



---

**Dynamically Adaptive Hybrid Nanoplasmonic Networks on Chips (NoCs)**

**Tarek El-Ghazawi**  
**THE GEORGE WASHINGTON UNIVERSITY**

---

**03/22/2019**  
**Final Report**

**DISTRIBUTION A: Distribution approved for public release.**

**Air Force Research Laboratory**  
**AF Office Of Scientific Research (AFOSR)/ RTB1**  
**Arlington, Virginia 22203**  
**Air Force Materiel Command**

DISTRIBUTION A: Distribution approved for public release.

**REPORT DOCUMENTATION PAGE***Form Approved  
OMB No. 0704-0188*

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

**PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

<b>1. REPORT DATE (DD-MM-YYYY)</b>		<b>2. REPORT TYPE</b>		<b>3. DATES COVERED (From - To)</b>	
<b>4. TITLE AND SUBTITLE</b>				<b>5a. CONTRACT NUMBER</b>	
				<b>5b. GRANT NUMBER</b>	
				<b>5c. PROGRAM ELEMENT NUMBER</b>	
<b>6. AUTHOR(S)</b>				<b>5d. PROJECT NUMBER</b>	
				<b>5e. TASK NUMBER</b>	
				<b>5f. WORK UNIT NUMBER</b>	
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b>				<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>	
<b>9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b>				<b>10. SPONSOR/MONITOR'S ACRONYM(S)</b>	
				<b>11. SPONSOR/MONITOR'S REPORT NUMBER(S)</b>	
<b>12. DISTRIBUTION/AVAILABILITY STATEMENT</b>					
<b>13. SUPPLEMENTARY NOTES</b>					
<b>14. ABSTRACT</b>					
<b>15. SUBJECT TERMS</b>					
<b>16. SECURITY CLASSIFICATION OF:</b>			<b>17. LIMITATION OF ABSTRACT</b>	<b>18. NUMBER OF PAGES</b>	<b>19a. NAME OF RESPONSIBLE PERSON</b>
<b>a. REPORT</b>	<b>b. ABSTRACT</b>	<b>c. THIS PAGE</b>			<b>19b. TELEPHONE NUMBER (Include area code)</b>

## INSTRUCTIONS FOR COMPLETING SF 298

**1. REPORT DATE.** Full publication date, including day, month, if available. Must cite at least the year and be Year 2000 compliant, e.g. 30-06-1998; xx-06-1998; xx-xx-1998.

**2. REPORT TYPE.** State the type of report, such as final, technical, interim, memorandum, master's thesis, progress, quarterly, research, special, group study, etc.

**3. DATE COVERED.** Indicate the time during which the work was performed and the report was written, e.g., Jun 1997 - Jun 1998; 1-10 Jun 1996; May - Nov 1998; Nov 1998.

**4. TITLE.** Enter title and subtitle with volume number and part number, if applicable. On classified documents, enter the title classification in parentheses.

**5a. CONTRACT NUMBER.** Enter all contract numbers as they appear in the report, e.g. F33315-86-C-5169.

**5b. GRANT NUMBER.** Enter all grant numbers as they appear in the report. e.g. AFOSR-82-1234.

**5c. PROGRAM ELEMENT NUMBER.** Enter all program element numbers as they appear in the report, e.g. 61101A.

**5e. TASK NUMBER.** Enter all task numbers as they appear in the report, e.g. 05; RF0330201; T4112.

**5f. WORK UNIT NUMBER.** Enter all work unit numbers as they appear in the report, e.g. 001; AFAPL30480105.

**6. AUTHOR(S).** Enter name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. The form of entry is the last name, first name, middle initial, and additional qualifiers separated by commas, e.g. Smith, Richard, J, Jr.

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES).** Self-explanatory.

**8. PERFORMING ORGANIZATION REPORT NUMBER.** Enter all unique alphanumeric report numbers assigned by the performing organization, e.g. BRL-1234; AFWL-TR-85-4017-Vol-21-PT-2.

