

Wireless Interconnect for Board and Chip Level

Gerhard P. Fettweis, Najeeb ul Hassan, Lukas Landau, and Erik Fischer
Vodafone Chair Mobile Communications Systems
Dresden University of Technology (TU Dresden), 01062 Dresden, Germany
Email: {fettweis, najeeb.ul.hassan, lukas.landau, erik.fischer}@ifn.et.tu-dresden.de

Abstract—Electronic systems of the future require a very high bandwidth communications infrastructure within the system. This way the massive amount of compute power which will be available can be inter-connected to realize future powerful advanced electronic systems. Today, electronic inter-connects between 3D chip-stacks, as well as intra-connects within 3D chip-stacks are approaching data rates of 100 Gbit/s soon. Hence, the question to be answered is how to efficiently design the communications infrastructure which will be within electronic systems. Within this paper approaches and results for building this infrastructure for future electronics are addressed.

I. INTRODUCTION

Future computing platforms will be dominated by massive parallelism in number of processing elements (processors of any kind). Today we are reaching on the order of 1000 processors on a single die for GPU implementations [1]. And today we see already multiple dies being stacked into a 3D chip-stack for high capacity flash memory realization [2]. As we will see a continuation of Moore's Law to 7nm technology, combined with the stacking of chips becoming mainstream, and further increase in building higher chip-stacks, it is foreseeable that the number of processors in a chip-stack package reaching far beyond multiple million elements. This creates the challenge of building a highly efficient and high-bandwidth intra-connect for this massive amount of processors in a chip-stack.

Many instances of this massive number of processors in a chip-stack package will be placed on a printed circuit board, e.g. of size 10cm x 10cm. Assuming 4-5 boards to be placed in a 1 liter box, a billion processors in a liter can be foreseen, which is an extraordinary large number in terms of today's systems. When connecting these up-to-a-billion processors via multiple boards with a backplane bus system, this requires massive bandwidth and switching capabilities. Again, as in chip-stack intra-connect, the major challenge lies in building a highly efficient interconnect architecture which can carry the bandwidth, enable switching and connectivity, as well as data rate requirements.

In all cases, designing communications links for delivering extreme data rates is of utmost importance:

- For intra-connects within a 3D chip-stack.
- For inter-connects between chip-stacks/packages on a board.

This work has been supported in part by the DFG in the CRC 912 "Highly Adaptive Energy-Efficient Computing" and European Social Fund in the framework of the Young Investigators Group 3D Chip-Stack Intra-connects".

978-3-9815370-0-0/DATE13/©2013 EDAA

- For the backplane of a multi-board system.

The backplane, as an aggregator of traffic and infrastructure provider of multiple simultaneously active connections, is a serious bottleneck for building systems of the future. Hence, we propose to take the load off the backplane by providing direct wireless links between boards, from chip-stack to chip-stack. These beams shall be using beam-steering antennas at carrier frequencies beyond 200 GHz. In this case e.g. a 4x4 antenna array can be realized in a 2mmx2mm real estate. For ensuring a best coupling out of the electro-magnetic wave, we propose at this time to use the interposer as a carrier for the arrays. This way the interposer can be designed with a permeability to achieve best coupling out of the wave. Within a 3D chip-stack multiple alternatives exist for communicating between the different chiplets, of which also at least two wireless alternatives exist:

- Inductive coupling.
- Capacitive coupling.

Each connection must be able to carry high data traffic. We propose today that we must develop solutions to achieve at least 100 Gbit/s, as e.g. [3]. In the coming years this data rate per link needs to be increased into the Tbit/s range.

In this paper we first analyze the wireless board-to-board link design challenge at a 200 GHz carrier frequency range. After measuring and calculating the link budget, targets which have to be met have been defined. The intra-connect within a 3D chip stack is addressed next, showing that a careful design of the analog/digital conversion needs to be carried out, to meet a very low power consumption target. The network design within a 3D chip-stack is addressed thereafter. And finally, new results on very low latency error correction coding for inter/intra-connects are presented.

II. LINK BUDGET FOR BOARD-TO-BOARD COMMUNICATIONS

Wireless board-to-board communications requires no routing delay and less material, which leads to spatial relaxation. Furthermore it is in general more flexible as compared to conventional communication methods in a system with multiple printed circuit boards (PCBs). In this section, we consider a scenario where two printed circuit boards are placed in parallel. Both PCBs are equipped with multiple wireless communication nodes. The board-to-board channel has been measured between 220-245 GHz. This data is used to derive a link budget, essentially for the design of a wireless link.

