Wireline and Wireless RF-Interconnect for next generation SoC systems

Sai-Wang Tam
TagArray, Inc
Palo Alto, USA

M. Frank Chang
Electrical Engineering Department, UCLA
Los Angeles, USA

Jongsun Kim
Electronic and Electrical Engineering, Hongik University
Seoul, Korea

***Abstract*— In the era of the nanometer CMOS technology, due to stringent system requirements in power, performance and other fundamental physical limitations (such as mechanical reliability, thermal constraints, overall system form factor, etc.), future SoC systems are relying more on ultra-high data rate scalable, re-configurable, highly compact and reliable interconnect fabric. To overcome such challenges, we explore the use of multiband wireline and wireless RF-Interconnect (RF-I) which can communicate simultaneously through multiple frequency bands with low power signal transmission, reconfigurable bandwidth and excellent mechanical flexibility and reliability.**

INTRODUCTION

In future VLSI system, it has been predicted chip-to-chip or on-chip interconnect to be the ultimate limitation of the overall system performance (as shown in Fig.1) in consumer electronics, high speed connector, high speed memory interface and future network-on-chip. Therefore, the interconnect system in future system on a chip (SOC) requires to have ultra-high data rates, scalable, re-configurable, highly compact and reliable interconnect fabric. In order to adopt such rapid evolution in future interconnect in both on-chip and chip-to-chip communication, we explore the use of multiband wireline and wireless RF-I which can communicate simultaneously through multiple frequency bands with low power, reconfigurable bandwidth and excellent mechanical flexibility and reliability. RF-I techniques based on CDMA or FDMA modulations have been applied to signaling technologies to resolve the problem of limited concurrency and increased latency of conventional on- and off-chip interconnects with less power consumption [2, 3, 9].

In this paper, we review recent advances in wireline and wireless RF-I in three different potential application domains, which include Network-on-Chips (NoCs), advanced memory interface and ultra-high speed contactless connectors.

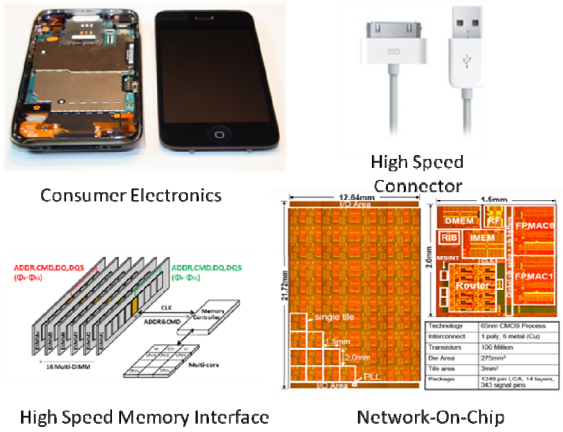


Fig 1 Interconnect becomes the ultimate limitation of the overall system performance in many systems such as consumer electronics, high speed connector, high speed memory interface and future network-on-chips

I. HOW CAN RF HELP SOC DESIGN?

One of the key benefits of the scaling of CMOS is that the switching speed of the transistor improves over each technology generation. According to ITRS [1], f_T and f_{max}, will be 600GHz and 1 THz, respectively, in 16nm CMOS technology. With the advance in CMOS mm-wave circuits, hundreds of GHz bandwidth will be available in the near future. In addition, compared with CMOS repeaters charging and discharging the wire, EM waves travel in a guided medium at the speed of light, which is about 10ps/mm on silicon substrate. The question here is: How can we utilize device f_T more efficiently for future SOC systems? The concept of RF-I [2, 3] is based on transmission of modulated RF signals which is *waves*, rather than conventional square wave voltage-mode signaling that usually requires switching energy. In the RF approach, an electro-magnetic (EM) wave

is continuously sent along the wire (treated as a transmission line) or over on-chip antenna. Data is modulated onto that carrier wave using amplitude and/or phase changes. By expanding the idea of the single carrier RF-I, it is possible to improve bandwidth efficiency using N-channel multi-carrier RF-I. In multi-carrier RF-I shown in Fig. 2, there are N mixers in the Tx. Each mixer up-converts individual base-band data streams into a specific channel. Those N distinct channels transmit N different data streams onto the same transmission line simultaneously. The total aggregate data rate (R_{Total}) through the shared transmission line equals to $R_{Total} = R_{baseband} \times N$, where the data rate of each base-band is $R_{baseband}$ and the number of virtual channels is N. A conceptual illustration of the six-carrier FDMA RF-interconnect is shown in Fig. 2.

