USC SDR, An Easy-to-program, High Data Rate, Real Time Software Radio Platform

Horia Vlad Balan   Marcelo Segura   Suvil Deora   Antonios Michaloliakos
Ryan Rogalin   Konstantinos Psounis   Giuseppe Caire
University of Southern California
{hbalan,mjsegura,deora,michalol,rogalin,kpsounis,caire}@usc.edu

ABSTRACT

We present USC SDR, a wireless platform designed for easy-to-program, high data rate, real time wireless experimentation. The design of our platform aims at removing most of the bottlenecks encountered in the design of current software radio architectures, e.g. the requirement to program new schemes in an FPGA, and the difficulty to run real-time experiments for a long time. The architecture combines generic PCI FPGA development boards with radio front-ends built as self-sufficient daughterboards. The daughterboards are connected to the FPGAs, which in turn are plugged into the PCIE slots of a general-purpose server.

Interestingly, the connection of the daughterboards to the FPGA cards is implemented through a standard FMC (FPGA Mezzanine Card) interface, such that the same RF front-end can be reused with future FPGA generations. In this way, USC SDR is not limited to a specific FPGA choice and does not require a complete re-design in order to accommodate for future more powerful FPGAs.

The hardware is supported by a real-time software architecture where signal processing tasks, PHY and MAC layer algorithms can be programmed as user-level applications. As an example, we will showcase a massive MIMO testbed built using a single server with multiple PCIE slots.

Categories and Subject Descriptors

C.2.1 [**Computer System Organization**]: Computer Communication Networks

General Terms

Design, Experimentation, Performance

Keywords

Software Radio, Wireless

1. INTRODUCTION

USC SDR is a research and education platform for real-time wireless multi-terminal communication systems, allowing the rapid prototyping of cutting-edge wireless technology while avoiding the limitations of current hardware/software solutions. The development of this particular platform has been prompted by the observation that in the past years software radios have evolved into convenient tools for experimental research into wireless system design, in academic and industrial settings. The fast prototyping/fast experimentation approach to system design that software radios offer mitigates two important problems in advancing and validating new technologies: industrial research and development is strongly standard-oriented, and therefore innovation occurs at the slow pace of standardization; academic research is constrained by the limitations of currently available experimental platforms and resources are more limited, such that demonstration of new fundamental ideas is impaired by the complexity of the development effort associated with the experimental hardware. Creating powerful, easy-to-use platforms is crucial for the advancement of fast-paced wireless systems research. If the ability to quickly adapt the platform to the requirements of various problems is also brought into the fold, the avenues for research applications multiply vastly.

On the other hand, wireless is facing critical, transformative and exciting times. Across the globe, public policy calls for unleashing the wireless broadband revolution, acknowledging the fact that wireless data traffic will explode in the next few years, under the pressure of advanced multimedia devices (smartphones and tablets) and of the forthcoming wireless-enabled infrastructure. Yet, short-term solutions such as releasing TV white spaces, while still relying on conventional cellular technology, are not going to meet this challenge. It is widely recognized that *new paradigmatic solutions must be developed*. While theory has made very significant and fundamental advances in the past few years, we claim that, ultimately, such novel ideas must be translated into technology, which needs to be *demonstrated*.

USC SDR aims at providing a cost-effective and flexible solution to enable such advances, and eventually help meet the *wireless spectrum crunch* challenge. For scalability, we base our platform on multiple FPGA cards directly connected to the PCIE bus of a host (e.g., a multi-core server to implement a base station, or a smaller machine to implement a mobile node). Each FPGA card can support multiple RF daughterboards, which are custom-designed for this purpose and include a newly designed integrated RF chip

capable of extensive frequency flexibility. This unique feature will be instrumental for the implementation of *White Space*-oriented dynamic spectrum sharing schemes. The proposed architecture transfers the baseband complex samples directly to/from the host's RAM (direct memory access), thus avoiding the bottlenecks of Ethernet or USB interfaces. While the most computationally demanding tasks can be executed by the FPGAs, most of the physical layer signal processing and the upper layer protocols are performed by the host's processor, making the development of real-time operations not more complicated than writing and compiling Matlab/C code.

2.   RELATED WORK

Established software radio research platforms include USRP [8], WARP [15], SORA [13] and OpenAir [1] and National Instruments' PXI platform[14].