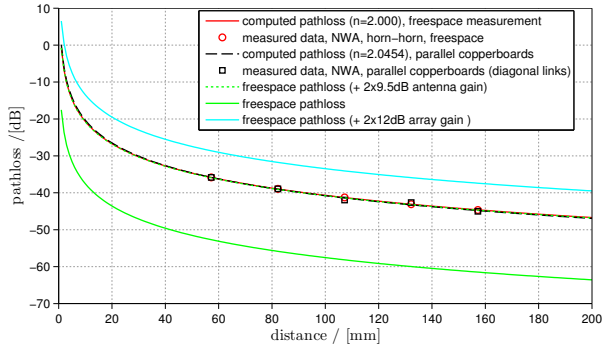


Fig. 1. Theoretical pathloss and measurement data from board-to-board communications.

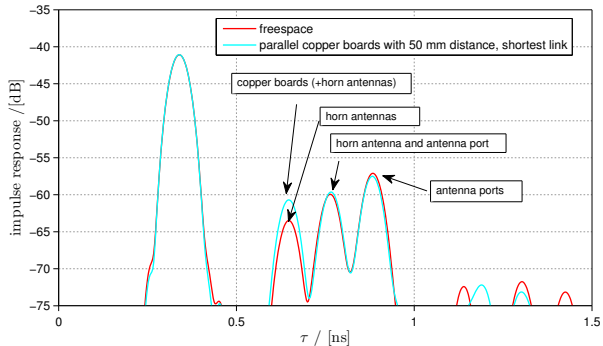


Fig. 2. Impulse response for a distance of 50mm antenna distance, freespace versus parallel copper boards.

A. Measurements with the Vector Network Analyser

For the measurements the network analyser R&S ZVA24 has been used with an extension for the frequency range between 220 GHz and 245 GHz. The channel is measured in frequency domain with 4096 samples. The system is calibrated with the direct connection of the waveguides. For the measurements, standard gain horn antennas have been installed on both measurement ports which provide approximately 10 dB gain at the considered frequency range. The distance between the measurement ports is controlled via a stepping motor. Two scenarios are considered:

- In the first setup, we consider freespace measurements with absorber material at the ground, for different distances. The purpose of this measurement is to identify the effective phase center and the effective antenna gain.
- In the second scenario, copper boards are included. This represents the worst-case of a printed circuit board. Notches are prepared for inserting the horn antennas. The distance between the two boards is fixed as 50mm, which shall be a lower bound on a board distance. Diagonal communications are modeled by a rotation of the boards on its z-axis, which also corresponds to different distances of the measurement ports.

Analysing the corresponding impulse response, obtained by applying discrete Fourier transformation leads to a unique identification of the reflecting objects. We conclude that the reflections presented in Fig. 2 are always at least 15 dB below the main signal path (line of sight), where we do

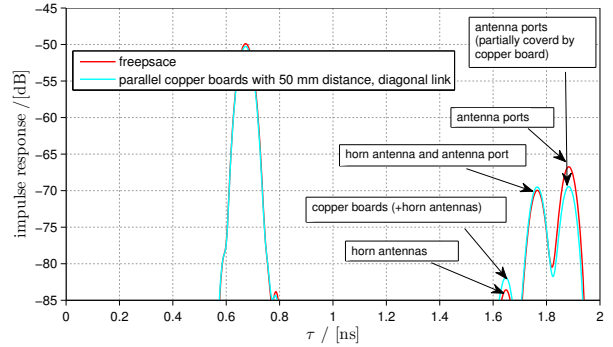


Fig. 3. Impulse response for a distance of 150mm antenna distance, freespace versus parallel copper boards (diagonal link).

TABLE I
LINK BUDGET PARAMETERS FOR BOARD-TO-BOARD COMMUNICATIONS.

	Unit	Value
RX noise figure	dB	10
Path loss exponent	-	2
Path loss for shortest link 0.1m (232.5 GHz)	dB	59.8
Path loss for largest link 0.3m (232.5 GHz)	dB	69.3
Array gain	dB	12
Butler matrix inaccuracy	dB	5
Polarization mismatch	dB	3
Implementation loss	dB	5
RX temperature	K	323

not distinguish between copper plate and the measurement equipment itself. This motivates us to have a detailed look at the line of sight component and the evaluation of a simplified freespace pathloss model which can be represented by

$$PL_d[dB] = PL_{d_0}[dB] + 10 n \log_{10} \left(\frac{d}{d_0} \right), \quad (1)$$

where d is the distance, PL_{d_0} is the reference path loss at $d = d_0$ and n is the pathloss exponent. After applying an effective phase center of the antennas and an antenna gain of 9.5 dB it can be seen in Fig. 1 that the pathloss model is in line with the measurements and especially also with the measurements including the copper boards.

