1

Prospects and Pitfalls of Modern Interconnect Technologies

1.1 Overview and Motivation

In the brief time, it takes you to read this sentence—about 10 seconds—US data centers will consume 700 GWh of energy, enough to power the entire country of Nigeria for a minute. This massive energy demand, driven largely by the recent AI boom, is projected to triple by 2030, posing a risk of nationwide grid failures (see Figure 1.1).

This highlights the urgent need to examine the physics of power dissipation and innovate device materials to significantly cut energy consumption and ensure a sustainable computing future. A significant portion of this power consumption and heat generation comes from servers of data centers, which include processors, memory, and interconnects. About 20% of a server's power usage is due to interconnects and switches. High-density interconnects in package-level integrations are especially power-intensive due to high parasitics, causing excessive power loss, signal delay, distortion, and crosstalk. For instance, next-generation 6G chips, designed to handle 4–8 Tb/s bandwidth through wide-lane data buses, could have interconnect densities up to 50,000/cm^2. At these densities, crosstalk among short-reach interconnects might increase by 100 dB, and power loss could rise 1,000-fold, leading to severe power penalties and potential device malfunctions. The excessive device parasitics originate from the capacitance among the interconnect networks. Capacitances are proportional to the dielectric constant of the material. Traditionally, capacitance has been mitigated using materials with low dielectric constants (low-k materials) such as porous silicon dioxide, organosilicate glass, and polymeric dielectrics. However, even with advancements in material processing, no

Artificial Plasmonics for VLSI Interconnects: Bridging the Gap between Electronics and Optics, First Edition. Soumitra R. Joy and Pinaki Mazumder.

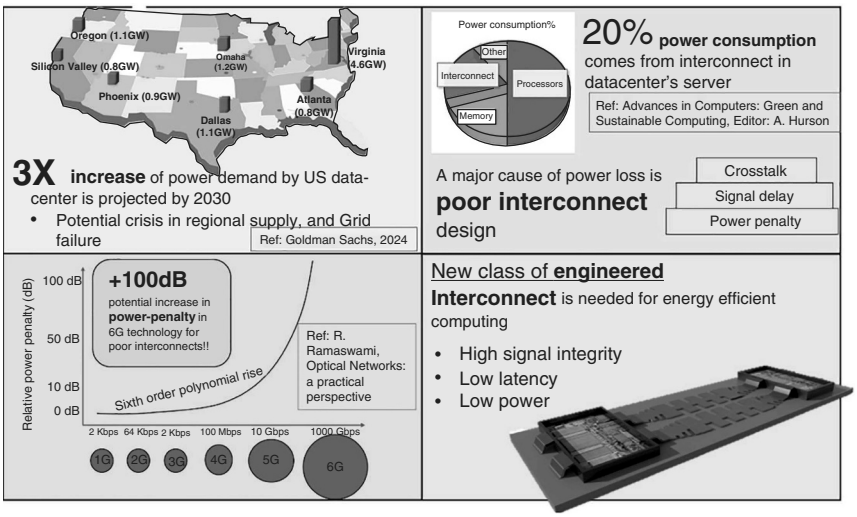


Figure 1.1 A dire power crisis has loomed in the advent of soaring demand of computing. A breakthrough technology for energy-efficient computing hardware is needed to address the issue. This proposal aims to introduce new technologies for chip-to-chip interconnects with potentially lower wire parasitic in order to realize energy-efficient, high performance computing.

natural material can achieve a dielectric constant less than 1, making it unlikely that current materials can reduce parasitic capacitances by more than an order of magnitude in silicon devices, which is nowhere near being sufficient. Therefore, a breakthrough technology for a new class of interconnect is essential to tackle the parasitic capacitance issue and thus mitigate the high-power penalties in micro-electronic devices.

1.1.1 Problem Specifics

Design of high-performance interconnects is mission-critical for next-generation IC packaging to achieve silicon scaling and heterogeneous integration, while allowing optimization of all five major design variables of a technology IP, namely power, performance, area, cost, and time-to-market. Design of short-reach interconnects ($\leq$1 cm) among chiplets in system-in-packages is particularly challenging due to high wire-parasitic among thousands of lanes, causing significant power loss and crosstalk.

Let us consider how catastrophic wire-parasitic effects can become in a dense interconnect fabric. In a state-of-the-art silicon die with 1,000 I/O pins packed into a 1 cm^2 area, mutual capacitance can create up to an astonishing 498,500

undesired parasitic paths through the connecting wires and silicon substrate. This vast number of parasitic paths can severely impair signal integrity and lead to significant energy loss.

There are two primary methods to mitigate parasitic capacitance: increasing the spacing between wires or reducing the dielectric constant of the substrate. Increasing spacing is impractical due to the loss of device density, while reducing the dielectric constant of the substrate is limited by the materials available. For example, replacing silicon ($\epsilon_r = 12$) with silicon dioxide ($\epsilon_r = 4$) only reduces capacitance by a factor of 3. Even using highly porous materials ($\epsilon_r \geq 1$) offers only a factor of 12 reduction, which is insufficient for the anticipated 100 dB increase in crosstalk in future technology nodes. This level of reduction falls short as we move from GHz to high-speed THz electronics (see Figure 1.2).

A potential solution could be the development of artificially engineered electromagnetic materials with an ultralow dielectric constant beyond what is achievable by natural materials. These engineered materials leverage deep subwavelength unit-cell spacing to create conjugate dipole moments that effectively cancel each other's electromagnetic induction, resulting in minimal parasitic capacitance. This approach depends more on geometric features than on the surrounding material properties and can be realized through advanced lithography.

This book seeks to discuss the development of a new technology for short-reach interconnects using standard complementary metal-oxide semiconductor (CMOS)-compatible materials and to engineer them in such a way as to behave as an artificial ultra-low-k material to minimize parasitics and improve power efficiency. The book will focus on modeling and system analysis of these interconnects to enable their integration into commercial design kits (PDKs), accelerating their adoption in microelectronics.

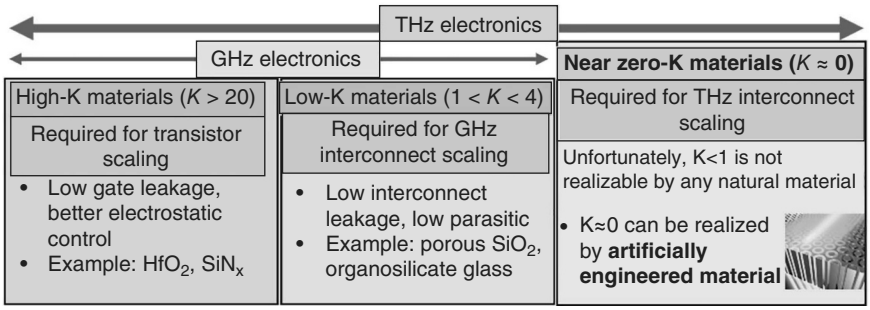


Figure 1.2 For data transmission at terabyte per second speed, we propose to design an artificially engineered interconnect with near-zero K value, with potential minimization of latency, crosstalk, and power usage.

1.2 Communications Challenges: Human-Level vs Machine-Level

When we deal with communication between person-to-person, typically the greater the distance of communication, the greater the complexity of the information transfer, owing to data loss, security and privacy loss, etc. But how about communication among machines? Does the information become safer from data leakage and crosstalk as we keep scaling down the length of communication distance? Well, quite counter-intuitively, no. In fact, the aggregate data density among machines increases exponentially with the scaling down of length of distance of interest. As the device area shrinks down, all pairs of point-to-point communication on the chip-scale begin to compete against each other, and signal integrity severely suffers from interference from neighboring channels, as illustrated in Figure 1.3. Keeping signal integrity unharmed from data leakage and interference is one of the Herculean challenges for modern chip design technology. For machines, especially machines at the scale of sub-micrometers and nanometers

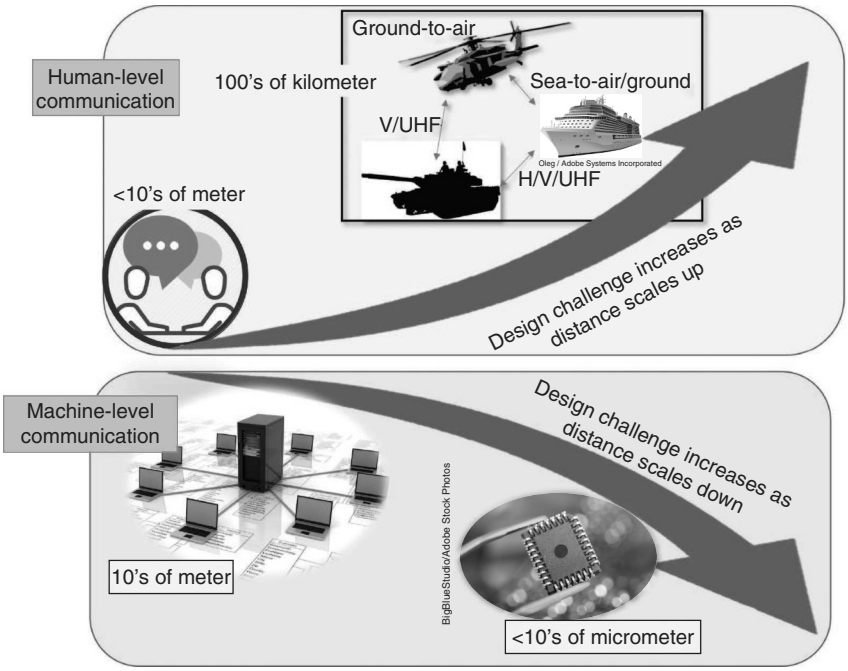


Figure 1.3 In human-level vs in machine-level communication, the degree of design complexity for preventing data leakage shows opposite trend with length scaling.

in a highly dense system, we need a very different kind of engineering strategy to prevent loss of signal integrity due to interference and crosstalk.

1.3 Modes of Interconnects: A Technology Gap

A modern computing system, by and large, is based on von Neumann architecture, and is made of three basic components: the processors, the memories, and a machine-to-human interface. The way all these critical units exchange information among themselves is called interconnects. Modern interconnect technologies used in digital systems have primarily emerged from two opposite ends of the electromagnetic spectrum [1–3]. On the electronic side of the spectrum, copper wires are employed to transfer data at the rate of nearly gigabits per second (Gbps). On the other hand, on the photonic side of the electromagnetic spectrum, optical interconnect technology with a wavelength of 1.55 μm (i.e., $\sim$ 200 THz) is used for communication between circuit boards and racks in large digital systems by deploying photons to transfer data at a much higher rate [4–7].

The two interconnect technologies, namely electrical and optical ones, rely on two different mechanisms of guiding electromagnetic modes through the channel. Electrical interconnects confine electromagnetic modes by setting a pair of conductors: each conductor carries equal current with opposite phase. As long as the spacing between the conductors is much smaller than the wavelength, the pair of wires guides electromagnetic waves along the longitudinal direction with negligible radiation to the surroundings. Optical interconnect, on the other hand, guides electromagnetic modes by using the difference in refractive index of the channel material with respect to its surroundings.

Figure 1.4 shows different types of interconnects, classified according to the distance they cover and the nominal data rates they are supposed to carry. As of its current standing, long distance, such as intercontinental or interdata center communication, is managed by optical interconnect, while the chip-level data transfer is supported by electrical interconnects. However, in the coming days, all the different levels of communication will keep shifting to their left in Figure 1.4, as the need for data rate will increase over time. At their current stature, it is clear that optical interconnects will remain the unchallenged candidate for long-haul communication. However, as the communication distance reduces down to chip scale, we will enter a gray area where neither optical nor electrical interconnect technology has a clear edge over the other. A great deal of research is going on in order to cover that coveted critical area of communication: some of the endeavors focus on pushing the limits of the present optical interconnect technology, some on pushing the limits of electrical ones. Meanwhile, a few researchers,

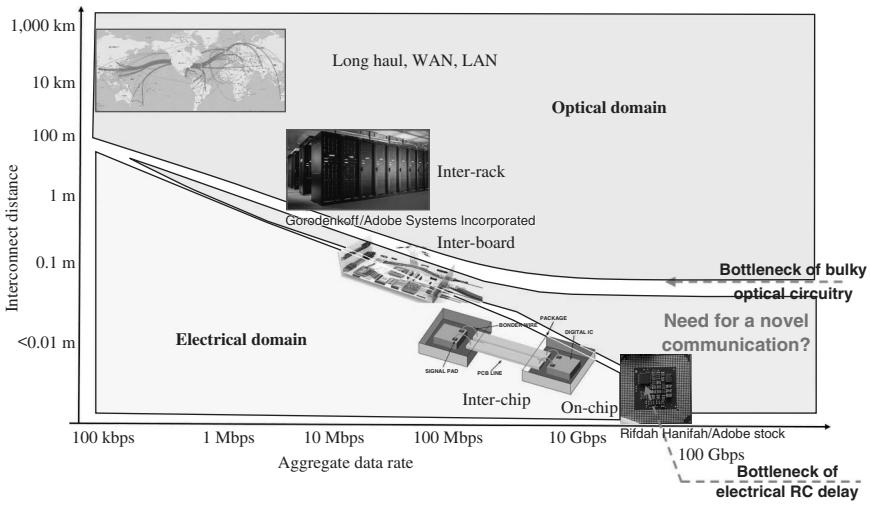


Figure 1.4 The illustration of how interconnects can be divided into multiple classes depending on what distance to cover and what data rate to support. While optical interconnects will remain unrivaled for long-haul communication for the foreseeable future, a technology gap between electrical and optical interconnect is emerging for centimeter-scale communication with data rate beyond 100 Gbps.

including the authors of the present book, have taken on venturing a novel, unforeseen interconnect technology with the aim to high-density information transfer while keeping energy consumption at a reasonable limit.

1.4 Innovations in Interconnect Frontier

From a bird's-eye view, we can categorize the approaches' innovation in interconnect domains into four different hierarchical levels, as shown in Figure 1.5.

- Innovation at the material level (cobalt, ruthenium are being considered for chip-scale communication; porous materials)
- Innovation at the architecture level (3D IC, reconfigurable differential pair channel)
- Innovation beyond von Neumann architecture (memory-in-processor)
- Innovation at the signal-guiding mechanism

1.4.1 Interconnect Research at Material Level

As chip scaling progresses below the 7 nm threshold, the interconnects at the chip level are undergoing significant transformations. The density of smaller transistors on a die is increasing, accompanied by a growing volume of data

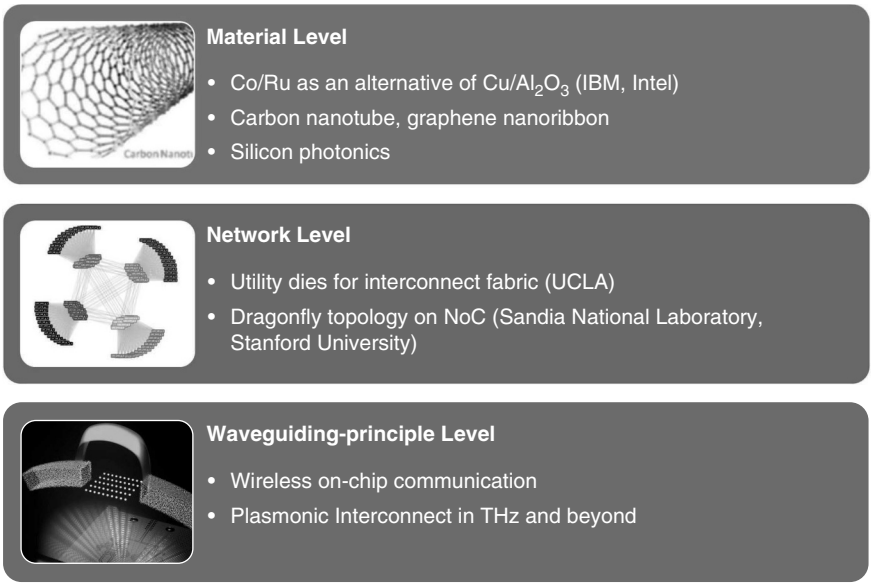


Figure 1.5 Different hierarchies of research venues on emerging interconnect technologies.

processing and transfer within and beyond chips or packages, as shown in Figure 1.26. Consequently, there are ongoing changes in the materials used for interconnects. Establishing reliable connections between different layers poses the primary challenge in interconnect fabrication. Copper, the long-standing material of choice for interconnects since the 130 nm node, has reached its limitations. One of the issues with copper is its tendency to diffuse into adjacent materials at advanced nodes. This necessitates the use of barrier layers, which must become thinner as scaling continues to 5 nm and beyond 3 nm while retaining conductivity. Additionally, copper encounters other challenges, including: (i) reduced reliability as interconnect dimensions decrease due to increased current density; (ii) increased resistivity at lower dimensions caused by grain-boundary scattering and surface scattering; (iii) rapidly rising resistivity due to Joule heating; and (iv) decreased conductivity at high operating frequencies resulting from the skin effect. Consequently, copper falls short of meeting the requirements for future technology demands in interconnect materials. The growing disparity between the needs of future on-chip interconnect materials and the current reliance on copper has spurred researchers and designers to seek innovative material solutions. Intel has already initiated the use of cobalt instead of copper for the local interconnect layers M0 and M1 starting from the 10 nm technology node. Cobalt is also being employed for liners in the back-end-of-line (BEOL) process, although ruthenium

is gaining momentum as an alternative. Furthermore, numerous other materials are being explored as potential options. At a more advanced stage of research, carbon nanotubes (CNTs) and graphene nanoribbons are being considered as potential alternatives to copper. CNTs consist of single layers of graphene sheets rolled into cylindrical structures with diameters ranging from 1 to 5 nm. Metallic CNTs exhibit ballistic electron transport, allowing electrons to move along the nanotube axis without scattering, resulting in a long mean free path (MFP) spanning micrometers. In contrast, the MFP of electrons in copper is limited to a few nanometers. Due to their large MFP and small diameter, electrons in CNTs experience fewer scattering events, leading to lower resistance. In the field of optical interconnect innovation, silicon photonics is on the verge of revolutionizing the datacom and telecom market for transceiver technology. Its primary appeal stems from its significantly lower cost compared to other alternatives, especially InP-based transceivers. Additionally, silicon photonics brings considerable advantages in terms of reduced upfront expenses, as numerous companies can take advantage of the established silicon photonics fabrication ecosystem.