USRP provides an advanced platform for MIMO experimentation, being capable of sending and receiving synchronously from a large number of front-ends. The interface to the host computer is realized over Ethernet, somehow limiting the real-time responsiveness of the platform. The main advantage of USRP is its perfect match to the GNU Radio software suite, allowing for easy application development.

WARP is a self-contained platform for MIMO experimentation, offering up to four commonly clocked RF front-ends. The interfacing to the host is again done over Ethernet. While it is possible to extend the platform for multiple host simultaneous transmission and for high data transfer rates through the Ethernet ports, these extensions are somehow more involved than in the case of USRP. However, the high number of logic slices available on the WARP FPGA make it a good candidate for very constrained real-time signal processing applications, such as distributed MIMO [4] or massive MIMO [17].

SORA and OpenAir are quite similar to our own platform design. The main difference is that both platforms are developed using custom FPGA board implementations, while we have chosen to interface the RF front-ends with generic FPGA boards.

PXI is a very advanced data-acquisition and transmission platform, however its cost makes it prohibitively expensive to most academic institutions. In comparison, in our case the cost of two daughterboards are comparable to the cost of the generic FPGA board (about $ 3500) and dominated by the cost of the computational back-end (i.e. the server machine). We believe that a typical installation of a massive MIMO system with four Xilinx FPGA cards support up to 16 antennas can be realized for about $20.000 to $30.000, depending on the availability of academic licensing terms for parts of the system.

3.   PLATFORM DESCRIPTION

The USC SDR platform is based on combining specially designed RF front-ends with generic FPGA boards configured as PCIE cards. The FPGA boards can be hosted in servers, with sufficient processing power needed by a multiple-channel SDR implementation of a massive MIMO base station. We envisage 16 antennas per server, with up to 4 servers connected into a cluster, for a total maximum of 64 jointly processed antennas. Also, the same configuration can be installed on smaller and inexpensive hosts in order to im-

plement user terminals with one or two antennas, emulating typical smartphone or tablet setup. The main components of our system are generic FPGA boards, custom-made RF daughterboards and servers, as detailed in the following.

The FPGA boards. We use generic Xilinx evaluation boards instead of dedicated designs created for SDR applications, as used in all the existing solutions such as WARP [15], USRP [8] or SORA [13]. This choice is motivated by the fact that, thanks to Moore's law and the large volumes of production, generic evaluation boards are available at significantly lower prices and are evolve rapidly in time, from generation to generation. Hence, USC SDR is be able to leverage the evolution of FPGA technology, avoiding future bottlenecks. As a matter of fact, there are almost no advantages in hosting an SDR design on a dedicated board. The reason why state of the art platforms have chosen this approach can be explained by the fact that, in the past, generic FPGA evaluation boards could not easily host dedicated RF daughterboards. However, this is now fully supported by the latest generation of Xilinx evaluation boards, through a standardized interface that effectively decouples the RF components from the baseband signal processing components, allowing a custom-designed RF daughterboard to be compatible with future generations of generic FPGA hosting boards. This standardized interface, called FMC (FPGA Mezzanine Card), was developed by Xilinx for the purpose of fostering the development of application-tailored, small daughterboards. Thanks to this standardization, the daughterboards are guaranteed to be compatible with the next generations of FPGA components, extending the lifetime and the relevance of the daughterboard design. In fact, the appearance of the FMC has spawned a whole cottage industry of daughterboard design for various DSP applications.

We have acquired Xilinx VC707 FPGA boards, hosting the latest generation of Virtex 7 FPGA with 2800 DSP slices (significantly more than any other software radio currently available) for our initial development. For the above reasons, we anticipate no problems in moving to newer generations of FPGAs, when they become available.

The Daughterboards. A central component of the USC SDR platform is a FMC-standard daughterboard hosting an RF chip custom-designed for LTE-compatible applications. The RF chip can operate over a channel bandwidth of 40 MHz at any carrier frequency from 50 MHz to 6 GHz. This key feature enables research, development and prototyping of dynamic spectrum allocation schemes and Cognitive Radio technologies. The board has been designed in-house, manufactured, tested and integrated with the FPGA digital design. This RF chip accepts an external clock reference in order to support multiple synchronous RF chains, and is programmable to a very large degree over a serial interface, achieving a large degree of flexibility and real-time reconfigurability.