These results shall be a careful justification for our pathloss model assumption, that will be used for the following link budget calculation.

B. Link Budget

For the board-to-board communication with multiple communication nodes on each board, we consider the extreme cases which are given by the ahead link (100mm) and the diagonal link (300mm). It is considered that each communication node uses a 4-by-4 antenna array, this corresponds to an array gain of each 12 dB for the transmitter and the receiver. We distinguish between beamforming/beamsteering, where we refer to the discrete realization of the beamforming vector investigated in [4] and butler matrix realization as a complexity trade-off which is investigated in [5]. In Table I the link budget parameters are summarized which are similar to those in [6]. In order to obtain wireless connections with data rates up to 100 Gbit/s (using dual polarization) the bandwidth is chosen as 25 GHz. Figure 4 shows the required transmit energy according a target SNR at the receiver, where it is

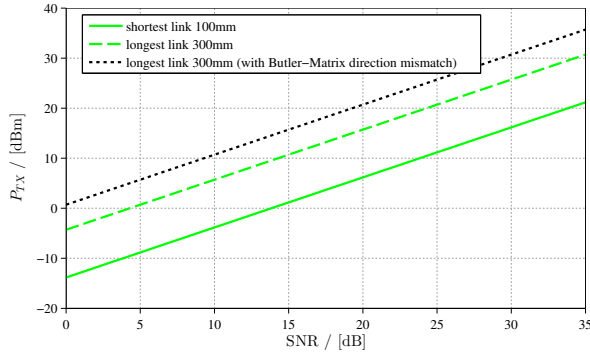


Fig. 4. Required transmit power for a desired SNR at the receiver.

assumed that only the worst-case links suffer from the butler matrix realization.

III. BANDWIDTH- AND ENERGY EFFICIENT MULTIGIGABIT/S COMMUNICATIONS BASED ON ONE BIT OVERSAMPLING RECEIVERS

When considering Multigigabit/s communication speeds over a short distance, the analog-to-digital conversion requires the main part of the total energy consumption. As a conclusion the considered conversion resolution has to be chosen as low as possible in order to save energy. To obtain a high spectral efficiency, advanced communication methods have to be applied. In this section we introduce an alternative scheme which is based on a simple one-bit oversampling receiver architecture [7]. When including an optimized intersymbol-interference (ISI) it can be shown that the information rate increases significantly [8].

For our investigations we have considered a regular 4-amplitude shift keying (ASK) modulation scheme and we found 5-fold oversampling as the smallest sampling rate, which enables unique detection. The investigated channel is the additive white Gaussian noise (AWGN) channel, which could be the discussed board-to-board channel. For simplicity the noise samples are considered to be uncorrelated within the oversampling vector. The ISI is represented by a linear filter which can overlap with another symbol. We allow for the design of this filter and proposed different strategies for different receiver architectures. On one hand we consider symbol-by-symbol detection where the ISI is an arbitrary distortion from the receiver point of view, which appears similar to dithering. For this case we use the information rate directly as the objective for the filter design optimization illustrated in Fig. 5(b).

On the other hand, it has been shown that it is beneficial to consider sequence estimation where the linear combination introduced by the ISI can be exploited even better. For this case we propose the design which maximizes the information rate shown in Fig. 5(c). We also propose a suboptimal filter design which is not based on the noise characteristics which might be unknown. In this case the information rate can't be computed and therefore the design is based on the unique detection property in the noise free case shown in Fig. 5(d).

