both the electrical and mechanical connections between devices. This could be a possible replacement for conventional bump bonds. In applications to date, the Cu/Sn bonds cover a large portion of the mating surface areas with about 10 μm of copper. Because 10 μm of copper is 0.07% of a radiation length, Fermilab is investigating procedures to reduce the copper coverage to less than 10% of the mating areas. Our goal is to have 7 μm bumps on a 20 μm pitch.
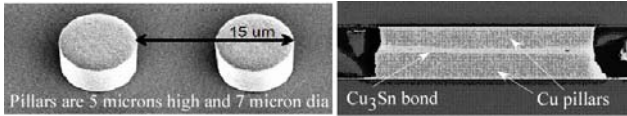
Figure 6: Cu pillar on 15 µm pitch and bond cross section.

RTI has made test structures, shown in Figure 6, which suggests that fine pitch Cu/Sn bonds are possible.

The process for the Cu/Sn bond is shown in Figure 7. A double handle transfer is required as shown in steps 2 and 3.
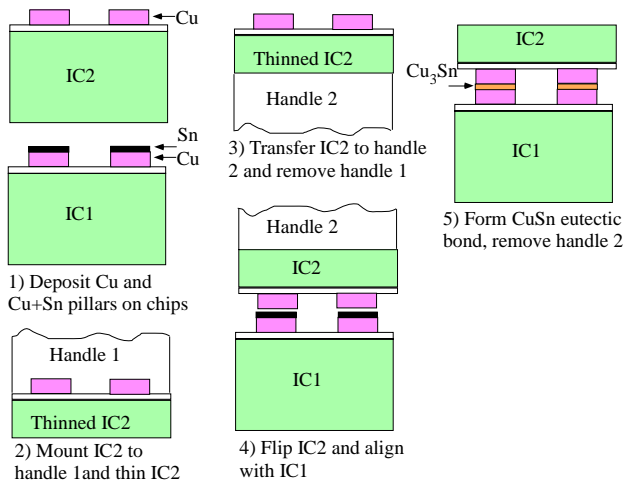

Figure 7: Face to face bond using Cu₃Sn

The IZM Fraunhofer Institute has used the Cu₃Sn bond technique to attach 10 µm thick chips to a target substrate [5]. In this case, the Cu bonds cover a large portion of the mating surfaces to provide mechanical support during processing. See figure 8.
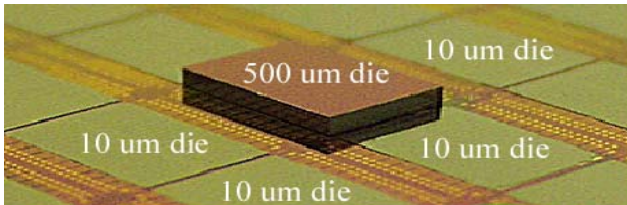

Figure 8: Chips 10 µm thick mounted on target substrate using Cu₃Sn along with unthinned 500 micron die for comparison.

2) Wafer to wafer bonding:

Wafers are frequently bonded together using an SiO₂ bond. For good bonding, the wafers must be very flat and the surfaces must be extremely clean. The techniques used for SOI detector wafer bonding and 3D wafer bonding are similar. For SOI detector design, the processing of each wafer is done <u>after</u> bonding. For 3D circuit design, processing of each wafer is done <u>before</u> bonding.

SOI wafers have several advantages for 3D wafer stacking. Each wafer can be easily thinned to the buried oxide layer (BOX) since the buried layer acts as an etch stop. Wafers have been thinned to 6 microns as shown in Figure 9.