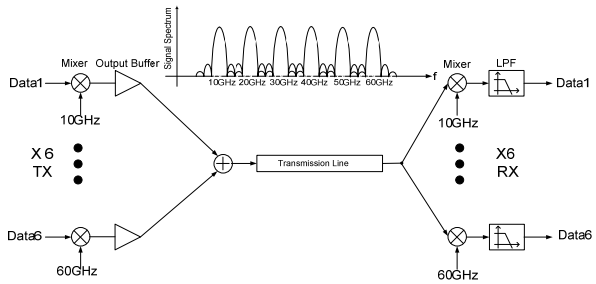


Fig 2 Conceptual schematic of the multi-band RF-Interconnect

II. RF-INTERCONNECT FOR FUTURE NETWORK-ON-CHIP (NOC)

In this section, we first illustrate the implementation of a simultaneous tri-band on-chip RF-interconnect [4] to demonstrate the feasibility of multi-band RF-interconnect for future network-on-chip. Furthermore, we proposed a micro-architectural framework [5, 6] that can be used to facilitate the exploration of scalable low power NoC architectures based on physical planning and prototyping.

The schematic of the proposed tri-band RF-I is shown in Fig. 3. The modulation scheme of each RF band is amplitude-shift keying (ASK), in which a pair of on-off switches directly modulates the RF carrier. The Baseband (BB) utilizes a low-swing capacitive coupling interconnect technique. The baseband data is transmitted and received using the common mode of the differential transmission line (TL). The transmitter and the receiver are connected by an on-chip 5mm long differential TL. In order to support simultaneous multi-band RF-I on a shared TL, RF band and BB are transmitted in differential mode and common mode, respectively. The tri-band on-chip RF-I was implemented in the 90 nm digital CMOS process. The tri-band RF-I achieves superior aggregate data rate (10Gb/s), latency (~6ps/mm) and energy per bit (0.09pJ/bit/mm and 0.125 pJ/bit/mm, for RF & BB, respectively). The measured BER across all channels is < $1x10^{-9}$.

While RF-I has dramatic potential in terms of low-latency, low-power, high-bandwidth operation, the key enabling component of RF-I for future microprocessor architectural design is re-configurability. As an example of this re-configurability, we recently proposed MORFIC (Mesh Overlaid with RF Inter-Connect), a hybrid NoC design which is composed of a traditional mesh of routers augmented with a shared pool of RF-I that can be configured as short-cuts within the mesh. From the simulation results of our in-house cycle-accurate simulator, we demonstrated a significant performance improvement of the Mesh Adaptive Shortcuts over the Mesh Baseline, an average packet latency reduction of 20-25%, through the reconfigurable RF-I [5, 6].