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES).** Enter the name and address of the organization(s) financially responsible for and monitoring the work.

**10. SPONSOR/MONITOR'S ACRONYM(S).** Enter, if available, e.g. BRL, ARDEC, NADC.

**11. SPONSOR/MONITOR'S REPORT NUMBER(S).** Enter report number as assigned by the sponsoring/monitoring agency, if available, e.g. BRL-TR-829; -215.

**12. DISTRIBUTION/AVAILABILITY STATEMENT.** Use agency-mandated availability statements to indicate the public availability or distribution limitations of the report. If additional limitations/ restrictions or special markings are indicated, follow agency authorization procedures, e.g. RD/FRD, PROPIN, ITAR, etc. Include copyright information.

**13. SUPPLEMENTARY NOTES.** Enter information not included elsewhere such as: prepared in cooperation with; translation of; report supersedes; old edition number, etc.

**14. ABSTRACT.** A brief (approximately 200 words) factual summary of the most significant information.

**15. SUBJECT TERMS.** Key words or phrases identifying major concepts in the report.

**16. SECURITY CLASSIFICATION.** Enter security classification in accordance with security classification regulations, e.g. U, C, S, etc. If this form contains classified information, stamp classification level on the top and bottom of this page.

**17. LIMITATION OF ABSTRACT.** This block must be completed to assign a distribution limitation to the abstract. Enter UU (Unclassified Unlimited) or SAR (Same as Report). An entry in this block is necessary if the abstract is to be limited.

## Final Report

El-Ghazawi (P.I.), Sorger (co-P.I.), Narayana (co-P.I.) George Washington University

Program Manager: **Dr. Gernot Pomrenke (AFOSR) and Dr. Erik Blasch**

### 1. Abstract

This project aims to explore the innovations required to build the communication infrastructure for next generation many-core computing chips, by hybridizing nanoplasmonic and conventional electronic technologies, with the objective of meeting the increasing bandwidth demands, achieving the desired low latency and power requirements, through the ability to adapt using DDDAS concepts to dynamically meet varying system requirements. Towards this end, our research is centered on the following three tasks:

- Design and benchmarking of novel, hybrid nanoscale plasmonic-photonic links that can be used as the building blocks for augmenting existing networks on chip (NoCs);
- Design and simulation of hybrid opto-electric NoCs that can configure the photonic links based on application traffic characteristics;
- Development of an adaptive network based on DDDAS-on-chip to address run-time demands and variations.

During the first year of the project we have developed the link level technology (HyPPI - Hybrid Plasmonics Photonic Interconnects) and demonstrated their benefits through our new unified figure-of-merit (FOM) termed CLEAR; we designed and simulated MorphoNoC – an electronic mesh augmented with configurable nanophotonic links; and we are studying the required router designs, performance monitoring mechanisms, and feedback strategies necessary for achieving DDDAS on-chip.

In the second year, we studied the fundamental scaling laws in nanophotonics to reveal the fundamental relations between the size, performance, and the device related cavities; we also investigated the possible nanophotonic devices enabled by HyPPI for on chip applications; we explored the design space by looking at how augmentation of a base mesh topology with HyPPI links of different lengths would affect the performance. These augmentations were implemented at design time. Moreover, we started designing a runtime reconfigurable system that can implement DDDAS on chip.

In the third year, we finalized the optical router and transceiver designs to aid DDDAS implementation, and examined techniques for dynamic data-driven adaptation. Furthermore, we verified the design using a FPGA platform.



## 2. Research Results

Here we discuss in further detail the research carried out and results obtained yearly.