3D wafer to wafer stacking has been used to build large pixel arrays. Lincoln Laboratories has built a two tier, 1024 x 1024 pixel array, CMOS image sensor with 8 µm square pixels [6]. Two 150 mm wafers were stacked using an oxide
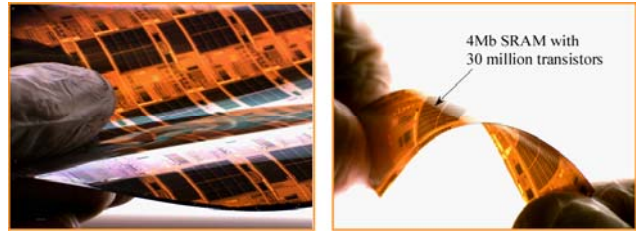

Figure 9: Wafer thinned to 6 microns and mounted to 3 mil kapton (courtesy of MIT Lincoln Labs)

bond. The first tier has the sensing diodes in a 3000 ohm-cm substrate, 50 micron thick. The second tier, which is only 7 microns thick, is made in a 0.35 µm SOI CMOS process. Vias (2 µm dia.) were dry etched and filled with tungsten. Each array has one million 3D vias with 99.999% yield. Figure 10 shows a cross sectional representation of the chip.
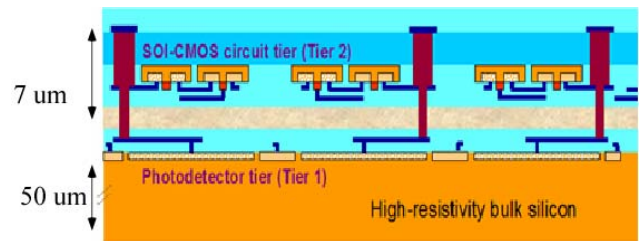

Figure 10: Cross section of CMOS imager showing 3D vias [7]

Stacking of three SOI layers has also been demonstrated by Lincoln Laboratory in a 3D Laser Radar Imager chip [8]. The chip is a 64 x 64 array of 50 µm square pixels. The first layer or tier is the high resistivity substrate with APD light sensors. The second tier is made in a 0.35 µm SOI process using high voltage devices for control of the APDs. The third tier uses a 0.18 µm SOI process for maximum density of the digital circuitry. There are six 3D vias in every pixel cell. Figure 11 shows a cross section of the chip.
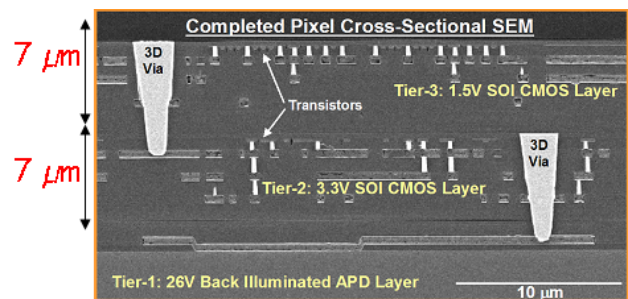

Figure 11: Cross section of 3 tier Laser Imager chip.

III. DESIGN OF A 3D ASIC FOR HEP

To show the capability of the 3D process for HEP, design of a readout chip has been undertaken. The ILC (International Linear Collider) vertex pixel detector has been chosen to demonstrate 3D design.

A. ILC Vertex Pixel Requirements

The ILC is expected to have a beam structure with 2820 crossings in a 1 msec beam train, occurring 5 times/sec. The pixel detector is comprised of 5 small concentric cylinders of pixels wrapped around the beam crossing point. It is assumed

that 0.03 particles will pass through each square millimeter on the surface of the cylinders for every beam crossing. To allow for charge spreading, hits between pixels, and magnetic field effects, it is assumed that there are 3 hit pixels for every particle. Thus the hit rate on each cylinder is 252 hits/beam train/mm [9].

Resolution for the pixel detector should be 5 microns or better. For 15 µm pixels, the resolution for a simple binary readout system would be 4.3 microns. For 20 µm pixels, the resolution would be 5.8 microns.

