Securing the Industrial-Tactile Internet of Things with Deterministic Silicon Photonics Switches

Ted H. Szymanski Dept. ECE, McMaster University,
Hamilton, Ontario, Canada

Abstract

Today's *Best-Effort* (BE) *Internet of Things* (IoT) faces challenges in providing the end-to-end-performance, security and energy efficiency needed for the *Smart Systems* of the 21-st century. These future smart systems will include *Smart Cities*, *Smart Transportation Systems* and *Smart Manufacturing*. This paper surveys the security weaknesses of the BE IoT. The BE-IoT cannot be partitioned into distinct interference-free *Virtual Networks*, which compromises performance, cyber-security and energy efficiency. The design of a *Secure Deterministic* (SD) *Industrial-Tactile* IoT core network which can embed millions of distinct *Secure Deterministic Virtual Networks* (SD-VNs) in layer 2 is then presented. Deterministic communications, combined with low-jitter scheduling, offers several benefits; (i) the removal of all congestion, interference and DOS attacks, (ii) a significant reduction in IoT router buffer sizes, (iii) a significant reduction in IoT energy use, (iv) a reduction of end-to-end IoT delays to the speed of light in fiber, and (v) deterministic packet-switches are relatively easy to synthesize using FPGA technologies. These benefits can apply to optical and 5G wireless networks. Future *Smart Systems* can reserve their own congestion-free SD-VNs in layer 2 to manage their traffic, with significantly improved performance, security and energy efficiency. A speed-of-light deterministic IoT core network can transform cloud services in the 21-st century by exploiting a new technology; FPGAs combined with silicon photonics transceivers to achieve Terabits/second of optical bandwidth. To illustrate the transformational potential, *Big Data* green cloud computing over a secure deterministic IoT spanning the European Union is explored.

[1] [2]

Index Terms

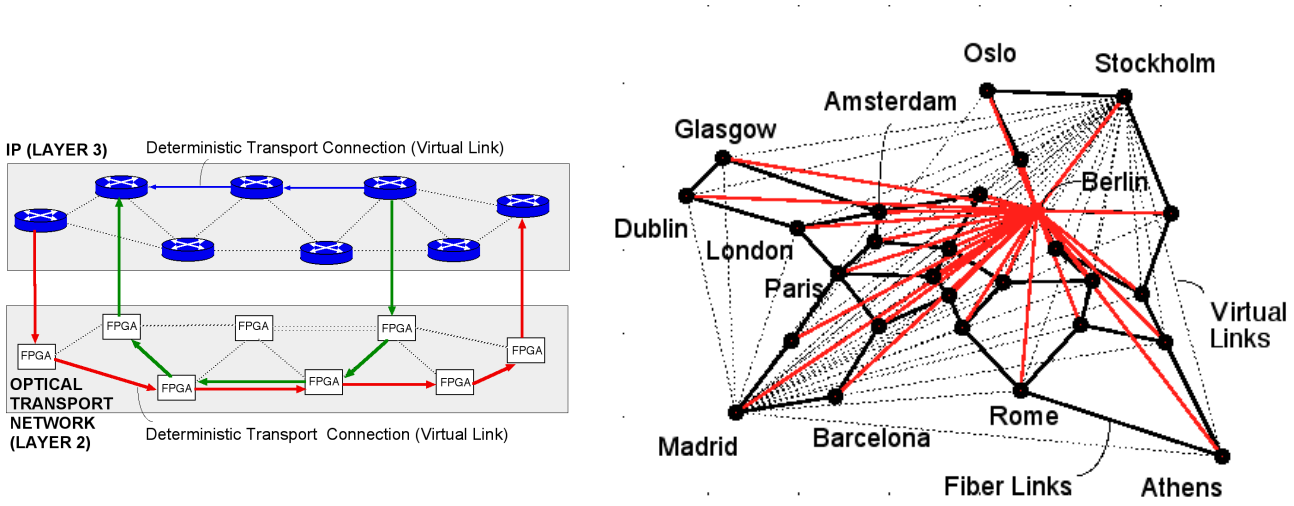
Security, Smart Systems, machine-to-machine (M2M), communication network, IoT, IIoT, Deterministic Virtual Private Networks, green Industrial Internet of Things, green Tactile Internet of Things, robots, Industrial Automation, Big Data, green cloud computing, exascale computing, SDN, SDN control plane, Silicon Photonics, FPGA, 5G wireless network, optical network, cloud, control, power efficiency, energy efficiency, green networks, layer 2 security

I. Introduction

Today's *Best Effort* (BE) IoT suffers from congestion, *Denial of Service* (DOS) attacks and targetted cyber-attacks, and provides only inefficient *Best-Effort* communications for cloud services [1]. This paper first surveys the security weaknesses of today's BE IoT. Today's BE-IoT can exhibit delays of 100s of milliseconds [2], [3], and it cannot be partitioned into distinct, interference-free *Virtual Networks* (VNs), which significantly compromises performance, cyber-security and energy efficiency. To reduce congestion and excessive delays, the BE-IoT is typically over-provisioned to operate at light loads, with typically less than 50% utilization. This over-provisioning can cost service providers an estimated $37 Billion per year in unnecessary capital and energy costs [4]. However, even with over-provisioning large delays still occur during times of congestion. As a result of congestion, DOS attacks and targetted cyber-attacks, the BE-IoT cannot support the demanding low-latency *Machine-to-Machine* (M2M) communications required in the *Smart Systems* of the 21-st century. These future smart systems will include *Smart Cities*, *Smart Manufacturing*, the *Smart Power Grid* [5], and *Smart Transportation Systems*.

[1] A related IEEE Communications magazine paper is available at: http://ieeexplore.ieee.org/document/7498096/?arnumber=7498096

[2] A related IEEE/ACM paper entitled *'An Ultra Low Latency Guaranteed-Rate Internet for Cloud Services'* is available at URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6917218



(a) (b)

Fig. 1. (a) A future SD-IoT network, with deterministic transport connections in layers 2 and 3. (b) An SD-IoT core network spanning the European Union with a Virtual Network originating at Berlin. Congestion-free Virtual Links are shown with bold red lines.

Several international efforts are defining the requirements for a *Future Internet* network with very low delays and very high reliability, to support the smart systems of the future. In March 2014, the *'Industrial Internet Consortium'* (IIC) was formed to develop the requirements for a very low-latency *'Industrial Internet'* network [6], [7], to revolutionize M2M communications. In August 2014, the ITU began exploring the high-level requirements of a very low-latency *'Tactile Internet'* network, to revolutionize M2M and *'Human-to-Machine'* (H2M) communications [8], [9]. Applications for both networks include *Smart Manufacturing and Industrial Automation*, the *Smart Power Grid*, *Smart Healthcare* systems and *Smart Transportation* systems. The *Industrial Internet* and *Tactile Internet* efforts share some similar goals; to achieve a future IoT with exceptionally low latency and high reliability. These international efforts are briefly reviewed in section III.

By definition, a *deterministic* traffic flow must receive a deterministic (or guaranteed) rate of transmission through a network, and therefore it must be immune to congestion and interference from other flows (please see section IV). Recent advances in the theory of deterministic networks have shown that the addition of deterministic traffic flows to today's BE-IoT, combined with low-jitter scheduling algorithms, can result in exceptionally low delays, jitters and packet-loss rates and exceptionally high reliability, in both optical and wireless networks [4], [10-14]. In 2015, the IETF approved the *Deterministic Networking* group to explore how to add deterministic communications to the BE-IoT, as work in progress [15], [16]. Unfortunately, the problem of scheduling deterministic traffic flows through the IoT network to minimize the delay and jitter is a well-known NP-hard scheduling problem which has remained unsolved since the 1990s [12,17-20]. These international efforts to achieve an exceptionally low-latency Internet, using wired or wireless switching technologies, will encounter this NP-Hard scheduling problem, and for this reason these international efforts specify only abstract high-level requirements.

In this paper, the security aspects of a converged *Secure Deterministic Industrial Internet of Things* (SD-IoT) core network are explored. (Given the similar goals of the Industrial and Tactile Internet efforts, a unified network can also be called an *'Industrial-Tactile Internet of Things'* core network.) The core network consists of many simple deterministic packet switches under the control of an SDN control plane. The converged SD-IoT network supports both the *Best-Effort* and *Deterministic* communications paradigms, and can achieve exceptionally low latencies, jitters and packet loss rates with exceptionally high reliability. According to recently developed theory on deterministic networks, the proposed SD-IoT core network exhibits the following properties:

- It can remove all congestion, interference and excessive delays associated with best-effort traffic

flows;
- It can operate all IoT links at 100% of their capacity;
- It can reduce IoT buffer sizes by a factor of 1000+ ;
- It can reduce end-to-end IoT delays to the speed of light in fiber.

Please see [10,11,12] for mathematical proofs of these properties. This paper establishes that the use of deterministic communications exhibits the following additional properties:

- It can remove DOS attacks and targeted cyber-attacks in layer 2 by admitting only authenticated deterministic traffic flows;
- Deterministic Silicon-Photonics packet-switches are relatively easy to synthesize using ASICs or FPGAs with optical Input/Output (IO);
- Deterministic Silicon-Photonics packet-switches can reduce the IoT energy use by 2...3 orders of magnitude.