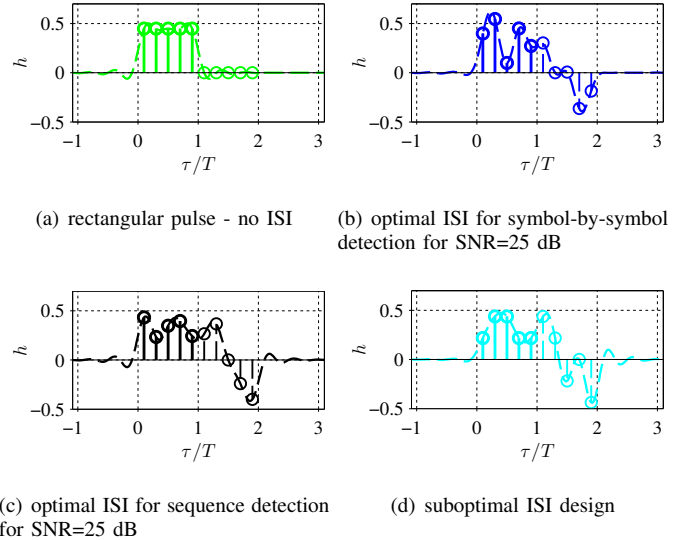


Fig. 5. Impulse response for different ISI filter designs.

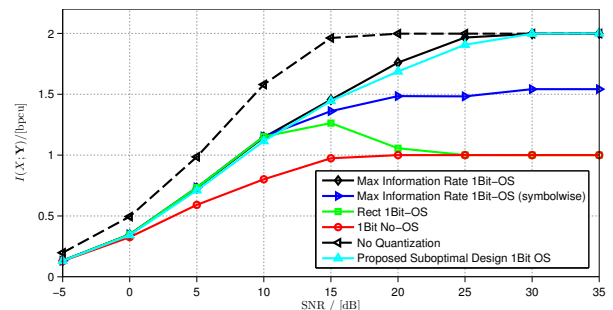


Fig. 6. Information rates considering 4-ASK communications; Comparison of different pulse designs for 5-fold oversampling and one bit quantization at the receiver.

We have compared our results with the ISI free case corresponding to the rectangular pulse [7]. Also, we consider two reference cases where no oversampling and no quantizations taken into account. Our results in Fig. 6 indicate a significant improvement of information rate when considering intersymbol-interference and especially when considering sequence estimation.

IV. 3D NiCS: A TOPOLOGY FOR FUTURE MANY-CORE SYSTEMS

In the last few years, a technology emerges to close the gap between today's multi-processor system-on-chips (SoCs) and future many-core SoCs [9], which implement thousands of processors, memories and interfaces on a single chip. The technology is called three-dimensional (3D) Network-in-Chip-Stack (NiCS) and allows the vertical stacking of multiple chips using, e.g., through silicon vias (TSVs), optical links or inductive or capacitive coupling [10]. Moreover, wireless chip-to-chip communication can provide a very flexible (even dynamic) solution for the interconnection. 3D NiCS enables a natural extension of the well-known concept of network-on-chip (NoC) [11] [12] for the interconnection of a large number of processors by exploiting the third dimension. Therefore, a high degree of freedom is provided for topology selection. This

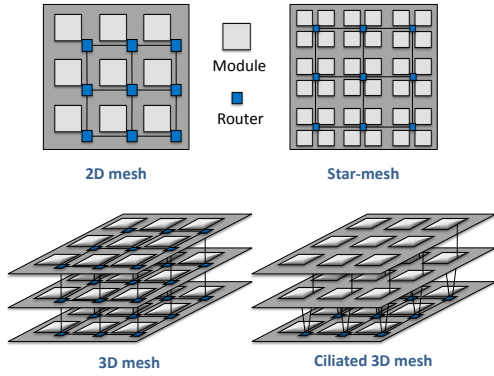


Fig. 7. Topology types: 2D mesh, star-mesh, 3D mesh and ciliated 3D mesh.

is especially the case, if wireless chip-to-chip communication is employed. New topologies can be explored that were not feasible or inefficient to be realized on a two-dimensional plane due to wiring constrains. Many 3D topologies have recently been discussed in literature, like 3D mesh, stacked mesh, ciliated 3D mesh, or tree-based topologies [13].