Sparsification is highly desirable to reduce the volume of data being transmitted off the chip and to reduce the digital power dissipated in the chip. Even though the expected occupancy per pixel is quite small, time stamping of hits is necessary to define when a hit occurred in a given pixel. The pixel hit time is used for reconstructing a particle track in association with data from other detectors.

As mentioned, the occupancy in a given pixel cell is expected to be very low. In a 15 micron square pixel, the number of hits/ bunch train would be 250 hits/mm^2 x (15 µm x 15 µm) = 0.056 hits/bunch train. Thus the chance of a pixel being hit twice in a bunch train is 0.056 x 0.056 or only 0.3%. If just 1 hit is recorded for each pixel, it follows that 99.7% of the hits are recorded unambiguously. Similarly, if the pixel is 20 microns, 99% of the hits are recorded unambiguously.

B. Demonstrator Chip Design

A number of design choices have been made which demonstrate the functionality that can be contained in a small pixel cell. The chip was originally designed for a binary readout, but an analog readout was added in the event that analog information proves useful. The hits for the ILC bunch train have been divided into 32 time slices. Expansion to 64 or more time slices is possible. The time slice information is stored within the pixel cell. A look-ahead token passing scheme was incorporated to sparsify the hit data on the chip. To reduce the size of the pixel, the pixel addresses are stored on the periphery. The chip design was divided into 3 tiers for fabrication in a MIT Lincoln Laboratory multi-project run. Since a detector layer is not available in this run, the intention is to mate the demonstrator chip to a detector at a later time. The chip is designed for a 1024 x 1024 array, but for economy has been laid out as a 64 x 64 array.
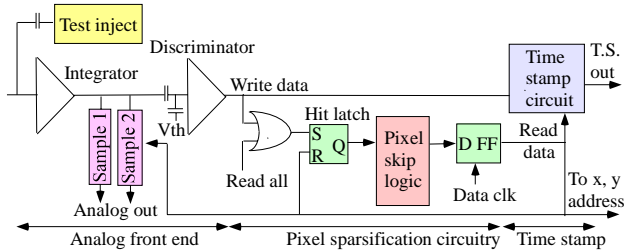

Figure 12: ILC Demonstrator Chip Pixel Cell

A simplified diagram of one pixel cell is shown in Figure 12. The front end consists of an integrator and a double correlated sample. The signal is capacitively coupled into a discriminator that has a chip-wide settable threshold. Output of the discriminator is fed to the time stamp circuit, and through an OR to a Hit Latch. The OR allows all cells to be

read regardless of the input signal. The Hit Latch along with the pixel skip logic and D flip flop performs the data sparsification. A programmable register is used to inject a test signal into every pixel cell.

The time stamp block shown in Figure 12 contains both a digital and an analog time stamp. This is done only for test purposes. The digital time is generated by a slow 5-bit Gray code counter on the perimeter which is clocked at about 30 µsec/step. The counter time is sent to all cells, where it is latched by the pixel hit signal. The hit time stored in the cell is read out on the same 5-bit bus after the beam train is over. The analog time stamp is formed from a slow voltage ramp that is sent to all pixels and sampled by the pixel hit signal. After the beam train, the stored analog voltage is read out on a separate bus to an external ADC. The digital time stamp has the advantage of being highly accurate while the analog time stamp has the advantage of requiring minimal silicon area.

A 1000 x 1000 pixel array is shown in Figure 13. During data acquisition, a latch is set in each pixel that is hit. After a beam train, sparse readout is performed row by row. To start readout, all hit pixels are disabled except the first hit pixel in the readout scan. The pixel being read out points to the X and Y addresses that are stored on the perimeter. At the same time the digital time stamp in the cell and all the analog information stored in the cell is sent to the edge of the chip.
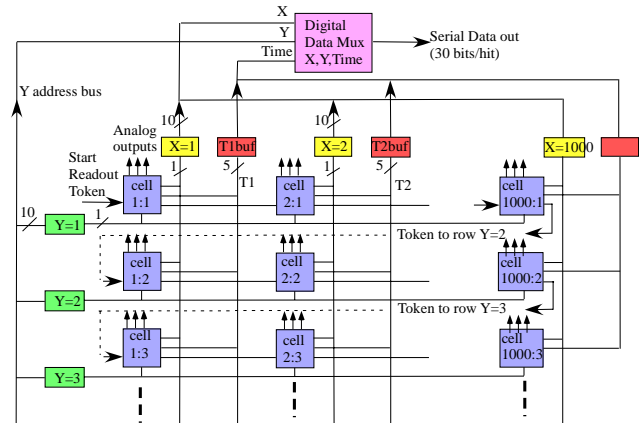