While deterministic communications will improve security in the IoT in layer 3, we argue that it is much easier to secure the IoT in layer 2, for both wired and wireless technologies. The proposed SD-IoT can embed millions of distinct, mutually-exclusive and interference-free *Secure Deterministic Virtual Networks* (SD-VNs) in layer 2, to significantly improve performance, cyber-security and energy efficiency. Each SD-VN is composed of many *Secure Deterministic Virtual Links* (SD-VLs), where each SD-VL is a congestion-free *Deterministic Transport Connection* (DTC) between 2 nodes in layer 2, as shown in Fig. 1a (bottom).

Every consumer-oriented cloud service provider (such as Netflix, Youtube, Amazon, Apple, or Google) can reserve its own distinct SD-VN in layer 2 to deliver its own traffic, with significantly-improved security, energy efficiency and performance. Governments entities, such as the US Dept. of Homeland Security or the US National Security Organization (NSA), can reserve their own distinct interference-free SD-VNs in layer 2, to deliver their traffic securely. Future *Smart Systems* such as *Smart Manufacturing*, *Smart Transportation* systems, the *Smart Power Grid*, and *Smart Healthcare* systems can all reserve their own distinct congestion-free SD-VNs in layer 2, to manage their own infrastructure securely.

Low-cost *Field Programmable Gate Array* (FPGA) devices which are integrated with many Silicon Photonics transceivers to provide Terabits/second of optical IO bandwidth will be available within a few years (see section VI). The deterministic packet-switching technologies proposed in this paper are ideally suited for synthesis in FPGA-based packet-switches, since deterministic packet-switches have exceptionally low hardware complexity. We show that a layer-2 deterministic SD-IoT using Silicon Photonics technologies should support *Big Data* green cloud computing with exascale performance within a decade, a goal which is not possible with todays inefficient BE-IoT.

In Feb. 2016, the IEEE's *Internet Initiative* hosted the *IEEE End-to-End Trust and Security Workshop*, consisting of several invited presentations, to explore new ways to secure the IoT [21]. (This paper explains the slides presented by this author.) This paper builds upon the deterministic IoT design presented in [4], [12], by adding security concepts presented by this author at the IEEE workshop [21], and by incorporating new and inevitable technologies, i.e., FPGAs integrated with Silicon Photonics transceivers.

This paper is organized as follows: Section II surveys the security problems of today's BE IoT. Section III discusses the evolution to low-latency networks for *Smart Systems*. Section IV presents the design of the SD-IoT core network. Section V presents experimental results. Section VI discusses FPGAs with optical IO. Section VII discusses secure *Big Data* green cloud computing over the EU core network. Section VIII concludes the paper. (A multimedia video presentation accompanies this paper.)

II. A Survey of Security Weaknesses of the Best-Effort Internet

A. Non-Deterministic Networks have Poor Performance

The BE-IoT suffers from congestion which causes: (a) excessively high end-to-end delays potentially as large as 100s of milliseconds [2], [3], and (b) packet loss rates which can approach 50%, unless the

Fig. 2. Several chassis of a Cisco CRS-3 router in 2016.

network is significantly over-provisioned [12]. The IETF recognizes that over-provisioning can lower the utilization of the BE-IoT links to typically below 50% [15], [16].

Current IoT routers typically use a *Bandwidth-Delay Product* (BDP) buffer sizing rule, which provides buffers for about 1/4 second of data per IO link to provide congestion-control for worst-case traffic [12], [22]. In the worst-case, a router with 800 Gbps links will have buffers for about 200 Gbits of data per Iink, equivalent to about 16.7 million maximum-size IP packets. The worst-case buffer-sizes grow larger as the link bandwidth increases.

Fig. 1b illustrates an IoT network spanning 28 cities in the European Union. Using the BDP buffer sizing rule, the BE-IoT router at Berlin with degree 6 will have a worst-case buffer size of about 100 million IP Packets (assuming 800 Gbps links). These large buffers increase the complexity, size, cost, power consumption and failure rates of BE-IoT routers. These large buffers are also responsible for the BE-IoT's excessive delays during times of congestion.

1) The Cisco CRS-3 Core Router: Cisco's largest BE-IoT core router chassis, the CRS-3, is shown in Fig. 2. One chassis has a capacity of 4.48 Tbps and consumes 7.66 kW of power, for an energy efficiency of 1.8 nanoJoules/bit (nJ/bit) in 2016 [23]. One chassis has dimensions (H,W,D) of 84, 24 and 42 inches respectively, and weighs 1,630 pounds (740 kilograms). According to Cisco, the CR3-3 system can scale to 322 Tbps of capacity, using a *"massive multi-chassis system"*. A single BE-IoT router with 322 Tbps capacity will require at least 72 chassis, with a combined weight of 117 thousand pounds and a combined power of 0.55 MegaWatts. A packet traversing one multi-chassis router could traverse several chassis, thereby consuming between 5...10 nJ/bit. According to Bolla et al [24], the software/hardware to process packet headers and perform routing functions accounts for about 60% of the power consumption of a BE-IoT router's data-plane. This example illustrates the state-of-the-art in *Best-Effort* (non-deterministic) IoT router technologies in 2016.

B. Non-Deterministic Networks are Prone to DOS Attacks

Non-deterministic networks are prone to *Denial of Service* (DOS) attacks. The vast majority of traffic in today's BE-IoT uses TCP (the *Transmission Control Protocol*) to regulate transmission rates and to reduce congestion [1][2]. However, the BE-IoT does not enforce users to adhere to TCP or any other flow control protocol at all [1]. Today's BE-IoT will allow any user(s) to transmit data at any rates to any destination(s) [1][21]. As a result, DOS attacks are relatively easy to create; Any adversary can cause several nodes to generate vast amounts of traffic directed to a single destination. The BE-IoT will attempt to deliver this traffic in layer 3, thereby causing the destination server to fail and creating a DOS attack.

The BE-IoT can deploy many *Firewall* middleboxes to stop unauthorized traffic from flooding a destination, but this creates a new set of security problems (please see the next subsection). The inability to

partition today's BE-IoT into mutually-exclusive interference-free SD-VNs allows for such DOS attacks, as well as targetted cyber-attacks, to occur.

C. Non-Deterministic Networks are a Security Threat

To combat the poor performance of the BE-IoT, the industry has deployed thousands of *'middleboxes'* into the BE-IoT infrastructure. Middleboxes perform several functions, including *'Load-Balancing'*, *'Network Address Translation'* (NAT) and *'Firewall'* functions. These middleboxes are unregulated; they often re-route packets and over-write packet headers [25]. As a result, it becomes difficult if not impossible to answer even simple questions on the performance or security of the BE-IoT.

A 2015 ACM SIGCOMM workshop on Internet verification outlined 6 basic questions on the BE-IoT that are difficult or impossible to answer [25]:

1) Which packets from source A can reach destination B?
2) Is traffic from source A isolated from traffic of source B?
3) Is the poor performance of a cloud service caused by the BE-IoT network or by the server providing the service?
4) Why is the backbone utilization so poor?
5) Is the load balancer distributing the load evenly?
6) What are the causes of mysterious packet losses?

It is difficult to provide cyber-security for *Smart Systems* operating over the BE-IoT in layer 3 when such basic questions cannot be answered. The networking community is now exploring *'Automated Theorem Proving'* (ATP) techniques to address network verification issues. The precise behaviour of each network node can be specified using logical modelling languages. The behaviour of the BE-IoT in layer 3 can then be verified by running complex theorem-proving programs. While this approach is important, it also raises a fundamental question: *"Is the BE-IoT infrastructure overly complex, such that it cannot be easily secured?"*. We believe the answer is 'Yes', and we argue that significant improvements in security can be achieved by adopting a much simpler *Deterministic Communications* paradigm operating in layer 2.

D. Security in Layer 3 - A Million Points of Vulnerability

Service-providers can currently create *Virtual Private Networks* (VPNs) in the BE-IoT in layer 3, to improve security. A VPN consists of one or more *Point-to-Point* (P2P) links between nodes in layer 3. A service-provider can use a technology such as *Internet Protocol Security* (IPsec) to secure the P2P links in the IP layer (layer 3) [26]. The *IPsec* technology uses a *Public Key* encryption algorithm, where the sender on a P2P link can encrypt an IP packet and encapsulate it into a new IPsec packet, which is transmitted to the destination. The destination receives the IPsec packet, extracts the encrypted IP packet, and decrypts it all in layer 3. Unfortunately the header of the IPsec packet is not encrypted, to allow the packet header to be processed at intermediate routers for routing decisions to be made in layer 3, which compromises security. In addition, the path taken through the layer 3 network is not necessarily fixed due to the use of unregulated middle-boxes, which also compromises security. The packets of a VPN could traverse a compromised router, to compromise the VPN security.

Another weaknesses of the IPsec model is that each router or middle-box in layer 3 is under autonomous control, which represents an extreme vulnerability. Each router or middle-box typically has a control processor which runs a Linux-based operating system, to run the BE-IoT control software. For example, Cisco routers run the Cisco IOS operating system, and Junniper routers run the Junos OS operating system. These routers communicate using socket-layer software such as the *Berkeley Socket Distribution* (BSD). Most implementations of IPsec use an open publically-available version of the BSD socket-layer software, and it has been established that this software has security failures [27].

It is extremely difficult to secure millions of routers or middle boxes in layer 3, each running BSD, since each node represents a security threat; The control processor in each node is vulnerable to cyber-attack, and can be compromised and controlled by a cyber-attacker. To achieve absolute cyber-security

in layer 3, each router or middle-box should be unconditionally secured, which represents an extreme challenge. Unfortunately, IPsec can provide only a modest level of security in the IP layer 3, which is insufficient for the *Smart Systems* of the future. Additional software such as the *Secure Shell* (SSH) can run on higher layers (4 and 5), to provide additional security. Unfortunately, even with several layers of security software in layers 3 and up, the BE-IoT security is weak [27], and cyber-attacks on the BE-IoT occur often and are growing in frequency.