Traditionally, silicon photonics has heavily relied on the silicon-on-insulator (SOI) platform as the primary means of integration, enabling the incorporation of various passive and active devices. However, certain passive devices, like wavelength-division multiplexing (WDM), present challenges related to process control and phase error due to the high index contrast of silicon material, resulting in stricter fabrication tolerances. To tackle these issues, a silicon nitride (SiN) platform has been proposed, utilizing silicon test wafers. While the SiN platform offers unique benefits such as extremely low propagation loss and wider fabrication tolerances, it lacks inherent active capabilities. As a solution, a SiN-on-SOI platform has been developed to leverage the advantages of both Si and SiN platforms. Figure 1.6 illustrates the cross-sectional schematic of a typical silicon photonics integration platform, encompassing SiN and SOI devices. The SiN waveguide layer is responsible for functions such as large-scale light routing, while all other passive and active devices can be manufactured within the SiN waveguide layers. However, it is important to note that silicon photonics does not fulfill all the requirements for a complete photonics integrated circuit, primarily due to the absence of a laser. Consequently, a conventional silicon transceiver experiences additional losses in coupling from the laser to the silicon photonics chip.

1.4.2 Interconnect Research at the Network Level

The growing computational demands of modern applications and services have prompted researchers to focus on enhancing semiconductor-based technology. As a result, there has been a significant evolution in on-chip networks, which are now receiving considerable attention due to their beneficial characteristics,

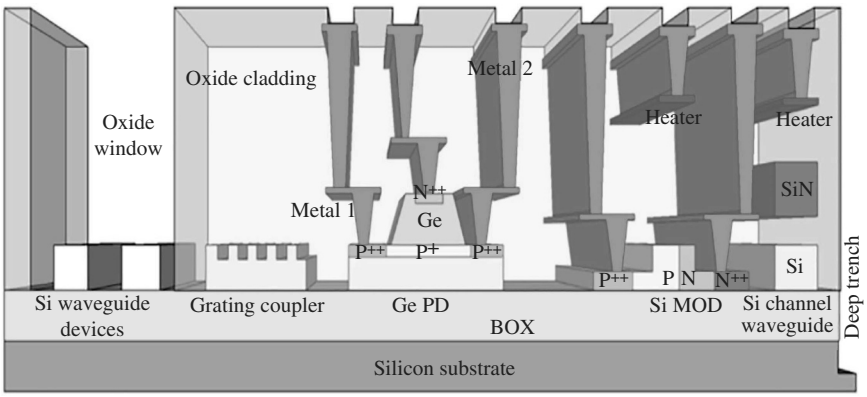


Figure 1.6 Different hierarchies of research venues on emerging interconnect technologies.

including small footprint and low power consumption. Noteworthy examples of on-chip architectures include System-on-Chip (SoC), Pasricha and Dutt [8], and multiprocessor-SoC (MPSoC), Wolf [9]. Additionally, various components of the on-chip network are interconnected using different interconnection networks, such as shared buses and regular buses. In a bus-based topology, devices primarily communicate through bus links, while wire collections are employed in shared bus topologies. Comparatively, the shared bus architecture offers a cost-effective solution with straightforward control features. These advantages make it the preferred choice for facilitating communication among the integrated processing units on the chip [10].

Topology, routing, and switching are prominent and promising research areas within on-chip design [11], see Figure 1.7. Active research is also focused on developing diverse application mapping algorithms to enhance on-chip performance by minimizing network latency. Thanks to advancements in integration technology, it is now possible to accommodate a significant number of heterogeneous cores on a single chip [12]. Consequently, when designing such chips, the interconnections between multiple cores play a vital role in system performance and communication. In addition to managing memory transactions, the on-chip communication architecture also handles I/O traffic and establishes a reliable channel for data sharing between processors. Therefore, the design of high-performance and scalable on-chip communications is critical for SoC design [13]. However, achieving efficient communication among on-chip components presents significant challenges. The network-on-chip (NoC) topology defines the physical organization of the architecture and serves as a key design criterion. It determines the interconnection scheme of the NoC elements. Traditionally, regular tile-based topologies have

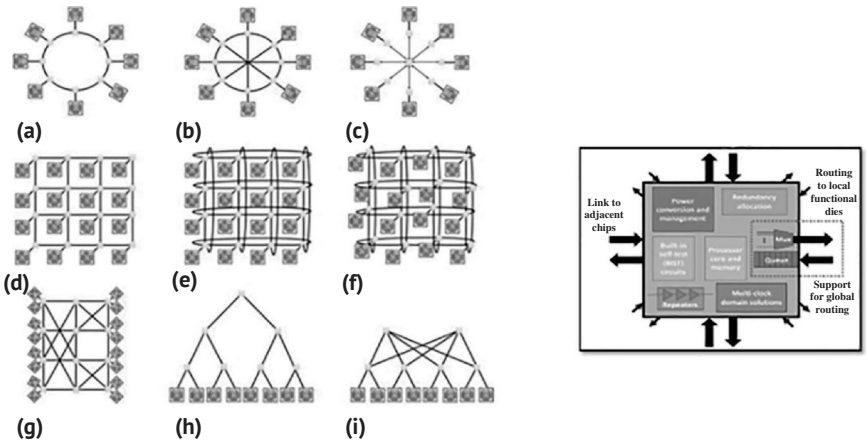


Figure 1.7 Network-on-chip (NoC) topologies, i.e. (a–i) ring, octagon, star, 4×4 mesh, 4×4 torus, 4×4 folded torus, butterfly, binary tree, and fat tree. The right-hand side figure shows the template for NoC chiplets organization. *Source*: Alimi et al. [11]/IntechOpen/CC BY 3.0.

been utilized to connect homogeneous cores. However, there is a growing focus on custom-based, domain-specific irregular topologies that can support heterogeneous cores with varying sizes, functionalities, and communication requirements. Figure 1.7 showcases a range of interconnect topologies and routing techniques currently being explored and experimented with for high-density chip-scale communication [11].

1.4.3 Interconnect Research at Waveguide Level

1.4.3.1 Wireless Communication at Chip Scale

The landscape of communication infrastructure within a chip is undergoing a transformative shift, transitioning from traditional bus-based systems to NoC architectures in response to the proliferation of many-core chips [15]. Simultaneously, interchip communication is rapidly evolving to meet the escalating bandwidth requirements while adhering to stringent power constraints. The conventional interchip interconnections encompass various methods, ranging from solder bumps or C4 interconnects in multichip modules within a System-in-Package (SiP), to more contemporary designs such as arrays of solder balls and silicon interposers [16]. In recent years, there has been a growing interest in leveraging wireless communication as a solution to address the limitations posed by wired metallic interconnections. The feasibility of incorporating on-chip antennas into a silicon die has been explored, leading to the proposition of wireless NoC (WiNoC) architectures [17], as illustrated in Figure 1.8. These WiNoCs

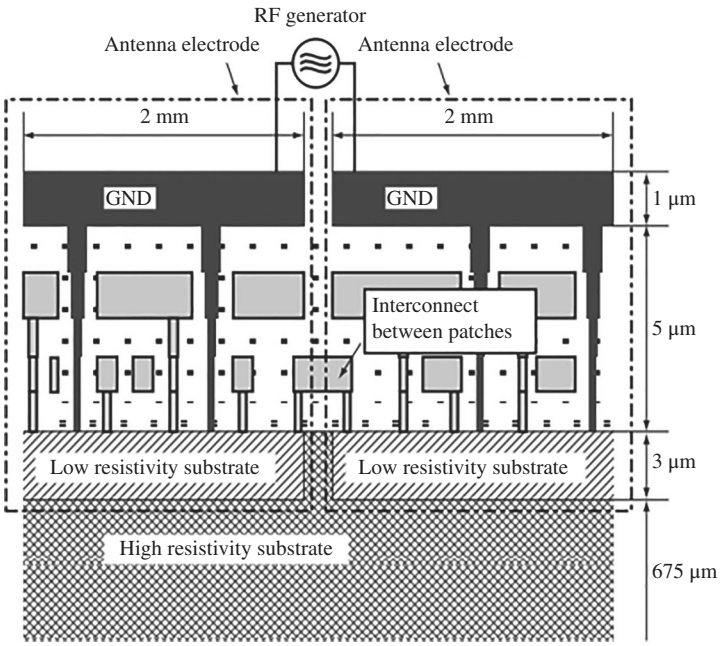


Figure 1.8 Schematic illustration of wireless network on chip. *Source*: Yordanov and Russer [14].

have demonstrated enhancements in energy efficiency and bandwidth for on-chip data communication within many-core chips, leveraging intra-chip wireless links [18]. Although extensive research has been conducted on WiNoCs, the potential of wireless interconnects for interchip communication has not been adequately explored, despite the potential advantages gained by eliminating chip-to-chip metallic interconnections. Researchers are currently investigating several emerging interconnect technologies, including silicon photonics and three-dimensional (3D) integration, to mitigate the challenges associated with metal interconnects. Wireless interconnects utilizing on-chip antennas tuned to millimeter-wave bands are considered the most mature interconnect technology, as these antennas can be fabricated using CMOS processes. Numerous on-chip antennas, such as meander dipoles, zigzag dipoles, slots, loops, inverted F, bowtie, and Yagi, have been designed and studied [19, 20]. Compactness is a critical requirement for on-chip antennas, as they need to fit within the limited footprint of the core, typically a few square millimeters, in current and future technology nodes. The concept of subwavelength interfacing between optics and electronics holds promise as a means to reduce the size of communication devices while maintaining the required speed and bandwidth [21]. However, this presents a formidable challenge, as existing

electronic technology struggles to respond to the rapid oscillating fields associated with optical frequencies, and the size of photonic components remains significantly larger than their electronic counterparts. A. Alù and N. Engheta proposed, in theory, an optical wireless channel between two nanoscale antennas [22]. Partial demonstrations of this idea have been achieved by directing an optical signal toward a distant luminescent receiver or by utilizing surface plasmons. Recent progress has showcased a proof-of-principle wireless link, where gold nano-antennas serve as transmitters, directing an out-of-plane laser signal toward an optical rectenna, where the transmitted radiation is rectified as direct current (DC).

1.4.3.2 Plasmonic Interconnects

To meet the ever-growing demand for enhanced bandwidth and performance, extensive research is being conducted to explore novel technologies. Among the numerous contenders, nanophotonics technology emerges as a promising candidate for revolutionizing integrated devices. However, the efficient guidance and modulation of light in these devices are constrained by the diffraction limit of light, which imposes a size limitation. Even with the utilization of high-index materials, achieving a minimum size of approximately 200 nm for telecommunication wavelengths remains challenging. Nevertheless, a subfield known as nanophotonics or plasmonics has emerged as a potential solution. By exploiting the momentum of free electrons at a metal-dielectric interface, referred to as surface plasmon polaritons (SPPs), metallic nanophotonics facilitates the confinement and guidance of light at the nanoscale [24], see Figure 1.9. Despite the notable advancement in integration offered by this approach, it is not without drawbacks. Metal nanophotonics inherently suffer from unavoidable losses, which are contingent on the level of modal confinement within the structure [25]. Consequently, although metal nanophotonics can confine light to dimensions as small as tens of nanometers, the associated losses pose significant challenges in the integration of these structures into current devices. However, researchers are actively engaged in studying potential compensatory measures, such as gain, to mitigate these losses.

The pursuit of low-loss metal nanophotonic structures initially focused on utilizing gold and silver as the metallic core. Through continuous refinements in fabrication techniques and optimization, metal nanophotonics has achieved sub dB/cm propagation, indicating a substantial reduction in losses [26]. Nevertheless, the larger size of the mode in comparison to corresponding photonic structures, approximately $10\,\mu m^2$, diminishes the Figure-of-Merit (FoM) associated with these structures. Recent efforts have explored various geometries aiming to reduce the mode size while maintaining low propagation loss. As a result, hybrid and dielectric-loaded waveguide structures have elevated the Figure-of-Merit (FoM)—defined as the inverse of the mode size multiplied by the attenuation—to values surpassing 1000 for mode sizes around $1\,\mu m^2$. Like the advancements

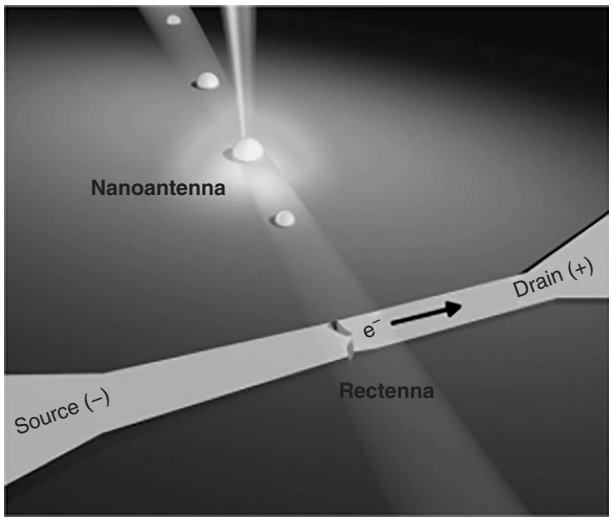


Figure 1.9 Wireless channel between nanoscale antenna. *Source*: Dasgupta et al. [23].

witnessed in photonic waveguides, further improvements in materials, fabrication techniques, and structural design are anticipated to drive the progress of metal nanophotonic waveguides in the future.

1.4.3.3 Terahertz Interconnects

The THz spectral region presents a highly promising opportunity to meet the ever-increasing demand for higher data transfer rates (adapted from Sun et al. [28]). Its extensive bandwidth is of paramount importance in the development of 6G mobile communication networks, where the wireless transmission of data at terabits per second over long distances is a prerequisite. Furthermore, this spectral band facilitates the establishment of an expansive "internet of things" network, connecting billions of devices. Additionally, the THz spectral band holds great potential in addressing the persistent challenge of achieving high-speed, energy-efficient, and cost-effective intrachip/interchip communication links within short distances, thereby enabling the realization of massive multicore processors, NoC architectures, and system-in-package solutions [29]. To realize these ambitious objectives, it is crucial to develop integrated and economically viable solutions for on-chip THz waveguiding and manipulation. Traditional approaches to THz waveguiding, including hollow metallic waveguides, metallic transmission lines, photonic crystals, metal wires, and THz fibers, suffer from inherent susceptibility to defects such as fabrication imperfections and significant losses at sharp corners due to bending. However, recent breakthroughs in the field of the topological phase of light have opened new possibilities for addressing these challenges

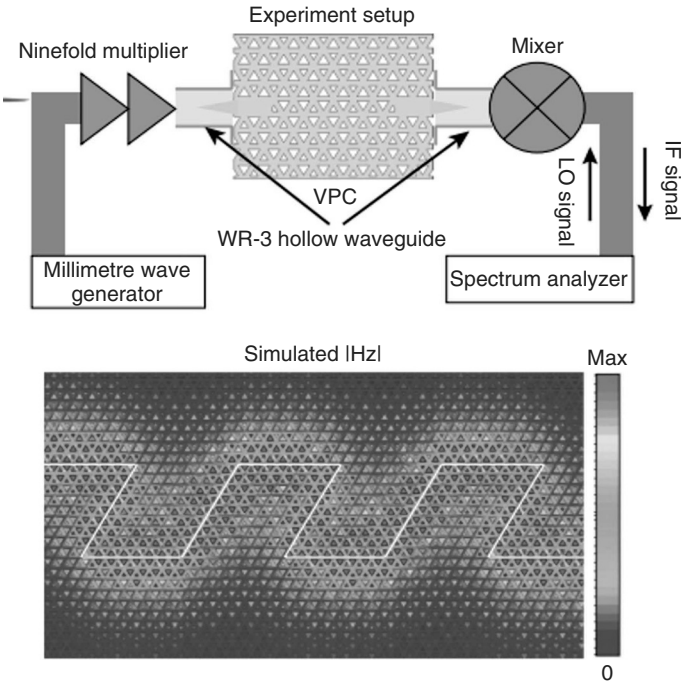


Figure 1.10 Terahertz signal propagation utilizing photonic topological insulators. *Source*: Yang et al. [27].

[27]. One promising solution lies in the utilization of photonic topological insulators (PTIs), which possess an "insulating" bulk while exhibiting "conducting" characteristics at their edges. These PTIs demonstrate robust edge transport, effectively suppressing backscattering even in the presence of disorder and sharp bends. Building upon the principles of the topological phase of light, Yang et al. successfully conducted an experimental demonstration of robust THz topological valley transport using an all-silicon chip; see Figure 1.10.

1.4.4 Chip-scale Interconnect Technologies: Major Industrial Steppingstones

The emergence of on-chip planar interconnect technology can be attributed to the innovative thinking of Robert Noyce, one of the co-founders of Fairchild Semiconductor. In 1959, Noyce introduced the concept of interconnecting silicon circuit components using aluminum metal lines placed on top of a protective oxide coating; see Figure 1.11. This advancement replaced the cumbersome practice of manually connecting wires, leading to the realization of Jack Kilby's solid circuit

Intel Free Press/ Wikimedia Commons / CC BY SA 2.0

Figure 1.11 Robert Noyce, a pioneer in integrated interconnect on chip.

through integrated metallic interconnects. Since its inception, interconnect technology has experienced a significant shift, marking a paradigmatic transformation.

1.4.4.1 Electrical Interconnects: Evolution Toward Nanoscopic, High Density Network

During the period spanning from 1970s to 1990s, transistors had relatively large gate lengths, exceeding 1 μm. At that time, interconnect delays were minimal in comparison to gate delays. However, in the latter half of the 1990s, aggressive miniaturization of transistors significantly reduced gate delays, resulting in interconnect delays becoming comparable to gate delays. This shift sparked an urgent need for research and development in high-speed interconnect technologies to propel the advancement of Moore's law and enable scaling. By the late 1990s, at the 0.25 μm technology node, the delays introduced by the aluminum/silicon dioxide (Al/SiO_2) interconnects became comparable to the intrinsic gate delay of transistors. Over time, as technology progressed, interconnect delays started dominating the total delays within a chip. As the slowest component of the circuit dictates the overall speed, interconnect delays emerged as a fundamental bottleneck in achieving high-performance devices.