The USC SDR daughterboard design includes the mentioned chip, ADC and DAC, all clocked to either a local oscillator or an external clock source. The board exports the sampling clock to the FPGA synchronous DSP operations. USC SDR allows a number of such boards to be driven by a common clock reference, with the RF front-ends, ADCs, DACs and signal processing designs in the FPGAs all operating phase coherently and timing synchronous. In this way, the FPGA boards would present to the hosting server mul-
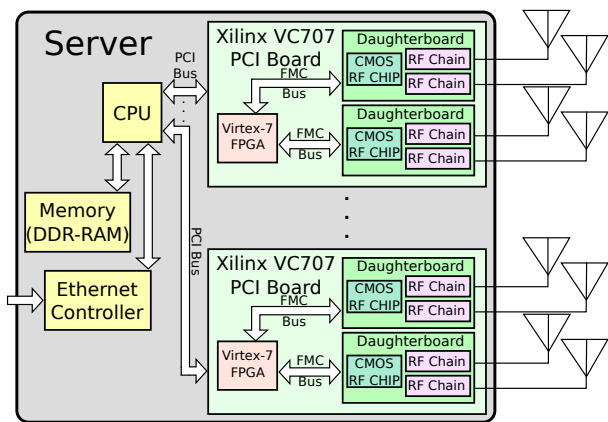
Figure 1: The USC SDR architecture leverages existing commercially- and institutionally-available hardware to reduce development time and costs. The use of general-purpose hardware allows USC SDR to scale to massive levels at a lower cost than existing SDR platforms.

tiple phase-coherent channels with commonly timestamped waveforms, as required in a MIMO system.

Each VC 707 board can host two daughterboards. Since the RF chip supports two parallel channels, it results that we could process up to four receive channels and four transmit channels per FPGA board.

The hosting servers. In order to implement a base station with up to 64 antennas, We acquired servers with up to 6 PCIE express slots each and a sufficient number of processing cores for receiving and sending the waveforms in real time. Previous research done with the SORA platform [13] has shown that 802.11a/g signals can be processed in real-time using two processor cores at 2.66 GHz. Since each machine would be processing up to 24 channels simultaneously, the number of required cores is 48. Such server machines are commonly available today. The design of the data transfer architecture tries to ensure that such a scaling is possible by avoiding all possible sources of bottlenecks. In particular, the PCIE transfers to the FPGA boards are implemented using scatter-gather transfers, which removes the necessity for copying the data to continuous kernel memory regions before DMA operations and allows the different processors and PCIE endpoints to operate independently. Benchmarking the throughput scalability brought by this architectural choice is an important point in the evaluation of our platform and we plan to publish the relevant data as soon as it becomes available. Initially, we have hosted the PCIE boards outside the server chassis and linked them with PCIE cables in order to avoid any spacing/interference problems. If the interference effects inside the chassis prove to be small and easily solvable, the boards will be installed in the chassis itself.

We are developing a Linux-based software stack consisting of a driver for the FPGA board which will move the waveforms samples to and from the main host RAM and expose them to userspace as a memory mapped device. The waveforms can thus be directly accessed by userspace programs through a zero-copy mechanism. The userspace programs

can be run without preemption (i.e. full real-time) if a Linux kernel scheduler extension such as RTLinux is used.

4. APPLICATIONS

In the past years our team has been active in the field of multiuser MIMO and distributed multiuser MIMO transmission. To support our work we have developed a large amount of Matlab and C code that implements OFDM modulation and MIMO OFDM transmission. We plan to reuse that body of code as the basis of physical layer applications realized on our platform. The platform offers new possibilities for experimenting at the MAC layer. The MAC layer functionality (some of which, such as ACKs, requires FPGA logic support) remains to be implemented on a per-application basis.

Massive MIMO in a multi-cell environment: One of the most remarkable recent success stories in information theory is the determination of the capacity region of the MIMO broadcast channel[5, 21, 10, 22]. Key to achieve the promised gains [19, 20, 5] is that sufficiently accurate channel state information is provided to the transmitter [7]. The possibility of providing such information depends critically on the predictability of the underlying wireless channel [18].

