Secure and Energy-Efficient Interconnects for Board-to-Board Communication

Bho Matthiesen†, Stefan Pfennig†, Mario Bielert†, Thomas Ilsche†, Andrew Lonnstrom†, Tao Li†,
Juan A. Cabrera†, Christian Scheunert†, Elke Franz†, Silvia Santini§, Thorsten Strufe†,
Eduard A. Jorswieck†, Wolfgang E. Nagel†, Giang T. Nguyen†, Frank H.P. Fitzek†,
†TU Dresden SFB 912 - HAEC, Email: {firstname.lastname}@tu-dresden.de
§Università della Svizzera Italiana, Email: silvia.santini@usi.ch

***Abstract*—To meet the consistently increasing computation demand while reducing the energy footprint, future computing architectures need to consider parallelism. In this paradigm, compute nodes with massive number of cores are directly interconnected in one board by short-range optical connections. At the same time, high-speed wireless connections of up to 100 Gbps are used on demand between compute nodes of different boards. That hybrid design not only enables flexibility and energy efficiency but also opens up new research questions to obtain secure and energy-efficient interconnects in various areas, such as communications, routing, distributed storage and especially security. This paper summarizes the state-of-the-art research findings in those areas, presents a novel key distribution scheme and expands current evaluation platforms with a novel testbed design and realization, leveraging the maturity of virtualization technologies.**

I. Introduction

To satisfy the consistently increasing demand for computing power while reducing the energy footprint, massively parallel processors with thin-cores have to be considered [1]. The Highly Adaptive and Energy-Efficient Computing (HAEC) project advocates this approach with a hybrid design employing both optical and wireless communications [2]. The illustration of the future HAEC Box consisting of four boards is given in Fig. 1. Assuming that each compute node is a 3D stacked processor chip with thousands of "thin" cores and local memory to offer massive intra-node parallelism, compute nodes can directly communicate via optical and wireless connections for intra- and inter-board communications respectively. In this way, the number of hops, especially for links between compute nodes of different boards, can be significantly reduced.

The HAEC Box will tackle energy efficiency and adaptivity in a holistic manner. At the hardware level, computing and communication components have to enable energy proportional operations, meaning that unused chips, optical or wireless links can also be turned off. Additionally, the software part has to be aware of energy as well, leveraging the adaptivity of computing and communication components to provide adaptive, context- and energy-aware software. Specifically, it has to include energy control loops to identify energy-critical states and trigger energy-saving hardware actions. This requires monitoring the system state of controlled objects such as computing, storage and networking resources.

This work is supported by the German Research Foundation (DFG) in the Collaborative Research Center 912 "Highly Adaptive Energy-Efficient Computing."

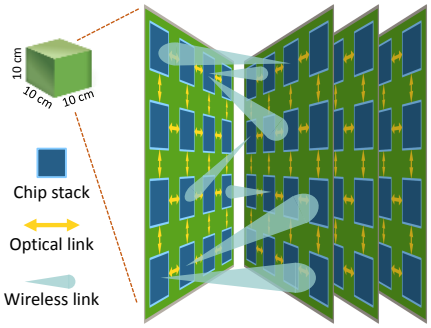


Fig. 1. Design of the HAEC Box.

Equally important is the design of a secure and energy-efficient interconnect to construct the topology of the HAEC Box, integrating computing, and optical and wireless connections with software's energy control loops. This task requires a collaborative effort of research in various areas. First, fundamental limits in terms of energy efficiency have to be understood for wireless inter-board communications. Second, research on novel routing protocols is required to ensure energy-efficient transport of data across compute nodes within the HAEC Box. Third, new distributed storage solutions have to be developed to enable a scale-out software architecture in which the HAEC Box will be used by various applications and users. Fourth, security issues such as protection and attack countermeasures required further research to ensure the security of the overall HAEC Box. Finally, unique evaluation platforms facilitating collaborative research are developed to enable the benchmark and demonstration of our research ideas and outcomes.

The contributions of this paper are two-fold. First, it summarizes the state-of-the-art studies and highlights our key findings in the areas of communications, routing, distributed storage and security. Second, the authors expand the current evaluation platforms with a state-of-the-art testbed design and realization to demonstrate and evaluate the HAEC Box topology in real world settings.

In the rest of the paper, Section II summarizes the latest

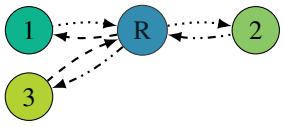


Fig. 2. System model of the 3-user multi-way relay channel with multiple unicast transmissions where node R is the relay and nodes 1 to 3 are the users. Messages travel along the different line styles.

findings in the area of interconnects, Section III describes our platforms for evaluation and Section IV summarizes the main content.

II. Secure and Energy-Efficient Interconnects

To enable secure and energy-efficient interconnects, research efforts have to be done individually and collaboratively in several areas such as communications, routing, distributed storage and security. This sections summarizes state-of-the-art studies and latest findings in those areas.