In 1997, IBM introduced copper interconnects with low-k dielectric materials, allowing processors to achieve unprecedented speeds. Copper exhibited 40% lower electrical resistance than aluminum, resulting in a 15% boost in chip performance. However, copper presented challenges due to its tendency to electromigrate into the silicon substrate. IBM's research focused on developing barrier materials, particularly tungsten, and implementing a unique metal deposition technique to isolate copper from silicon and enhance its performance. To reduce parasitic RC delay in interconnect routes, low-k materials such as SiOF (with $k \leq 2.2$) were paired with copper interconnects. In 1998, GlobalFoundries Inc. filed a patent for a semiconductor interconnect structure that incorporated air gaps to reduce intralayer capacitance in metal layers. Various research groups pursued the idea of including air or porosity in the silicon substrate to address the mechanical fragility

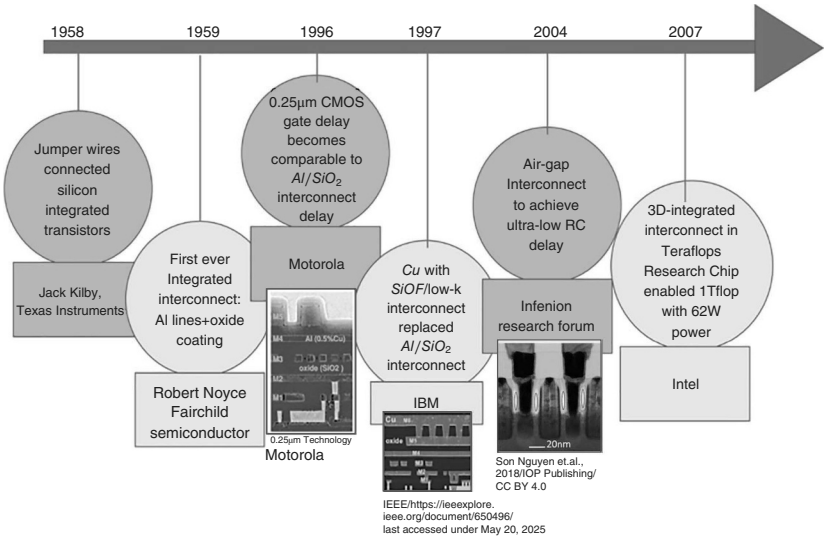


Figure 1.12 Major milestones over the time toward ultrafast chip-scale interconnect technologies.

of these structures. For instance, Infineon Research Forum demonstrated the use of air-gap interconnects in 2004, while D. Kioussis et al. discussed optimization techniques for processing ultra-low-k films to reduce RC delay and power consumption in 2011. The advent of 3D integration using copper interconnect technology opened up new possibilities for high-performance chips. Examples include Intel's Teraflop Research chip, unveiled in 2007, which achieved 1TFLOP operation while consuming significantly less than 100 W of power. The 3D integration of wiring systems, along with proper placement within the circuit, promises higher clock rates, lower power dissipation, and increased integration density. Through-silicon vias (TSVs) have already been implemented in the production of CMOS image sensors, where the active silicon area is bonded into glass and accessed from the backside. Figure 1.12 illustrates some major milestones in realizing high-performance commercial interconnects.

1.5 Scaling Issue of System Level Interconnect

Over the past decade, we have seen an exponential growth of IC speed and integration levels, as the IC fabrication technology, coupled with aggressive circuit design, has advanced significantly. However, as the gate delay reduced with the

shrinkage of transistor size at newer process nodes, the interconnect delay, unfortunately, increased for two reasons: as interconnects are getting narrower, it causes metallic resistance to rise, and higher interconnect density causes parasitic capacitance to rise. The effect of the interconnect delay bottleneck is presented in Figure 1.13. In order to get the benefit of the improvement of the chip-level speed at the overall system performance, the communication bandwidth among different ICs and sub-systems must scale accordingly. Currently, communication links in various applications approach Gbps data rates. These applications include computer-to-peripheral connections, local area networks, memory buses, and multiprocessor interconnection networks. Figure 1.14 shows the cross-section of Integrated Circuit (IC) modules on board to reveal the wire pitch of different levels of interconnects.

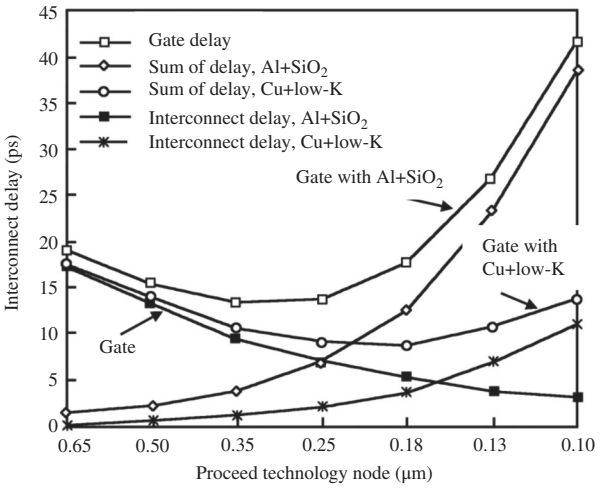


Figure 1.13 Trend of interconnect delay as it evolves through every process node.

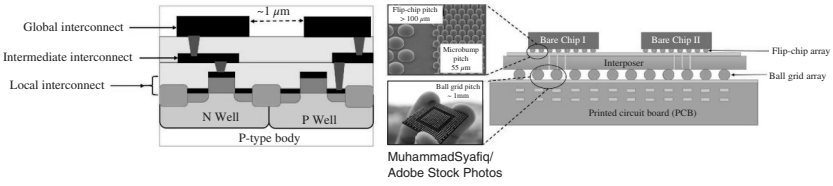


Figure 1.14 Cross-sectional schematic of a very large scale Integration chip, revealing interconnect at various hierarchical levels.

1.5.1 Chip Package Signaling

Board-based interconnects have been used for relatively long-range interconnects ($\sim 30 - 100$ cm) between modules. To keep pace with the chip functionality that doubles every 2 years or so, the off-chip bandwidth also multiplies by a factor of 2 every other year, thanks to the increases in signal pin count and per-pin bandwidth. According to the International Technology Roadmap for Semiconductors (ITRS) prediction, the chip package bandwidth would have to surpass 10 Tbps in the timeframe of 2030 and beyond, refer to Figure 1.15. This unprecedented bandwidth enhancement relies on two factors: an increase in pin-count per package, and the increase in data rate per pin. Figure 1.16 summarizes the growth in pin count and data rate per-pin over time. While pin count per package increases at a linear rate, the data rate per pin has been increasing at an exponential rate over the last 20 years. For example, a state-of-the-art chip can support several hundred differential I/O ports, where each of the channels can carry 10 Gbps data. Unfortunately, neither of these trends of growth can continue too far into the future, as they are subject to fundamental barriers such as power, area, and frequency wall. The increases in the bandwidth of system-level interconnects do not scale at the same rate as on-chip core frequencies [30]. This is primarily because of the inherent limit of the minimum channel pitch achievable on the motherboard. For instance, high-density differential stripline pair typically has a pitch around ~ 500 μm. Therefore, with a 10 Gbps channel data rate, the maximum bandwidth density that we can achieve on-board-based interconnect is in the ballpark of ~ 15 Gbps/mm. Figure 1.17, reproduced from Yamada et al. [31], shows the trend in the wire/pad pitch of the Very Large Scale Integration (VLSI) technology at different levels of interconnects, where wire pitch is defined as the

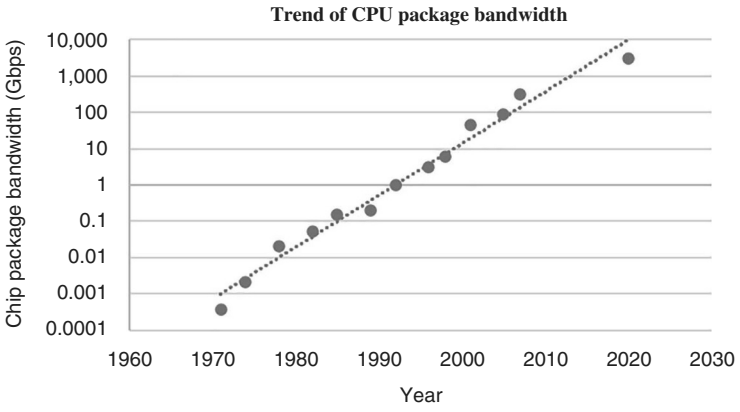


Figure 1.15 Chip package bandwidth over the course of the generation in microprocessor since its inception.

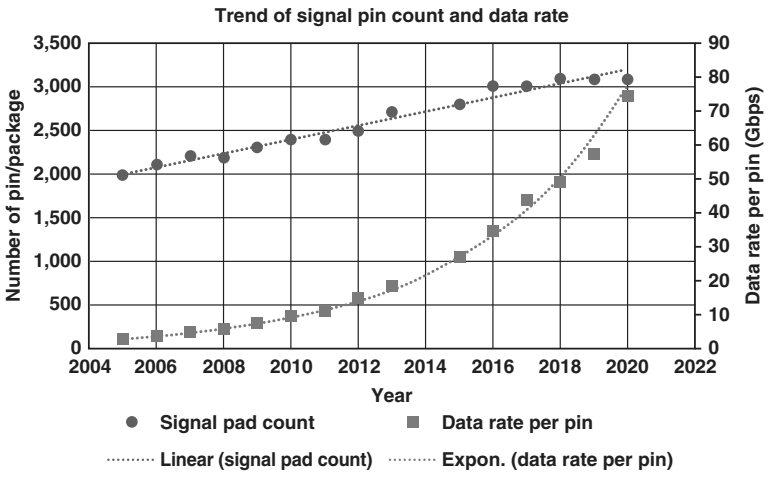


Figure 1.16 ITRS projections of signal pin count and per-pin bandwidth.

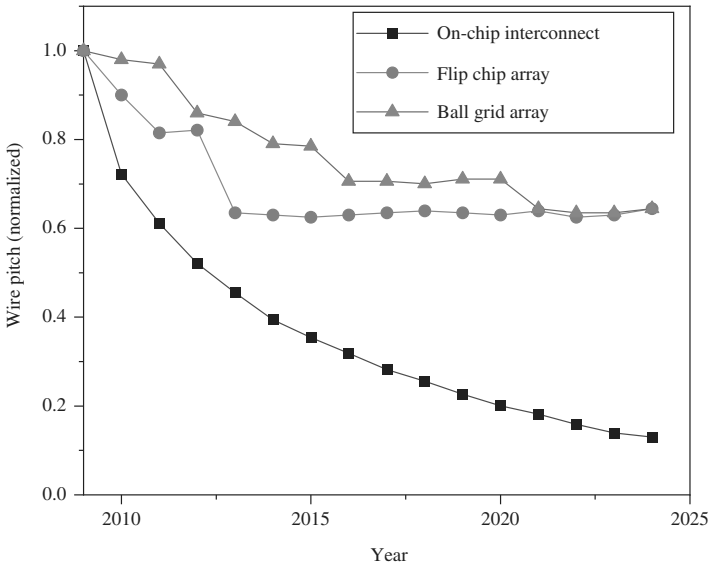


Figure 1.17 Trend of wire pitch/pad pitch at different levels of interconnect. *Source*: Adapted from Yamada et al. [31].

spacing between neighboring interconnects. It is evident that for on-chip interconnects, the wire pitch is being scaled down by a factor of 2 every 6 years, whereas the pad pitch on the PCB has barely scaled for the last few years.

For the best performance in data transfer, a bidirectional, differential channel would have to be designed to occupy eight bumps on the system-on-a-chip (SoC), where each directional link is associated with two signal lines, power, and ground. In addition, the pins will have to be designed as close as possible to the bump to reduce the parasitic capacitance to overcome the RC delay. The pad capacitance is related to the physical geometry, and unfortunately, the geometric size of the pad cannot be aggressively scaled down. Signal pad capacitance is made up of the capacitance of the physical landing pad as well as the gate capacitance of driver/receiver transistors of the interconnect, and the capacitance of the ESD circuitry. None of these capacitances are amenable to exponential scaling over generation. To elaborate the reason, the capacitance of the physical landing pad derives from the size of the landing pad, which is dictated by C4 bumping packaging technology and the separation from the ground plane. This separation between signal wire and ground plane is related to the technological metal stack-up. On the contrary, the geometric width of the off-chip interconnect cannot be scaled too much, as it must provide a low-impedance path to ESD circuitry that combats any issue with current surge. The reliability issue, such as electromigration, also sets a lower bound for the feature size of the interconnects. All these factors combine to raise the parasitic capacitance of the channel, which takes a toll on the data rate.

1.5.2 Issue of Crosstalk

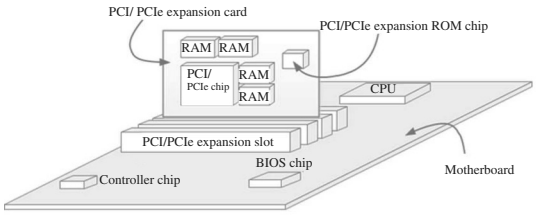
A second category of interference is crosstalk, where signaling channels in proximity capacitively couple energy to one another in a direct fashion or inductively couple energy through return path loops [32]. Forward crosstalk in the direction of wave propagation results in far-end crosstalk (FEXT) at the sensitive receivers of a unidirectional bus. Reverse crosstalk in the opposite direction of signal propagation results in near-end crosstalk (NEXT) at the transmitters of a unidirectional bus. FEXT and NEXT components increase with frequency, and their energy can approach that of the signal at extreme frequencies.

Board-based interconnects have been used for relatively long-range interconnects ($\sim$ 30–100cm) between modules.

Apparently, there are two ways of increasing the bandwidth of system-level interconnects to keep pace with the speed improvement of on-chip interconnects. One way is to increase the pin-count of the IC package to scale up the data throughput of the package. The main problem is that simply adding pins to an off-chip bus does not necessarily increase the throughput in a linear manner [33]. As pins are added to a package to increase the bus size and throughput, the noise induced by the signals that are simultaneously switching leads to a significant increase in noise. This means that as signals are added, the per-pin data rate of each pin needs to be decreased to meet the system noise limits. As a result, the overall throughput of an off-chip bus approaches an asymptotic limit as signals are added [34].

Wide parallel buses undergo several signal integrity issues associated with simultaneous switching of digital signals [35, 36]. The noise induced in the package is proportional to the number of off-chip signals that are switching simultaneously. The other possible option of increasing the throughput of system-level interconnects is to drive each pin at a higher data rate, or to deploy the mechanism of signal modulation, coupled with frequency division multiplexing (FDM) so that each physical channel on the system level can support multiple virtual channels. The problem of this approach is that the electromagnetic coupling among the adjacent wires at the elevated frequencies would be too high to maintain signal integrity. One solution to this problem is to increase the number of power and ground pins in the bus configuration so that they can provide electromagnetic shield between two nearby channels. For instance, PCI-Express data bus, which is a successor of PCI parallel bus, intersperses ground pins more frequently among the I/O data pins in order to facilitate at least an order of increase in the speed of data transfer with high signal fidelity. Figure 1.18 shows the connector pin-out from 18 to 32 pins of PCI-Express ×4 data bus, where we can see that each differential pair of input and output channels is shielded by two ground pins (one on each side of the channel) to maintain signal integrity. However, the inclusion of more ground pins among the data bus increases the cost of the package because the number of I/O pins increases.

Among the existing technologies of parallel buses are [35] Rapid I/O, PCI Express, and Hyper Transport. The factors that cause an upper limit of the information transfer capacity of a parallel data bus are propagation loss in the channel due to channel resistance, channel latency (i.e., minimum rise/fall time of signal at the channel terminal) due to electrical capacitance, noise (such as thermal and flicker noise), and crosstalk due to electromagnetic coupling among adjacent channels. Among them, the resistance derives from the material properties,



(a)

Pin number	Side B	Side A	Description
18	Ground	Ground	
19	HSOp(1)	Reserved	Transmitter Lane 1, differential pair
20	HSOn(1)	Ground	
21	Ground	HSIp(1)	Receiver Lane 1, differential pair
22	Ground	HSIn(1)	
23	HSOp(2)	Ground	Transmitter Lane 2, differential pair
24	HSOn(2)	Ground	
25	Ground	HSIp(2)	Receiver Lane 2, differential pair
26	Ground	HSIn(2)	
27	HSOp(3)	Ground	Transmitter Lane 3, differential pair
28	HSOn(3)	Ground	
29	Ground	HSIp(3)	Receiver Lane 3, differential pair
30	PWRBRK	HSIn(3)	
31	PRSNT2	Ground	
32	Ground	Reserved	

(b)

Figure 1.18 Illustration of (a) an exemplary parallel bus (PCI module expansion) on motherboard, (b) I/O pin-out configuration of PCIe-×4 data bus, showing frequent use of ground pins to prevent crosstalk among data bus.

specifically from the conductivity of metal. Whereas crosstalk mostly owes to the geometry/layout of the wires. If the metal traces are assumed to be almost ideal (i.e., metal with very large conductivity), then the propagation loss as well as channel latency would be insignificant. Under such conditions, the only limiting factor for information transfer would then be the crosstalk among neighboring wires in the bus.

1.5.3 High Power Consumption in System-level Interconnect

Despite significant research efforts, the energy cost of information transfer through interconnects did not scale down the same way as the energy consumption of transistors [37]. As a result, most of the energy dissipation inside modern electronic devices is associated with data communication through interconnects and data links rather than data processing. The energy cost of different operations in short-reach wired communication technology is shown in a pie chart in Figure 1.19 [38–40]. It shows that interconnects may cause energy dissipation three orders of magnitude larger than switching and memory operations.

With the scaling down of feature sizes of CMOS transistors, the chip itself continues to scale up both in size and complexity. While the performance of local interconnects for wiring at the first metal layer(s) of the chip between individual transistors can be improved by judicious uses of repeaters and wire engineering, the biggest challenge lies in keeping up the performance of global interconnects and system-level interconnects. The wire delay in system-level interconnects is getting aggravated, and has already exceeded gate delay by more than an order of magnitude [41], making interchip communications increasingly challenging [42].