### 2.1. Year 1

#### 2.1.1. Link-level Technology Investigations - Hybrid Photon-Plasmon Interconnects (HyPPI)

Moore's Law for traditional electric integrated circuits is coming to an end due to challenges based on physics, process technology, and economics. Among those challenges is the fact that the bandwidth-per-compute is unable to keep up with increasing demands, whereas the total energy needed for data movement keeps rising. As such, innovations in the link technology used for Network-on-Chips is essential to reduce the energy-per-bit. With this aim, we conducted first fundamental physics and device-based benchmark of multi-technology link options to include electronics, photonics, plasmonics, as well as hybrid photon-plasmon interconnects (HyPPI). HyPPI makes use of synergistic properties of materials and devices; that is all active optoelectronic devices are plasmon-based whereas all passive photon-routing building blocks constitute of low-loss photonic solutions such as Silicon photonics. Moreover, two modulation strategies are demonstrated in this hybridization interconnect technology termed 'HyPPI- extrinsic' which uses an external electro-optic modulator to modulate the light and 'HyPPI- intrinsic' which modulate the light source directly with electronic driver (**Fig. 1**). Hybridizing these two technologies uses the best of both worlds; polaritonic (matter-like) optical modes for efficient active devices (laser, modulator, detector, switch), and low-loss and economically viable Silicon platform such as SOI. Furthermore, such hybridization enables photonic power- saving solutions such as gating the source for direct modulation as supposed to externally via an EO modulator. Our analysis shows that such hybridization will overcome the shortcomings of both pure photonic and plasmonic links. Furthermore, it shows superiority in a variety of performance parameters such as point-to-point latency, energy efficiency, throughput, energy delay product, crosstalk coupling length, and bit flow density, which is a new metric that we defined to reveal the tradeoff

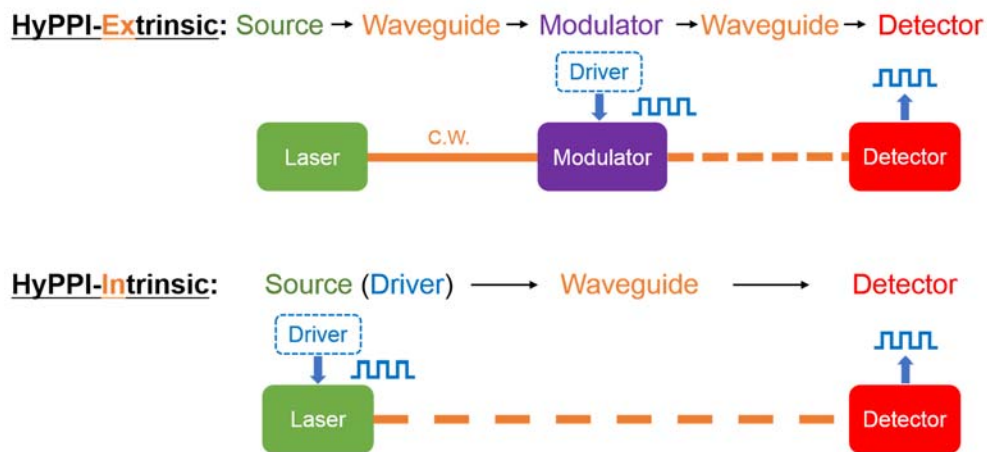


Figure 1. Schematic design of HyPPI with two modulation strategies.

between the footprint and performance. As such HyPPI demonstrates significantly superior performance compared with other links.

A comparison of point-to-point links based on electronics, photonics, plasmonics, as well as our proposed HyPPIs (**Fig. 2**). HyPPI-extrinsic refers to links wherein modulation is carried out by using a plasmon modulator. HyPPI-intrinsic achieves even further efficiency by directly modulating the plasmon laser. We show that an electrical link is able to provide extremely low delay and energy efficiency over micron-scale distances due to its length dependent RC characteristics but loses its advantages for longer link lengths beyond 10's of micrometers. The latency for the photonic and HyPPI links is dominated by both the various active devices and the waveguide propagation timescales in short link lengths and by the waveguide only for long propagation distances. The plasmonic link latency and energy inefficiency both grow due to