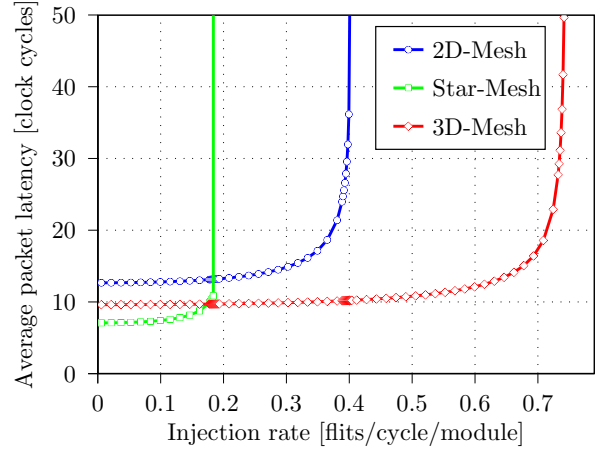
In this section, we focus on the 3D mesh topology with the objective to demonstrate its performance potential compared to classical 2D topologies and study its properties when scaling to many-core SoC. For investigating the network performance, an analytic model based on queuing theory is employed [14]. The model is very flexible and allows for fast and accurate simulation of large NoC topologies. A classical two-dimensional (2D) mesh, as well as a hierarchical star mesh (also called concentrated mesh) serve as 2D reference topologies [15]. They are compared with a 3D mesh. Note that a star mesh topology can also be applied to a 3D layered architecture, which yields a ciliated 3D mesh as shown in [13]. Figure 7 illustrates these four topology types.

The following advantages of the 3D mesh are expected.

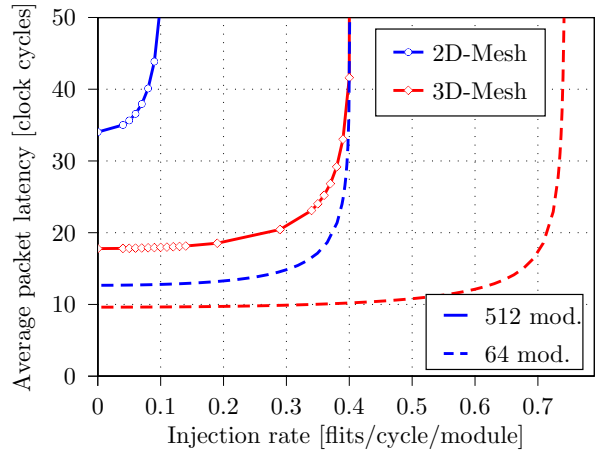
- *Low latency*: The high network concentration and short wire lengths promise for low routing latencies.
- *High throughput*: A high degree of interconnection, i.e., a high bisection bandwidth, combined with low routing latencies provide a high network throughput.
- *Short wires*: The small distance of the vertical layers and the regular structure of the 3D mesh result in short wires.

The results of the performance analysis for the case of 64 modules (8×8 2D mesh vs. $4 \times 4 \times 4$ star-mesh vs. $4 \times 4 \times 4$ 3D mesh) are shown in Fig. 8(a). Therein, the mean latency in the network is analyzed. A global uniform traffic pattern is assumed with Poisson arrival streams. Different injection rates are considered, ranging from 0.01 to 0.8 flits/cycle/module.

We clearly find that the classical 2D mesh is a bad choice w.r.t to latency (13 clock cycles at low traffic) due to the low network concentration and long routing paths. The point where the latency tends towards infinity is called network saturation point. This determines the capacity, i.e., the maximum throughput, of the network. It can be seen that the 2D mesh provides a medium throughput of 0.41 flits/cycle/module in this case.



(a) 64 cores



(b) 512 cores

Fig. 8. Performance analysis of average latency in network for 3D mesh topology.

The star-mesh topology provides a very good latency at low traffic (7 clock cycles) due to the high network concentration. However, this advantage comes at the cost of low network throughput (0.19 flits/cycle/module). To improve the low bisection bandwidth of this topology a common technique is to employ multiple inter-router links (IRLs) or using express channels [15]. The drawback of this approach is the high area consumption of the routers due to the big number of ports. In addition, the star-mesh topology does not provide an inherent (natural) scaling. I.e., the number of IRLs has to be adapted manually with increasing network size and concentration factor.

Finally, the 3D mesh shows a very good tradeoff between latency and achievable throughput. We observed a good latency (10 clock cycles) combined with a very high throughput limit (0.75 flits/cycle/module), as depicted in Fig. 8(a). Furthermore, the 3D mesh offers very good scaling abilities, as Fig. 8(b) shows for the case of a NoC with 512 modules (32×16 2D mesh vs. $8 \times 8 \times 8$ 3D mesh). We observe that the latency gap between these two topologies increases significantly.

We conclude that 3D NiCS using a 3D mesh topology is

a promising approach for closing the gap to many-core SoC. Further investigation of 3D topologies is still necessary. E.g., the large area of TSVs will probably not allow to equip every router with a vertical link. Furthermore, the vertical inter-chip links are expected to offer a higher bandwidth compared to on-chip links. Therefore, irregular topologies with heterogeneous links should be investigated more closely.