In [12], a very large number of antennas at each base station is considered and TDD reciprocity is exploited in order to estimate the channel from the training symbols sent by the users in the uplink direction. In the limit of infinite base station antennas per user, this "massive" MIMO scheme yields very attractive performance. In addition, such performance can be achieved by the simplest of all linear precoders: the so-called conjugate beamforming. The recent work [9] provides a more refined theoretical analysis of the performance and system optimization of a massive MIMO systems in the regime of large but finite number of antennas per user. In particular, it shows that in this realistic regime the choice of the multiuser MIMO precoding scheme is *very important*, and that the gain of regularized zero-forcing beamforming over conjugate beamforming is very significant, unless the number of base station antennas per user grows to impractically large values.

In the recent work [17], a team at Rice university demonstrated the implementation of a 64-antenna base station based on WARP called "Argos". Conjugate beamforming is a reasonable option for real-time operations, since transferring the estimated channel coefficients from the uplink training phase, computing the precoder and transferring back the precoded signals to all the antennas in real-time requires significant backplane throughput. However, it is known that a 16 antenna base station with regularized zero-forcing precoding can significantly outperform a 64-antennas base station with conjugate beamforming [9]. We plan to use USC SDR in order to implement multiple massive MIMO base stations, with TDD reciprocity. Thanks to the high backplane capacity of USC SDR, we can demonstrate centralized multiuser MIMO precoding schemes in real time (such as zero-forcing precoding) on a single base station and, in the future, in a multi-cell environment with inter-cell interference.

In an early demonstration, we plan to perform multiuser transmission aided by downlink channel estimation in order to evidentiate the above gains. To this end, we plan to use a server with four to eight antennas and a number of single-antenna clients.

Distributed multiuser MIMO: Distributed multiuser MIMO (also known as "Network MIMO" or "Cooperative Multi-Point") consists of coordinating a large number of antennas at different and spatially separated nodes, which can share data at high speed through a wired backhaul, but without common sampling and RF clock and common timing information at the symbol level. While the theory of distributed multiuser MIMO is well developed for ideal synchronous models (see for example [23, 11]), the key implementation difficulty here consists to provide frequency and timing synchronism at a large number of spatially separated transmitters, each driven by its own clock, in a scalable and robust way. We have significant work in progress in the theoretical study and SDR implementation of distributed multiuser MIMO (see [4, 3, 16]). USC SDR shall enable us to demonstrate a truly real-time large scale distributed multiuser MIMO setting, implementing the synchronization schemes [16].

Cognitive MIMO radio systems: Cognitive radio is a general paradigm advocating dynamic spectrum reuse, such that an unprecedented density of "localized" links can co-exist by intelligently sharing the same frequency band [6]. As this "super-dense" spatial reuse of the wireless spectrum is the key to meet the spectrum crunch challenge of the nest 5-10 years. We will implement a cognitive scenario formed by a "legacy" base station (modeling the primary system) and several point-to-point MIMO cognitive links (secondary system), trying to establish their point to point connections by exploiting the holes in the time-frequency and space domain. In particular, as far as exploiting the antenna domain is concerned, we aim at demonstrating our recently proposed *Reverse TDD* scheme [2] showing that, when the secondary nodes have multiple antennas, they can sustain high transmission rates without significantly impacting the primary system performance.

5. DEMO DESCRIPTION

We will demonstrate the functioning USC SDR platform capable of transmitting and receiving waveforms over multiple channels and exposing them to user-level software for real-time processing.

To our knowledge, USC SDR offers a unique combination of low-cost, high availability hardware foundation, high data transfer rates, real-time programability at the user-level, MIMO capabilities and ease of use. We intend to demonstrate all these aspects of the platform through example applications.

Our software stack allows us to perform real-time downlink transmission based on a tight channel estimation loop. While the timing of the current software does not allow for the implementation of a CSMA/CA MAC protocol, we will demonstrate real-time channel adaptation in the multiuser MIMO downlink.

We plan to use in the demonstration a single FPGA card with two RF daughterboards and four antennas, operating over 20 MHz wide channels.

6. REFERENCES

[1] The OpenAir interface website.

[2] A. Adhikary and G. Caire. On the coexistence of macrocell spatial multiplexing and cognitive femtocells. *arXiv preprint arXiv:1202.1449*, 2012.

