A Comparison of CMOS Chip to Chip Interconnections:
Vertical Optical Through-Si Links Versus High Density Electrical Interconnect
Martin Brooke, Nan Marie Jokkerst, Steven W. Bond, D. Scott Wills
School of Electrical and Computer Engineering, Georgia Institute of Technology
Atlanta, GA 30332-0250, Tel. 404-894-3304, email: martin.brooke@ece.gatech.edu

Silicon CMOS integrated circuit electrical interconnection density is increasing, and poses a communications bottleneck which will worsen with time. The 1997 Semiconductor Industry Association (SIA) Roadmap identifies interconnect as one of the critical technology problems which must be addressed for future gigascale Si CMOS circuits to be successful. At this point in time, the SIA Roadmap identifies the interconnection densities and lengths projected for the years 2009 (10,000 m/chip at projected 70 nm CMOS linewidths) and 2012 (24,000 m/chip at projected 50 nm CMOS linewidths) as problems for which no solution has been identified [1]. Three dimensional interconnection and optoelectronic (OE) interconnection have been identified by the SIA as possible interconnect approaches to solve these problems. In this paper, we compare two types of competing chip to chip interconnection for Si CMOS, namely, high density electrical interconnect (HDI), and vertical optical through-silicon communication links for the realization of 3D stacked Si circuits.

Some computational challenges, such as real-time high resolution imaging, real-time video quality virtual reality rendering, and automatic target recognition for automated vehicles require compact, low cost, massively parallel processing systems [2][3]. One electronic interconnection solution which is currently being investigated is high density interconnect (HDI). However, these multilayer dielectric/metal electrical interconnections can suffer from high crosstalk, high latencies, and impedance mismatch [4]. An alternative interconnection solution utilizes vertical optical interconnections through layers of standard foundry silicon circuits (which are transparent to wavelengths longer than approximately 1.2 μm) to realize three dimensional stacked foundry Si CMOS structures. Optoelectronic interconnections have been previously analyzed as an interconnection option, and offer some solutions to electrical interconnect issues [5][6].

Current and future trends in electronic interconnects will meet the needs of many current and future systems, but for massively parallel systems, which are ideally suited, for example, to image processing and generation, electronic interconnects are inadequate. First, consider size and latency. Stacked 3D Si CMOS circuits with vertical OE interconnections will have a smaller volume, shorter interconnection lengths, and lower latency than the inherently 2D format of HDI, which mounts (with electrical interconnections such as ball grid array or chip scale packaging) circuits onto a high density interconnect board. As an example, consider a 3D computational mesh structure with 32x32x4 (4096) nodes (arranged in 4 tiles of 32x32 processors on an HDI substrate), shown in Figure 1a. Each processor node operates at 300 MIPS (million instructions per second), and occupies 1x1 cm² in area. Assuming 0.3 processor operands are needed from the network per operation, a processor operand size of 64 bits, and 8 bits of network overhead, yields a processor I/O bandwidth of 6.48 Gbps. This 4096 node system can compute a maximum likelihood reconstruction of a positron emission tomography medical image scan in seconds, but yields a thin film technology HDI substrate over 2x2 ft² in size. With advanced thin film dielectrics and copper metallization, (65 ps/cm delay time), this substrate yields a 2.15 ns (3.89 ns with more common alumina/molybdenum with 118 ps/cm delay) [7] transmission media latency (how long the signal propagates down the transmission media) for adjacent processor/processor communication from layer to layer (which is across the 32x32 array in the 2D HDI format), as shown in Figure 1. This is a major limitation to computational throughput. Because the 2D HDI interconnections are inherently nonscalable, this latency increases to 4.23 ns (7.67 ns for a ceramic substrate) for a 64x64x8 processor array (32,768 nodes), and the electrically interconnected system size rises to nearly 6x6 ft². Backplanes could be used to reduce the size of the electrical system, however, for the 64x64x8 processor module, this would require a 26.7 Tbps aggregate data transfer rate (over 26,000 links operating at 1 Gbps) for modest 300 MIPS processors [2].

Progress in the area of 3D optically interconnected systems is expanding and developing. Vertical optical communication links operating at 40 Mbps directly through stacked Si foundry CMOS circuits using direct hybrid integration of InP-based emitters and detectors bonded to Si CMOS transmit and receive circuits has been demonstrated [8]. Wafer bonding and flip-chip techniques have been used toward a demonstration of 3D memory architectures with an external light source [9]. For multi-chip module stacking, InGaAs/GaAs vertical cavity surface emitting laser arrays have been coupled with InGaAsP/InP photoreceiver arrays and microlenses into a 3D module with laser drilled holes for the interconnection path [10].

The scalability of 3D vertical OE interconnection technology may ultimately yield effective massively
parallel computational systems. For example, a $32 \times 32 \times 4$ processor system using the simple LED link demonstrated in [8] would yield a $12 \times 12$ in$^2$ module. The transmission media latency, based upon the speed of light, is 3 ps (negligible) for adjacent processor/processor communication from layer to layer in any size system. The transceiver circuits and emitter/detector delay will dominate the optical link latency.

Power dissipation is another important system concern. The electrically connected $32 \times 32 \times 4$ HDI system, assuming an aggressive 2 pF/cm [7], 3 V signal levels, a channel data rate of 100 Mbps, and a 64 bit bus, has interconnect power consumption of 97 kW. This power dissipation is largely due to charging and discharging the 33 pF average line capacitance. Based upon measured power dissipation data at 40 Mbps for a vertical OE through-Si link extrapolated to 100 Mbps, we expect each integrated OE transceiver link to dissipate 223 mW. The channel data rate will be 100 Mbps, with 64 channels per processor (which corresponds to an 8x8 array of LEDs and detectors, which has been demonstrated through hybrid integration onto Si CMOS with 100% yield [11, 12]), with a 1% optical coupling per through-Si link. The theoretical $32 \times 32 \times 4$ system would then use 60 kW of power for interconnect. Long wavelength (>1.2 µm) VCSELs are being vigorously developed, and one can, based upon GaAs-based VCSELs, forecast that, using a 20 µm diameter, 1 mA threshold VCSEL that a nearly 10 times improvement in the coupling efficiency for the 3D through-Si OE vertical link is probable. By assuming reasonable link operation with VCSELs at 600 Mbps with a total link power dissipation of 5 mW, and by conservatively assuming a factor of 10 improvement in the coupling efficiency, the optical interconnect power for the $32 \times 32 \times 4$ processor system would drop to approximately 250 W; a factor of nearly 400 times improvement over the HDI system.

In conclusion, HDI electrical interconnections and vertical OE through-Si links for 3D computational systems have been modeled, and compared and contrasted in the areas of latency, size, and power dissipation. Based upon SIA Roadmap projections, which indicate interconnection as a challenge area for coming generations of electronics, optoelectronics may serve a role in chip to chip interconnections for massively parallel computational systems.

![Figure 1: a) HDI Substrate Interconnect b) 3D OE Through-Si Interconnect](image-url)