1) Communication: Communication within the HAEC Box is done optically and wirelessly. Each node is equipped with both communication interfaces and connected to its four nearest on-board neighbors through optical links. Communication between adjacent nodes is thus possible at very high speed and low energy consumption. However, on-board connections to more distant nodes via the optical interface must be routed across one or more other nodes. This requires unused bandwith on all involved nodes, costly optical switching, and introduces additional delay. Instead, utilizing the wireless interface to route the traffic across a node on an adjacent board might be favorable. In addition, communication between boards is only done wirelessly. Understanding the fundamental limits in terms of throughput, energy efficiency (EE) and delay in those wireless multi-hop networks is of paramount importance for designing the HAEC Box. We have identified and analyzed several typical communication scenarios in it and for two of these we present our results.

The first model is a multi-way relay channel (MWRC) [3] with multiple unicast transmissions depicted in Fig. 2. Throughput and EE for various relaying schemes and symmetric channels are analyzed in [4]. The considered relay operations are decode-and-forward (DF), amplify-and-forward (AF), and noisy network coding (NNC). Results show that DF achieves the sum capacity for SNRs up to 8 dB and NNC and AF achieve the sum capacity within 0.877 bit/s/Hz and 1.5 bit/s/Hz, respectively.

In future communication networks like the HAEC Box, EE [5] is the key performance metric. It is defined as the ratio of the total amount of data transmitted reliably during a time T to the associated total energy consumption of the communication interface during that time, i.e., for a total number of K nodes, $\mathrm{EE} = \frac{TBR_\Sigma(P_1,\ldots,P_K)}{T\sum_{i=1}^{K}(\phi_i P_i + P_{c,i})}$ where B is the communication bandwidth, $R_\Sigma(P_1,\ldots,P_K)$ the achievable sum rate in bit/s/Hz, P_i is the ith node's transmit power, ϕ_i is a constant to model energy users depending linearly on the transmit power, e.g., the power amplifier, and $P_{c,i}$ models the static circuit power consumption of node i. Maximizing the EE leads to non-convex, fractional problems which often have high computational complexity.

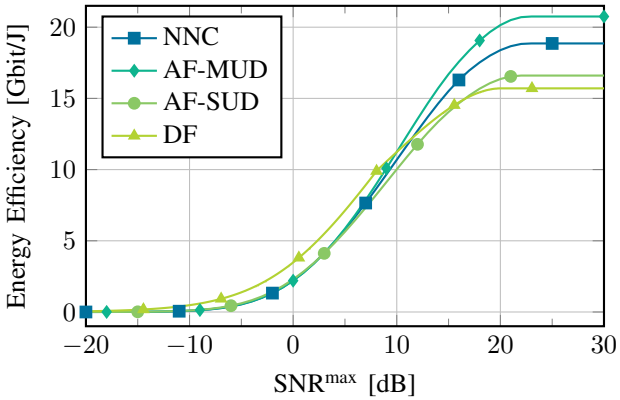


Fig. 3. Energy efficiency in the 3-user MWRC of 1) NNC, 2) AF with multi-user decoding (AF-MUD), 3) AF with single-user decoding (AF-SUD), and 4) DF as a function of the SNR for the HAEC Box.

Assume for the moment that, despite different hardware complexities, we would use the same $P_{c,i}$ for all transmission schemes. Then the scheme with the highest throughput would also achieve the highest EE. Obviously, this does not result in meaningful results. Instead, we model the different hardware complexities with different constants $P_{c.i}$ to obtain a fair and meaningful comparison of the EE. This is done in [4] with measurements and estimates from the HAEC hardware group. Simulation results taking the different hardware complexities into account are reported in Fig. 3 for symmetric channels. It can be seen that up to approximately 10 dB DF performs best whereafter AF with multi-user decoding achieves the highest EE. This continues to hold for higher SNRs, where DF is outperformed by all other schemes. The surprisingly good EE of AF motivated us to look further into this relaying scheme. In [6], the largest known achievable rate region for the channel at hand with AF relaying is derived and in [7] we begin to extend the presented results to arbitrary channels.

Data traffic in the HAEC Box may be loosely categorized in non-elastic traffic with high quality of service (QoS) demands and much less time critical elastic traffic. Examples of such types of traffic are control signaling and memory migration, respectively. Often, this non-elastic traffic does not fully utilize its available spectrum. Thus, a higher system spectral efficiency might be achieved by utilizing the same spectrum by a secondary link transmitting elastic data while ensuring that QoS demands of the primary (non-elastic) link are still met. For the HAEC Box, the most relevant approaches to this spectrum sharing problem are overlay and underlay cognitive radio [8]. In the overlay approach, the primary transmitter cooperates with the secondary transmitter to enable transmission in the secondary link while also increasing its own rate. In underlay operation, the primary transmitter does not cooperate with the secondary, thus reducing cooperation delay and overhead. Instead, it tolerates a reduction of its achievable transmission rate as long as its QoS constraints are still met.

Energy-efficient resource allocation for both modes of op-

eration are addressed in [9]. We formulate the problem as the maximization of the secondary EE subject to a minimum rate requirement for the primary user. This leads again to challenging non-convex, fractional problems. We obtain the global optimal solution for the underlay scenario and propose two low complexity algorithms yielding first-order optimal resource allocations for the overlay approach.