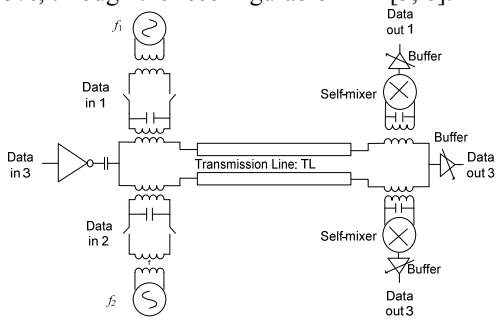


Fig 3 Schematic of the on-chip tri-band ASK RF-I

III. ADVANCED MEMORY INTERFACE

As the required aggregate data rate of the off-chip high speed I/O link such as memory interface keeps increasing, the overall power consumption and total number of I/O pins have been increasing as well, and such stringent design requirements eventually become the ultimate bottleneck of the overall system. Therefore, we proposed a dual-band low power RF-I memory interface system [7] that can support multiband and reconfigurable data communication by adding a lower energy per bit RF band on top of existing baseband high speed memory interface shown in Fig. 4. This simple prototype is implemented together with the conventional baseband (BB) mobile memory interface and RF transceivers which enable simultaneously BB and RF-band (22GHz RF carrier) communication over a shared transmission line on the printed circuit board. This advanced RF-I memory interface transceiver chip is fabricated in a 65 nm digital CMOS technology. The test board with dual chips achieves an aggregate data throughput of 8.4Gb/s/pin over a 10cm FR4 differential transmission line with energy efficiency of 2.5pJ/b.

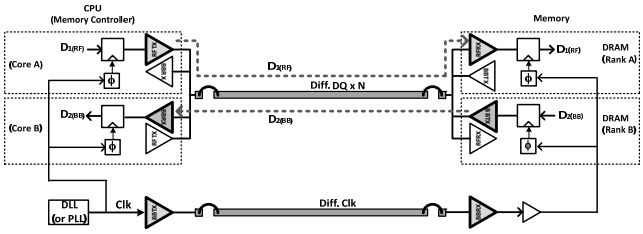


Fig. 4 Schematic of the advanced RF-I memory interface system

IV. WIRELESS RF-I: MILLI-METER-WAVE-WIRLESS-INTERCONNECET (M2W2-I)

In this section, a new wireless RF-I technique using milli-meter-wave-wireless-interconnect (M2W2-I) with asynchronous ASK modulation scheme [8] is proposed for applications in multi-gigabit ultra short distance contactless connector. M2W2-I modulates data at milli-meter wave

frequencies and transmits modulated data over the air in a very short distance (1cm or less) without any physical contact, in contrast to conventional high speed broad to broad connector, which is poor in both mechanical reliability and signal integrity. The main advantages of M2W2-I include high data rate, low power, scalable bandwidth density, no mechanical contact and simple transceiver architecture. Fig. 5 shows the schematic of a single channel M2W2-I using on-chip antenna. The M2W2-I is implemented in 65nm digital CMOS process, and the maximum data rate is 5.7Gbps with BER $<1x10^{-12}$. The carrier frequency is 63GHz, and the energy per bit is 8pJ/bit. The maximum measured bandwidth density and maximum communication distance is 7.25Gbps/mm and up to 3mm, respectively.

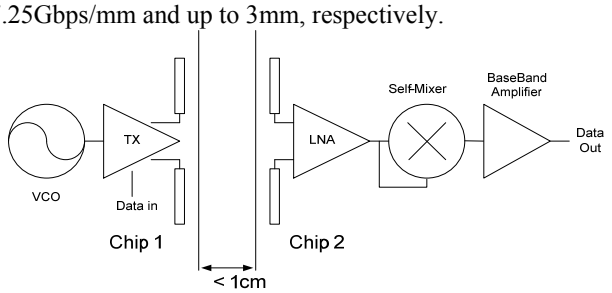
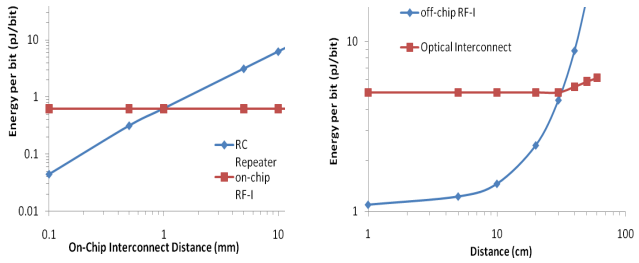