III. Low Latency Networks for Smart Systems

A. The Industrial Internet Consortium

General Electric (GE) initiated the *Industrial Internet* effort to support the low-latency M2M communications required in the *Smart Systems* of the future [6], [7]. GE envisions that the transformation to *Smart Manufacturing and Industrial Automation* using M2M communications will enable the 4-th wave of the *Industrial Revolution* which began in the 18th century. GE argues that *Smart Manufacturing and Industrial Automation* should increase world GDP by $15 Trillion by the year 2030, by improving manufacturing processes and by reducing costs and waste. GE also argues that the *Industrial Internet* should control the production of about half of the world's GDP by 2030, or about $82 Trillion of world GDP.

In March 2014, the *Industrial Internet Consortium* (IIC) was formed by 5 companies (GE, IBM, Cisco, AT&T and Intel) to develop standards for the future *Industrial Internet*. By Feb. 2016 the IIC included over 237 companies, indicating significant industrial support. The IIC released a first draft of the *Industrial Internet* reference architecture in Oct. 2015. The draft specifies only the high-level requirements, and it is broad rather than technically deep. It does not discuss deterministic M2M communications, thereby avoiding the NP-Hard problems associated with deterministic communications.

B. The ITU and IEEE Tactile Internet Efforts

In Aug. 2014, the ITU published a 'Technology Watch' paper on the *Tactile Internet* [8]. Like the *Industrial Internet*, the *Tactile Internet* is expected to offer extremely low latency with high availability, reliability and security. Applications include *Industrial Automation*, *Smart Transportation Systems*, and *Smart Healthcare* systems. From a technical perspective, the goals of the *Industrial Internet* and *Tactile Internet* are similar, to achieve a future IoT with exceptionally low latency and high reliability. The ITU *Tactile Internet* effort emphasizes the use of 5G wireless communications, and focusses on applications with latencies of 1 millisecond (ms) or less. The speed of light in fiber is about 200 kilometers (km) per millisecond. Hence, *Smart Systems* which are distributed over distances larger than about 200 km will require a low-latency IoT core network, as proposed in this paper. The ITU has not released any technical description of a *Tactile Internet* reference architecture as of 2016, and makes no mention of deterministic M2M communications.

In March 2016, the IEEE approved the P1918.1 *Tactile Internet* project (www.ieee.org). There are no publications yet from this project, but it also plans to develop the reference model framework and architecture of the *Tactile Internet* for mission-critical applications, such as *Smart Manufacturing*, *Smart Transportation* systems and *Smart Healthcare* systems.

C. IETF Activities in Deterministic Networks

In Oct. 2015, the IETF approved the *Deterministic Networking* working group to explore how to add deterministic communications to the BE-IoT to achieve low latency. The IETF presented the *Deterministic Forwarding Per Hop Behaviour (PHB)* RFC which states that an application can explicitly reserve time for the transmission of the packets of a deterministic flow through a router [15], [16]. However, the draft explicitly states that it is a set of abstract high-level requirements: for maximum flexibility the draft does not specify which router/switch architecture should be used, or which routing/scheduling algorithms

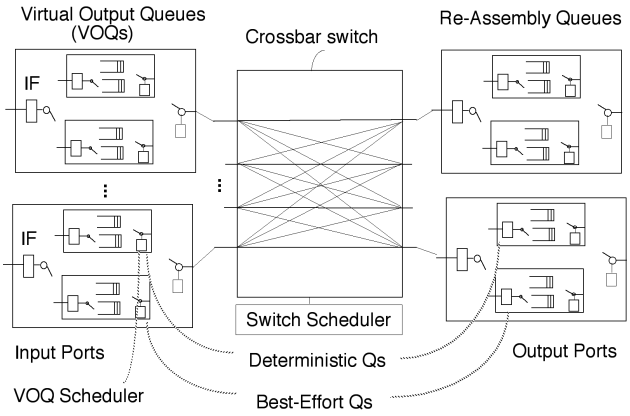


Fig. 3. Basic CIOQ switch with 2 traffic classes, the Deterministic and Best-Effort classes.

should be used, thereby avoiding several NP-Hard problems associated with transporting deterministic traffic flows in practical networks [4].

The IETF is allowing for relatively low minimum link utilizations of 50% for deterministic traffic flows, to simplify the scheduling problem [15], [16]. Reference [4] shows that these low utilizations can cost the industry up to $37 Billion per year in unnecessary capital costs. As of 2016, the IETF *Deterministic Networking* working group did not address security issues in deterministic networks.

D. Is a Cooperative International Effort Worthwhile?

All of these 4 international efforts have been initiated in the last few years, reflecting the urgent global desire to improve the current IoT, but each effort has not yet acknowledged the existence of any of the other international efforts. For example, the ITU's *Tactile Internet* effort does not acknowledge the much larger *Industrial Internet* effort, and the IEEE's *Tactile Internet* effort does not acknowledge the ITU's effort or the *Industrial Internet* effort. Given the common goals of these international efforts, we argue that some cooperation between these international efforts may be worthwhile.

IV. A Secure Deterministic Industrial-Tactile IoT

By definition, a *Deterministic Transport Connection* (DTC) must receive a deterministic (or guaranteed) transmission rate through the IoT, which cannot be affected by other traffic flows [10][11][12]. DTCs can be established in layers 2 or 3, in wired or wireless networks, as shown in Fig. 1a (top and bottom). In this section, we assume the DTCs are established in layer 2 to provide enhanced security. It would be difficult to achieve the same security in layer 3 as explained in section II. The following properties hold for DTCs established in layer 2.

Property 1: Every DTC has a source node s, a destination node d, and a deterministic data-rate to be supported by the network from s to d. A deterministic data-rate can also be called a 'guaranteed data-rate'. The data-rate can be expressed as a guaranteed number of packet transmissions within a scheduling-interval called the scheduling-frame [11][12]. To enhance security, a DTC will always transmit packets at its guaranteed rate (transmitting very short packets if necessary to maintain its guaranteed rate; please see section IV-C ahead).

Property 2: Every DTC is routed along a fixed path of layer-2 switches and links in the network, from node s to node d [11][12]. The DTC along a fixed path in layer 2 is also called a *Secure Deterministic Virtual-Link* (SD-VL).

Property 3: To provision a SD-VL, an SDN control plane will reserve buffer space (i.e., a virtual flow-queue [12]) and the guaranteed data-rate in every layer 2 switch traversed by the SD-VL. The SDN control plane will reserve the guaranteed data-rate on every layer 2 link traversed by the SD-VL [11][12].

Property 4: By definition, a DTC can transmit data at a deterministic data-rate, and it is therefore immune to congestion and interference from any other traffic flow. Otherwise it could not possibly be a DTC, which would lead to a contradiction [11][12]. It is clear that a key requirement of a DTC is the *reservation* of sufficient buffer-space and bandwidth in every switch and every link, so that every DTC will never experience congestion and interference from any other traffic flows.

Property 5: In the steady-state, a deterministic switch will receive packets from a SD-VL at its guaranteed data-rate unconditionally (i.e., with probability 1). Similarly, in the steady-state a deterministic switch will transmit the packets of a SD-VL at its guaranteed data-rate unconditionally [11][12]. Property 5 follows by implication from properties 1-4. The steady-state is achieved within a few scheduling frames [34], i.e., typically within 10-100 microseconds.

Property 6: Any unauthorized packet from a cyber-attacker cannot be transmitted through any layer 2 deterministic switch. Property 6 follows by implication from properties 1-5. An unauthorized packet will not have any reservations for buffer-space or bandwidth in any switch (or link) in layer 2. A deterministic switch in layer 2 will not dedicate a virtual flow-queue to store packets for an unauthorized traffic flow [12], and it will not receive nor forward any packets of an unauthorized traffic flow. As a consequence, a SD-VL is immune to DOS attacks and targetted cyber-attacks from unauthorized traffic flows. (Please note that the same properties do not hold for the layer 3 IoT. According to section II, it is difficult to secure BE IoT routers operating in layer 3).

Property 7: To provide protection against router or link failures, an application can reserve one or more edge-disjoint SD-VLs between nodes s and d in layer 2. An application can for example transmit a redundant data stream along a redundant path(s), to provide protection from link or switch failures.

Property 8: A collection of SD-VLs can be logically combined into a layer 2 *Secure Deterministic Virtual-Network* (SD-VN). By Property 4, the traffic within one SD-VN is immune to congestion and interference from the traffic in any other SD-VN. By Property 6, the traffic within one SD-VN is immune to cyber-attacks from unauthorized traffic sources.

Property 9: To achieve enhanced security in layer 2, an application can encrypt the packets of a SD-VL at the source node s, and decrypt the packets at the destination node d, using a *Public Key* encryption algorithm. The data transmitted over a layer 2 SD-VL is never observed in the layer-2 deterministic switches [12], so that the entire packet can be encrypted at the source and decrypted at the destination.