V. LOW-LATENCY ERROR CORRECTION CODING

One of the most important issue in terms of latency and link performance is the selection of a suitable channel coding scheme. In [16] and [17], it has been shown that convolutional codes are most favorable for low latency applications, whereas strong codes like low-density parity-check (LDPC) codes have a better bit-error-rate (BER) performance when higher latency can be tolerated. LDPC convolutional codes (LDPC-CCs) can combine both advantages [18], which makes them suitable for latency constrained high-performance error correction applications. We consider here the structural latency of the code and is defined as the time that the en/decoder has to wait for the input bits before the mapping of input bits can take place is due to the structure of the code. The structural latency is a feature of the coding scheme itself, regardless of current and future ways of implementation. Hence, as pointed out in [16], it provides an ultimate lower bound on the actual delay of the code.

A. Low-Density Parity-Check Convolutional Codes

Consider the transmission of a sequence of L codewords $\mathbf{v}_t, t = 1, \dots, L$. Unlike block encoding, these L blocks are coupled over various time instants t with m_{cc} determines the maximal distance between a pair of coupled blocks. Here we restrict ourselves to protograph based codes due to there ability to facilitate low complexity hardware implementation. A protograph consists of n_c check nodes and n_v variable nodes and is represented by its bi-adjacency matrix \mathbf{B} , called *base matrix*. The edges are spread according to the component base matrices $\mathbf{B}_0, \mathbf{B}_1, \dots, \mathbf{B}_{m_{cc}}$. In order to maintain the degree distribution and structure of the original ensemble, a valid edge spreading should satisfy the condition [19]

$$\sum_{i=0}^{m_{cc}} \mathbf{B}_i = \mathbf{B} . \quad (2)$$

The resultant ensemble of terminated LDPC-CCs can be described by means of a convolutional protograph with termination length L

$$\mathbf{B}_{[1,L]} = \begin{bmatrix} \mathbf{B}_0 & & & \\ \vdots & \ddots & & \\ \mathbf{B}_{m_{cc}} & & \mathbf{B}_0 & \\ & \ddots & \vdots & \\ & & \mathbf{B}_{m_{cc}} & \end{bmatrix} . \quad (3)$$

$(L+m_{cc})n_c \times Ln_v$

The last m_{cc} additional check nodes result in the rate-loss due to termination. This can be decreased by increasing L , which increases the resultant structural latency. In the following, we

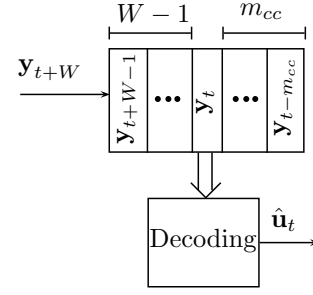


Fig. 9. Schematic diagram of a window decoder. The decoding unit here represents the belief propagation decoder for an LDPC block code.

introduce an elegant yet natural way to decode the LDPC-CC with large L . The parity-check matrix \mathbf{H} of the LDPC-CC can be obtained by replacing every 1 in $\mathbf{B}_{[1,L]}$ by a permutation matrix of size $N \times N$, with N being the lifting factor.

B. Window Decoding

The sequence of L blocks in (3) corresponds to a coupled codeword with $\mathbf{v} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_t, \dots, \mathbf{v}_L]$. The decoding can be performed by applying the belief propagation over the lifted matrix \mathbf{H} , but this results in large structural latency. A sliding window decoder of size W operates on W consecutive coupled code blocks \mathbf{v}_t [20]. The size W of the window can vary from $m_{cc}+1$ to $L-1$. Consider the decoding of a received block \mathbf{y}_t at time instant t . Based on the results of the Sec. II, we consider AWGN channel between the nodes on the boards. Figure 9 shows the schematic block diagram of the window decoder when symbols in the received block \mathbf{y}_t are the target symbols. The decoding of \mathbf{y}_t can only start once the succeeding $W-1$ blocks are available. Each of these code blocks contains Nn_v code bits. Furthermore, window decoder also requires read access to the m_{cc} previously decoded blocks due to the memory of the code as shown in Fig. 9. Hence for a code with rate R , the structural latency of the window decoder T_{WD} depends on the window size W and is expressed in terms of number of information bits as [18] [20]

$$T_{WD} = W \cdot Nn_v \cdot R \text{ [information bits]}. \quad (4)$$

Note that the latency for the window decoder in (4) is independent of L .