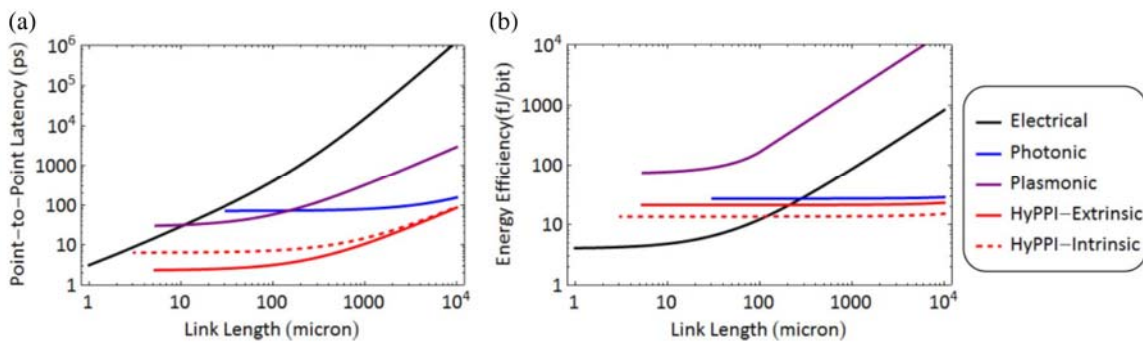


Figure 2. Comparison of latency and energy efficiency of point-to-point links for different technology options, including our proposed HyPPIs [2].

repetition every 100  $\mu\text{m}$ , as plasmonic suffers from large ohmic losses.

HyPPI links use the classical photonic SOI waveguides, and show good cross-talk lengths, extending to the chip-scale (centimeters). The spacing between waveguides as well as the cross-talk and propagation lengths impose important constraints on the total number of bits that can be carried over a given chip area. Towards gaining more insights for on-chip applications of the various interconnect technology options, we define Bit Flow Density (BFD) as the number of bits transmitted through a certain chip width (cross-section) to reach a specific required communication length, which is highly related to the size of each device and spacing. The bit flow density for the four different kinds of interconnects reveals waveguide geometry-based performance regions driven by both crosstalk and data throughput shown as contour plots wherein higher (or lower) bit flow density is indicated by increasing red (or blue) color (**Fig. 3**). Here HyPPI demonstrates over 10-1000x higher bit flow density compared with both traditional photonic interconnects and plasmonic interconnects. Between the two HyPPI options, intrinsic provides even more superior performance due to the lower area and power overhead required from the omitted modulator. HyPPIs combines the best of two worlds - small energy consumption from plasmonic devices and long propagation from photonic waveguides.

In summary, key highlights from our link investigations are as follows; HyPPI point-to-point links show significant improvements in performance relative to that of pure photonic or pure plasmonic links in P2P latency ( $< 100 \text{ ps/cm}$ ), energy ( $\sim 20 \text{ fJ/bit}$ ), and combined metrics such as Energy Delay Product and Energy Delay Squared Product. Plasmonic links are limited by their high optical

losses whereas; photonic links are bound by high power consumption due to overheads in modulation. Moreover, the sufficient crosstalk length (over 1 cm chip-size) of HyPPI enables dense integration schemes leading to high bit flow density (0.1~0.5 Gbps/ $\mu\text{m}^2$ ), and higher area efficiency. Such high-performance results from technology hybridization appears to be the only technologies at this point in time that can supply both energy and bit-flow-density requirements matching roadmaps for any communication range requirements. Finally, at the top view, the