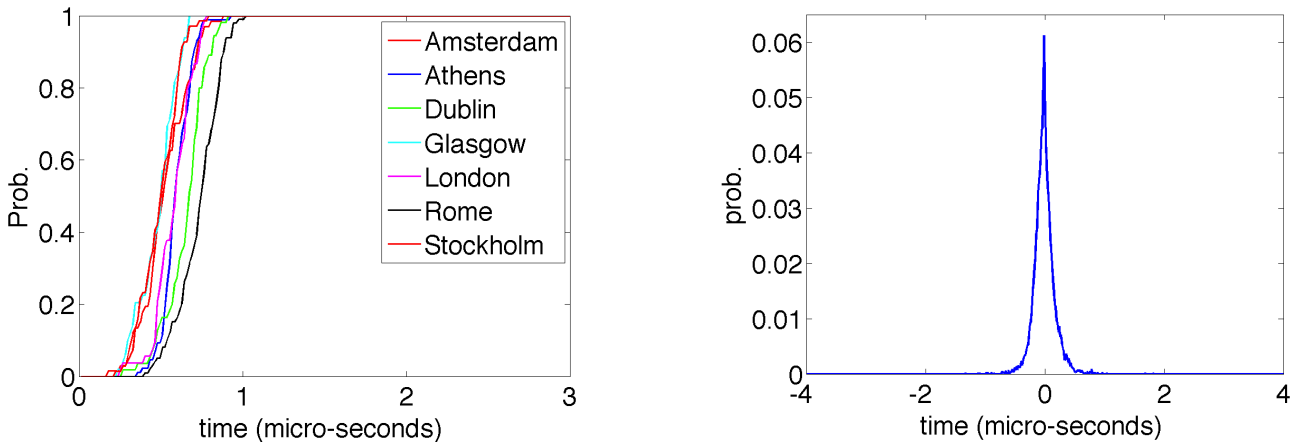
The remainder of this section will explain these properties in more detail.

A. A SD-IoT for the European Union

Fig. 1b illustrates a *Secure Deterministic IoT* core network for the European Union (EU), with 28 nodes (cities) and 82 edges. The bold lines represent fiber-optic links between cities. The dotted lines represent congestion-free DTCs between cities in layer 2. The DTCs in layer 2 are also called *Secure Deterministic Virtual Links* (SD-VLs), since an Internet router in layer 3 treats these links as very-low latency virtual one-hop layer 2 links between remote cities.

Our SDN control plane can program millions of VNs into the core network in layers 2 or 3, as shown in Fig. 1a and 1b. Internet routers are interconnected at the 'Internet Protocol' (IP) layer 3, and perform routing functions to determine the routes that packets take through the Internet. Our SDN control plane can embed VNs directly into the layer 3 IP network, using predetermined routes, as shown at the top of Fig. 1a. In our SDN control plane, the predetermined routes are found using a *Maximum Flow Minimum Cost* routing algorithm [28]. Traffic flows can be distributed over multiple paths to improve reliability and availability [11], and backup paths can also be provisioned.

However, the VNs can also be embedded into a layer 2 *'Optical Transport Network'* (OTN) network, as shown at the bottom of Fig. 1a. In the OTN, each DTC will bypass several routers in layer 3 to significantly improve energy efficiency. Current IoT routers consume between 1...10 nJ/bit per bit transmitted [29], [30]. The FPGA-based packet-switches with optical IO proposed ahead are expected to use about 1-10 pJ/bit in 2019, resulting in a potential energy savings of about 1,000 times.



(a) (b)

Fig. 4. (a) End-to-end queuing delay CDF on Virtual Links from Berlin. Destination cities are shown. (b) Age jitter distribution for all flows.

In Fig. 1b, three SD-VNs are embedded into the network. Three cities (Madrid, Stockholm and Berlin) each have a SD-VL directed to every other city in the network. A SD-VN with 27 SD-VLs originating at Berlin and directed to every other city is shown by the bold red lines. *Smart Systems* such as *Smart Cities*, the *Smart Power Grid*, *Smart Transportation Systems*, and *Smart Healthcare* systems can all reserve their own distinct, secure, and interference-free SD-VNs in layer 2, through the SDN control plane.

B. A Deterministic Packet Switch Design

There are several different packet switch designs, including switches with

- *Input Queueing* (IQ) or *Output Queueing* (OQ),
- *Combined Input and Output Queueing* (CIOQ), and
- *Combined Input, Crosspoint and Output Queueing* (CIXOQ).

Large OQ switches are considered to be intractable and are not often used [12]. The performance of a deterministic CIXOQ switch for the IoT has been explored in [4]. In this paper, we explore a simpler deterministic CIOQ switch with a more complex scheduling algorithm for the IoT.

Fig. 3 illustrates a simple CIOQ switch which supports deterministic traffics flows with 100% throughput and with deterministic delay and jitter guarantees [12]. Variable-size Internet packets arrive at the Input Ports (IPs). The may be re-formatted into smaller fixed-sized blocks for transmission through the switch. The variable-size Internet packets may be reassembled at the Output Ports (OPs). In an NxN CIOQ switch, each input port has N *Virtual Output Queues* (VOQs), where VOQ(i,j) stores data which arrives at IP(i) and which departs on OP(j). Each VOQ in Fig. 2 supports 2 traffic classes, the *Deterministic* and *Best-Effort* classes. However, a VOQ can support many different prioritized traffic classes, such as a new Deterministic class and the 3 existing traffic classes (EF, AF, DE) in the IETF's *Differentiated Services* (DiffServ) model.

Each NxN CIOQ switch can have an NxN matrix of deterministic traffic rates to be supported between the input and output ports. Ref. [12] presents a very fast recursive scheduling algorithm, which can schedule the transmission of packets through a CIOQ switch with near-minimal delay and jitter and with 100% throughput. The algorithm recursively decomposes the NxN matrix to achieve an extremely low-jitter transmission schedule with 100% throughput. The algorithm in [12] provides a very fast approximate solution to the NP-hard problem of scheduling deterministic traffic flows through CIOQ switches to minimize delay and jitter without speedup. This algorithm allows our SD-IoT network to achieve 100% link utilizations, rather than the relatively low 50% link utilizations targeted by the IETF.

The scheduling algorithm yields a deterministic transmission schedule called the *TX-Schedule* for each IP [12]. At each IP, the *TX-Schedule* identifies a VOQ with a reservation to transmit data for each time-slot of the scheduling frame. The *TX-Schedule* provides each input-port with a deterministic rate of transmission, to the output ports of the switch. For each time-slot when an IP receives service, a second *IP-Class-Schedule* can specify the traffic class (or traffic flow) to be serviced. Once packets are re-assembled at the output ports, an optional *OP-Class-Schedule* can specify the traffic class (or traffic flow) to be serviced. The switches in Fig. 1b do not need to be synchronized, as the schedules can be circularly rotated and remain valid. This point is very important, because the routers/switches in the EU network shown in Fig. 1b need not be synchronized to microseconds or nanoseconds of precision [4], which would be very hard to implement over distances of thousands of miles/kilometers.

C. Secure Deterministic VNs

Each SD-VL s provably immune to congestion, DOS attacks and targeted cyber-attacks, as stated in Properties 1-9. The behaviour of each layer 2 switch is deterministic and independent of the contents of the packets it transports, since it uses deterministic hardware. As stated in Property 6, unauthorized packets generated by a cyber-attacker cannot be forwarded out of a deterministic layer-2 switch.

A scheduling frame consists of a F time-slots, where a maximum-sized packet can be transferred through a switch in one time-slot. Our SDN control plane will configure every deterministic switch in layer 2 with several deterministic transmission schedules, which completely determines the movement of packets through the switch for every time-slot in a scheduling frame. Each deterministic switch reserves an exact number of time-slot reservations for the packets of each deterministic traffic flow, and monitors the number of packet transmissions per traffic flow. According to Property 5, a deterministic layer-2 switch will receive (and transmit) packets at the exact guaranteed data-rate of an SD-VL. Any inconsistency, such as receiving a packet from an unauthorized traffic flow, or receiving too few or too many packets per authorized traffic flow, will be detected and generate an alarm message to the SDN control plane for corrective action. The security of the layer 2 SD-IoT can be summarized with the following properties:

Property 10: In the steady-state, the appearance of any unauthorized packets in layer 2, even 1 extra packet, can be detected within one scheduling frame by the SDN control plane. The duration of one scheduling frame is typically about 15 microseconds.

Property 11: Each deterministic switch in layer 2 does not contain a micro-processor to process packet headers. Such a microprocessor usually executes machine code which stored in RAM *(Random Access Memory)* which can be compromised by a cyber-attacker. (If a micro-processor is used in a deterministic switch, for example to decrypt packets from the SDN control plane, then its program must execute out of *Read-Only Memory* (ROM), to yield a deterministic system which cannot be compromised.)

Property 12: The security challenges of all unregulated middleboxes operating in layer 3 have been removed from the SD-VNs operating in layer 2.

D. A Secure SDN Control-Plane

As stated in Properties 1-12, the behaviour of each deterministic packet-switch in layer 2 is completely determined by the SDN control plane. SD-VNs are typically established for longer periods of time, i.e., for hours, days or years. The SDN control plane must configure all the deterministic transmission schedules in each switch periodically when the traffic rates change (i.e., every few minutes of hours). Assume that the *Deterministic Ethernet* packet format [31] is used in the layer-2 SD-IoT network. This packet format allocates 3 extra bytes for a VN label to the basic Ethernet packet, which allows for the specification of 16 million different VNs in layer 2.

Let the VN with label 0 be the *Control-VN*, which the SDN control plane uses to configure each packet-switch in layer 2. The SDN control plane can encrypt all packet transmissions over the *Control-VN* in layer 2, by using a *Public Key* encryption algorithm. Every packet-switch can have a *Public Key* and a *Private Key*. The SDN control plane can send encrypted packets to any switch using the public key. Each

switch can decrypt the packets it receives over the *Control-VN* in layer 2 using its private key. (The SDN control plane can also configure the packet-switches using a purely Private Key encryption algorithm.)