Fig 5. Schematic of a single channel M2W2-I using on-chip antenna

V. FUTURE RESEARCH DIRECTIONS

In this final section, we compare the performance and the proper communication range for all three types of interconnects, including the traditional repeater-based wire bus, the RF-I and the optical interconnect. As CMOS continues to scale toward 16 nm, traditional on-chip RC repeated wires are more suitable for local interconnects with short communication distance due to further increased physical density through the use of minimum-feature-width metal wires. Fig.6(a) illustrates the projected power/performance of both repeater-based RC wires with optimal delay and RF-I with a 16nm CMOS process. Under approximately 1mm, the RC repeater is able to provide superior energy efficient communication, but beyond 1mm, the repeater buffers become less efficient than those of RF-I. The RF-I is expected to maintain its performance advantages for global interconnect on-chip due to its total compatibility with the CMOS technology. However, can it maintain the same superiority to an extended distance off-chip? Especially, to what range can it compete with the optical interconnect which is clearly superior for longer-distance communications? We offer the answer to those questions by comparing the energy efficiency between the off-chip RF-I and optical interconnect in Fig. 6(b). Accordingly, the RF-I actually exhibits better energy efficiency at mid-range distances of 30 cm or below. As the communication distance increases, RF-I energy efficiency decreases rapidly due to the excessive power required to compensate for the severe loss from the on-board transmission lines, while the power consumption of optical interconnect remains almost constant. Therefore, despite substantial disadvantages in integration and cost, the optical interconnect becomes more beneficial at interconnect distances beyond 30 cm. That is to say, in between traditional RC repeater buffer and optical interconnects, there is an obvious technology gap for achieving cost/performance-effective communications in mid-distance range from a few millimeters to several tens of centimeters. The CMOS compatible RF-I may be the right technology to fill in such a technology gap, as shown in Fig. 7.



(a) (b)

Fig 6. (a) RF-I will crossover the energy efficient curve of the RC repeater, and become more energy efficient above a 1mm interconnect distance; (b) RF-I has much better energy efficiency in the mid-range distance of 30 cm

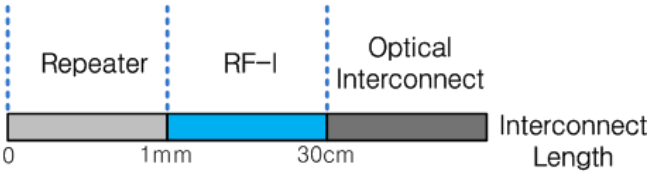


Fig. 7. Communication range versus interconnect technologies as CMOS process continuously to scale toward 16nm.

Acknowledgments

This work (Grant No. 000419970110) was supported in part by Business for Cooperative R&D between Industry, Academy and Research Institute funded Korea Small and medium Business Administration in 2010.

References

[1] "ITRS," Semiconductor Industry Association, 2010

[2] M.F. Chang, et al., "RF/Wireless Interconnect for Inter- and Intra-chip Communications," Proceedings of the IEEE, April 2001

[3] J Ko, et al, "An RF/Baseband FDMA-Interconnect Transceiver for Reconfigurable Multiple Access Chip-to-Chip Communication," ISSCC 2005

[4] Sai-Wang Tam, et al., "A Simultaneous Tri-band On-Chip RF-Interconnect for Future Network-on-Chip," VLSI Circuits Symp, 2009

[5] M.F. Chang, et al, "CMP Network-on-Chip Overlaid With Multi-Band RF-Interconnect," IEEE HPCA Symp. , Feb. 2008

[6] Sai-Wang Tam, et al, "RF-Interconnect for future network-on-chip, Low Power Network-on-Chip," (Book Chapter 10), Springer, 2010

[7] Gyung-Su Byun, et al, "An 8.4Gb/s 2.5pJ/b Mobile Memory I/O Interface Using Bi-directional and Simultaneous Dual (Base+RF)-Band Signaling,", ISSCC 2011

[8] Sai-Wang Tam, et al, "A Multi Gbps Short-range Milli-Meter-Wave-Wireless-Interconnect (M2W2-Interconnect)," submitted to IEEE J. of Solid-State Circuits

[9] J. Kim, et al, "Design of an interconnect architecture and signaling technology for parallelism in communication," IEEE Trans. VLSI Systems, Vol.15, no. 8, pp.881-894, Aug. 2007