2) Routing: The overall goal of routing in the HAEC Box is the energy-efficient transportation of messages (or packets) between nodes. The low-power design of the optical and wireless communication links are not the only features to save energy. They are also designed to be adaptive and configurable during the runtime to achieve different performance and energy consumption levels. The basic requirement of adaptivity is the capability to power on/off each individual link including their associated hardware. When one communication link is under low utilization or even unused for a period of time, powering it off and redirecting the messages transmitted through it to another one can further reduce energy consumption of the whole network. This methodology is first introduced in the area of Green Internet and data center networks (DCNs) [10]–[12]. The design of the HAEC Box brings additional challenges to the direct application of this energy-saving methodology. Compared with the traditional network equipped solely with wired links, the available candidate links for transmitting messages include both optical and wireless links which increases the size of the solution search space. In addition, interference can occur if one node serves as receiver of multiple wireless links at the same time. Thus, we have designed a joint scheduling mechanism of link activation and routing for the HAEC Box networking architecture [13].

The lack of optical routing devices suggests that each node performs packet switching. We advocate the principle of Software Defined Networking (SDN) to support the joint scheduling mechanism, which assumes the existence of a centralised controller in the network [14]. Fig. 4 illustrates the basic architecture of our proposed joint scheduling mechanism. The core component is the network controller that accepts *QoS request messages*, e.g, bandwidth requirements, either directly from deployed distributed applications or from a high-level network resource negotiator [15] in a round- or time-based fashion. It determines which links should be powered on and selects the path for communication between two nodes so as to maximize energy savings and fullfill QoS requirements. Power mode decisions are propagated with so-called *link state configuration messages* to the link state manager of each node and *forwarding rule messages* are used to install routing decisions into each node's routing table. The nodes then act accordingly. Our simulation results reported in [13] show that the proposed optimization model and heuristic algorithm for the described joint scheduling problem achieves relatively low energy consumption compared to the fully operational mode of the HAEC Box or the existing dimensional routing algorithm.

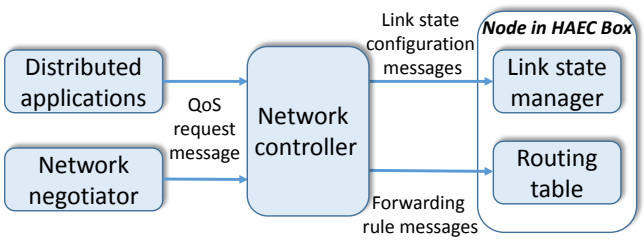


Fig. 4. Design of joint scheduling of link activation and routing.

3) Distributed Storage: Applications intended to run within the HAEC box (e.g., database management systems) require the storage, retrieval, and processing of data. One storage node, provided with enough storage capacity, could take care of the storage of the information. However, due to the highly-adaptive nature of the HAEC box, some of the nodes might be in sleep mode and unable to provide access to the stored data, causing information unavailability. Furthermore, within the HAEC box architecture, if an application requests or modifies data stored at a node several communication hops away, the operations might suffer long delays from the latency overhead added at each communication hop. Moreover, if several applications request information simultaneously, the data can be served with a throughput limited by the capacity of the serving node. To guarantee the availability and reliability of the information as well as a high throughput and low latency, the system can store the information distributedly among different nodes. By distributing the data and adding redundancy, the system can guarantee that the information is available when some of these nodes are unavailable. Furthermore, data distribution causes the total throughput of data requests to be the aggregated value of each upload throughput at the individual nodes. Similarly, by distributing the information in nodes physically located in different places in the HAEC box, the system can reduce the delays of data requests by reducing the number of hops in the communication.

The advantages of distributed storage systems are not new. Systems like RAID [16] are known since the 80's and have shown the benefits of using block codes over simple data replication. Block codes, by forming linear combination of the original data, increase the reliability of storage systems while reducing the storage costs when compared with replication schemes. However, traditional codes, such as Reed-Solomon codes, present a disadvantage over replication when some nodes become unavailable and the system repairs the lost redundancy in newcomer nodes (the repair problem). If the data is replicated, the information stored in an unavailable node is copied from other nodes into a newcomer node. On the other hand, when using traditional block codes, repairing lost redundancy involves the transfer of the whole data and a process of decoding and re-encoding. For years it was thought that the bandwidth overhead of the repairs was unavoidable when using block codes, but [17] proved that the use of linear network coding, allows distributed storage systems to operate over the optimal curve of the trade-off between storage costs and repair bandwidth.

However, the literature focuses on solutions for distributed storage systems for peer-to-peer networks or within data centers. Furthermore, the repair problem is usually addressed

for scenarios where the repair of the lost redundancy is performed in newcomer nodes that are available within a short time after a node becomes unavailable. Our goal is to investigate different distributed storage protocols and techniques to disseminate data within the context of the HAEC box architecture. For instance, how to distribute the data within the HAEC Box topology in order to understand the intrinsic trade-offs between storage, energy consumption, throughput, latency and reliability when using replication, traditional block codes and network coding schemes. Furthermore, we investigate the repair problem when the data losses due to the unavailability of the nodes are not catastrophic but rather temporary, and the system that performed a repair will need to remove extra redundancy once the missing node becomes available again. Our results include [18] where we studied the repair problem evaluating the trade-offs between storage costs and repair bandwidth in systems without newcomer nodes available. We also investigated the feasibility of managing the storage and transport of the information with Random Linear Network Coding (RLNC) as a single code structure. In [19], we investigated techniques to optimize the operations involved in network coding when performed in multicore architectures, namely the parallelization of matrix multiplication and matrix inversion over Galois fields in order to reduce the computation latency.