Figure 13: Demonstrator Pixel Array for ILC

While a pixel is being read, the token scans ahead, looking for the next pixel to read. To assure finding the next hit pixel before the token arrives, the chip is set to always read out at least 1 pixel per row regardless of the hit pattern. The token passing speed is 0.2 nsec per pixel. If there are 1000 cells in a row, the maximum time to reach the next cell to be read is 200 nsec. All the digital information is serialized on the perimeter of the chip for transmission off the chip. It is assumed that the serializer runs at a very reasonable 50 MHz. Thus, allowing 10 bits for the X and 10 bits for the Y addresses, along with 10 bits for the digital time stamp and status, a hit pixel is read out in 600 nsec, which is far more than the time required to find the next hit pixel.

For the ILC pixel vertex detector, there is 200 msec to read out the hits in a bunch train before the next train arrives. Assuming a 1000 x 1000 array of 15 µm pixels, the maximum number of hits in the hottest part of the detector is calculated to be about 56,250. If one extra pixel is readout in each row, the maximum number of pixels to be read is 57,250. To read

57,250 pixels with 30 bits/pixel at 50 MHz takes 34 msec. For a Megapixel chip with 20 micron pixels, the readout time would be 60 msec. Thus the readout time is far less than the ILC allowed 200 msec. Therefore the readout clock can be even slower, or several chips can share the same readout bus. Since the digital outputs are CMOS, the power is only dependent on the number of bits being read, and not the length of time needed for read out.
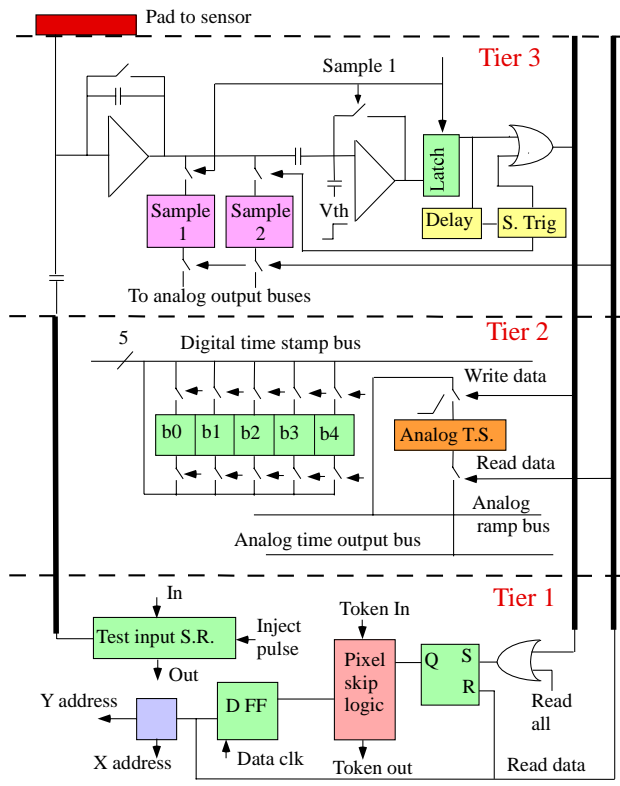

Figure 14: Division of ILC pixel design into 3 tiers.