Figure 10 shows the required E_b/N_0 to achieve a BER of 10^{-5} as a function of decoding latency. The decoding latency for the window decoder depends on W and N . The window size is the property of the decoder and can be varied to reduce the required E_b/N_0 for a given code. The window size can be adjusted in the decoder depending on the requirements of the application at the given time without changing the encoder. This provides a flexibility in terms of latency and performance on the decoder side. For example, consider the curve with $N=40$ in Fig. 10. The performance improves by increasing W , ($W=3 \rightarrow 4$) but eventually the rate of this improvement decreases ($W=7 \rightarrow 8$). To cope with this, the lifting factor N has to be increased. The lifting factor determines the constrained length of the code, hence increasing N increases the constraint length and thus the strength of the code. This is

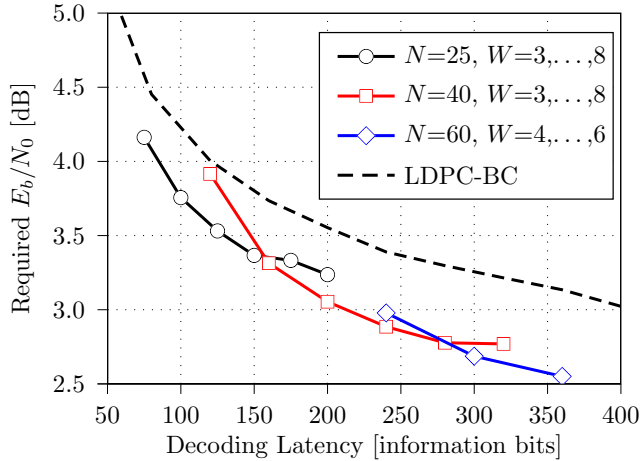


Fig. 10. Required E_b/N_0 for (4,8)-regular LDPC-CCs to achieve BER of 10^{-5} as a function of decoding latency. The component base matrices used here for LDPC-CC are $\mathbf{B}_0 = [2, 2]$, $\mathbf{B}_1 = \mathbf{B}_2 = [1, 1]$, (4,8)-regular with $\mathbf{B} = [4, 4]$ is used for LDPC-BC.

demonstrated in Fig. 10 when the window size is varied for different code with $N=25, 40$ and 60 . The required E_b/N_0 to achieve a BER of 10^{-5} for LDPC block code (LDPC-BC) is also plotted. The structural latency of a block code is equal to the number of information bits in one block and is given as

$$T_B = Nn_v \cdot R \text{ [information bits]}. \quad (5)$$

Figure 10 shows that for the complete range of latency, LDPC-CC outperform the corresponding block codes from which they are derived. For example, consider the operating value for an $E_b/N_0=3$ dB. LDPC-CC requires $T_{WD}=200$ information bits, whereas LDPC-BC requires the latency of $T_B=400$ information bits to achieve BER of 10^{-5} . This provides the gain of 200 information bits in terms of latency compared to the LDPC-BC.

VI. CONCLUSION

We propose a new system consisting of wireless links between boards, where each node is a 3D chip stack. We analyzed the wireless board-to-board links operating at the carrier frequency of 200 GHz range. The channel measurements suggests that the channel can be assumed to be static and largely frequency flat. In terms of quantization, one bit quantization together with oversampling allows the possibility of achieving ultra-high data rates at low-power consumption of one bit analog to digital converters. The results also indicate a significant improvement of information rate when considering ISI and sequence estimation. Moreover the use of 3D NiCS using a 3D mesh topology is shown to be a promising approach for fulfilling the requirements of future data rates. In the end an LDPC-CC have been analyzed, which is suitable for providing flexibility between latency and performance of the system. This provides adaptability to the system depending on the application requirements.

ACKNOWLEDGMENT

The authors would like to thank Prof. D. Plettemeier, M. Jennings and K. Wolf from Technische Universität Dresden for

undertaking the board-to-board measurements.