4) Security: To ensure total security of the interconnects, we consider a comprehensive approach including both protection, i.e. prevent attackers in the first place, and attack countermeasures.

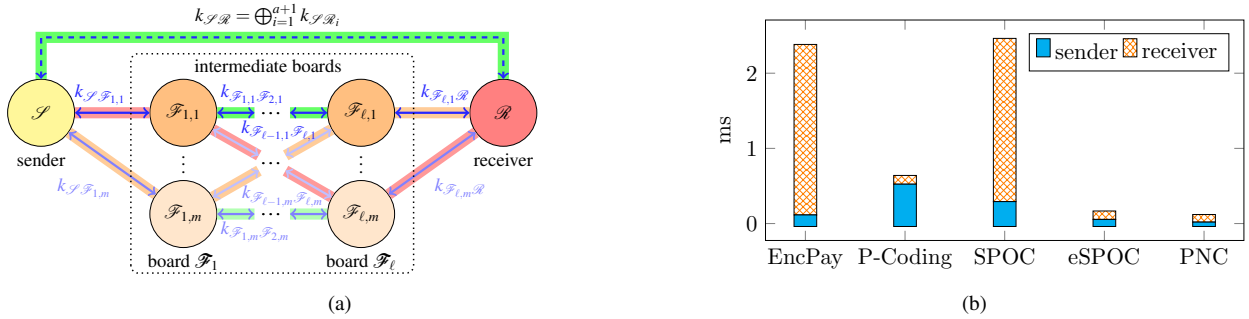
a) Protection: We apply network coding [20] within the HAEC Box, because it offers fast and efficient communication. Instead of sending parts of the data consecutively, network coding allows for sending linear combinations of the data (coded packets) and for recoding of packets at intermediate nodes instead of simply relaying them. A receiver does not depend on certain packets but only needs "enough" coded packets to decode. Communication is only valuable as long as the messages are undisclosed, unmodified, and available. There exist many state-of-the-art security measurements for network coding, but we found two main issues that have to be solved. Since all efficient systems are based on symmetric cryptography, we need an efficient and secure way to exchange symmetric keys between all nodes of the HAEC Box. Second, whereas most schemes try to optimize only throughput, we also want to optimize for high efficiency and low latency.

One opportunity for an efficient key distribution is to utilize physical layer key generation (PLKG) [21]. In this process, communication partners agree on a key by exploiting the characteristics of a wireless channel. An eavesdropper at a different location is not able to obtain the key, because he has different channel characteristics to the communication partners. However, to generate a physical layer key the two nodes need a direct link. Thus, we tried to find a way for a secure and efficient key establishment over multiple hops.

In [22], [23], we provide solutions for end-to-end key exchange under the consideration of different attacker models and evaluate the costs. In case of only passive attackers (eavesdroppers), we propose to use different paths to transmit partial keys between sender and receiver, where each direct link is protected by means of the generated physical layer keys. Sender and receiver locally compute the end-to-end key from the partial keys, and as long as at least one path used for the exchange of a partial key is trustworthy, an attacker cannot learn the end-to-end key. Fig. 5(a) shows an example where sender and receiver want to establish a common key. By sending partial keys over different paths (e.g., red and orange) the receiver as well as the sender can calculate the XOR of all partial keys. A single eavesdropper, however, cannot gain any knowledge. For active attackers, we propose to use robust secret sharing [23] to prevent modifying attackers from hindering the key exchange. We show that a key exchange is possible if the number of nodes per board is larger than the sum of eavesdroppers and modifying attackers on a board. Overall, this illustrates a possible way for an initial key distribution between the nodes of a HAEC Box.

For secure network coding approaches, we have to distinguish between two protection goals. Because of the recoding at intermediate nodes, integrity is of high importance. Otherwise, the recoding of a bogus message and regular messages will result in corrupted messages that will pollute the communication and harm the availability of the system (pollution attacks [24]). The other protection goal is confidentiality of the data. Instead of applying end-to-end encryption, there are possibilities to exploit the inherent (algebraic) security of network coding. In SPOC [25], only the encoding coefficients are encrypted instead of the data. In P-Coding [26], all symbols of a packet are permuted to prevent an eavesdropper from gaining any information. Our goal is to analyze the efficiency of secure network coding schemes in the HAEC Box. First, we want to find appropriate schemes and a way to adaptively choose the most efficient scheme for given parameters. Second, we want to increase the efficiency of the secure approaches.