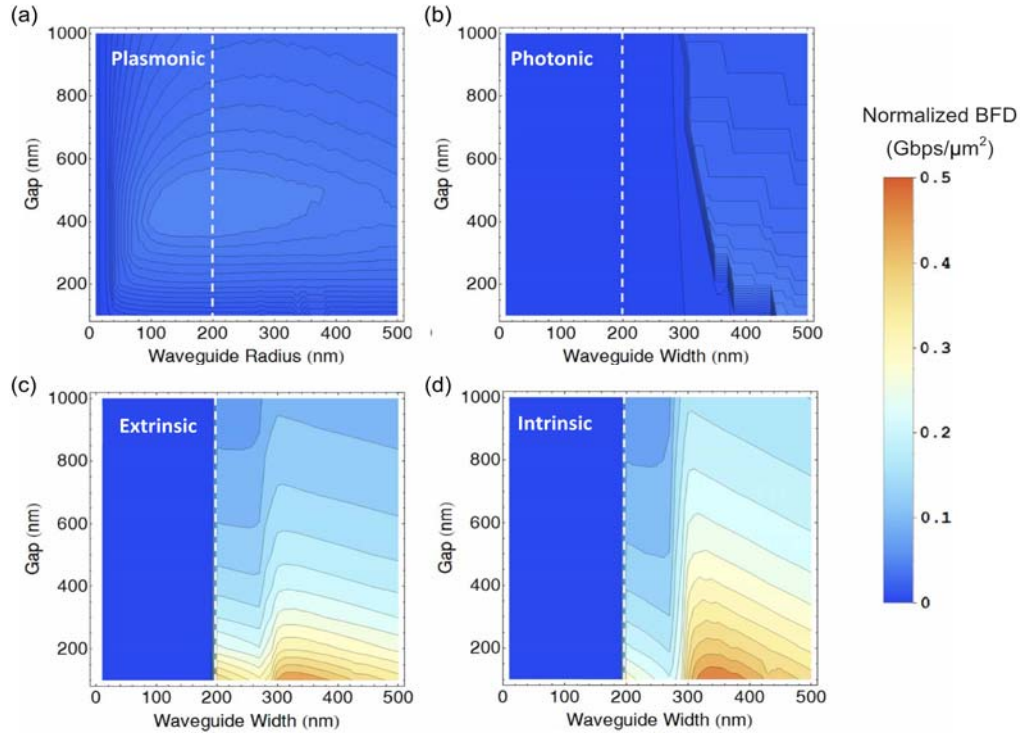


Figure 3. Bit flow density comparison for the four different kinds of interconnect [2].

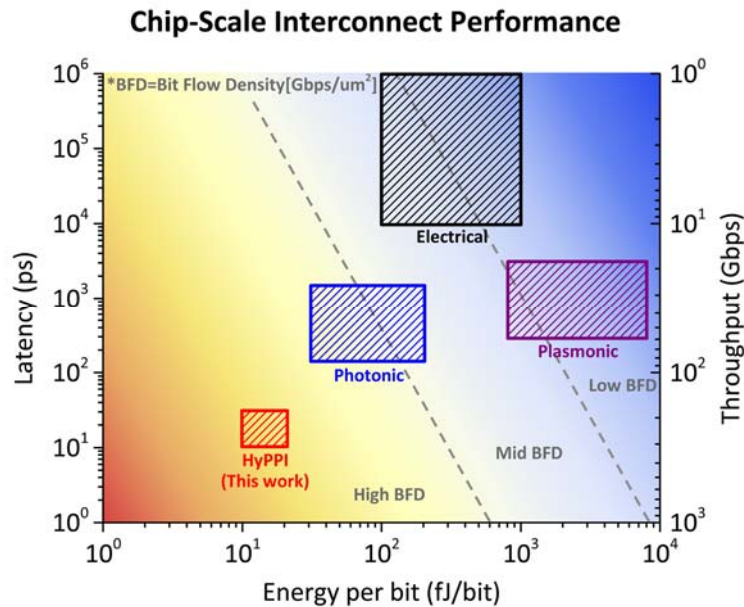


Figure 4. The overview of chip-scale (1 cm) length interconnects performance comparison for the four different kinds of interconnect with i) latency, ii) energy efficiency, iii) throughput, and iv) bit flow density breakdowns [2].

Hybrid Photonic Plasmonic Interconnects show the best potential with 10× energy efficiency, 100× throughput and delay improvements, which results in 1~3 orders of magnitude higher BFD than other interconnects at on-chip integration (**Fig. 4**).