The US government has approved the *Advanced Encryption Standard* (AES) algorithm for secure US government communications [32]. The SDN control plane can also use the AES algorithm to encrypt the packet transmissions in layer 2, to conform with the US government security standards. Cyber-security in the SD-IoT is enhanced by several key concepts:

- The SD-IoT can embed millions of distinct, interference-free SD-VNs in layer 2, which cannot transport unauthorized traffic flows, such as traffic from a cyber-attacker;
- Deterministic packet transmissions can be encrypted and decrypted in layer 2, something which is not possible in the current VPN technology used in the BE-IoT;
- All control-processors which execute software which resides in RAM have been removed from the deterministic switches in layer 2, to avoid the points of vulnerability described in section IIC.

The security of the SD-IoT relies upon the security of the SDN control plane. The issue of how to secure an SDN control plane is the subject of much current research [33]. In our SD-IoT, the security of the SDN control plane can be improved by using a *Triple Modular Redundancy* (TMR) voting system, to have 3 distinct Control-VNs running in parallel. In TMR, 3 modules perform every function and a voting system is used to detect abnormal behaviour of any module. For example, with 3 Control-VNs running in parallel, a cyber-attacker could not compromise any one Control-VN without being detected. The security of each Control-VN is achieved by using a distinct AES encryption key. It would be very difficult for a cyber-attacker to compromise one Control-VN at all, and it would be virtually impossible for a cyber-attacker to compromise 3 Control-VNs simultaneously to avoid detection (each with a distinct AES key).

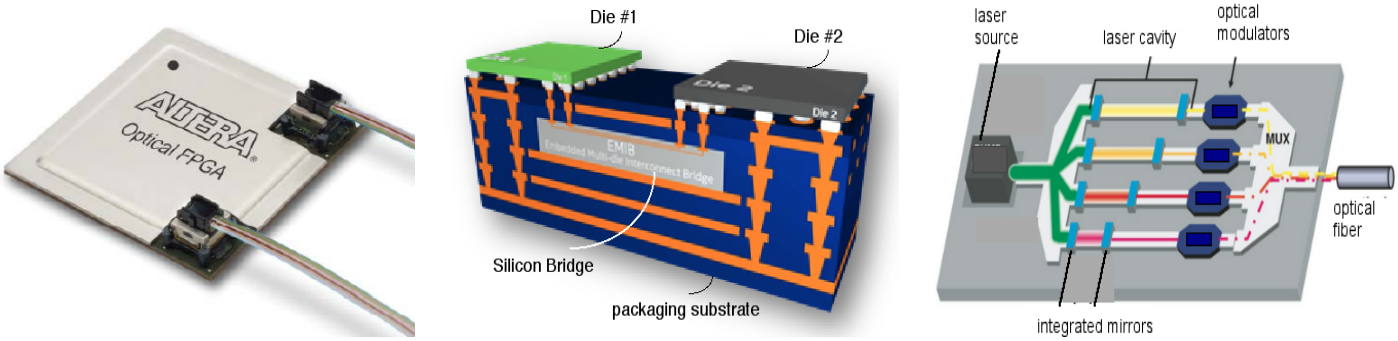
V. Experimental Results for a EU Core Network

Our SDN control plane programmed 744 VLs in the EU core network shown in Fig. 1b to achieve a link utilization of 100%, and the performance was determined. A Scheduling-Frame with 1,024 time-slots was used. Each time-slot was sufficient to transmit a maximum-size IP packet over a VL. Assuming 1500 byte IP packets and 800 Gbps edges, a time-slot has a duration of 15 nanoseconds. (A 800 Gbps edge could be partitioned into 2 parallel 400 Gbps channels, in which case a time-slot consists of 30 nanoseconds).

The network performance is deterministic, and was determined using a software simulator which traces the deterministic system states and computes the system performance. The software simulator has been validated by comparing its performance with several hardware testbeds which were synthesized on an FPGA [4], [34]. Several deterministic core networks with between 10 and 30 simple deterministic routers were synthesized in hardware on an Altera FPGA. The FPGA hardware testbeds transmit between 100 million and 400 million packets per second [4], [34]. The performance of the hardware testbeds and the software simulator are in perfect agreement. Also, reference [12] presents theoretical bounds on the end-to-end latencies and jitters.

References [11][12] establish the theory that deterministic packet switching, combined with low-jitter scheduling, can reduce the end-to-end delays in the IoT to the speed of light in fiber. Fig. 4a illustrates the end-to-end IoT queueing delays along several SD-VLs originating at Berlin, expressed in microseconds (μsec). Fig. 4a does not include the fiber latency. The queueing delays in Fig. 4a are less than 1 μsec. Recall that the speed of light in single-mode fiber is about 200 kilometers (km) per millisecond. Consider the SD-VL between Berlin and Dublin. The length of the SD-VL in Fig. 1b is at least 1315 km, depending upon which path is taken through the network. The fiber latency therefore exceeds 6 milliseconds. The end-to-end queueing delay along the VL (less than 1 μsec) is a few thousand times smaller than the end-to-end fiber delay (6 milliseconds), corroborating the theory presented in [11][12].

References [10][12] establish the theory that deterministic packet switching, combined with low-jitter scheduling, can reduce the packet jitter to very small values. Fig. 4b illustrates the jitter of the packets leaving a SD-VL (which is averaged over all SD-VLs in the EU network). The jitter is defined as the



(a) (b) (c)

Fig. 5. (a) An Altera FPGA with optical IO (www.altera.com). (b) Intel's Embedded Multi-Die Interconnect Bridge (EMIB) Technology. (c) Intel's silicon photonics WDM modulator design.

deviation in the age of the departing packets of a deterministic connection, relative to the mean age of the packets. In Fig. 4b, the jitter is less than 1 μsec. Each packet exits the EU network with an age that deviates from its perfect age by less than 1 μsec. These jitters are exceptionally small when compared to the end-to-end fiber delays in the EU core network, which are measured in milliseconds, corroborating the theory presented in [10][12].

Reference [12] establishes the theory that deterministic packet switching, combined with low-jitter scheduling, can reduce the worst-case buffer sizes of IoT routers by a factor of 1000 times of more. Our experiments show that a maximum buffer size of 104 packets occurs in the Berlin node of the EU core network shown in Fig. 1b, when operating at 100% load with 800 Gbps links. Recall that the worst-case *BDP* buffer-sizing rule would allocate about 100 million packet buffers in the Berlin node (using 800 Gbps links). The use of deterministic packet-switching has reduced the worst-case buffer sizes from about 100 million packets down to 104 packets, representing a reduction by a factor of about one million times. Reference [12] states that buffer-size reductions of a factor of 1,000 times or more are achievable with deterministic communications, but our experimental results indicate the buffer-size reductions can be much larger for Terabit capacity optical links. These dramatic buffer-size reductions allow for the creation of single-chip layer-2 deterministic photonic packet switches using FPGAs or ASICs with optical IO, which are described next.

VI. FPGAs WITH OPTICAL IO

The potential of Silicon Photonics transceivers to transform the communications industry is described in [35]. Silicon Photonics transceivers can be manufactured using CMOS technologies; they can be produced in quantities of millions with very low costs. FPGAs (*Field Programmable Gate Arrays*) are CMOS *Integrated Circuits* whose functionality can be programmed dynamically in the field by using *Computer Aided Design* (CAD) tools. Current FPGAs can contain up to a few million programmable logic gates, a few hundred megabits of high speed memory, and can reach performances of several Teraflops per second (for single-precision floating-point arithmetic). Given their extreme flexibility, FPGAs are produced in quantities of millions with very low costs. Unfortunately, the impressive on-chip performance of FPGAs is severely limited by the inability to move vast amounts of data onto and off the chip easily.

The electrical IO bandwidth of FPGAs is currently limited to about 1 or 2 Terabits per second, using high-power electronic IO signalling technologies which can consume up to 80 W of power. The integration of electrical FPGAs with optical IO technologies represents a very attractive low-cost method to introduce optical technologies into the communications and computing industry. Several integrated FPGAs with optical IO have been demonstrated years ago [36], [37], before the availability of low-cost Silicon Photonics transceivers (these are described ahead). FPGAs which are integrated with low-cost Silicon Photonics transceivers to provide Terabits/second of optical IO bandwidth should be available

within a few years. These devices will have a significant impact on the design of cloud data-centers, supercomputers and the IoT.

In 2015 Intel purchased Altera for $16.7 Billion US, and Altera announced the development of FPGAs with optical IO using Intel's *Embedded Multi-Die Interconnect Bridge* (EMIB) technology. An FPGA with optical IO will consist of one or more FPGA die, and several Silicon Photonics transceiver die, which are interconnected and packaged onto a traditional *Integrated Circuit* (IC) package as shown in Fig. 5a. A 'Silicon Bridge' is a small integrated circuit which can be inverted and bonded to 2 or more die, thereby forming a 'bridge' with numerous high-speed low-energy wires between the die, as shown in Fig. 5b. FPGAs integrated with Silicon Photonics transceivers are ideally suited to implement the simple deterministic packet-switches proposed in this paper. As shown in the last section, the use of deterministic packet switches can reduce the worst-case buffer sizes by factors of about 1 million times for terabit-capacity optical links, thereby allowing a simple deterministic packet-switch to fit on a single FPGA.