For these goals, we provided analysis, implementation, and measurements of selected schemes. We also implemented performance models within our parallel simulation framework [27] to test the approaches on a large scale. Based on these studies, we enhanced existing schemes in order to achieve better efficiency and lower latency. In the field of confidential network coding, we extended the work of [25] and developed a scheme called eSPOC [28] that provides less communication overhead, lower energy consumption, and less latency. Fig. 5(b) shows the latencies for different lightweight confidentiality schemes (P-Coding, SPOC, and eSPOC) in comparison to an end-to-end encryption of the payload applied before network coding (EncPay). As a baseline, the latency of network coding without security measures (practical network coding (PNC) according to [29]) is reported as well. We split the latencies in the time needed by a sender (blue) until he can send the first packet and the time a receiver (orange) needs to decode after receiving the last packet. As can be seen, eSPOC nearly achieves the latency of insecure PNC.



(a) (b)

Fig. 5. (a) End-to-end key exchange with the help of PLKG. (b) Evaluation of latencies for PNC, EncPay, (e)SPOC, and P-Coding for generation size of 16 and a packet size of 1360 B.

b) Attack countermeasure: Wyner, in his seminal paper [30], laid the foundation for evaluating information theoretic security for the wire-tap channel without the use of secret keys. More recent investigations have been e.g. into the secrecy rate for multiple-input single-output (MISO) channels with channel state information (CSI) [31]. One of the goals of our research is to provide an information theoretic analysis of massive MIMO with regards to the secure goodput of the system defined as the amount of bits which can be reliably and information theoretic securely received.

We consider an $n \times m$ complex-valued MIMO system where n is the number of transmit antennas at Alice and m is the number of receive antennas at Bob as shown in Fig. 6. In our system model, Bob is the legitimate receiver for a message transmitted by Alice and Eve is the eavesdropper. The received signal for Bob is $\mathbf{y_b} = \mathbf{Hx}+\mathbf{n_b}$ and for Eve is $\mathbf{y_e} = \mathbf{Gx}+\mathbf{n_e}$, where $\mathbf{x} \in \mathbb{C}^n$ is the input signal, $\mathbf{H}, \mathbf{G} \in \mathbb{C}^{m \times n}$ are the channel matrices, and $\mathbf{n_b}, \mathbf{n_e} \in \mathbb{C}^m$ are zero-mean circularly symmetric complex Gaussian noise vectors with powers σ_b^2 and σ_e^2 respectively. The SNR is defined as $\rho = \frac{p}{\sigma^2}$ where p is the average transmit power.

The ordered eigenvalues of the effective channels to Bob are $\lambda_l(\mathbf{HH}^\dagger)$, where $\lambda_1 \geq ... \geq \lambda_n$. Similarly, the ordered eigenvalues of the effective channel to Eve are $\gamma_l(\mathbf{GG}^\dagger)$, where $\gamma_1 \geq ... \geq \gamma_n$.

An information theoretic analysis of the secure goodput for massive MIMO systems as well as multi-mode fiber optic systems is done in [32]. Allowing k out of n total streams to be compromised, the following optimization problem is derived for the secure goodput

$$\max_{1 \leq k \leq n} \max_{\Delta > 0} (n-k) \Pr(\lambda_n \geq \gamma_{k+1} + \Delta) \log\left(1 + \frac{\rho\Delta}{1+\rho(\lambda_n - \Delta)}\right)$$

where ρ is the SNR of the system and Δ is the advantage (i.e. difference) between γ_{k+1} and λ_n such that $\lambda_n \geq \gamma_{k+1} + \Delta$.

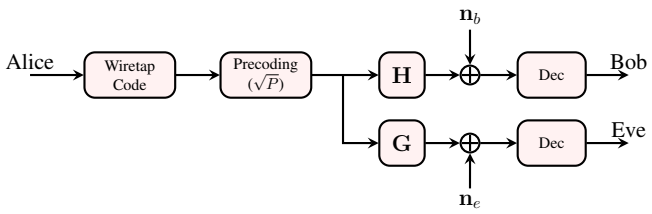


Fig. 6. Wiretap channel model for communication between Alice and Bob with eavesdropper (Eve).

Due to the complex nature of the optimization problem it was solved with an exhaustive search.

The method and results obtained in [32] are a useful part in solving the larger problem of information theoretically securing the board-to-board wireless communication within the HAEC box.

III. Evaluation Platforms

This sections describes our evaluation platforms, including the simulation framework and our testbed. They work as key components for the evaluation of research ideas.

1) A Parallel Simulation Framework: Simulation plays an important role during the research and design of the HAEC Box and its components. The challenge is to connect research on different hardware components as well as software runtime environments and applications. Since the focus is on future generation components and architecture, it is not feasible to build integrated hardware prototypes at the target scale of the HAEC Box. Analytical models only reveal static properties for simple assumptions instead of complex dynamic workloads.

The HAEC simulation framework [33] allows to study how relevant parallel applications perform on the envisioned components. Simulation enables the evaluation of alternative design choices early during research. By providing an integrated framework, it is possible to see how multiple components interact with each other under realistic workloads. For HAEC, it is essential to not only predict performance, but also energy consumption as a result of the simulation.