### 2.1.2. CLEAR – A Holistic Figure-of-Merit Applicable from Devices to Architecture

We introduced a holistic and multi-hierarchical figure-of-merit (FOM) to compare performance-to-cost ratios of a) multiple technology options to include electronics, photonics, plasmonics, and hybrids, and b) showed that this FOM is able to capture performance spanning multiple hierarchies to include the device, link, circuit, and system levels. This CLEAR (Capability-to- Latency-Energy-Amount-Resistance) is a multi-hierarchical FOM allowing post-dict computer developments accurately and predicts photonics-based networks as a logical continuation for next generation compute systems. In short, the definition is:

$$\text{CLEAR} = \frac{\text{Capability}}{(\text{Latency}) \times (\text{Energy}) \times (\text{Amount}) \times (\text{Resistance})}$$

The individual factors in CLEAR are defined based on the hierarchy levels it applies to; for instance, at the compute system level, CLEAR breaks down as follows: the capability (C) is the system performance given by million-instructions- per-second (MIPS); the minimum latency (L) relates to the clock frequency and is limited by the temporal window between two adjacent clock cycles; the energy efficiency (E) represent energy cost for operating each bit in the units of joule-per-bit; the amount (A) represents the spatial volume of the system and is a function of the process dimensionality; the resistance (R) quantifies the economic resistance against a new technology adoption. It is derived from experience models and includes macroeconomic effects. Our investigations on using CLEAR to capture computer evolution trends indicate that it is the most accurate metric compared with other FOMs such as Moore’s law (which only tracks transistor count), Makimoto’s FOM, and Koomey’s Law (**Fig. 5a**). As such, CLEAR is able to capture historic growth trends of computer development. This is unlike other FOMs, which eventually deviate from the actual development pace of the semiconductor industry. We further were interested to answer the question why photonics link solutions, if apparently superior, have not yet been adopted in compute technology on-chip, and used CLEAR to predict performance into the future (**Fig. 5b**). We found that despite HyPPI offering a superior technological performance the economic factors need to be considered as well. For instance, the production costs of a transistor today only costs one-billionth of a photonic device price or less. As technology and manufacturing processes improve, the performance-per-cost (i.e. CLEAR) break- even-distance shortens, due to a flatter cost curve of electronics compared to photonics, the latter following a power law with time. Interestingly, the CMOS-based silicon photonic chip demonstrated by IBM in 2015, indicated as a yellow star in Fig. 3b, is close to the break-even area of two technologies. This point in time signifies an important juncture for photonics and HyPPI to becoming mainstream technology options in the near future. In fact, looking at the network level, our further studies indicate that larger flit sizes (~128-bits) gives rise to significant area overheads for electrical wires, positioning HyPPI to CLEARly outperform electrical link- based NoCs (see below).