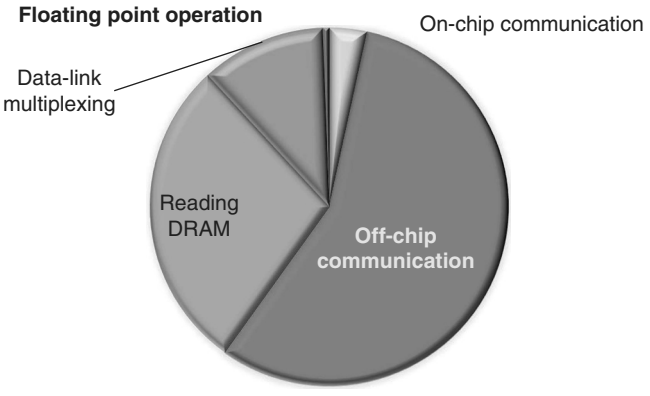


Figure 1.19 Comparison of power dissipation in different components of short-reach wired communication.

As an illustration, the bandwidth of information carried by the internet has reached 280 Tb/s by 2016 [43]. At the same time, a typical graphic processor chip has a peak data rate of 1.4 Tb/s [38], and mere 200 such chips can generate as much data as the entire global long-distance internet traffic. Thus, a major challenge in high-performance cloud computing is to meet the demand for big data exchange in-between electronic devices. While ITRS roadmap is aiming to reach 0.5 THz speed with existing silicon technology [44, 45] and there have been reports on demonstration of high-speed III-V semiconductor material-based transistors operating at THz [46–50], existing data transfer technologies cannot fully utilize the data generation capacity provided by the electronic switching devices [1].

1.6 Optical Interconnect: Evolution Toward Chip-scale Communication

Over the last three decades, optical interconnects have been increasingly adopted into communications systems, ranging from cross-country distance coverage to small-scale distances on electrical backplane. Figure 1.20 shows the commercial realization of optics over time. The approximate trend of shrinking distance and increasing bandwidth for optical interconnect indicates a staggering bandwidth-distance product exceeding 10^6 Gbps-m with single-mode solutions. To put this figure-of-merit into perspective, the electrical interconnects such as copper links have usually been limited within 100 Gbps-m bandwidth-distance product bar. If this trendline continues, many researchers, with somewhat qualified optimism,

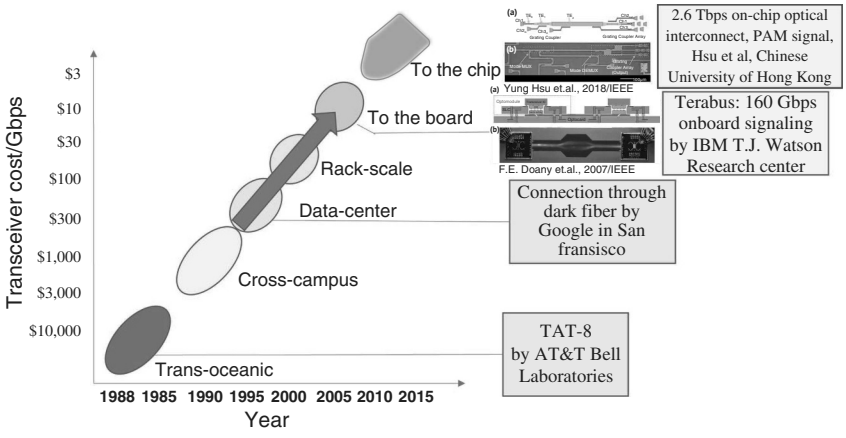


Figure 1.20 Timeline of adoption of optical interconnect toward smaller scale and lower cost/bandwidth.

believe that optical interconnect would eventually substitute interchip and intrachip electrical interconnect, representing link distance below 1 m.

Compared to electrical links, optical channels offer two prime advantages: First, their substantially greater bandwidth and response uniformity over a wide spectrum, which allow designers to avoid the complexity and costs of channel equalization; and second, optical channels using parallel fibers and wavelength division multiplexing (WDM) can significantly increase the number of effective channels to a chip. Thus, the bandwidth of a chip package is no longer bound by the count of physical I/O pins, which reduces the complexity of the clock-data recovery problem and lowers its energy cost. Because of the inherent low material loss of optical waveguides, the main sink of power loss is the electro-optic converter, i.e., the laser driver, modulator, and the receiver circuit. Hence, the prospect of achieving low energy/bit in optical interconnects is critically dependent on designing low-power transceivers.

A low-power transmitter requires intimate integration of modulator and driver circuit components with low parasitics. Although monolithic integration is ideal for low-parasitic integration, photonic devices and electrical circuits currently have contradictory requirements on substrate, precluding a low-energy budget for short-distance coverage. In (Progress in Low-power switched optical interconnects), Krishnamoorthy et al. demonstrated a low-power driver using a 90 nm CMOS process with an effective voltage swing of up to 2 V. They achieved error-free transmission at 5 Gb/s while the total transmitter electrical power consumption was only 2 mW, yielding an energy efficiency of 400 fJ/bit.

Recently, Hsu et al. [51] in Chinese University of Hong Kong combined the use of dense wavelength division multiplexing and mode division multiplexing (MDM) with advanced modulation formats for on-chip optical interconnects. A three-mode MDM multiplexer (MUX) and demultiplexer (DEMUX) was designed and used. Each wavelength carrier had an aggregate data rate of 192 Gbit/s using 4 level pulse amplitude modulation. The mode crosstalk of the MDM MUX-DEMUX was measured to be below −20 dB in the 40 nm wavelength band from 1520 to 1560 nm. Fourteen wavelength channels are multiplexed and successfully transmitted with an aggregate transmission capacity of 2.6 Tbit/s (14 wavelengths × 3 modes × 64 Gbit/s) with all channels satisfying the hard decision forward error correction threshold.

1.6.1 Integrated Photonic Circuits on Silicon

Silicon photonics is an innovative endeavor that leverages the knowledge and expertise acquired from silicon CMOS processing techniques to advance the realms of photonics and opto-electronics. The primary objective is to achieve substantial enhancements in yield, cost-effectiveness, and time-to-market while

minimizing additional investments [52]. At its core, silicon photonics revolves around manipulating light within silicon, encompassing crucial functions such as light generation, routing, modulation, processing, and detection [53], as illustrated in Figure 1.21. These integrated functionalities collectively form the optical equivalent of electronic integrated circuits, commonly referred to as photonic integrated circuits (PICs). The fundamental principles of silicon photonics can be traced back to the influential work of Bennett and Soref [54]. Their paper introduced the concept of free carrier dispersion in silicon, highlighting how the introduction or removal of carriers within the material can induce alterations in its optical properties, specifically the refractive index (n) and extinction coefficient (k). This pivotal breakthrough paved the way for integrating PN junctions into waveguide structures, enabling the development of the initial crucial component of a photonic integrated circuit: the phase modulator [55]. Subsequently, numerous academic institutions and research centers embarked on an extensive exploration, dedicated to realizing various device types essential for a comprehensive silicon PIC solution, thereby instigating a wave of innovation and progress in the field.

The development of silicon-on-insulator (SOI) wafers by Bruel enabled advancements in silicon photonics [56]. This allowed for the confinement of light in the vertical direction, enabling the definition of waveguides within the plane using lithography techniques. Initially, the SOI waveguides were relatively large, but the research community later adopted a standard size of 220 nm, resulting in low losses and tight bending radii. These standardized waveguides facilitated the implementation of various passive devices like splitters and crossings. Ongoing discussions revolve around the use of thicker silicon, which offers advantages such as lower waveguide losses and improved vertical light coupling.

In the field of silicon photonics, significant progress has been made with the emergence of various components. Directional couplers, multiplexer/demultiplexer devices, and active components like tunable micro-ring resonators

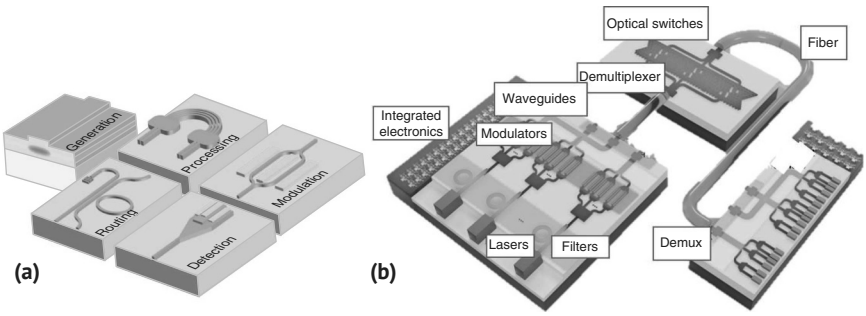


Figure 1.21 (a) Circuit components for silicon photonics; (b) a functional photonic circuit integrated with electronics on silicon.

have become prominent. High-speed Mach–Zehnder modulators capable of achieving a bandwidth of 1 GHz based on the plasma-dispersion effect have been developed. Additionally, the integration of strained SiGe on silicon into waveguide photodetectors has been successful. While the silicon photonic chip is nearing completion, the absence of a silicon laser remains a significant challenge. Efforts are being made to address this challenge through laser integration into silicon via packaging and the exploration of silicon-based lasers.

Packaging solutions for silicon photonics have gained attention as the technology progresses toward commercialization. Integrating silicon photonic chips with optical fibers poses challenges due to the refractive index mismatch, leading to reflections and low coupling efficiencies. However, innovative approaches such as high-efficiency grating couplers and edge couplers have shown promise in achieving reasonable coupling losses. It is anticipated that advancements in packaging will continue, with the expectation that 2.5D or 3D integrated photonics will be commercially deployed in the near future. 2.5D integration involves interconnecting chips through an interposer or substrate, while 3D integration entails stacking optical and electrical dies.

A significant milestone in silicon photonics was the development of commercially viable transceivers. In 2007, Luxtera introduced silicon photonics transceivers for the datacom market, debuting the world's first 40G active optical cable with a reach of up to 300 meters. In 2014, Acacia unveiled their silicon photonics solutions, including the first 100G coherent transceiver suitable for metro research applications.

However, like any emerging technology, silicon PICs are not without their shortcomings and limitations. The following account for some of the key challenges and drawbacks associated with silicon photonic integrated circuits [57].

- **Integration Complexity**: One of the primary challenges in silicon PICs lies in the complexity of integration. While silicon offers excellent compatibility with existing CMOS fabrication processes, integrating different photonic components on a single chip remains a formidable task. The integration of active and passive components, such as lasers, modulators, detectors, and waveguides, requires precise alignment and optimization, resulting in increased fabrication complexity and cost.
- **Absence of On-Chip Laser Source**: One of the fundamental limitations of silicon PICs is the lack of a monolithic on-chip laser source. Pure crystalline silicon lacks a direct bandgap, making it incapable of emitting light efficiently. As a result, the integration of a laser source with silicon PICs necessitates additional off-chip elements or hybrid integration techniques, which can compromise the overall performance, scalability, and cost-effectiveness of the integrated circuit.

- **Limited Wavelength Range**: Silicon PICs primarily operate in the near-infrared wavelength range, typically around 1.3 and 1.55 μm. This limited wavelength range restricts their applicability in certain fields, such as telecommunications, where other wavelength bands are preferred for specific applications. Expanding the operational wavelength range of silicon PICs requires the incorporation of alternative materials, such as III–V semiconductors, which introduces further integration challenges and complexities.
- **High Propagation Losses**: Silicon, despite its compatibility with CMOS processes, suffers from intrinsic optical losses due to its relatively high refractive index and material absorption characteristics. These losses become more pronounced as the dimensions of the waveguides and optical components shrink. Consequently, high propagation losses limit the overall efficiency and reach of silicon PICs, necessitating the use of additional amplification techniques or expensive low-loss fabrication platforms.
- **Thermal Management**: Silicon PICs generate significant amounts of heat due to their compact form factor and high integration density. Efficient thermal management is crucial to ensure the reliable operation and longevity of the integrated circuits. However, silicon's poor thermal conductivity poses challenges in dissipating heat effectively, leading to potential performance degradation, reliability issues, and increased power consumption.
- **Packaging and Fiber Coupling Challenges**: The successful integration of silicon PICs with external fiber-optic systems is vital for practical deployment. However, the high refractive index of silicon presents challenges in achieving efficient coupling between the chip and optical fibers. Mode and refractive index mismatches lead to significant reflections and coupling losses, reducing the overall system performance. Innovative packaging techniques, such as grating couplers and edge couplers, have shown promise but require further development and optimization.

1.6.2 Is Optical Interconnect Viable at Short Range?

By now, we have argued how crosstalk among the channels in data bus can become the ultimate limiting factor in achieving high bandwidth density, given the fact that electrical wires succumb to electromagnetic interference as we move toward higher frequencies. This may ignite the thought of replacing electrical data bus with optical data bus, since optical waveguides suppress crosstalk at high frequencies: an opposite trend of electromagnetic interference with frequencies compared to that of electrical ones. While the optical data bus can indeed outperform the electrical data bus in terms of dealing with crosstalk, it comes with a price: optical waveguides always require sources of carrier frequency, modulators, and demodulators, which add to the overhead energy budget of information transfer. Usage of

a given interconnect technology over the other alternatives is justifiable when the technology in question can result in the minimum energy consumption for per bit of transferred data. Therefore, unless the rate of data transfer exceeds a threshold value, optical interconnects give in to their electrical counterparts in the ability of data transfer. It turns out that, for short-range interconnects, the rate of data traffic in a channel may vary in such a wide range over time that employing an optical interconnect may often remain underutilized of its full capacity of information transfer, and therefore cost unnecessary loss of energy attributed to its overhead power budget.

Because of the above-mentioned system complexity and overhead cost, fiber-optic communication systems couldn't make their niche in short-distance communication. They have primarily been installed in long-distance applications, where they can be used to their full transmission capacity, offsetting the set-up cost. Finally, plasmonic interconnects, which grabbed researchers' interest over the last decades due to their ability of subwavelength confinement, are plagued with metallic losses and the high attenuation factor limits the signal propagation length within a few microns.

1.6.2.1 Electrical vs Optical Interconnect: Cost Comparison

The cost per data rate in a communication link tends to increase as the distance between the transmitter and receiver becomes longer. This is primarily due to factors such as longer cables and the need for additional circuitry to compensate for signal loss. However, the cost dynamics change significantly at critical transition points dictated by packaging limitations. On-chip interconnects are cost-effective because of their integration within the chip. However, once the interconnect length exceeds the chip's typical dimensions, the cost rises as additional packaging is required. Similarly, at the board level, the cost of the communication channel gradually increases with the channel length until it surpasses the board's typical dimensions, necessitating additional connectors and receptacles. The cost pattern continues as link lengths exceed the size of a room or building. Interestingly, Figure 1.21 shows that optical interconnects, due to lower material loss compared to copper interconnects, can minimize signal compensation circuitry and thus reduce the cost per unit length. However, optical links require additional circuit components like lasers and photodetectors, which add to the overhead cost. Electrical links, on the other hand, do not require such components and remain preferable for chip-scale communication. As we move to board-level communication and longer links, the advantages of optical links become evident. The extra cost of electro-optic conversion circuitry in optical links is offset by the significant material loss experienced in electrical links. The crossover point between optical and electrical links depends on the data rate, with higher data rates leading to smaller crossover lengths, as illustrated in Figure 1.22.

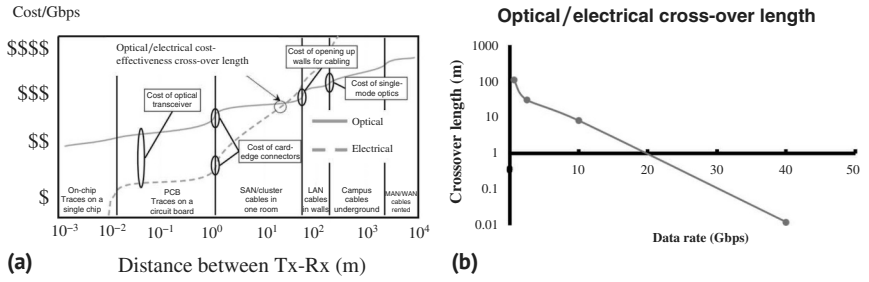


Figure 1.22 Link cost versus distance plot showing the factors determining effective optical/electrical crossover length.

In summary, electrical transmission is cost-effective for shorter chip-scale links, while optical means offer benefits for longer distances. The cost dynamics in communication links are influenced by factors such as distance, packaging limitations, material loss, and the specific requirements of optical and electrical components.

1.6.3 Wireless Network on Chip

Decades of dedicated research in the millimeter-wave (mmWave), THz, and optical frequency bands have positioned wireless technology as a viable solution for interconnecting Systems-on-Chip (SoC). By integrating numerous antennas and transceiver circuitry into intricate networks, wireless data transfer within microprocessor modules on a single chip becomes achievable, as illustrated in Figure 1.23. Particularly in the context of many-core architectures, where the speed of core-to-core communication is primarily limited by latency rather than interconnect bandwidth, the adoption of wireless NoC (WNoC) shows promise in facilitating direct and unobstructed communication among multiple processor cores, schematically shown in Figure 1.24. At the opposite end of the electromagnetic spectrum, the optical bands have emerged as an attractive option for wireless communication within chips, thanks to significant advancements in integrated silicon photonics, giving rise to microscale lasers [58], photodetectors [59], and optical antennas [60]. Furthermore, the THz band is gaining attention as a potential WNoC solution, despite being relatively nascent, as progress is being made in complementary electronic, photonic, and plasmonic approaches to bridge the THz technology gap. However, the benefits of wireless techniques in efficient and low-latency multicasting of data come at a considerable cost: the amplified occurrence of multipath interference among channels. This necessitates the deployment of complex circuitry for equalization, filtering, and the implementation of protocols for time and frequency multiplexing, raising concerns about the overall cost-effectiveness of wireless data recovery circuits.

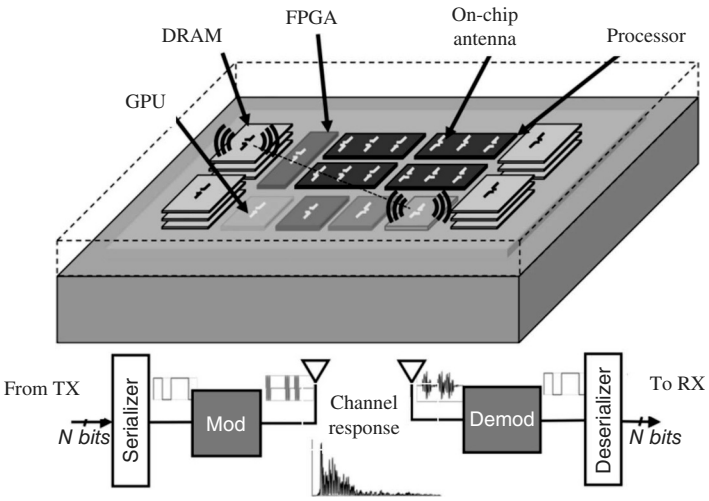


Figure 1.23 A general view on wireless communications at the chip scale within a heterogeneous computer architecture with detail on the wireless transmission process.