There is a good amount of literature about simulating applications on parallel architectures. One approach to simulation is discrete event simulation, which focuses on processing events rather than advancing time. For most existing simulators, the focus is on performance as a resultant metric. Dimemas [34] simulates the execution of parallel MPI or multi-threaded programs based on application traces. However, the Dimemas simulator itself is sequential. BigSim [35] is also based on traces of MPI applications and focuses on large target systems. xSim [36] uses a lightweight approach for simulating a large system using oversubscription on a smaller system.

The simulation workflow for HAEC begins with recording the runtime behavior of a parallel application on an existing system as a event trace. The simulation itself processes these events in parallel and creates an output trace that describes the

performance and energy consumption of the application on the target platform, e.g. the HAEC Box. Analysis and visualization of the simulation output can be performed using exiting tools for performance analysis on application traces.

Score-P [37] is used to instrument the parallel application. The resulting trace file contains events such as a call to a specific function, on a given thread or a message between two processes with a certain size. All events are locally ordered and contain a timestamp. The parallel simulation uses OTF2 [38] as file format and library both for the input and output trace. Each thread of the input application trace is handled by a worker process in the simulation. Further processes manage the state of shared resources (e.g. links) for the target architecture. The resulting trace is visualized with Vampir [39], revealing the dynamics of the simulated application execution, e.g. the simulated time of a specific message or the power consumption of a hardware component over time. Vampir also allows the comparison of multiple traces, e.g. from two simulations of different configurations of the HAEC Box architecture.

An important aspect for HAEC is modeling of network coding schemes for error-prone transmissions [40]. We have implemented models for the wireless and optical links as well as resource management to model contention for shared links. The energy aspect is covered by power models for CPUs that are based on recorded performance monitoring counters as well as power models for interconnect that are based on the utilization state of links. One focus during model building and simulation, is verification in order to ensure reliable results [41].

Simulation in HAEC provides valuable feedback to researchers that are working on hardware components and interconnect techniques, by revealing the impact of design descisions on the performance and energy consumption of relevant workloads.

2) Testbed: To present our research findings in real-world environment, we need a testbed which also works as a proof-of-concept of the HAEC Box architecture. With the testbed, we would like to flexibly create HAEC reference topologies as well as state-of-the-art ones for performance comparison. Simultaneously, the testbed should allow for energy measurement and calibration at a reasonable cost and has to provide compatible hardware interfaces capable of running energy-aware software from collaborative research groups.

Our approach to build the testbed is a combined solution of both hardware and software. The former provides computation power and physical connectivities internally between computing elements as well as externally to the Internet. The latter provides us the flexibility to create networks of arbitrary topologies. In that sense, we leverage the maturity of virtualization technologies for computing and networking to create virtually arbitrary topologies under test. Specifically, for the software part, we decide to deploy a cloud management software to reuse its helpful functionalities such as resource aggregation, automated networking services (e.g., building virtual switches or routers, IP address assignment,

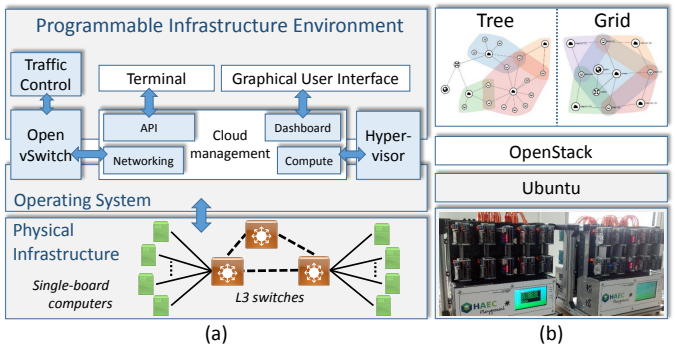


Fig. 7. HAEC Playground testbed: (a) Conceptual design; and (b) Realization consisting of SCB Odroids assembled into movable hardware modules, OpenStack software deployed to aggregate computation resources, two exemplary topologies (tree and grid).

etc.) and managing multiple simultaneous projects and users. For the underlying infrastructure, we use single-board computers (SBCs) due to their small sizes, yet reliable operation and economical matter. Fig. 7 illustrates the design and realization of our testbed, HAEC Playground.

We organize the SBCs in a star topology to enable connectivity between all compute nodes in the testbed. Furthermore, to facilitate a modular and extensible setup, we organize SBCs into subgroups interconnected by a switch. Each subgroup can support a small to medium scale experiment. When a larger setup is needed, several subgroups should be aggregated by interconnecting switches into a ring. For that reason, we select L3 switches which are stackable to support interconnection of multiple switches. Since our focus is not on the chip technology at compute nodes themselves, but rather on the interconnects between them, we advocate the use of off-the-shelf hardware. For SBC, we select Odroid XU-4 from Hardkernel[1], each is equipped with 8 ARM CPUs with the big.LITTLE architecture, meaning 4 high performance Cortex-A15 cores and 4 energy-efficient Cortex-A7 ones. Additionally, we design and manufacture a tailor-made rack to make the testbed's modules more mobile. This design is also extensible to future need, e.g., one can connect a high-performance computer with large-volume storage and fast network connections to the testbed, supporting resource intensive operations.