A Silicon Photonics WDM transmitter with 4 optical channels is shown in Fig. 5c [38]. A laser source can be split into 4 channels, where each channel can be tuned to a distinct wavelength and modulated with digital data. All 4 channels can then be combined and transmitted onto a single fiber. Currently, both Mellanox and Luxtera manufacture Silicon Photonics transceivers with energy efficiencies of 35 pJ/bit, or equivalently 35W per Tbps [39], [40]. Recently, more energy-efficient transceivers have been announced. In 2015, Fujitsu demonstrated a transceiver with an energy efficiency of 5 pJ/bit [41]. In 2016, researchers have demonstrated transceivers with energy efficiencies of 2 pJ/bit [42]. Using these transceivers, each Terabit of optical bandwidth will require only 2 watts of power.

According to [43], a high-performance *Integrated Circuit* has a maximum power of about 200W, and a reasonable design can allocate about 40W for optical IO and 40W for electrical IO. Our proposed photonic packet-switches will have very little electronic IO, and we therefore assume that 80W can be allocated for Silicon Photonics transceivers, leaving 120W available for electronic functions within the FPGA. (However, it is also feasible to allocate for example 50W for optical IO and 150W for electronic functions.)

Table 1 illustrates our projections for the Silicon Photonics transceivers used for integrated FPGA-based packet switches, assuming (i) 80W of power is available for the transceivers, and (ii) the cost for optical IO is limited to under $1,500. The top row illustrates our assumed optical bandwidth capacity per FPGA. We assume a FPGA with 20 Tbps of optical capacity will be available by 2022. The 2nd row illustrates the available energy per bit, assuming 80W is available for optical IO. The available energy per bit in 2022 is 4 pJ/bit (a target which can be met today). The next 2 lines (denoted with a *) illustrate Intel's goals for Silicon Photonics technologies [44]. Intel hopes to achieve an energy-efficiency of 250 femtoJoules/bit (fJ/bit) in 2022, and a cost for optics of $20 per Tbps.

The projected energy use and cost of several switch designs are shown in the bottom 2 rows, assuming Intel's energy efficiency and cost targets are met. In 2022, a switch with 20 Tbps of optical bandwidth should be feasible, at a reasonable cost and power for the optical IO ($400 and 5W). In the future, higher optical bandwidths should be available at lower costs, as projected by Intel [44].

Tucker has shown that CMOS switches have comparable energy efficiency to all-optical switches [29], [30]. According to [45], the power used for a 20 Tbps electronic packet switch implemented in an *Application Specific Integrated Circuit* (ASIC) with a 10-nm CMOS process is under 10W. The use of an FPGA rather than an ASIC typically increases the power by a factor of ≤ 8, and hence the power of a 20 Tbps electronic switch implemented in an FPGA using 10-nm CMOS will be $\leq$ 80W. (This figure does not include the optical IO power, which will be about 5W).

To summarize, assuming Intel's targets are met, a *simple deterministic* FPGA-based Silicon Photonics packet-switch with 20 Tbps capacity will require about 200W of power and can reside on one *Integrated Circuit* package as shown in Fig. 5a. It can replace about 5 chassis of the Cisco CRS-3 router shown in Fig. 2, which would consume about 38 KW of power and weight about 8,150 pounds.

VII. Secure Big Data Green Cloud Computing

Big Data analytics are fundamental to enable the *Smart Systems* of the future. The billions of smart devices to be connected to the future IoT are expected to generate vast amounts of data, which will be stored in cloud data-centers distributed around the globe. This vast amount of data will be processed using *Big Data* cloud computing systems, to optimize the *Smart Systems* of the future.

The concept of *Big Data* cloud computing with exascale capacity was first explored in [45]. Reference [45] assumed that a Iow-latency IoT core network could be constructed using ASICs or FPGAs integrated with optical IO. Several integrated FPGA-based packet-switches with optical IO have been demonstrated using similar technologies [36], [37]. In this section, we explore the topic of *Big Data* green cloud computing, using FPGAs combined with Silicon Photonics transceivers.

Governments in the USA, European-Union, China, India and Asia are all pursuing Exascale computing initiatives, aiming to achieve Exascale computing performance by 2020. An Exascale super-computer can perform approximately 1 Exa floating-point operations per second (one billion GigaFlops/sec). An Exascale super-computer will require about 1 million desktop computers in the year 2016, (or fewer processors performing at a higher rate). Several governments are building dedicated super-computers to meet these goals, incurring capital costs measured in the 100s of $Millions.

Cloud Computing offers an alternative approach to achieve *Big Data* analytics. A primary advantage of cloud computing is reduced cost. A cloud computing system can utilize existing cloud data-centers, thereby removing the capital costs of building a dedicated super-computer. Cloud computing can also reach larger aggregate computing capacities potentially measured in the tens of Exa-Flops per second, which can be shared by a large number of cloud computing tasks running simultaneously. Cloud data-centers are often under-utilized with typically $\leq 50\%$ utilization [46]. A *Big Data* cloud computing system can utilize the idle capacity of existing cloud data-centers, thereby improving data-center utilizations and energy efficiencies, while accomplishing useful work at the same time.

In 2013, Cycle Computing reported that it completed a cloud computing task with about 1.21 Petaflops/sec of peak aggregate computing power in about 18 hours, using 156,314 processing cores distributed over 8 regions of Amazon Web Services (AWS) around the globe (www.cyclecomputing.com). This task represents about 264 years of computing time on a single standard desktop workstation. A stand-alone supercomputer with the same performance would cost about $68 million to build, whereas this cloud task cost $33,000. Clearly, cloud computing offers significant financial incentives for industry, the availability of supercomputer-class performance on a *'pay-as-you-go'* basis, without the upfront capital costs of buying a large supercomputer.

A. The Bandwidth Challenge

The US *Department of Energy* (DOE) has summarized several technical challenges to achieving *Big Data* computing with exascale performance [47], [48]. Some key challenges include:

- Scalability to millions of CPU cores;

TABLE I
Silicon-Photonics Switch Projections

	2016	2019	2022
Optical BW	800 Gbps	4 Tbps	20 Tbps
Avail. E/bit	100 pJ/bit	20 pJ/bit	4 pJ/bit
Intel* E/bit	11 pJ/bit	1.7 pJ/bit	250 fJ/bit
Intel* Cost	$1000/Tbps	$160/Tbps	$20/Tbps
Optical Cost	$800	$640	$400
Optical Power	8.8 W	6.8 W	5 W

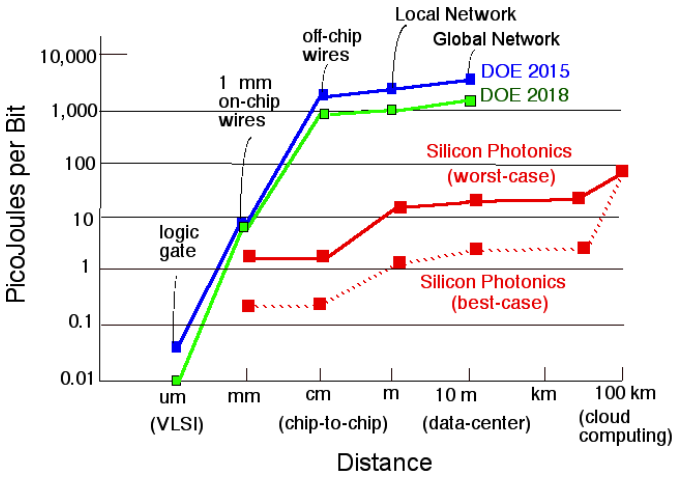


Fig. 6. Energy per bit vs. distance requirements.

- Development of energy-efficient computational units (CPUs and GPUs);
- Development of a higher-performance energy-efficient communications network to support exascale computations.

In the 2011 Mallegan Report on Cloud Computing [47], [48], the US DOE concluded that cloud computing could not compete with stand-alone super-computers, primary due to

- The very high latencies of the BE-IoT;
- The very high energy use of the BE-IoT.

In this section, we explore the impact of an ultra-low latency deterministic IoT core network on *Big Data* cloud computing.

According to *Moore's law*, the performance of electronic integrated circuits (ICs) has been increasing exponentially for the last few decades. The number of transistors on an integrated circuit (and its performance) typically doubles every 24 months. According to [24], the bandwidth capacity of IoT routers is also increasing exponentially, with a faster growth rate. The bandwidth capacity of a IoT router typically increases by a factor of 2.5 every 18 months. Recall that Cycle Computing has achieved roughly 1 Petaflop/sec performance in 2013. Assuming that (i) the performance of cloud computing systems is limited by the IoT communications bandwidth, and (ii) that the long-term growth rate in IoT bandwidth capacity remains unchanged, then *Big Data* cloud computing systems with exascale performance should be achievable in the year 2024, i.e., within a decade. Clearly, these long-term growth rates favour cloud computing over stand-alone supercomputers.

Silicon Photonics technologies are expected to result in a dramatic boost to communication capacity and a dramatic reduction in cost in a short period of time (i.e., the next 6 years), which will improve the performance of cloud computing systems even faster than the capacity growth rates presented in [24]. The bandwidth of a SD-IoT spanning the EU can be increased dramatically by using Silicon Photonics technologies. Assume that an FPGA-based packet-switch with 20 Tbps of optical bandwidth is available in 2022, as shown in Table 1. A single FPGA-based deterministic packet-switch with 20 Tbps of capacity could replace a Cisco CRS-3 router, which would consume about 38 KW of power and weigh about 8,150 pounds, as shown in the last section. The low cost of FPGAs with optical IO favours the use of cloud computing; it can be considerably cheaper to interconnect existing under-utilized data-centers with layer 2 deterministic IoT core network, than it is to construct a dedicated super-computer.