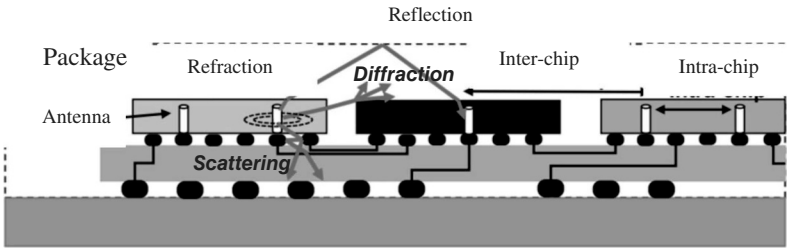


Figure 1.24 Schematic representation of wave propagation in an interposer system with flip chip package excited with vertical monopole antennas, distinguishing between intra- and interchip regions, and exemplifying different propagation phenomena.

One of the key obstacles in wireless technology within chip packages is the significant loss of signal propagation. Studies have demonstrated that even at short distances of a few millimeters, carrier frequencies in the tens of gigahertz range can experience substantial attenuation of up to 50 dB. High attenuation levels have been observed in wireless links within standard chips and packages. A visual representation in Figure 1.25 showcases FDTD simulations, illustrating that the loss of a 60 GHz carrier within a standard flip-chip package is significantly higher compared to guided wave propagation in a similar medium. In response to this challenge, researchers have explored various approaches to mitigate the excessive loss. For instance, Zhang et al. [61] have investigated the use of low-doped silicon as a means to reduce losses induced by the substrate, resulting in a 20 dB reduction in

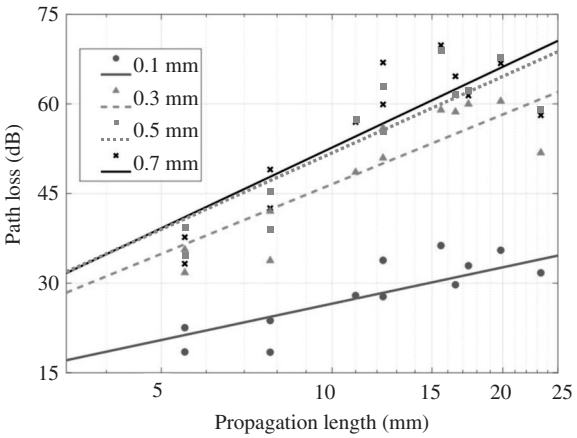


Figure 1.25 Intrachip propagation loss within a flip-chip package at 60 GHz with variable silicon thickness, heat spreader of 0.2 mm propagation loss in silicon.

loss. Similarly, incorporating an undoped silicon layer beneath the die substrate is proposed, which can achieve attenuation levels of 15–30 dB. Another alternative avenue involves leveraging metamaterials integrated on the chip, enabling stronger coupling to surface waves and offering protection against radiation losses [62]. However, it is important to note that these unconventional approaches introduce compatibility issues with standard silicon technology. While improvements in signal propagation within high-resistivity silicon can be achieved, they may come at the expense of circuit performance, such as slower transistor switching speeds. Additionally, the utilization of metamaterial-assisted waveguiding techniques poses challenges in the manufacturing process of integrated chip packaging.

1.6.3.1 Dispersive Nature of the Intrachip/Interchip Wireless Channels

The group delay dispersion is a crucial metric that affects the data transfer capability. It arises from the multiple reflections of incident waves on various interfaces within flip-chip packages. It was demonstrated that a 60 GHz carrier signal traveling through a silicon chip package can experience delay dispersion on the order of a fraction of a nanosecond over a distance of a few tens of units [63], as shown in Figure 1.26. Consequently, the bandwidth limited by dispersion within a standard package is typically a few Gbps at best. To mitigate delay dispersion, one approach is to design thinner packages or utilize carrier frequencies with longer wavelengths, which minimize the phase shift caused by multiple reflections at package interfaces. Timoneda et al. capitalized on this concept by developing a flip-chip package design that effectively controlled dispersion-induced delay

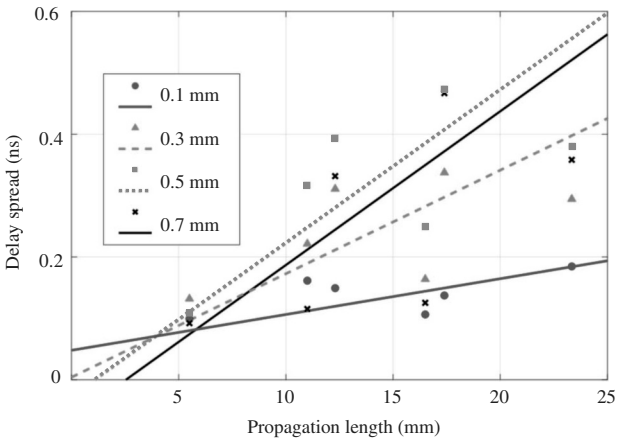


Figure 1.26 Delay dispersion in intra-chip channel at 60 GHz with variable silicon thickness. *Source*: Abadal et al. [63]/IEEE/CC BY 4.0.

spread to below 100 ps Timoneda et al. [64]. This design achieved moderate propagation loss while providing a chip-wide bandwidth exceeding 10 GHz. By reducing delay dispersion through optimized packaging, the limitations on data transfer caused by dispersion were alleviated, leading to improved performance in terms of bandwidth and data transmission capabilities.

1.6.3.2 Signal Fidelity: Gigahertz Carrier vs Optical Carrier

In order to implement low power data transfer over relatively longer distances, propagation loss has to be minimized. It has been computationally shown that the optical free space carrier (10–300 THz) is more susceptible to path loss than carrier with a few tens to a few hundred of gigahertz, as shown in Figure 1.27. Moreover, the propagation constant periodically changes over frequencies, where the periodicity derives from the frequencies between two neighboring surface wave modes. It has been shown that a thinner underfill layer and a bottom layer between the silicon substrate and the heat sink would be able to reduce the path loss. Besides, the thickness of the silicon substrate must be cautiously designed to aid signal immunity to propagation loss.

1.6.4 Carbon Nanotube Interconnect

Another material gaining attention for interconnect applications is carbon nanotubes (CNTs). As integrated circuits (ICs) continue to scale down, the size of interconnects becomes increasingly smaller, reaching nanometric dimensions. This reduction in interconnect size results in delays with copper (Cu) interconnects, as discussed earlier. CNTs have emerged as a potential replacement for

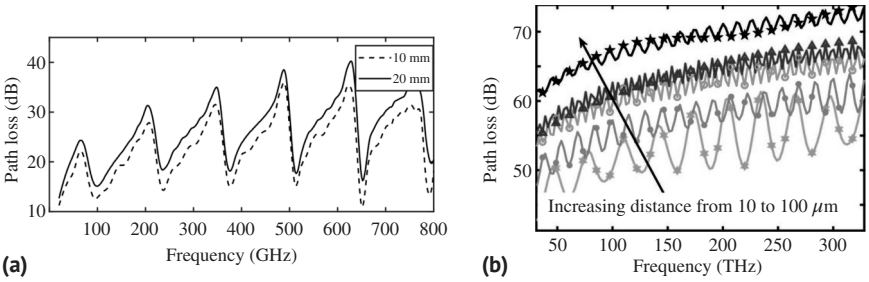


Figure 1.27 Propagation loss of the THz chip-scale wireless channel at (a) sub-THz frequency and at (b) THz frequency. *Source*: Abadal et al. [63]/IEEE/CC BY 4.0.

on-chip Cu interconnects [65]. Compared to traditional Cu interconnects, CNTs exhibit significantly higher current-carrying capabilities, with reported current densities around 10^9 A/cm^2, whereas Cu interconnects typically have current densities in the order of 10^7 A/cm^2. In 2003, ballistic transport in single-walled CNTs (SWCNTs) was reported [66]. CNTs demonstrate low resistivity, high current density tolerance, and scalability to nanoscale dimensions. Numerous research efforts have focused on CNTs as future interconnect materials. Bailey et al. provided a comprehensive summary of recent advances in CNT interconnects and highlighted the challenges that the research community needs to address, particularly in terms of modeling technology [67]. Brun et al. described CNT models for high-frequency interconnects using electromagnetic and circuit approaches [68]. CNTs have been recognized as a promising technology for sub-nanometer scales. Since their discovery in 1991 by Iijima, global research endeavors have been made to grow CNTs for specific applications. While CNTs have also been considered for mainstream microelectronics, there are various integration challenges with VLSI components that must be properly addressed to achieve desired outcomes. The use of CNTs for interconnections is still in the research phase. An equivalent transmission line model for CNTs is proposed [69], and it was found that, at the local interconnect level, CNT interconnects do not outperform Cu wires. However, at the intermediate and global interconnect levels, CNTs may exhibit equal or even better performance compared to Cu interconnects.

1.7 Complexity and Dilemma in Data Transfer

We identify two key features prevailing in interchip communications:

- Centimeter-scale distance is the most critical range of communication, where no single means of standard communication (between electronics and optics) stands out to be distinguishingly superior to the other.

- The nature of data-traffic rate statistically varies from time to time rather than being constant. Hence, any fixed mode of communication link for centimeter-scale distance of coverage may run the risk of remaining largely under-utilized or over-utilized for most of the time, resulting in a severe strain on energy dissipation.

In the following sections, we will exposit the term "last centimeter" in conjunction with "statistical nature of data-traffic" in some greater detail.

1.7.1 The Last Centimeter Barrier: The Most Critical Distance in Data Transfer

To discuss the trade-off between electrical and optical interconnect in terms of power dissipation at 10's of centimeter distance, we provide Figure 1.28a, which compares the power consumption of electrical and optical interconnects at 4 and 6 Gb/s, where the diagram is reproduced from Cho et al. [70]. It can be observed that power consumption in electrical interconnect rises with length and bit rate due to a larger attenuation and fixed noise sources. At higher bit rates, a smaller rise time would further increase noise sources.

The optical interconnect power dissipation also rises with length due to loss in optical power, although this rise is slower than that for the electrical interconnect. Beyond a critical length, optical interconnect yields lower power. This critical length reduces at higher bit rates. The "critical length" divides the space into two regions in Figure 1.28a. The region to the left of "critical length" is favorable for

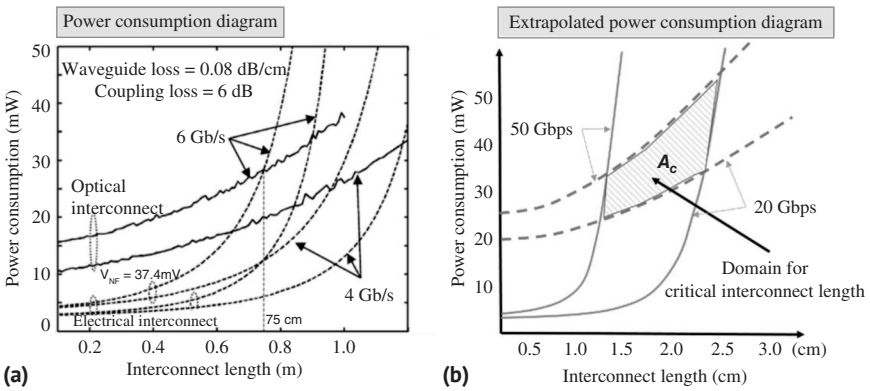


Figure 1.28 (a) Power comparison between optical and electrical interconnects, as predicted by Cho et al.; (b) power consumption diagram by extrapolating data in Cho et al. [70]/IEEE, showing that the critical interconnect length, depending on the data rate at a given time, can be anywhere in region A_c.

electronic data transmission, while the region to the right is favorable for optical data transfer. Figure 1.28b illustrates the power consumption diagram plotted by extrapolating the data taken from Figure 1.28a. It shows that, if the data rate varies in a wide range of "20–50 Gbps" over the course of time, then the "critical length" could be anywhere in the area A_c. In other words, for a given length of interconnect in the order of centimeters and given that the data rate varies with time, there is no fixed mode of interconnect technologies that can ensure minimum power consumption for all possible time.

1.7.2 Time Variation in Data Traffic: A Dilemma in Selecting Interconnect Technology

The dilemma in the selection of interconnect for energy economization becomes obvious once we consider the time-dependent statistical nature of the volume of data traffic. Traffic can be characterized by assuming that the packet arrivals obey the Poisson distribution. For example, at IT data centers, it witnesses a large temporal variation in data flow over the course of a day [71]. The time variation of traffic gets even more prominent as we move down the network hierarchy: from data center core to the individual user terminals. Benson et al. [72] showed that link utilization is one-fourth in the terminal layer than in the core layer, with the full utilization not exceeding 10%. Because of the temporal distribution of data traffic, especially at the range of centimeter-scale interconnect length at the board-level, it is most difficult to choose an optimum interconnect technology that can ensure full utilization of link capacity and minimize energy cost for a highly random volume of data flow. This is why it would be immensely beneficial to overcome the "last centimeter interconnect barrier" with a technology that, rather than supporting a fixed mode, can alternate between two different modes of communication.

1.7.3 Do We Have Any Alternative Interconnect Technology in Hand?

As discussed in the preceding sections, there remains a range of distance of 0.1–10 cm, dubbed as the "last centimeter barrier," where state-of-the-art interconnect technologies compromise on critical performance metrics such as bandwidth and energy efficiency [73]. The unresolved challenges in short-range communications between integrated circuits (ICs), therefore, warrant discoveries of new interconnect technologies having ultrafast data transfer rate, high energy-efficiency, and compatibility with CMOS technology [41, 42, 74]. As shown in Figure 1.29, between the electromagnetic spectral ranges for electronic and optical technologies, an emerging terahertz regime (0.1–10 THz) in the underexploited and underutilized electromagnetic spectral band is now garnering attention for wired and wireless communication systems [75–79].

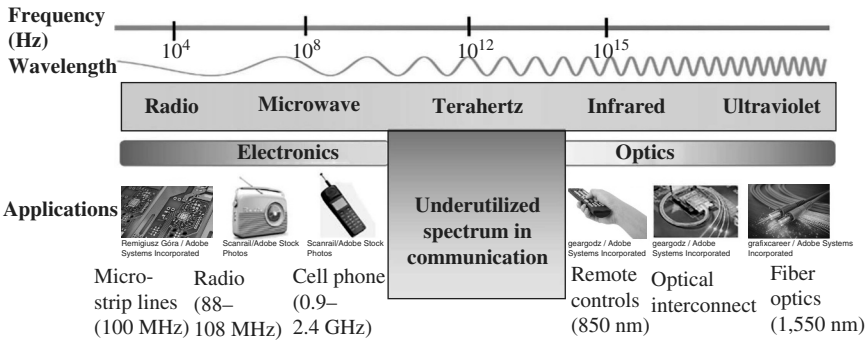


Figure 1.29 Communication device design in microwave and optical frequencies left terahertz regime underexploited.

1.7.4 Spoof Plasmon Interconnect: A New Paradigm in Communication Technology

In the framework of the current book, we have investigated the concept of implementing a communication network that can transfer information by a new type of electromagnetic mode, called spoof surface plasmon polariton (SSPP) (Figure 1.30) for the future high-speed data transfer between integrated circuit (IC) chips, overcoming the limitations of state-of-the-art interconnects. SSPP is a special, surface-bound electromagnetic wave in patterned metal [80, 81]. Artificially engineered interconnects offer a potential breakthrough by combining the benefits of both electrical and optical technologies, see Figure 1.31. On the one hand, they could provide an energy-efficient alternative for short-reach data transfer (10 GHz–10 THz) by minimizing the wire parasitics to suppress crosstalk and ground bounce, and on the other hand, they offer a scalable solution to short-reach interconnects by avoiding the extra energy overhead of transceivers.

It may be noted that the precursor of spoof plasmon technology was surface plasmon polariton (SPP) or plasmonic mode of signal propagation in the optical and near-infrared frequency range (300–800 THz).

1.7.4.1 Plasmonics: An Early Precursor for Spoof Plasmonic Technology

The emergence of plasmonics-based technology occurred in the early 2000s, aiming to harness the unique optical properties of nanoscale metallic structures for the manipulation and routing of light at the nanoscale. Researchers envisioned integrating plasmonic, electronic, and traditional photonic devices on a single chip to leverage the strengths of each technology. The idea was that metal nanostructures could possess the ideal combination of electronic and optical properties to address existing challenges and achieve significantly enhanced processing speeds.

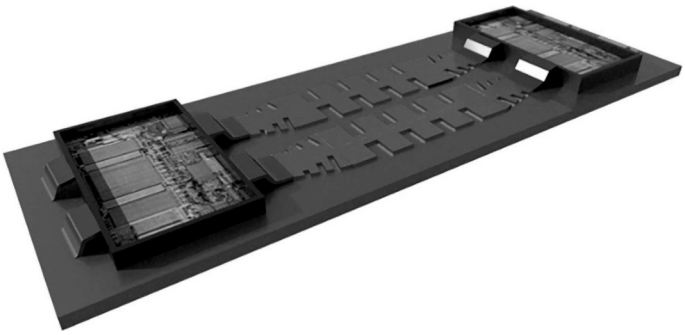


Figure 1.30 An interchip communication network employing SSPP channels. The typical free space wavelength corresponding to spoof plasmon resonance is 300 microns.

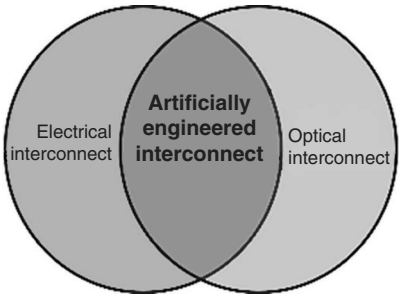


Figure 1.31 Artificially engineered interconnect can bring the advantage of both the electrical and optical interconnect technology to fill the gap for and energy-efficient, high bandwidth, low latency short-reach interconnects.