The cloud management software deployed on the testbed is the key to provide the programmable infrastructure environment. We decide to deploy OpenStack[2] in our testbed since it is the most mature and active open-source project of its kind. In addition to its capability to create arbitrary network topologies, the software allows for modifying Quality-of-Service metrics (such as delay and packet loss), with add-ons, to emulate characteristics of a wireless connection. The overall advantage of OpenStack is that it facilitates an emulation environment, meaning that our networking setup is fully compatible with real-world networks. Furthermore, the environment provides both a graphical user interface to visualize aggregated resources, instantiated networks and computing nodes as well as an API for repeatable setups via a command-line interface.

[1] www.hardkernel.com
[2] www.openstack.org

All in all, the testbed allows us to instantiate networks of arbitrary topologies, including the HAEC Box for evaluation purposes. More importantly, it also works as a proof-of-concept to demonstrate the idea of the visionary HAEC Box.

IV. Summary

We summarized in this paper state-of-the-art studies and highlighted latest outcomes to obtain secure and energy-efficient interconnects in various areas, such as communications, routing, distributed storage and especially security. Furthermore, we presented a novel testbed design and realization, leveraging virtualization trend in cloud computing with off-the-shelf single-board computers, introducing various advantages such as flexibility, mobility and reasonable cost. The testbed enables the construction of arbitrary topology for interconnect research as well as demonstration.

References

[1] A. Marowka, "Back to thin-core massively parallel processors," *Computer*, vol. 44, no. 12, pp. 49–54, Dec. 2011.

[2] G. Fettweis, W. Nagel, and W. Lehner, "Pathways to servers of the future," in *Design, Automat. and Test Eur. Conf. and Exhibition (DATE)*, Dresden, Germany, Mar. 2012, pp. 1161–1166.

[3] A. Chaaban and A. Sezgin, *Multi-way Communications: An Information Theoretic Perspective*, ser. Found. and Trends Commun. and Inf. Theory. Now Publishers, 2015, vol. 12, no. 3-4.

[4] B. Matthiesen, A. Zappone, and E. A. Jorswieck, "Resource allocation for energy-efficient 3-way relay channels," *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4454–4468, Aug. 2015.

[5] A. Zappone and E. Jorswieck, *Energy Efficiency in Wireless Networks via Fractional Programming Theory*, ser. Found. and Trends Commun. and Inf. Theory. Now Publishers, 2015, vol. 11, no. 3-4.

[6] B. Matthiesen and E. A. Jorswieck, "Instantaneous relaying for the 3-way relay channel with circular message exchanges," in *Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2015.

[7] ——, "Global sum rate optimal resource allocation for non-regenerative 3-way relay channels," in *Workshop Broadband Wireless Commun. Comput. Boards (Atto-Nets), IEEE Int. Conf, Ubiquitous Wireless Broadband (ICUWB)*, Salamanca, Spain, Sep. 2017.

[8] J. Mitola and G. Q. Maguire, "Cognitive radio: making software radios more personal," *IEEE Personal Commun.*, vol. 6, no. 4, Aug. 1999.

[9] A. Zappone, B. Matthiesen, and E. A. Jorswieck, "Energy efficiency in MIMO underlay and overlay device-to-device communications and cognitive radio systems," *IEEE Trans. Signal Process.*, vol. 65, no. 4, pp. 1026–1041, 2017.

[10] M. Zhang, C. Yi, B. Liu, and B. Zhang, "GreenTE: Power-aware traffic engineering," in *Int. Conf. Network Protocols (ICNP)*, 2010.

[11] T. Wang *et al.*, "Towards bandwidth guaranteed energy efficient data center networking," *J. Cloud Comput.*, vol. 4, no. 1, p. 1, 2015.

[12] G. Lin, S. Soh, and K.-W. Chin, "Energy-aware traffic engineering with reliability constraint," *Comput. Commun.*, vol. 57, pp. 115–128, 2015.

[13] T. Li and S. Santini, "Energy-aware coflow and antenna scheduling for hybrid server-centric data center networks," in *IEEE Int. Conf. Commun. (ICC)*, Paris, France, 2017.

[14] B. A. A. Nunes, M. Mendonca, X.-N. Nguyen, K. Obraczka, and T. Turletti, "A survey of software-defined networking: Past, present, and future of programmable networks," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1617–1634, 2014.

[15] R. Soulé *et al.*, "Merlin: A language for provisioning network resources," in *ACM Conf. emerging Networking Experiments and Technol. (CoNEXT)*, 2014.

[16] D. A. Patterson, G. Gibson, and R. H. Katz, *A case for redundant arrays of inexpensive disks (RAID)*. ACM, 1988, vol. 17, no. 3.

[17] A. G. Dimakis, P. B. Godfrey, Y. Wu, M. J. Wainwright, and K. Ramchandran, "Network coding for distributed storage systems," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4539–4551, 2010.

[18] J. A. Cabrera, D. E. Lucani, and F. H. P. Fitzek, "On network coded distributed storage: How to repair in a fog of unreliable peers," in *Int. Symp. Wireless Commun. Syst. (ISWCS)*, Sep. 2016, pp. 188–193.