B. The Energy Efficiency Challenge

Fig. 6a illustrates our projections on the energy needed to transmit a bit, versus the distance travelled. The data for the two upper energy curves (2015 and 2018) for distances $\geq$ 1 mm is from the US DOE [47].

The data for these two upper curves, for distances < 1 mm, is based on projections for 10-30 nm VLSI technologies from [29], which cites the *International Technology Roadmap for Semiconductors* (ITRS). The US DOE estimates that the energy needed to transmit each bit out an integrated circuit package in 2018 is just under 1 nJ/bit. It also estimates that the energy needed to transmit each bit through the local and global networks in 2018 are about 1 and 10 nJ/bit respectively.

The two lower curves show our *best-case* and *worst-case* projections for Silicon Photonics transceivers. In 2016, Silicon Photonics transceivers with energy efficiencies of about 2 pJ/bit are feasible [42]. For the worst-case analysis, we assume no further improvement in energy efficiency relative to [42]; every Terabit/second of optical bandwidth will require 2 watts of power. The energy use is relatively independent of the distance travelled for short distances ranging from a few centimeters to 1 meter. For the local and global networks, we assume a packet traverses 5 and 10 switches each with an energy use of 4 pJ/bit, for total energies of about 20 and 40 pJ/bit. Silicon Photonics transmissions can travel up to a few 10s of kilometers over single-mode fiber, but for transmission over longer distances (i.e., 100 km), Fujitsu Flashwave transponders will be required to re-format the optical signals for long distance transmission. The Fujitsu Flashwave 9500 transponders require about 91 pJ/bit [49], and are responsible for the jump in power in Fig. 6a for distances $\geq$ 100 km.

The lower red curve shows the best-case targets for Silicon Photonics transceivers assuming Intel's projections from Table 1 are reached. For short distances $\leq$ 100 km, the best-case energy use is about 1 order of magnitude lower relative to the worst-case. For transmission over long distances, Fujitsu Flashwave transponders will likely be required and the power for the transponders dominates for distances $\geq$ 100 km.

We observe that the worst-case Silicon Photonics energy projections, which are achievable in 2016, are about 1...2 orders of magnitude better that the DOE energy targets for the global network performance needed in 2018. Silicon Photonics will certainly revolutionize the design of data-centers, super-computers and the IoT, even with a worst-case analysis. If Intel's best-case targets shown in Table 1 are met by 2022, then Silicon Photonics will offer about 2...3 orders of magnitude improvement in energy efficiency over the US DOE energy targets for the year 2018 shown in Fig. 6.

C. The Scalability Challenge

Fiber-optic cables typically consist of about 100 strands of fiber. Typically only a few fibers are active, since the capital costs of adding layer-3 BE-IoT routers (such as the Cisco CRS-3 router shown in Fig. 2) to activate all the fibers can grow to 10s or 100s of millions of dollars. By constructing a SD-IoT core network in using low-cost FPGA-based switches in layer 2 as shown in Fig. 5a, the aggregate bandwidth capacity of the IoT can be increased significantly, potentially by 1 or 2 orders of magnitude, while simultaneously reducing energy use. (Multi-threading can also be used to hide the fiber latencies of cloud computing, which are comparable to hard disk latencies.)

D. Secure Green Cloud Computing using SD-VNs

Our SDN control plane can embed millions of SD-VNs into layer 2, to support secure, distinct, mutually-exclusive and non-interfering *Big Data* cloud computing tasks. Fig. 7 illustrates the embeddings of 2 SD-VNs into the EU core network. According to Property 4, the traffic in an SD-VN is immune to congestion and interference from traffic in other SD-VNs. According to Property 6, an SD-VN will not transport any unauthorized packets from a cyber-attacker, and the SD-VN is therefore immune to DOS attacks and targetted cyber-attacks from unauthorized traffic sources. According to Property 9, each cloud computing task can encrypt its packets before transmission over its SD-VN in layer 2, to achieve a level of cyber-security which is not possible with today's BE-IoT. According to Property 10, the appearance of even a single un-authorized packet from a cyber-attacker can be detected within approximately 15 microseconds, and corrective action can be taken by the SDN control plane.

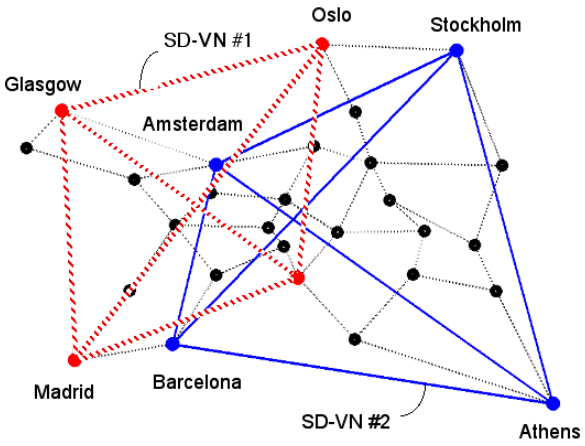


Fig. 7. 2 SD-VNs embedded into the EU core network. First VN uses bold dotted lines, second VN uses bold solid lines. Fiber optic links are shown with light dotted lines.

VIII. Conclusion

The future *Internet of Things* network is expected to support the demanding *Smart Systems* of the 21-st century. Future smart systems will include *Smart Cities*, *Smart Manufacturing*, *Smart Transportation* systems, *Smart Healthcare* systems, and the *Smart Power Grid*. Cyber-security remains an outstanding challenge in the design of the future IoT. The design of a *Secure Deterministic Industrial Internet of Things* (SD-IoT) core network which can embed millions of distinct *Secure Deterministic Virtual Networks* (SD-VNs) in layer 2 has been presented. The SD-VNs in layer 2 are immune to congestion, interference and DOS attacks. The use of SD-VNs which are managed by an SDN control plane can also remove all targeted cyber-attacks by admitting and forwarding only authorized deterministic traffic flows in layer 2. Future *Smart Systems* can reserve their own distinct mutually-exclusive SD-VN in layer 2, to achieve significantly improved security, performance and energy-efficiency. All transmissions over an SD-VN can also be encrypted by the application before transmission over layer 2 to achieve a level of cyber-security which is not possible with today's BE-IoT.

Finally, low-cost energy-efficient FPGAs combined with Silicon Photonics transceivers will be available soon, driven by Intel's purchase of Altera in 2015 for $16.7 Billion. These devices will provide Terabits/second of optical IO capacity, and are ideally suited to realize a low-cost energy-efficient SD *Industrial-Tactile* IoT core network in layer 2. We argue that a Silicon Photonics SD-IoT can revolutionize the capability of the cloud. Such a deterministic network can reduce end-to-end IoT delays to the speed of light in fiber, and can reduce the energy used for IoT communications by a factor of 1,000 times relative to today's BE-IoT. A secure deterministic *Industrial-Tactile Internet of Things* core network can enable *Big Data* green cloud computing with exascale performance levels within a decade, to optimize the *Smart Systems* of the 21-st century.

References

[1] A. Afanasyev, N. Tilly, P. Reiher and L. Kleinrock, "Host-to-Host Congestion Control for TCP", IEEE Comm. Surveys and Tutorials, 3Q (12.3), 2010, pp. 304-342
[2] M. Ford, "Workshop Report: Reducing Internet Latency 2013", ACM SIGCOMM CCR, Vol. 44, No. 2, April 2014, pp. 80-86
[3] A. Singla, B. Chandrasekaran, P.B. Godfrey, B. Maggs, "The Internet at The Speed of Light", ACM Hotnets 2014, Oct. 2014, LA, USA, pp. 1-7
[4] T.H. Szymanski, "Supporting Consumer Services in a Deterministic Industrial Internet Core Network", IEEE Communications Magazine, Vol. 54.6, June 2016, pp. 110-117
[5] RQ Hu, Y Qian, HH. Chen and H. Mouftah, "Cyber Security for Smart Grid Communications: Part 1, Guest Editorial", IEEE Communications Magazine, Aug. 2012, pp. 16-17.
[6] GE and Accenture Corp., "Industrial Internet Insights Report 2015", 2015, pp. 1-36.
[7] World Economic Forum, "Industrial Internet of Things: Unleashing the Potential of Connected Products and Services", Jan. 2015, pp. 1-40.