Commonly used metals for electrical interconnections, such as Cu and Al, could excite surface plasmon polaritons (SPPs), which are electromagnetic waves propagating along the interface between a metal and a dielectric, coupled to the free electrons in the metal.

From an engineering perspective, SPPs can be seen as a distinct type of light wave. The metallic interconnects supporting these waves act as miniature optical waveguides known as plasmonic waveguides. The concept that the optical mode, or "light beam," diameter perpendicular to the metal interface could be significantly smaller than the wavelength of light generated great excitement and fostered the vision of eventually interfacing nanoscale electronics with optical (plasmonic) devices of similar size.

One of the most promising applications of plasmonics is in sensing and imaging, exemplified by its use in a scanning tunneling microscope (STM). Brongersma et al. from Stanford University developed a photon scanning tunneling

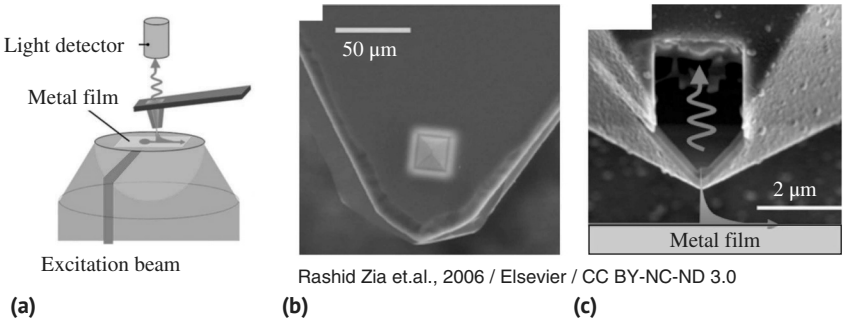


Figure 1.32 (a) Illustration of the functioning of a PSTM (photon scanning tunneling microscope), enabling the investigation of surface plasmon polariton (SPP) propagation along metal film surfaces. The arrow indicates the launch of an SPP from an excitation spot onto the surface of a metal film, achieved using a high numerical aperture microscope objective. The scanning electron microscopy (SEM) image (b) displays the near-field optical cantilever probe utilized by Brongersma et al. It comprises a microfabricated, hollow glass pyramid coated with a thick layer of aluminum (Al). The Al film on the top of the pyramid features a tiny aperture (50 nm) through which light can be collected or emitted. Additionally, a cross-sectional view (c) of the same hollow pyramidal tip is depicted after a focused ion beam (FIB) was employed to remove a significant portion of the sidewall. *Source*: (b,c) Zia et al. [82]/Elsevier/CC BY-NC-ND 3.0

microscope (PSTM) by modifying a commercially available scanning near-field optical microscope. PSTMs are widely employed for studying the propagation of SPPs along extended films and metal stripe waveguides. Figure 1.32 demonstrates how a microscope objective within the PSTM system can focus a laser beam onto a metal film at a precisely defined angle, effectively launching an SPP along the top surface of the metal. This technique for exciting SPPs utilizes the well-known Kretschmann geometry, which enables phase matching between the free space excitation beam and the SPP.

To locally tap into the guided SPP wave and scatter light toward a far-field detector, a sharp pyramidal tip coated with metal is employed (as depicted in Figure 1.32b,c). These tips possess a nanoscale aperture at the apex of the pyramid through which light can be collected. The scattered light is then detected using a photomultiplier tube. By measuring the intensity of the local light directly beneath the tip and scanning the tip across the metal surface, the propagation of SPPs can be imaged and visualized.

Because of the unique properties of surface plasmon polariton, plasmonics were advertised to be the proper conglomeration for electronics and optics. Figure 1.33 captures the envisioned role of plasmonic devices on a chip, illustrating their operating speeds and critical dimensions in comparison to other chip-scale device technologies. In the past, devices were limited by their sluggishness and bulkiness.

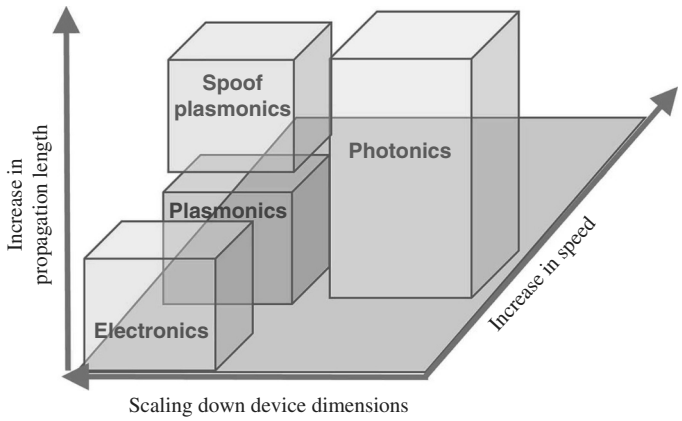


Figure 1.33 The proposed perspective in this book toward different technologies that entails a holistic view of figures of merit for evaluating the prospect of chip-scale communication. While a simple two-dimensional diagram in Figure 1.6 propounded for plasmonics technology to merge the gap between electronics and optics, a more sophisticated three-dimensional diagram incorporating signal propagation length may help understand the limit of plasmonics and how we can push the boundary further by adopting spoof plasmonics, a version originated from but different from its precursor, plasmonics.

However, the semiconductor industry has made remarkable progress in scaling electronic devices to nanoscale dimensions. Despite this achievement, challenges arise in terms of interconnect delay times, hindering the realization of purely electronic circuits operating above approximately 10 GHz.

In contrast, photonic devices possess immense bandwidth, enabling high-capacity data transmission. Nevertheless, dielectric photonic components face limitations in size due to the laws of diffraction, impeding their scalability comparable to electronics. This is where plasmonics were hoped to offer a distinct advantage, combining the size characteristics of electronics with the speed capabilities of photonics. Plasmonic devices have the potential to naturally interface with photonic devices operating at similar speeds and electronic components of comparable sizes. By bridging the gap between these two technologies, which currently struggle to communicate effectively, plasmonics were envisioned to serve as the interface between two disparate technologies.

Although plasmonic technology can potentially enhance the capability of nanoimaging, the potential of plasmonics for light propagation and routing appears rather limited. What makes it so is the inherent material loss associated with plasmonic systems. The metals commonly used for plasmonic devices, such as gold and silver, exhibit considerable absorption and scattering losses, limiting the efficiency and performance of these devices. Material loss directly impacts the signal

intensity and propagation distance, reducing the overall effectiveness of plasmonic systems.

To comprehensively investigate and quantify the signal propagation properties of plasmonic systems, a research team in Mazumder Lab (K. Song, P. Mazumder) embarked on a study in the early 2000s. The focus of Mazumder Lab was on analyzing the power flow characteristics in metallic nanowires (MNWs) composed of silver (Ag). The choice of nanowire radii within the tens of nanometers range aimed to ensure compatibility with interconnects in VLSI chips. In order to assess the performance of MNWs in conjunction with nanoelectronics, a transmission line model was employed, transforming the TM mode of the nanowires governed by Maxwell's electrodynamics. This approach enabled to express the properties of the MNWs using standard signal transmission parameters such as resistance, inductance, capacitance, and conductance (referred to as RLCG parameters) within the SPICE circuit model. As a result, Mazumder Lab successfully characterized the propagation wave vector and attenuation constant of Ag MNWs. Through a comprehensive investigation, two significant findings emerged:

First, Mazumder Lab found that the frequencies at which the electromagnetic mode couples with plasmonic nanowires to create nonradiative surface-bound modes were situated in the ultraviolet range, with angular frequencies on the order of 10^{15} rad/s, see Figure 1.34. This frequency range proved incompatible with both baseband transmission in gigahertz electronics and the 1.55 μm technology commonly utilized in optical fiber transmission.

Second, Mazumder Lab observed that the length of signal propagation, limited to −3 dB power transmission, barely extended to a few hundred nanometers.

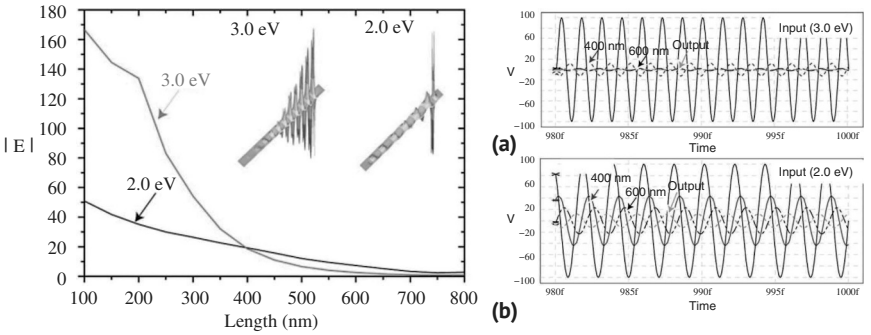


Figure 1.34 (a) Average *E*-field density in the Ag MNW (radius = 10 nm) performed by FDTD. These values are calculated at the positions at 15 nm offset positions (*z*-directed) from the axis of MNW. (b) HSPICE transient analysis of Ag MNW (radius: 10 nm, length: 800 nm) obtained by HSPICE simulator: blue optical frequency ($\omega \sim 4.56 \times 10^{15}$ rad/s: 3.0 eV), and red optical frequency ($\omega \sim 3.04 \times 10^{15}$ rad/s: 2.0 eV).

This distance fell at least 5000 times shorter than the size (1 mm) of a contemporary chip.

These discoveries revealed the shortcomings of plasmonic technology to replace or enhance the bandwidth capacity of nanoscale electronics in VLSI. Consequently, it became evident that a true candidate for high-performance data transmission in chip-scale applications needed to be sought in a regime where metallic loss is more manageable and less detrimental. This realization led Mazumder and his colleagues to develop a keen interest in spoof surface plasmon polaritons, an electromagnetic mode inspired by plasmonics but superior in many ways, particularly in terms of manageable material loss.

In this book, the authors aim to introduce the potential of the alternative technology based on spoof surface plasmon polariton, particularly in the domain of chip-scale ultrafast data transfer, as illustrated in Figure 1.30. SSPP waves are strongly confined subwavelength modes [80]. Therefore, the parallel data bus of SSPP can be densely packed, and data can be safeguarded from crosstalk interference.

In addition, SSPP modes can propagate a considerable distance (a few centimeters) by circumventing ohmic loss resulting from penetration of the mode into the metal. On one hand, SSPP interconnect possesses the advantage of CMOS compatibility [83] and its energy-economic quasi-electrical mode can transfer data without requiring excessive energy. On the other hand, its quasi-optical mode with deep subwavelength confinement can provide high bandwidth overcoming the conventional 1/RC-limit. In our research, we demonstrated that spoof-plasmon-based data bus can provide nearly 300 Gbps data transfer rate per channel with 1 Tbps/mm bandwidth density, while the signal can propagate up to a wire length of 1 cm [84]. These results indicate a remarkable potential of SSPP interconnects to become the most viable technology for the next-generation high-speed interchip short-range communications, as illustrated in Figure 1.35. Therefore, in this book, we propose a visual illustration in Figure 1.33, which is an extended version of the vision of the early developers of plasmonic technologies represented in Figure 1.36, yet simple enough to understand and appreciate how different technologies posit in terms of their speed, size, and signal propagation properties, so that we can find any gap in between and work toward mitigating the technology gap. In particular, Figure 1.33 expands the two-dimensional (2D) view of Figure 1.36 into a 3D view by adding one more critical parameter for signal transmission in the other axis, namely signal propagation length. The elaborated view of Figure 1.33 points to the limitation of plasmonics within 1 μm of signal propagation, and shows how we can push the frontier even further by adopting spoof plasmonic technology.

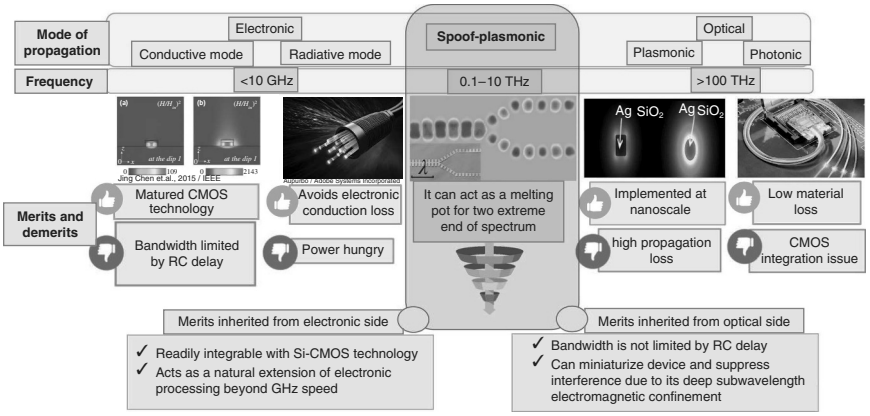


Figure 1.35 Illustration of how the spoof plasmonic technology posits in a spot between electronics and optics, bringing in the advantages of both the domain, and merge the technology gap in between.

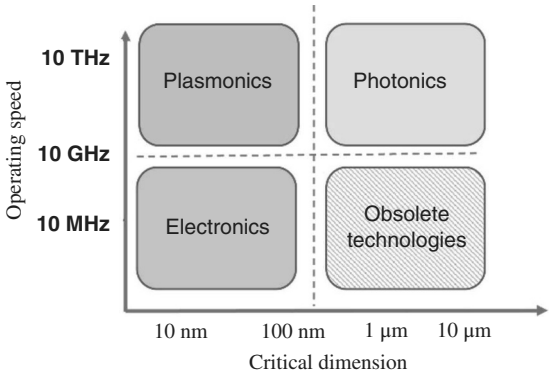


Figure 1.36 Early adopters of plasmonic technologies championed the idea that plasmonics might bring together the advantages of both nanoscale electronics and tera-speed photonics. Operating speeds and critical dimensions of various chip-scale device technologies, highlighting the strengths of the different technologies.
Source: Zia et al. [82]/with permission of Elsevier.

1.8 Research on Spoof Plasmon Wave: Toward CMOS Compatibility

Spoof plasmon wave has become a new exciting field of research [85–90], as illustrated in Figure 1.37. Researchers in Mazumder Laboratory have developed the theory and demonstrated multifaceted applications of SSPP-based systems [81, 91–99]. Pioneered by Sir John Pendry's group at the Imperial College,

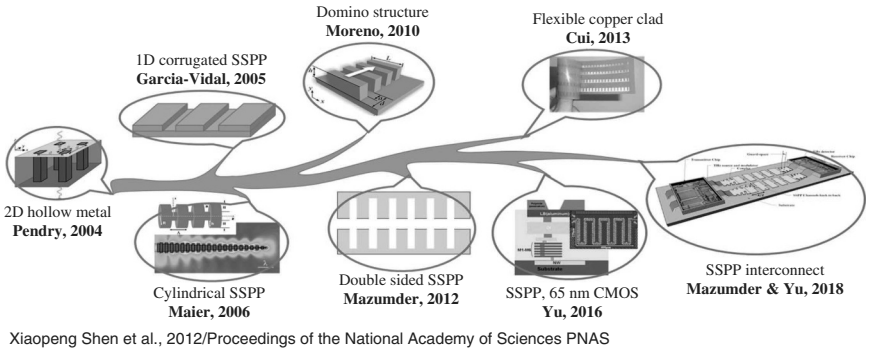


Figure 1.37 Brief illustration of the evolution in SSPP research. The concept and realization of THz surface wave evolved from complex structures toward on-chip, CMOS technology-compatible thin-film structures.

UK [80, 100], several researchers such as Stefan Maier at Imperial College, UK [101–104], Federico Capasso at Harvard University, US [86, 89, 105], Harry Atwater at California Institute of Technology, US [106], Tei Cui at Southeast University, China [83, 90, 107, 108], Pinaki Mazumder at University of Michigan, US [81, 92, 98, 99], and Hao Yu at Nanyang Technological University, Singapore [83, 109, 110] have advanced the field of THz surface-wave engineering.

Starting from a large 2D periodic hollow metal structure proposed by Pendry et al. [80], THz SSPP morphed into 1D corrugated structures by the work of Garcia-Vidal et al. [87] (on rectangular SSPP structures) and Maier in 2006 (on cylindrical SSPP structures) Maier et al. [101]. Subsequently, a single-sided SSPP structure was designed by Martin-Cano et al. [111] and a double-sided SSPP structure was introduced by Xu et al. [112] to increase the length of signal propagation with low attenuation. In order to adapt the spoof THz surface wave technology (SSPP) to thin-film interconnect technology, surface-wave-based structures are fabricated on a silicon substrate using commercial 65 nm CMOS technology [83, 84]. Their work has now opened a greater opportunity to improve the bandwidth of the interconnect technology by an order of magnitude, while making THz surface-wave propagation on an SSPP structure compatible with commercial CMOS technology.

1.8.1 Leading Researchers on Spoof Surface Plasmon Technology

Important milestones in the research on spoof plasmonics: spoof plasmonics, a groundbreaking subfield of photonics, has emerged as a powerful alternative to traditional plasmonic systems. By mimicking the plasmonic behavior usually found in the ultraviolet regime at low frequencies through 2D hollow metal

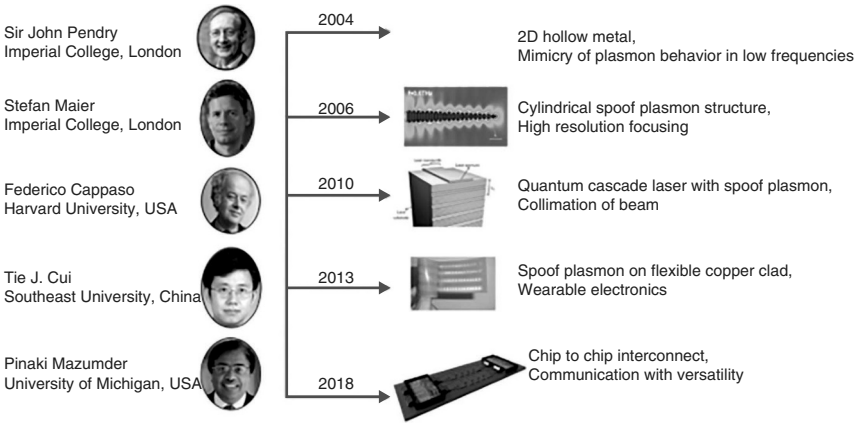


Figure 1.38 Some of the key researchers who pioneered the field of spoof plasmonics.

structures, researchers have unlocked a myriad of applications in high-resolution imaging, laser technology, wearable electronics, and chip-to-chip communication. In the following, we will mention some of the notable contributions of leading researchers in spoof plasmonics to date and explore how their work has led to the progress in this field, refer to Figure 1.38.