[3] H. V. Balan, R. Rogalin, A. Michaloliakos, K. Psounis, and G. Caire. AirSync: Enabling Distributed Multiuser MIMO with Full Spatial Multiplexing. Submitted to IEEE Transactions on Networking, July 2012.

[4] H. V. Balan, R. Rogalin, A. Michaloliakos, K. Psounis, and G. Caire. Achieving high data rates in a distributed mimo system. In *Proceedings of the 18th annual international conference on Mobile computing and networking*, Mobicom '12, pages 41–52, New York, NY, USA, 2012. ACM.

[5] G. Caire and S. Shamai. On the achievable throughput of a multiantenna Gaussian broadcast channel. *IEEE Trans. Info. Theory*, 49(7):1691–1706, July 2003.

[6] N. Devroye, P. Mitran, and V. Tarokh. Achievable rates in cognitive radio channels. *IEEE Trans. Info. Theory*, 52(5):1813–1827, May 2006.

[7] P. Ding, D. J. Love, and M. D. Zoltowski. Multiple antenna broadcast channels with shape feedback and limited feedback. *IEEE Trans. Signal Processing*, 55(7):3417–3428, July 2007.

[8] Ettus Research LLC. Universal Software Radio Peripheral. In *http://www.ettus.com/index.html*.

[9] H. Hoon, G. Caire, H. C. Papadopoulos, and S. A. Ramprashad. Achieving "Massive MIMO" Spectral Efficiency with a Not-so-Large Number of Antennas. *IEEE Trans. Wireless Commun.*, 11(9):3226 – 3239, 2012.

[10] N. Jindal and A. Goldsmith. Dirty-paper coding versus TDMA for MIMO broadcast channels. *IEEE Trans. Info. Theory*, 51(5):1783–1794, May 2005.

[11] P. Marsch and G. Fettweis, editors. *"Coordinated Multi-Point in Mobile Communications"*. Cambridge University Press, 2011.

[12] T. Marzetta. Noncooperative cellular wireless with unlimited numbers of base station antennas. *Wireless Communications, IEEE Transactions on*, 9(11):3590–3600, november 2010.

[13] Microsoft Corporation. Sora - Microsoft Research. In *http://research.microsoft.com/en-us/projects/sora/*.

[14] National Instruments. National instruments PXI platform.

[15] Rice University. Wireless open-Access Research Platform. In *http://warp.rice.edu/trac/*.

[16] R. Rogalin, O. Bursalioglu, H. Papadopoulos, G. Caire, A. Molisch, A. Michaloliakos, H. V. Balan, and K. Psounis. Scalable Synchronization and Reciprocity Calibration for Distributed Multiuser MIMO.

[17] C. Shepard, H. Yu, N. Anand, L. Li, T. Marzetta, R. Yang, and L. Zhong. Argos: Practical many-antenna base stations. *SIGCOMM Comput. Commun. Rev.*, Aug. 2012.

[18] H. Shirani-Mehr and G. Caire. Parametric channel estimation and prediction, with applications to channel state feedback for MIMO downlink schemes. In *Proc. Asilomar Conf. Signals, Systems, and Computers*, Oct. 2008.

[19] S. S. (Shitz) and A. D. Wyner. Information-theoretic considerations for symmetric, cellular, multiple-access

fading channels- Part I. *IEEE Trans. Info. Theory*, 43(6):1877–1894, Nov. 1997.

[20] S. S. (Shitz) and A. D. Wyner. Information-theoretic considerations for symmetric, cellular, multiple-access fading channels- Part II. *IEEE Trans. Info. Theory*, 43(6):1895–1911, Nov. 1997.

[21] P. Viswanath and D. N. C. Tse. Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality. *IEEE Trans. Info. Theory*, 49(8):1912–1921, Aug. 2003.

[22] H. Weingarten, Y. Steinberg, and S. S. (Shitz). The capacity region of the Gaussian multiple-input multiple-output broadcast channel. *IEEE Trans. Info. Theory*, 52(9):3936–3964, Sep. 2006.

[23] A. D. Wyner. Shannon-Theoretic approach to a Gaussian cellular multiple-access channel. *IEEE Trans. Info. Theory*, 40(6):1713–1727, Nov. 1994.