Figure 14 shows how the pixel readout chip is divided into 3 tiers. Tier 3 contains most of the analog circuitry (38 transistors + storage capacitors), tier 2 contains all of the time stamp circuitry (72 transistors), and tier 1 primarily contains the sparsification logic (65 transistors). The thick vertical lines show the vias that are formed to interconnect the tiers. Figure 15 shows roughly how much area was required by the different circuit elements on each of the three tiers in the design. Each pixel cell is 20 µm x 20 µm.

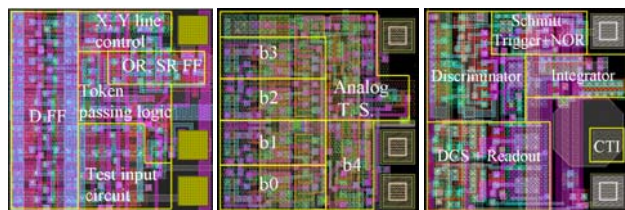Figure 15: Tier layouts from left to right: sparsification, time stamp, analog.

The demonstrator chip is a multifunctional device to be used a proof of the 3D principle for HEP. It is a 64 x 64 array that can be expanded to 1000 x 1000. There are 175 transistors in each 20 micron pixel. The active thickness of the 3 combined layers is only 22 microns. A choice will be made for future applications that will select analog or binary

readout and one of the two time stamp approaches. The current chip has provision for a test input signal which can be expanded to include a pixel disable circuit with little extra circuitry. The power dissipated by a full scale version of this chip is consistent with air cooling of the ILC pixel vertex detector. The support logic around the perimeter of this chip is small. In future designs, this can be reduced further. The chip is to be submitted for fabrication in a multi-project run at MIT Lincoln Labs on October 1, 2006.

IV. CONCLUSION

Industry is moving toward three dimensional integrated circuits. 3D is a natural progression for higher performance and higher functionality and is opening new approaches for HEP applications such as pixel arrays. Several approaches, such as die to wafer bonding, wafer to wafer bonding, or a combination of the two, are possible for pixels. Fermilab is working on a number of these ideas for 3D integration.

V. ACKNOWLEDGEMENTS

VI. REFERENCES

[1] J. Marczewski, et. al., "SOI Active Pixel Detectors of Ionizing radiation – Technology and Design Development," IEEE Trans. On Nucl. Sci., vol 51, No 3, June 2004, pp. 1025 – 1028.

[2] Y. Arai, et. al., "First Results of 0.15 um CMOS SOI Pixel Detector," SNIC Symposium, Stanford, Califorina, April 3-6, 2006.

[3] R. Yarema, "3D Integrated Circuits for HEP," Sixth International Meeting on Front End Electronics, Perugia, Italy, May 17- 20, 2006.

[4] C. Bower, et. al., "High Density Vertical Interconnects for 3D Integration of Silicon ICs," 56th Electronic Components and Technology Conference, San Diego, May 30-June 2, 2006.

[5] A. Klumpp, "3D System Integration," Sixth International Meeting on Front End Electronics, Perugia, Italy, May 17-20, 2006.

[6] V. Suntharalingam, et. al., Megapixel CMOS Image Sensor Fabricated in Three-dimensional Integrated Circuit Technology," IEEE SSCC 2005, pp. 356-7.

[7] C. Keast, et. al., "MIT Lincoln Laboratory's 3D Circuit Integration Technology Program", 3D Architectures for Semiconductor Integration and Packaging, Tempe, Arizona, June 13-15, 2005.

[8] B. Aull, et. al., "Laser Radar Imager Based on 3D Integration of Geiger-Mode Avalanche Photodiodes with Two SOI Timing layers," IEEE SSCC 2006, pp. 26-7.

[9] C. Baltay, "Monolithic CMOS Pixel Detectors for ILC vertex Detection," 2005 International Linear Collider Workshop, Stanford, CA, March 18-22, 2005.