- **John Pendry—Imperial College London**: In 2004, John Pendry, based at Imperial College London, proposed the concept that plasmonic behavior, typically confined to the ultraviolet regime, could be emulated at low frequencies using a 2D hollow metal structure. This marked a pivotal moment in the field of spoof plasmonics, as it laid the foundation for subsequent research and development of practical applications.
- **Stefan Maier—Imperial College London**: Building upon Pendry's work, Stefan Maier, also from Imperial College London, introduced a cylindrical spoof surface plasmon structure in 2006. This innovative design showcased promising potential for high-resolution imaging and focusing. Maier's contributions solidified the viability of spoof plasmonics in real-world applications.
- **Federico Capasso—Harvard University**: In 2010, Federico Capasso, a professor at Harvard University, demonstrated an augmentation of the photoemission efficiency of quantum cascade lasers. His laboratory achieved this feat by incorporating spoof plasmons on the emitting surface of the laser. This research is expected to bridge the gap between traditional plasmonics and practical applications, opening new avenues for laser technology.
- **Tie J. Cui—Southeast University China**: Tie J. Cui, a professor at Southeast University China, and his lab demonstrated in 2013 that spoof plasmons could be supported on a flexible 2D copper clad. This demonstration

opened the prospect of wearable electronics aided by spoof plasmon technology. Cui's work showed how this technology could be integrated to potentially transform the landscape of electronic wearables.

- **Mark Stockman—Georgia State University**: Mark Stockman has played a critical role in developing the fundamental understanding of the way in which metallic nanostructures can be used to concentrate and manipulate light at the nanoscale through the excitation of surface plasmons. Research work on plasmonic waveguides, sources, and plasmonic resonator antennas was inspired by his ability to formulate basic plasmonic design principles that can capture the essence of their operation. His research on plasmon hybridization in metallic nanoparticle dimers continues to stimulate new nanophotonics developments. In these systems, hybridization can lead to a desirable redistribution of optical fields and an intense light concentration. The extreme light confinement renders the dimer's optical response also very sensitive to minute changes in gap size.
- **Nader Engheta—University of Pennsylvania**: In 2016, Nader Engheta, a professor at the University of Pennsylvania, showed theoretically and numerically, that the structural dispersion of waveguides, when augmented with the presence of thin metallic wires judiciously located at interfaces, may provide a useful platform, particularly in the microwave, millimeter-wave, and THz regimes, to demonstrate a variety of plasmonic phenomena, which in the infrared and optical wavelengths can be done with natural materials with negative or near-zero permittivity that may be lossy. However, since the proposed structures involve only positive dielectric materials and metallic walls and wires, which exhibit very little loss at these lower frequencies, it could enable achieving various plasmonic features with a low level of dissipation. This may open doors to possibilities in the design of low-loss metamaterials that offer exciting phenomena that have otherwise been limited by material loss. His laboratory also introduced the concept of "waveguide metatronics," an advanced form of optical metatronics that uses structural dispersion in waveguides to obtain the materials and structures required to construct this class of circuitry. Using numerical simulations, his lab has demonstrated that the design of a metatronic circuit can be carried out by using a waveguide filled with materials with positive permittivity. This includes the implementation of all "lumped" circuit elements and their assembly in a single circuit board. In doing so, they extended the concepts of optical metatronics to frequency ranges where there are no natural plasmonic materials available.
- **Andrea Alu—University of Texas at Austin**: Andrea Alu and his group have contributed to the theoretical development of THz metamaterial-based applications. Starting from the theoretical and methodological motivations

of their research and covering macro-scale performance-driven design of volumetric and planar metamaterials, they introduced advanced task-oriented modeling approaches. Their approaches include specific reference to their multi-scale/multi-physics customization in recent metamaterial science and engineering. In the introduction of these concepts, they illustrated the physical mechanisms and phenomena at the basis of the field manipulation capabilities enabled by metamaterials. These could benefit a wide set of recent applications in microwave, terahertz, photonics, and optics scenarios, highlighting the logical and methodological interconnections among the different domains.

- **Vladimir Shalaev—Purdue University**: Vladimir Shalaev and his group have demonstrated dynamically controllable optical properties of metasurfaces in both space and time. Gigahertz-to-terahertz-speed optical transistors for telecom (spanning the visible to mid-infrared wavelengths), beam-steering devices for smart cars and lidar, and active cloaking for defense applications are but a few examples of devices that need dynamically configurable optical properties. For metasurfaces, such dynamic control is achieved by tuning the surface's dielectric permittivity or changing its topology (geometry). Amplitude modulation—where the intensity of light is changed by controlling the metasurface's reflectance, transmittance, and absorbance—finds use in optical computation and data transfer. They demonstrated wavefront manipulation through phase control, which could allow beam-steering for lidar and autonomous vehicles, flat-tunable lenses for 3D imaging and endoscopy, and holograms for augmented and virtual reality (AR/VR).
- **Pinaki Mazumder—University of Michigan, US**: In 2019, Soumitra Joy and Pinaki Mazumder at the University of Michigan demonstrated a spoof plasmon-based interconnect system for chip-to-chip communication. This demonstration gave a proof-of-concept validation of the feasibility of employing spoof plasmons in chip-to-chip communication systems, spanning over several centimeters. This demonstration can open the possibility of on-package integration of silicon chips interfaced by high-speed short-reach interconnects to enable spoof-plasmon-based VLSI technology for high-performance computing.

The advancements made by these leading researchers in spoof plasmonics have reshaped the landscape of photonics and paved the way for a myriad of applications in various domains. From high-resolution imaging and laser technology to wearable electronics and chip-to-chip communication, spoof plasmonics has become a versatile and promising field of study. As the technology continues to evolve, future researchers will be able to build upon these contributions, further pushing the boundaries of what is possible with spoof plasmons. This book would

encourage the research communities and pave the path toward solving the long-standing problem of "last centimeter barrier" of interconnect by revealing the full potential of spoof plasmonics.

1.9 Summary of the Chapter

This chapter presents a comprehensive overview of interconnect technologies and their significance in advancing next-generation computing. It delved into various hierarchical levels of interconnect research, identifying key areas of focus. The chapter evaluated the performance and advantages of conventional interconnect technologies while exploring lesser-known research frontiers that hold promise for future developments. Additionally, the chapter introduces readers to a novel interconnect technology centered around spoof surface plasmon polaritons. This innovative technology has the potential to address longstanding interconnect challenges and complement existing solutions toward addressing the "Last Centimeter Barrier," a term indicative of the critical length of interconnect where neither electronics nor optics can provide an energy-efficient optimal solution. By combining theoretical insights with practical applications, the chapter shed light on the transformative impact of this emerging interconnect approach in shaping the future of computing.

References

1 Ammar Karkar, Terrence Mak, Kin-Fai Tong, and Alex Yakovlev. A survey of emerging interconnects for on-chip efficient multicast and broadcast in many-cores. *IEEE Circuits and Systems Magazine*, 16(1):58–72, 2016.

2 J Michel and LC Kimerling. Electronics and photonics: Convergence on a silicon platform. In *33rd European Conference and Exhibition of Optical Communication*, pages 1–2, VDE - VERLAG, 2007. https://doi.org/10.1049/ic:20070261.

3 Yutaka Urino, Takanori Shimizu, Makoto Okano, Nobuaki Hatori, Masashige Ishizaka, Tsuyoshi Yamamoto, Takeshi Baba, Takeshi Akagawa, Suguru Akiyama, Tatsuya Usuki, et al. First demonstration of high density optical interconnects integrated with lasers, optical modulators, and photodetectors on single silicon substrate. *Optics Express*, 19(26):B159–B165, 2011. https://doi.org/10.1364/OE.19.00B159.

4 M Haurylau, G Chen, H Chen, J Zhang, NA Nelson, DH Albonesi, EG Friedman, and PM Fauchet. On-chip optical interconnect roadmap: Challenges and critical directions. *IEEE Journal of Selected Topics in Quantum Electronics*, 12(6):1699–1705, 2006. https://doi.org/10.1109/JSTQE.2006.880615.

5 Gerd Keiser. Optical fiber communications. In *Wiley encyclopedia of telecommunications*. John Wiley & Sons, Inc., 2003. https://doi.org/10.1002/0471219282.eot158.

6 René-Jean Essiambre, Gerhard Kramer, Peter J Winzer, Gerard J Foschini, and Bernhard Goebel. Capacity limits of optical fiber networks. *Journal of Lightwave Technology*, 28(4):662–701, 2010. https://www.osapublishing.org/abstract.cfm?uri=jlt-28-4-662.

7 AEJ Lim, J Song, Q Fang, C Li, X Tu, N Duan, KK Chen, RPC Tern, and TY Liow. Review of silicon photonics foundry efforts. *IEEE Journal of Selected Topics in Quantum Electronics*, 20(4):405–416, 2014. https://doi.org/10.1109/JSTQE.2013.2293274.

8 Sudeep Pasricha and Nikil Dutt. *On-chip communication architectures: System on chip interconnect*. Morgan Kaufmann, 2010.

9 Wayne Wolf. Multiprocessor system-on-chip technology. *IEEE Signal Processing Magazine*, 26(6):50–54, 2009.

10 Krishna Sekar, Kanishka Lahiri, Anand Raghunathan, and Sujit Dey. Dynamically configurable bus topologies for high-performance on-chip communication. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 16(10):1413–1426, 2008.

11 Isiaka A Alimi, Romil K Patel, Oluyomi Aboderin, Abdelgader M Abdalla, Ramoni A Gbadamosi, Nelson J Muga, Armando N Pinto, and António L Teixeira. *Network-on-chip topologies: Potentials, technical challenges, recent advances and research direction*. IntechOpen, 2021.

12 Tao Li, Jie Hou, Jinli Yan, Rulin Liu, Hui Yang, and Zhigang Sun. Chiplet heterogeneous integration technology—status and challenges. *Electronics*, 9(4):670, 2020.

13 Paolo Mantovani. *Scalable system-on-chip design*. Columbia University, 2017.

14 Hristomir Yordanov and Peter Russer. Wireless inter-chip and intra-chip communication. In *2009 European Microwave Conference (EuMC)*, pages 145–148. IEEE, 2009.

15 A Vijaya Bhaskar and TG Venkatesh. Performance analysis of network-on-chip in many-core processors. *Journal of Parallel and Distributed Computing*, 147:196–208, 2021.

16 Yongrong Shi, Dengyun Shao, Wenjie Feng, Junzhi Zhang, and Ming Zhou. Silicon interposer package for MMIC heterogeneous integration based on gold/solder ball flip-chip technique. *IEEE Transactions on Components, Packaging and Manufacturing Technology*, 9(8):1659–1662, 2019.

17 Jacob Murray, Paul Wettin, Partha Pratim Pande, and Behrooz Shirazi. *Sustainable wireless network-on-chip architectures*. Morgan Kaufmann, 2016.

18 Md Shahriar Shamim, Jagan Muralidharan, and Amlan Ganguly. An interconnection architecture for seamless inter and intra-chip communication

using wireless links. In *Proceedings of the 9th International Symposium on Networks-on-Chip*, pages 1–8. Association for Computing Machinery, 2015.

19 Sanjukta Mandal, Sujit Kumar Mandal, and Ashis Kumar Mal. On-chip antennas using standard CMOS technology: A brief overview. In *2017 International Conference on Innovations in Electronics, Signal Processing and Communication (IESC)*, pages 74–78. IEEE, 2017.

20 Mohammad Alibakhshikenari, Esraa Mousa Ali, Mohammad Soruri, Mariana Dalarsson, Mohammad Naser-Moghadasi, Bal S Virdee, Caslav Stefanovic, Anna Pietrenko-Dabrowska, Slawomir Koziel, Stanislaw Szczepanski, et al. A comprehensive survey on antennas on-chip based on metamaterial, metasurface, and substrate integrated waveguide principles for millimeter-waves and terahertz integrated circuits and systems. *IEEE Access*, 10:3668–3692, 2022.

21 Jiaxin Chen, Jinzhao Wang, Jiewen Li, Yong Yao, Yunxu Sun, Jiajun Tian, Yi Zou, Xiangjie Zhao, and Xiaochuan Xu. Subwavelength structure enabled ultra-long waveguide grating antenna. *Optics Express*, 29(10):15133–15144, 2021.

22 Francesco Monticone, Christos Argyropoulos, and Andrea Alu. Optical antennas: Controlling electromagnetic scattering, radiation, and emission at the nanoscale. *IEEE Antennas and Propagation Magazine*, 59(6):43–61, 2017.

23 Arindam Dasgupta, Marie-Maxime Mennemanteuil, Mickaël Buret, Nicolas Cazier, Gérard Colas-des Francs, and Alexandre Bouhelier. Optical wireless link between a nanoscale antenna and a transducing rectenna. *Nature Communications*, 9(1):1992, 2018.

24 M Fukuda, Y Tonooka, T Inoue, and M Ota. Feasibility of plasmonic circuits for on-chip interconnects. *Solid-State Electronics*, 156:33–40, 2019.

25 Shaloo Rakheja and Vachan Kumar. Comparison of electrical, optical and plasmonic on-chip interconnects based on delay and energy considerations. In *Thirteenth International Symposium on Quality Electronic Design (ISQED)*, pages 732–739. IEEE, 2012.

26 Isaac Suárez, Albert Ferrando, Jose Marques-Hueso, Antonio Díez, Rafael Abargues, Pedro J Rodríguez-Cantó, and Juan P Martínez-Pastor. Propagation length enhancement of surface plasmon polaritons in gold nano-/micro-waveguides by the interference with photonic modes in the surrounding active dielectrics. *Nanophotonics*, 6(5):1109–1120, 2017.

27 Yihao Yang, Yuichiro Yamagami, Xiongbin Yu, Prakash Pitchappa, Julian Webber, Baile Zhang, Masayuki Fujita, Tadao Nagatsuma, and Ranjan Singh. Terahertz topological photonics for on-chip communication. *Nature Photonics*, 14(7):446–451, 2020.

28 Yiwen Sun, Zhijie Mei, Xuejiao Xu, Qingxuan Xie, Shuting Fan, Zhengfang Qian, and Xudong Liu. The perspective of topological photonics for on-chip terahertz modulation and sensing. *APL Photonics*, 8(11): 110901, 2023.

29 Xiang Yi, Cheng Wang, Zhi Hu, Jack W Holloway, Muhammad Ibrahim Wasiq Khan, Mohamed I Ibrahim, Mina Kim, Georgios C Dogiamis, Bradford Perkins, Mehmet Kaynak, et al. Emerging terahertz integrated systems in silicon. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 68(9):3537–3550, 2021.

30 Mark Horowitz, Chih-Kong Ken Yang, and Stefanos Sidiropoulos. High-speed electrical signaling: Overview and limitations. *IEEE Micro*, 18(1):12–24, 1998.

31 Koji Yamada, Yutaka Urino, Takahiro Nakamura, and Yasuhiko Arakawa. Integrated silicon-based optical interconnect for fast, compact, energy-efficient electronic circuit systems. *NTT Technical Review*, 11(2):41–47, 2013. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201302fa7_s.html.

32 Yi-Ting Tsou, I-Huai Wang, and Sung-Mao Wu. Discussion of the signal transmission crosstalk. In *2018 IEEE 20th Electronics Packaging Technology Conference (EPTC)*, pages 785–788. IEEE, 2018.

33 Actel Application Note. Simultaneous switching noise and signal integrity. *Actel Corporation*, 2003.

34 Chunjie Duan, Brock J LaMeres, and Sunil P Khatri. *On and off-chip crosstalk avoidance in VLSI design*. Springer, 2010.

35 Bryan K Casper, Matthew Haycock, and Randy Mooney. An accurate and efficient analysis method for multi-gb/s chip-to-chip signaling schemes. In *2002 Symposium on VLSI Circuits. Digest of Technical Papers (Cat. No. 02CH37302)*, pages 54–57. IEEE, 2002.

36 Hyungsoo Kim, Youchul Jeong, Jongbae Park, Seok Kyu-Lee, Jong Kuk-Hong, Youngsoo Hong, and Joungho Kim. Significant reduction of power/ground inductive impedance and simultaneous switching noise by using embedded film capacitor. In *Electrical performance of electrical packaging (IEEE Cat. No. 03TH8710)*, pages 129–132. IEEE, 2003.

37 David AB Miller. Attojoule optoelectronics for low-energy information processing and communications. *Journal of Lightwave Technology*, 35(3):346–396, 2017.

38 Stephen W Keckler, William J Dally, Brucek Khailany, Michael Garland, and David Glasco. Gpus and the future of parallel computing. *IEEE Micro*, 31(5):7–17, 2011.

39 Keren Bergman, John Shalf, and Tom Hausken. Optical interconnects and extreme computing. *Optics and Photonics News*, 27(4):32–39, 2016.

40 Sébastien Rumley, Robert P Polster, Simon D Hammond, Arun F Rodrigues, and Keren Bergman. End-to-end modeling and optimization of power consumption in HPC interconnects. In *2016 45th International Conference on Parallel Processing Workshops (ICPPW)*, pages 133–140. IEEE, 2016.

41 Krishna C Saraswat and Farrokh Mohammadi. Effect of scaling of interconnections on the time delay of VLSI circuits. *IEEE Transactions on Electron Devices*, 29(4):645–650, 1982.

42 John D Owens, William J Dally, Ron Ho, DN Jayasimha, Stephen W Keckler, and Li-Shiuan Peh. Research challenges for on-chip interconnection networks. *IEEE Micro*, 27(5):96–108, 2007.

43 V Cisco. *The zettabyte era: trends and analysis*. CISCO, 2017.

44 Wolfgang Arden. Future semiconductor material requirements and innovations as projected in the ITRS 2005 roadmap. *Materials Science and Engineering: B*, 134(2–3):104–108, 2006.

45 Michel Brillouet. Emerging technologies on silicon. In *IEDM Technical Digest. IEEE International Electron Devices Meeting, 2004*, pages 17–24. IEEE, 2004.