[19] S. Wunderlich, J. Cabrera, F. H. P. Fitzek, and M. V. Pedersen, "Network coding parallelization based on matrix operations for multicore architectures," in *IEEE Int. Conf. Ubiquitous Wireless Broadband (ICUWB)*, Oct. 2015.

[20] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, 2000.

[21] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge University Press, 2011.

[22] S. Pfennig, E. Franz, S. Engelmann, and A. Wolf, *End-to-End Key Establishment using Physical Layer Key Generation and Specific Attacker Model*. Springer, 2016, ch. 6, pp. 93–110.

[23] S. Pfennig, S. Engelmann, E. Franz, and A. Wolf, "Robust secret sharing for end-to-end key establishment with physical layer keys under active attacks," in *Workshop Communi. Security (WCS), Annu. Int. Conf. Theory and Appl. Cryptographic Techn. EUROCRYPT*, 2017.

[24] M. Krohn, M. Freedman, and D. Mazières, "On-the-fly verification of rateless erasure codes for efficient content distribution," in *IEEE Symp. Security and Privacy*, 2004.

[25] J. P. Vilela, L. Lima, and J. Barros, "Lightweight security for network coding," in *IEEE Int. Conf. Commun.*, 2008.

[26] P. Zhang, Y. Jiang, C. Lin, Y. Fan, and X. Shen, "P-Coding: Secure network coding against eavesdropping attacks," in *IEEE Int. Conf. Comput. Commun. (INFOCOMM)*, 2010.

[27] F. M. Ciorba *et al.*, "Analysis of applications on a high performance–low energy computer," in *Workshop Unconv. High Performance Comput. (UCHPC), Int. Conf. Parallel Process. (Euro-Par)*, 2014.

[28] S. Pfennig and E. Franz, "eSPOC: enhanced secure practical network coding for better efficiency and lower latency," in *Workshop Network Coding and Appl. (NetCod), IEEE GLOBECOM*, Washington, DC, 2016.

[29] P. A. Chou, Y. Wu, and K. Jain, "Practical network coding," in *Annu. Allerton Conf. Comm., Control, and Comput.*, 2003.

[30] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[31] S. Gerbracht, C. Scheunert, and E. A. Jorswieck, "Secrecy outage in MISO systems with partial channel information," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 2, pp. 704–716, Apr. 2012.

[32] A. Lonnstrom, E. Jorswieck, D. Haufe, and J. Czarske, "Robust secure goodput for massive mimo and optical fiber wiretap channels," 2017, submitted to *IEEE Int. Workshop Signal Process. Advances Wireless Commun. (SPAWC)*.

[33] M. Bielert, F. M. Ciorba, K. Feldhoff, T. Ilsche, and W. E. Nagel, "HAEC-SIM: A simulation framework for highly adaptive energy-efficient computing platforms," in *Conf. Simulation Tools and Techn.*, 2015.

[34] J. Labarta, S. Girona, and T. Cortes, "Analyzing scheduling policies using Dimemas," *Parallel Comput.*, vol. 23, no. 1-2, pp. 23–34, Apr. 1997.

[35] E. Totoni *et al.*, "Simulation-based performance analysis and tuning for a two-level directly connected system," in *IEEE Intl. Conf. Parallel and Distributed Syst.*, 2011, pp. 340–347.

[36] S. Böhm and C. Engelmann, "xSim: The extreme-scale simulator," in *Intl. Conf. High Perf. Comput. and Sim. (HPCS)*. Istanbul, Turkey: IEEE Computer Society, Los Alamitos, CA, USA, Jul. 2011, pp. 280–286.

[37] A. Knüpfer *et al.*, "Score-P: A joint performance measurement run-time infrastructure for Periscope, Scalasca, TAU, and Vampir," in *Tools High Performance Comput.*, H. Brunst, M. S. Müller, W. E. Nagel, and M. M. Resch, Eds. Springer, 2011, pp. 79–91.

[38] D. Eschweiler, M. Wagner, M. Geimer, A. Knüpfer, W. E. Nagel, and F. Wolf, *Open Trace Format 2: The Next Generation of Scalable Trace Formats and Support Libraries*, ser. Advances in Parallel Computing, 2012, vol. 22, pp. 481–490.

[39] A. Knüpfer *et al.*, "The Vampir performance analysis tool-set," in *Tools High Performance Comput.*, M. Resch, R. Keller, V. Himmler, B. Krammer, and A. Schulz, Eds. Springer, Jul. 2008, pp. 139–155.

[40] S. Pfennig, E. Franz, F. M. Ciorba, T. Ilsche, and W. E. Nagel, "Modeling communication delays for network coding and routing for error-prone transmission," in *Int. Conf. Future Generation Commun. Technol. (FGCT)*, 2014.

[41] S. Pfennig *et al.*, "Simulation models verification for resilient communication on a highly adaptive energy-efficient computer," in *High Performance Comput. Symp. (HPC), SCS Spring Simulation Multi-Conf. (SpringSim)*, Pasadena, CA, 2016.