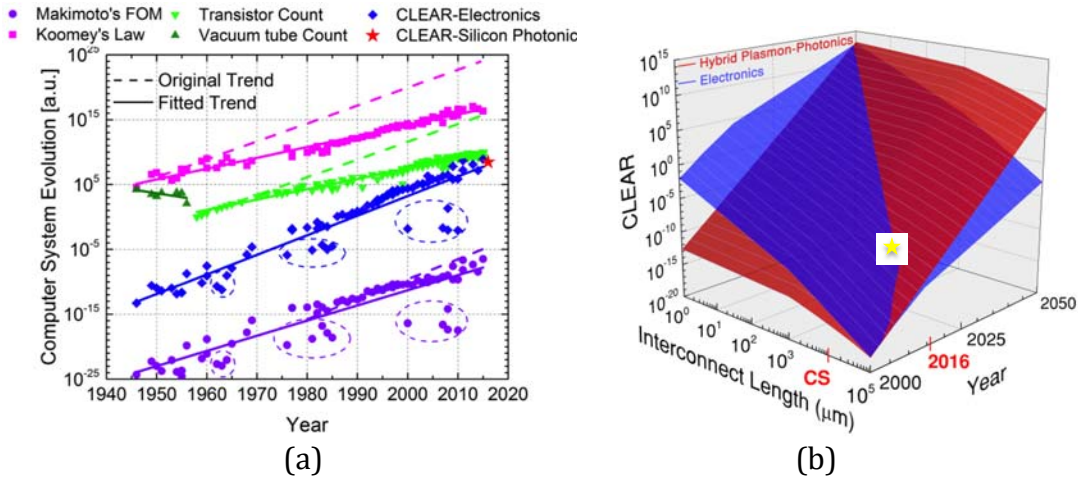


Figure 5. (a) Computer systems evolution trend compared using CLEAR, Moore’s law (component count), Makimoto’s FOM, and Koomey’s Law. Solid lines represent the linear fitting of each set of data points. Dashed lines represent the predicted  $2\times$ /year growth rate for each set of date points start from their deviation year. (b) CLEAR-based comparison of HyPPI and conventional electronic links.

### 2.1.3. MorphoNoCs – Configurable Networks-on-Chip using Nanophotonics

Contemporary Network-on-Chips (NoC) are designed as regular architectures that allow scaling to hundreds of cores. However, the lack of a flexible topology gives rise to higher latencies, lower throughput, and increased energy costs. With our aim towards designing reconfigurable NoCs incorporating DDDAS, we explore MorphoNoCs - scalable, configurable, hybrid NoCs obtained

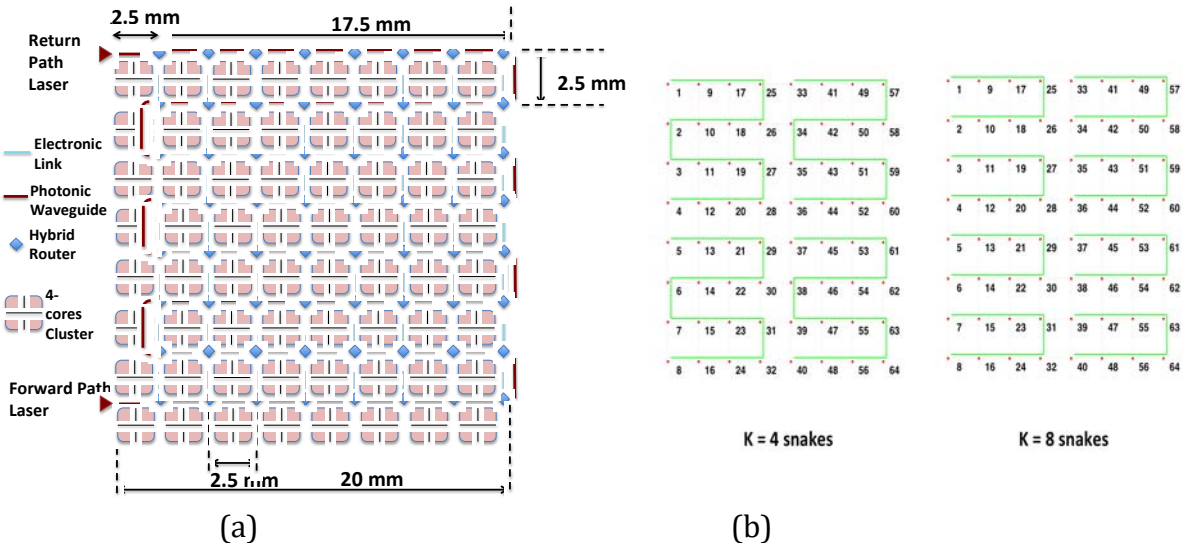


Figure 6. (a) One version of MorphoNoC for an 8x8 Mesh. In this incarnation, a set of waveguides snake around all of the electronic router nodes and thus provide connectivity between any pair of nodes. The number of wavelengths and number of waveguides restricts how many nodes can be connected together with these “express links”. Furthermore, each router node needs to limit the number of photonic links that can be sourced at its location, to limit the router size; (b) Different flavors of MorphoNoCs, achieved by splitting the snaking waveguides, thereby reducing laser power (lower losses) but reducing the potential for long-range links [7].