46 Robert Chau, Suman Datta, and Amlan Majumdar. Opportunities and challenges of III-V nanoelectronics for future high-speed, low-power logic applications. In *IEEE Compound Semiconductor Integrated Circuit Symposium, 2005. CSIC'05*, pages 4. IEEE, 2005.

47 Gilbert Dewey, Mantu K Hudait, Kangho Lee, Ravi Pillarisetty, Willy Rachmady, Marko Radosavljevic, Titash Rakshit, and Robert Chau. Carrier transport in high-mobility III–V quantum-well transistors and performance impact for high-speed low-power logic applications. *IEEE Electron Device Letters*, 29(10):1094–1097, 2008.

48 Suman Datta, G Dewey, JM Fastenau, MK Hudait, D Loubychev, WK Liu, M Radosavljevic, W Rachmady, and R Chau. Ultrahigh-speed 0.5 v supply voltage $in_{0.7}ga_{0.3}as$ quantum-well transistors on silicon substrate. *IEEE Electron Device Letters*, 28(8):685–687, 2007.

49 Ravindiran Munusami and Shankar Prabhakar. Group III–V semiconductor high electron mobility transistor on Si substrate. In *Different Types of Field-Effect Transistors: Theory and Applications*, page 89, Intechopen, 2017.

50 Jun Wu. *Vertical III-V/High-k nanowire MOS capacitors and transistors*. Lund University, 2016.

51 Yung Hsu, Chun-Yen Chuang, Xinru Wu, Guan-Hong Chen, Chin-Wei Hsu, Yuan-Chia Chang, Chi-Wai Chow, Jyehong Chen, Yin-Chieh Lai, Chien-Hung Yeh, et al. 2.6 tbit/s on-chip optical interconnect supporting mode-division-multiplexing and pam-4 signal. *IEEE Photonics Technology Letters*, 30(11):1052–1055, 2018.

52 David Thomson, Aaron Zilkie, John E Bowers, Tin Komljenovic, Graham T Reed, Laurent Vivien, Delphine Marris-Morini, Eric Cassan, Léopold Virot, Jean-Marc Fédéli, et al. Roadmap on silicon photonics. *Journal of Optics*, 18(7):073003, 2016.

53 Siqi Yan, Jeremy Adcock, and Yunhong Ding. Graphene on silicon photonics: Light modulation and detection for cutting-edge communication technologies. *Applied Sciences*, 12(1):313, 2021.

54 Brian R Bennett and Richard A Soref. Electrorefraction and electroabsorption in InP, GaAs, GaSb, InAs, and InSb. *IEEE Journal of Quantum Electronics*, 23(12):2159–2166, 1987.

55 Chia-Chi Wang, Marc Currie, Sotiris Alexandrou, and Thomas Y Hsiang. Ultrafast, all-silicon light modulator. *Optics Letters*, 19(18):1453–1455, 1994.

56 Michel Bruel. Separation of silicon wafers by the smart-cut method. *Materials Research Innovations*, 3(1):9–13, 1999.

57 Shawn Yohanes Siew, Bo Li, Feng Gao, Hai Yang Zheng, Wenle Zhang, Pengfei Guo, Shawn Wu Xie, Apu Song, Bin Dong, Lian Wee Luo, et al. Review of silicon photonics technology and platform development. *Journal of Lightwave Technology*, 39(13):4374–4389, 2021.

58 Stephanie Law, Viktor Podolskiy, and Daniel Wasserman. Towards nano-scale photonics with micro-scale photons: The opportunities and challenges of mid-infrared plasmonics. *Nanophotonics*, 2(2):103–130, 2013.

59 Guanyu Chen, Yu Yu, Yang Shi, Nanxi Li, Wei Luo, Lin Cao, Aaron J Danner, Ai-Qun Liu, and Xinliang Zhang. High-speed photodetectors on silicon photonics platform for optical interconnect. *Laser & Photonics Reviews*, 16(12):2200117, 2022.

60 Pablo Ginel-Moreno, Alejandro Sánchez-Postigo, José de Oliva-Rubio, Abdelfettah Hadij-ElHouati, Winnie N Ye, J Gonzalo Wangüemert-Pérez, Íñigo Molina-Fernández, Jens H Schmid, Pavel Cheben, and Alejandro Ortega-Moñux. Millimeter-long metamaterial surface-emitting antenna in the silicon photonics platform. *Optics Letters*, 46(15):3733–3736, 2021.

61 Yachao Zhang, Shiteng Liu, Jinbang Ma, Yixin Yao, Baiqi Wang, Shengrui Xu, Bin Hou, Zhihong Liu, Jincheng Zhang, and Yue Hao. More than 60% RF loss reduction and improved crystal quality of GaN-on-SI achieved by in-situ doping tert-butylphosphorus. *Journal of Crystal Growth*, 625:127443, 2024.

62 Lei Zhu, Fan-Yi Meng, Liang Dong, Jia-Hui Fu, and Qun Wu. Low-loss magnetic metamaterial at THz frequencies by suppressing radiation losses. *IEEE Transactions on Terahertz Science and Technology*, 3(6):805–811, 2013.

63 Sergi Abadal, Chong Han, and Josep Miquel Jornet. Wave propagation and channel modeling in chip-scale wireless communications: A survey from millimeter-wave to terahertz and optics. *IEEE Access*, 8:278–293, 2019.

64 Xavier Timoneda, Sergi Abadal, Albert Cabellos-Aparicio, Dionysios Manessis, Jin Zhou, Antonio Franques, Josep Torrellas, and Eduard Alarcón. Millimeter-wave propagation within a computer chip package. In *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 1–5. IEEE, 2018.

65 Kyung-Hoae Koo, Hoyeol Cho, Pawan Kapur, and Krishna C Saraswat. Performance comparisons between carbon nanotubes, optical, and Cu for future high-performance on-chip interconnect applications. *IEEE Transactions on Electron Devices*, 54(12):3206–3215, 2007.

66 Phaedon Avouris, Joerg Appenzeller, Richard Martel, and Shalom J Wind. Carbon nanotube electronics. *Proceedings of the IEEE*, 91(11):1772–1784, 2003.

67 Chris Bailey, Stoyan Stoyanov, Hua Lu, Tim Tilford, Chunyan Yin, and Nadia Strusevich. Modelling technologies and applications. *Nanopackaging: Nanotechnologies and electronics packaging*, pages 45–82. Cham, Springer, 2018.

68 Christophe Brun, Chin Chong Yap, Dunlin Tan, Stéphane Bila, Sebastien Pacchini, Dominique Baillargeat, and Beng Kang Tay. Flip chip based on carbon nanotube–carbon nanotube interconnected bumps for high-frequency applications. *IEEE Transactions on Nanotechnology*, 12(4):609–615, 2013.

69 Carlo Forestiere, Antonio Maffucci, Sergey A Maksimenko, Giovanni Miano, and Gregory Ya Slepyan. Transmission-line model for multiwall carbon nanotubes with intershell tunneling. *IEEE Transactions on Nanotechnology*, 11(3):554–564, 2012.

70 Hoyeol Cho, Pawan Kapur, and Krishna C Saraswat. Power comparison between high-speed electrical and optical interconnects for interchip communication. *Journal of Lightwave Technology*, 22(9):2021, 2004. https://www.osapublishing.org/jlt/abstract.cfm?uri=jlt-22-9-2021.

71 Brandon Heller, Srinivasan Seetharaman, Priya Mahadevan, Yiannis Yiakoumis, Puneet Sharma, Sujata Banerjee, and Nick McKeown. Elastictree: Saving energy in data center networks. In *NSDI*, volume 10, pages 249–264, USENIX, 2010.

72 Theophilus Benson, Ashok Anand, Aditya Akella, and Ming Zhang. Understanding data center traffic characteristics. *ACM SIGCOMM Computer Communication Review*, 40(1):92–99, 2010.

73 QJ Gu. THz interconnect: The last centimeter communication. *IEEE Communications Magazine*, 53(4):206–215, 2015. https://doi.org/10.1109/MCOM.2015.7081096.

74 Man Lung Mui, Kaustav Banerjee, and Amit Mehrotra. A global interconnect optimization scheme for nanometer scale VLSI with implications for latency, bandwidth, and power dissipation. *IEEE Transactions on Electron Devices*, 51(2):195–203, 2004.

75 JM Chamberlain. Where optics meets electronics: Recent progress in decreasing the terahertz gap. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 362(1815):199–213, 2004. https://doi.org/10.1098/rsta.2003.1312.

76 Carlo Sirtori. Applied physics: Bridge for the terahertz gap. *Nature*, 417(6885):132–133, 2002. https://doi.org/10.1038/417132b.

77 R Han and E Afshari. Filling the terahertz gap with sand: High-power terahertz radiators in silicon. In *2015 IEEE Bipolar/BiCMOS Circuits and Technology Meeting – BCTM*, pages 172–177, IEEE, 2015. https://doi.org/10.1109/BCTM.2015.7340574.

78 Daryoosh Saeedkia. *Handbook of Terahertz Technology for Imaging, Sensing and Communications*. Elsevier, 2013.

79 Thomas Kleine-Ostmann and Tadao Nagatsuma. A review on terahertz communications research. *Journal of Infrared, Millimeter, and Terahertz Waves*, 32(2):143–171, 2011.

80 JB Pendry, L Martín-Moreno, and FJ Garcia-Vidal. Mimicking surface plasmons with structured surfaces. *Science*, 305(5685):847–848, 2004. https://doi.org/10.1126/science.1098999.

81 Mikhail Erementchouk, Soumitra Roy Joy, and Pinaki Mazumder. Electrodynamics of spoof plasmons in periodically corrugated waveguides. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 472(2195):20160616, 2016.

82 Rashid Zia, Jon A Schuller, Anu Chandran, and Mark L Brongersma. Plasmonics: The next chip-scale technology. *Materials Today*, 9(7-8):20–27, 2006.

83 Yuan Liang, Hao Yu, Jincai Wen, Anak Agung Alit Apriyana, Nan Li, Yu Luo, and Lingling Sun. On-chip sub-terahertz surface plasmon polariton transmission lines with mode converter in CMOS. *Scientific Reports*, 6:30063, 2016. https://doi.org/10.1038/srep30063.

84 SR Joy, M Erementchouk, H Yu, and P Mazumder. Spoof plasmon interconnects—communications beyond RC limit. *IEEE Transactions on Communications*, 67(1):599–610, 2019. https://doi.org/10.1109/TCOMM.2018.2874242.

85 Anastasia Rusina, Maxim Durach, and Mark I Stockman. Theory of spoof plasmons in real metals. In *SPIE NanoScience+ Engineering*, pages 77572R–77572R. International Society for Optics and Photonics, 2010.

86 Mikhail A Kats, David Woolf, Romain Blanchard, Nanfang Yu, and Federico Capasso. Spoof plasmon analogue of metal-insulator-metal waveguides. *Optics Express*, 19(16):14860–14870, 2011. https://doi.org/10.1364/OE.19.014860.

87 FJ Garcia-Vidal, L Martin-Moreno, and JB Pendry. Surfaces with holes in them: New plasmonic metamaterials. *Journal of Optics A: Pure and Applied Optics*, 7(2):S97, 2005.

88 Linfang Shen, Xudong Chen, and Tzong-Jer Yang. Terahertz surface plasmon polaritons on periodically corrugated metal surfaces. *Optics Express*, 16(5):3326–3333, 2008. https://doi.org/10.1364/OE.16.003326.

89 Nanfang Yu, Qi Jie Wang, Mikhail A Kats, Jonathan A Fan, Suraj P Khanna, Lianhe Li, A Giles Davies, Edmund H Linfield, and Federico Capasso. Designer spoof surface plasmon structures collimate terahertz laser beams. *Nature Materials*, 9(9):730–735, 2010. https://doi.org/10.1038/nmat2822.

90 Hui Feng Ma, Xiaopeng Shen, Qiang Cheng, Wei Xiang Jiang, and Tie Jun Cui. Broadband and high-efficiency conversion from guided waves to spoof surface plasmon polaritons. *Laser & Photonics Reviews*, 8(1):146–151, 2014. https://doi.org/10.1002/lpor.201300118.

91 Zhao Xu, Kyungjun Song, and Pinaki Mazumder. Analysis of doubly corrugated spoof surface plasmon polariton (DC-SSPP) structure with sub-wavelength transmission at THz frequencies. *IEEE Transactions on Terahertz Science and Technology*, 2(3):345–354, 2012b.

92 K Song and P Mazumder. Active terahertz spoof surface plasmon polariton switch comprising the perfect conductor metamaterial. *IEEE Transactions on Electron Devices*, 56(11):2792–2799, 2009. https://doi.org/10.1109/TED.2009.2030838.

93 Z Xu and P Mazumder. Bio-sensing by Mach Zehnder interferometer comprising doubly-corrugated spoofed surface plasmon polariton (DC-SSPP) waveguide. *IEEE Transactions on Terahertz Science and Technology*, 2(4):460–466, 2012. https://doi.org/10.1109/TTHZ.2012.2202811.

94 K Song and P Mazumder. Dynamic terahertz spoof surface plasmon–polariton switch based on resonance and absorption. *IEEE Transactions on Electron Devices*, 58(7):2172–2176, 2011. https://doi.org/10.1109/TED.2011.2135370.

95 K Song and P Mazumder. Design of highly selective metamaterials for sensing platforms. *IEEE Sensors Journal*, 13(9):3377–3385, 2013. https://doi.org/10.1109/JSEN.2013.2260143.

96 N Zheng, M Aghadjani, K Song, and P Mazumder. Metamaterial sensor platforms for Terahertz DNA sensing. In *2013 13th IEEE International Conference on Nanotechnology (IEEE-NANO 2013)*, pages 315–320, 2013. https://doi.org/10.1109/NANO.2013.6720831.

97 Kyungjun Song and Pinaki Mazumder. Nonlinear spoof surface plasmon polariton phenomena based on conductor metamaterials. *Photonics and Nanostructures - Fundamentals and Applications*, 10(4):674–679, 2012. https://doi.org/10.1016/j.photonics.2012.06.008.

98 M Aghadjani, M Erementchouk, and P Mazumder. Spoof surface plasmon polariton beam splitter. *IEEE Transactions on Terahertz Science and Technology*, 6(6):832–839, 2016. https://doi.org/10.1109/TTHZ.2016.2599289.

99 Soumitra Roy Joy, Mikhail Erementchouk, and Pinaki Mazumder. Spoof surface plasmon resonant tunneling mode with high quality and purcell factors. *Physical Review B*, 95(7):075435, 2017. https://doi.org/10.1103/PhysRevB.95.075435.

100 Anders Pors, Esteban Moreno, Luis Martin-Moreno, John B Pendry, and Francisco J Garcia-Vidal. Localized spoof plasmons arise while texturing closed surfaces. *Physical Review Letters*, 108(22):223905, 2012.

101 Stefan A Maier, Steve R Andrews, L Martín-Moreno, and FJ García-Vidal. Terahertz surface plasmon-polariton propagation and focusing on periodically corrugated metal wires. *Physical Review Letters*, 97(17):176805, 2006. https://doi.org/10.1103/PhysRevLett.97.176805.

102 Zhen Liao, Yu Luo, Antonio I Fernández-Domínguez, Xiaopeng Shen, Stefan A Maier, and Tie Jun Cui. High-order localized spoof surface plasmon resonances and experimental verifications. *Scientific Reports*, 5:9590, 2015.

103 AI Fernandez-Dominguez, L Martin-Moreno, FJ Garcia-Vidal, Steven Andrews, and SA Maier. Spoof surface plasmon polariton modes propagating along periodically corrugated wires. *IEEE Journal of Selected Topics in Quantum Electronics*, 14(6):1515–1521, 2008.

104 Jianqiang Gu, Ranjan Singh, Xiaojun Liu, Xueqian Zhang, Yingfang Ma, Shuang Zhang, Stefan A Maier, Zhen Tian, Abul K Azad, Hou-Tong Chen, et al. Active control of electromagnetically induced transparency analogue in terahertz metamaterials. *Nature Communications*, 3(1):1–6, 2012.

105 David Woolf, Mikhail A Kats, and Federico Capasso. Spoof surface plasmon waveguide forces. *Optics Letters*, 39(3):517–520, 2014. https://doi.org/10.1364/OL.39.000517.

106 Victor W Brar, Min Seok Jang, Michelle Sherrott, Josue J Lopez, and Harry A Atwater. Highly confined tunable mid-infrared plasmonics in graphene nanoresonators. *Nano Letters*, 13(6):2541–2547, 2013.

107 Xiaopeng Shen, Tie Jun Cui, Diego Martin-Cano, and Francisco J Garcia-Vidal. Conformal surface plasmons propagating on ultrathin and flexible films. *Proceedings of the National Academy of Sciences of the United States of America*, 110(1):40–45, 2013.

108 Xi Gao and Tie Jun Cui. Spoof surface plasmon polaritons supported by ultrathin corrugated metal strip and their applications. *Nanotechnology Reviews*, 4(3):239–258, 2015.

109 Yuan Liang, H Yu, Junfeng Zhao, Wei Yang, and Yuangang Wang. An energy efficient and low cross-talk CMOS sub-THz I/O with surface-wave modulator and interconnect. In *2015 IEEE/ACM International Symposium on Low Power Electronics and Design (ISLPED)*, pages 110–115, 2015a. https://doi.org/10.1109/ISLPED.2015.7273499.

110 Yuan Liang, Hao Yu, Hao Chi Zhang, Chang Yang, and Tie Jun Cui. On-chip sub-terahertz surface plasmon polariton transmission lines in CMOS. *Scientific Reports*, 5:14853, 2015b. https://doi.org/10.1038/srep14853.

111 D Martin-Cano, ML Nesterov, AI Fernandez-Dominguez, FJ Garcia-Vidal, L Martin-Moreno, and Esteban Moreno. plasmons for subwavelength terahertz circuitry. *Optics Express*, 18(2):754–764, 2010. https://doi.org/10.1364/OE.18.000754.

112 Z Xu, K Song, and P Mazumder. Analysis of doubly corrugated spoof surface plasmon polariton (DC-SSPP) structure with sub-wavelength transmission at THz frequencies. *IEEE Transactions on Terahertz Science and Technology*, 2(3):345–354, 2012a. https://doi.org/10.1109/TTHZ.2012.2188558.