REFERENCES

- [1] [Online]. Available: http://www.theregister.co.uk/2012/09/18/nvidia_tesla_k20_benchmarks/
- [2] Cadence White paper, "3D ICs with TSVs Design Challenges and Requirements."
- [3] D. Walter, S. Hoppner, H. Eisenreich, G. Ellguth, S. Henker, S. Hanzsche, R. Schuffny, M. Winter, and G. Fettweis, "A source-synchronous 90Gb/s capacitively driven serial on-chip link over 6mm in 65nm CMOS," in *Proc. of the 59th International Solid-State Circuits Conference (ISSCC)*, Feb. 2012, pp. 180–182.
- [4] J. Israel and A. Fischer, "An approach to discrete receive beamforming," in *Proc. 9th International ITG Conference on Systems, Communications and Coding (SCC)*, Munich, Germany, Jan. 2013.
- [5] J. Israel, A. Fischer, and J. Martinovic, "Optimal antenna positioning for wireless board-to-board communication using a Butler matrix beamforming network," in *17th International ITG Workshop on Smart Antennas (WSA)*, Stuttgart, Germany, Mar. 2013.
- [6] S. Krone, F. Guderian, G. Fettweis, M. Petri, M. Piz, M. Marinkovic, M. Peter, R. Felbecker, and W. Keusgen, "Physical layer design, link budget analysis and digital baseband implementation for 60 GHz short-range applications," *EuMA International Journal of Microwave and Wireless Technologies (IJMWT)*, vol. 2, no. 3, 2011.
- [7] S. Krone and G. Fettweis, "Achievable rate with 1-bit quantization and oversampling at the receiver," in *Proc. IEEE Communication Theory Workshop*, Cancun, Mexico, May. 2010.
- [8] L. Landau, S. Krone, and G. Fettweis, "Intersymbol-interference design for maximum information rates with 1-bit quantization and oversampling at the receiver," in *Proc. 9th International ITG Conference on Systems, Communications and Coding (SCC)*, Munich, Germany, Jan. 2013.
- [9] S. Borkar, "Thousand core chips - a technology perspective," in *Proc. of DAC*, 2007.
- [10] TU-Dresden, "ESF Young Investigators Group; 3D Chip Stack Intraconnects - 3DCSI," last visited on 15/10/2012. [Online]. Available: http://tu-dresden.de/die_tu_dresden/fakultaeten/fakultaet_elektrotechnik_und_informatik/3dcsi
- [11] W. Dally and B. Towles, "Route packets, not wires: On-chip interconnection networks," in *Proc. of Design Automation Conference (DAC)*, 2001, pp. 684–689.
- [12] L. Benini and G. De Micheli, "Networks on chips: a new SoC paradigm," *Computer*, vol. 35, no. 1, pp. 70–78, Jan. 2002.
- [13] B. Feero and P. Pande, "Networks-on-chip in a three-dimensional environment: A performance evaluation," *IEEE Transactions on Computers*, vol. 58, no. 1, pp. 32–45, Jan. 2009.
- [14] E. Fischer, A. Fehske, and G. Fettweis, "A flexible analytic model for the design space exploration of many-core network-on-chips based on queueing theory," in *Proc. of The Fourth International Conference on Advances in System Simulation (SIMUL)*, 2012, pp. 119–124.
- [15] J. Balfour and W. J. Dally, "Design tradeoffs for tiled cmp on-chip networks," in *Proc. of the 20th annual international conference on Supercomputing (ICS)*, 2006, pp. 187–198.
- [16] T. Hehn and J. Huber, "LDPC codes and convolutional codes with equal structural delay: A comparison," *IEEE Transactions on Communications*, vol. 57, no. 6, pp. 1683–1692, Jun. 2009.
- [17] S. Maiya, D. Costello, T. Fuja, and W. Fong, "Coding with a latency constraint: The benefits of sequential decoding," in *48th Annual Allerton Conference on Communication, Control, and Computing*, Oct. 2010, pp. 201–207.
- [18] N. Ul Hassan, M. Lentmaier, and G. Fettweis, "Comparison of LDPC block and LDPC convolutional codes based on their decoding latency," in *Proc. 7th International Symposium on Turbo Codes & Iterative Information Processing*, Aug. 2012, pp. 225–229.
- [19] M. Lentmaier, M. Prenda, and G. Fettweis, "Efficient message passing scheduling for terminated LDPC convolutional codes," in *Proceedings of IEEE International Symposium on Information Theory (ISIT)*, Aug. 2011, pp. 1826–1830.
- [20] M. Papaleo, A. Iyengar, P. Siegel, J. Wolf, and G. Corazza, "Windowed erasure decoding of LDPC convolutional codes," in *IEEE Information Theory Workshop (ITW)*, Jan. 2010, pp. 1–5.