[8] G. Fettweis, H. Boche, et al, "The Tactile Internet", ITU-T Technology Watch Report, Aug. 2014, pp. 1-24.
[9] M. Maier, M. Chowdhury, B.P. Rimal, D.P. Van The, "Tactile Internet: Vision, Recent Progress, and Open Challenges", IEEE Communications Magazine, Vol. 54.5, May, 2016, pp. 138-45.
[10] T.H. Szymanski, "A Low Jitter Guaranteed Rate Scheduling Algorithm for Packet Switched IP Routers", IEEE Trans. on Communications, Vol. 57, No. 11, Nov. 2009, pp. 3446-3459.
[11] T.H. Szymanski and D. Gilbert, "Provisioning Mission-Critical Telerobotic Control Systems over Internet Backbone Networks with Essentially-Perfect QoS", IEEE JSAC, Vol. 28, No. 5, June 2010, pp. 630-643.
[12] T.H. Szymanski, "An Ultra Low Latency Guaranteed-Rate Internet for Cloud Services", IEEE Trans. on Networking, Vol. 24, No. 1, Feb. 2016, pp. 123-36.
[13] T.H. Szymanski, "Method to Achieve Bounded Buffer Sizes and Quality of Service Guarantees in the Internet Network", U.S. Patent No. 8,665,722, Mar. 2014, pp. 1-45.
[14] T.H. Szymanski, "Delay and Jitter Limited Wireless Mesh Network Scheduling", US Patent, US 20140119347 A1, May 1, 2014, pp. 1-48.
[15] N. Finn, P. Thubert, "Deterministic Networking Problem Statement (03)", IETF Internet-Draft, Standards Track, May 2015, pp. 1-16.
[16] S. Shah, P. Thubert, "Deterministic Forwarding PHB (03)", IETF Internet Draft, March 2015, pp. 1-9.
[17] V. Anantharam, N. McKeown, A. Mekittikul and J. Walrand, "Achieving 100% Throughput in an Input Queued Switch", Trans. Comm., Vol. 47, No. 8, Aug. 1999, pp. 12601267.
[18] W.J. Chen, C-S. Chang, and H-Y. Huang, "Birkhoff-von Neumann Input Buffered Crossbar Switches for Guaranteed-Rate Services, IEEE Trans. Comm., Vol. 49, No. 7, July 2001, pp. 1145-1147.
[19] Chang, C. S., Lkeslassyee, D. S., and Yue, C. Y., "Providing Guaranteed Rate Services in the load balanced Birkhoff-von Neumann Switches", IEEE/ACM Trans. Networking, Vol. 14. No. 3, June 2006, pp. 644-656.
[20] I. Keslassy, M. Kodialam, T.V. Lakshman and D. Stilliadis, "On Guaranteed Smooth Scheduling for Input-Queued Switches", IEEE/ACM Trans. Networking, Vol. 13, No. 6, Dec. 2005, pp. 1364-1375.
[21] T.H. Szymanski, "A Secure Industrial Internet of Things using Deterministic Photonic Packet Switches", Invited Presentation, IEEE End-to-End Trust and Security Workshop, Washington DC, Feb. 4, 2016, pp. 1-15. (slides available at: http://standards.ieee.org/events/iot/index.html)
[22] S. Iyer, RR. Kompella, N. Mckeown, "Designing Packet Buffers for Router Linecards", IEEE Trans. Networking, Vol. 16, No. 3, June 2008, pp. 705-717.
[23] Cisco Systems, "Cisco CRS Carrier Routing System Multishelf System Description", Sept. 2015, pp. 1-4, (www.cisco.com)
[24] R. Bolla, R. Bruschi, F. Davoli, and F. Cucchietti, "Energy Efficiency in the Future Internet: A survey of existing approaches and trends in energy-aware fixed network infrastructures", IEEE Communications Surveys & Tutorials, Vol. 13, No. 2, May 2011, pp. 223-244.
[25] G. Varghese, N. Bjorner, 'ACM SIGCOMM Tutorial on 'Network Verification', London, UK, August 17, 2015, (http://conferences.sigcomm.org/sigcomm/2015/tutorial-nwverif.php)
[26] S. Kent, K. Seo, IETF RFC 4301, "Security Architecture for the Internet Protocol", Dec. 2005, pp. 1-101.
[27] Adrian D, Bhargavan K, Durumeric Z, Gaudry P, Green M, Halderman JA, Heninger N, Springall D, Thom E, Valenta L, VanderSloot B., "Imperfect Forward Secrecy: How Diffie-Hellman Fails in Practice", 22nd ACM SIGSAC Conf. on Computer and Communications Security, Oct. 2015, pp. 5-17.
[28] T.H. Szymanski, "Max-Flow Min-Cost Routing in a Future Internet with Improved QoS Guarantees", IEEE Trans. Comm., Vol. 61, No. 4, April 2013, pp. 1485-1497.
[29] R.S. Tucker, "Green optical communications - Part II: Energy limitations in networks", IEEE Journal of Selected Topics in Quantum Electronics, Vol. 17, No. 2, 2011, pp. 261-274.
[30] R.S. Tucker, R. Parthiban, J. Baliga, K. Hinton, R.W.A. Ayre, and W.V. Sorin, "Evolution of WDM Optical IP Networks: A Cost and Energy Perspective", OSA JLT, Vol. 27, No. 3, 2009, pp. 243-252.
[31] IEEE 802 Tutorial, "Deterministic Ethernet: 802.1 Standards for Real-Time Process Control, Industrial Automation, and Vehicular Networks", 2012, pp. 1-72. (www.ieee802.org)
[32] US Government, National Institute of Standards and Technology (NIST), Federal Information Processing Standards (FIPS), Publication 197, "Announcing the Advanced Encryption Standard (AES)", 2001, pp. 1-51. (http://csrc.nist.gov/publications)
[33] Network World, "SDN Security Attack Vectors and SDN Hardening", Oct. 28, 2014, pp. 1-3. (http://www.networkworld.com)
[34] T.H. Szymanski and M. Rezaee, "An FPGA Controller for Deterministic Guaranteed-Rate Optical Packet Switching", IFIP/IEEE Integrated Management (IM) Conf., Ottawa, Canada, May 2015, pp. 1177-1183.
[35] Y.A. Vlasov, "Silicon-CMOS Integrated Nano-Photonics for Computer and Data-Communications Beyond 100G", IEEE Comm. Mag., Feb. 2012, pp. 67-72.
[36] S.S. Sherif, S.K. Griebel, A. Au, D. Hui, T.H. Szymanski, H. Scott Hinton, "Field Programmable Smart Pixel Arrays: Design, VLSI Implementation and Applications", OSA Applied Optics - Information Processing, Feb. 1999, pp. 838 - 846.
[37] T.H. Szymanski, M. Saint-Laurent, M., V. Tyan, A. Au, B. Supmonchai, B., Field Programmable Logic Devices with Optical I/O", OSA Applied Optics - Information Processing, Feb. 10, 2000, pp. 721 - 732.
[38] S. Koehl, V. Krutul, M. Paniccia, Intel Corp., "Continuous Silicon Laser", White paper, pp. 1-6. (www.intel.com)
[39] Mellanox Corp. Release, "Mellanox Introduces Next Generation 100Gb/s Silicon Photonics Transceivers", March 23, 2015, (http://ir.mellanox.com/releasedetail.cfm?releaseid=902825)
[40] Business Wire, "Luxtera Debuts 1310nm 100G-PSM4 QSFP28 Module and Silicon Photonics Chipset at OFC 2015", March 23, 2015, (http://www.businesswire.com/news/home/20150323005646/en/Luxtera-Debuts-1310)
[41] Fujitsu Press Release, Fujitsu, PETRA, and NEDO Achieve World's Lowest Energy Requirements of 5 mW per 1 Gbps for High-Speed Inter-Processor Data Transmissions, Feb. 23, 2015, pp. 1-2. (www.fujitsu.com)
[42] Li, J., Zheng, X., Krishnamoorthy, A. V., and Buckwalter, J. F., "Scaling Trends for Picojoule-per-Bit WDM Photonic Interconnects in CMOS SOI and FinFET Processes", IEEE Journal of Lightwave Technology, Vol. 34, No. 11, June 2016, pp. 2730-2742.

[43] D.A.B. Miller, "Device Requirements for Optical Interconnects to Silicon Chips", Proc. IEEE, Vol. 97, No. 7, July 2009, pp. 1166-1185.
[44] J. Hruska, "IBM Announces Silicon Photonics Breakthrough: Set to Break 100 Gb/s Barrier", ExtremeTech, May 14, 2015, pp. 1-4.
[45] T.H. Szymanski, "Impact of Future Trends on Exascale Grid and Cloud Computing", 29th Int. Supercomputing Conf. (ISC 2014), Springer, June 2014, Leipzig, Germany, pp. 215-231.
[46] E. Masanet, R.E. Brown, A. Shehabi, J.G. Roomey, and B. Nordman, "Estimating the Energy Use and Efficiency Potential of U.S. Data-Centers", Proc. IEEE, 2011, pp. 1440-1453.
[47] US Dept. of Energy, "SOS 14: Challenges in Exascale Computing ", 2014, pp. 1-31. (www.csm.ornl.govworkshops)
[48] K. Yelik, S. Coghlan, B. Draney, R.S. Canon, "Mallegan Report on Cloud Computing for Science", US Dept. of Energy, Office of Science, Dec. 2011, pp. 1-170. (http://science.energy.gov/)
[49] Fujitsu Corp., "Flashwave 9500 Packet Optical Networking Platform", 2016, pp. 1-5.
[50] Greentouch White Paper, "Greentouch Green Meter Research Study: Reducing Net Energy Consumption in Communications Networks by up to 90% by 2020", June 2013, pp. 1-25. (www.greentouch.org)

Ted H. Szymanski completed the PhD degree at the University of Toronto. From 2001 to 2011, he held the Bell Canada Chair in Data Communications in the Dept. of ECE at McMaster University. Previously, he was a professor at Columbia University and its *Center for Telecommunications Research*, and McGill University and the *Canadian Institute for Telecommunications Research*. He participated in a 10-year national research program within the *Networks of Centers of Excellence* of Canada, which demonstrated a free-space 'intelligent optical backplane' using CMOS/SEED photonic packet-switches with about 1K optical channels. Collaborators included Nortel Networks (Ericsson), Newbridge Networks (Alcatel), Lockheed-Martin/Sanders, and McGill, McMaster, Toronto and Heriot-Watt Universities. His group also demonstrated the first FPGA with optical IO, using the US ARPA/Lucent/Coop CMOS/SEED smart-pixel foundry service. His interests include security, energy efficiency, deterministic communications, smart systems, and the Industrial-Tactile Internet of Things.