by augmenting regular electrical networks with configurable nanophotonic links. This class of networks can establish additional photonic links based on application traffic characteristic available a priori, and serves as the basis for developing HyPPI-based configurable NoCs. An overview of MorphoNoCs is provided in Fig. 6. As a basic component, it uses a serpentine waveguide that covers all of the electronic routers, and thus provides a potential direct connection between routers located several hops away on the electronic mesh. Different flavors of MorphoNoCs that we studied included those with 1 “snake” (one serpentine waveguide, Fig. 6a), 2 snakes, 4 snakes and 8 snakes (e.g. Fig 6b). The motivation for these variants is the lowered optical propagation losses in smaller length snakes, thus offering power reduction by trading off performance. Similarly, we also study such tradeoffs by changing the stride, namely, the routers which are connected to the waveguide; for instance, a stride=2 indicates every second router along the serpentine path is actually attached to the waveguide using a hybrid opto-electric router.

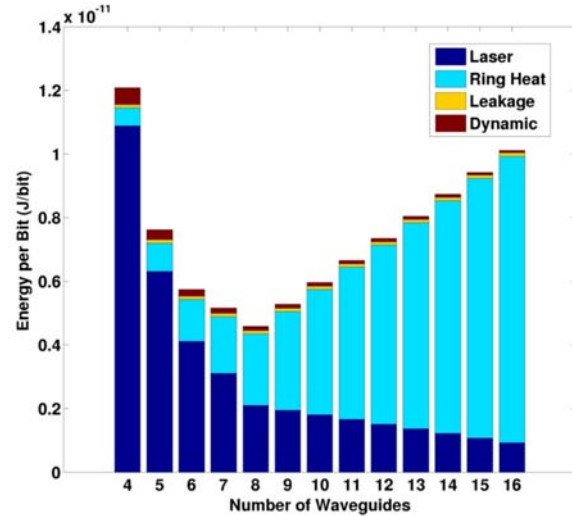


Figure 7. Energy components for MWMR waveguides that support sixteen 128 Gb/s links, at length 70mm, and injection rate=0.1 [7].

In order to design MorphoNoCs, we first carried out a detailed study of the design space for Multi-Write Multi-Read (MWMR) nanophotonics links. For instance, we varied the number of waveguides, number of wavelengths per waveguide, as well as the stride (points at which the modulator/detector are attached). For example, the optimum energy-efficient configuration requires 8 waveguides with 128 Gb/s bandwidth each at a fixed data rate per wavelength (for illustration) at 8 Gb/s in order to support 16 links (Fig. 7). These results were obtained by modifying DSENT tool to support MWMR links.

After identifying optimum MWMR design points, we then explored a suitable router architecture for deploying them in hybrid electronic-photonic NoCs. Next, we investigated the design space at the network level, by varying the waveguide lengths and the number of hybrid routers. We achieve this by varying the number of snakes (=K) discussed above, as well as the stride value (=S). This allowed us to carry out energy-latency trade-offs.

Finally, for our evaluations, we adopted traces from synthetic benchmarks as well as the NAS Parallel Benchmark suite, to compare MorphoNoCs with regular networks (Fig 8). Our results indicate that MorphoNoCs can achieve latency improvements of up to 3.0× or dynamic energy improvements of up to 1.37× over the base electronic network. Note that energy improvements are achieved even with additional hardware, indicating the promise of using hybrid NoCs. A latency improvement of 3x is significant, which can potentially improve the total execution time of applications by the same amount. Higher energy improvements (including lowered static power shown in Fig 8a) are expected as we extend this network with HyPPIs, as detailed next.

### 2.1.4. CHyPPI NoCs – Configurable HyPPI-based Network-on-Chip

Building on MorphoNoCs, we explored the use of configurable HyPPI (CHyPPI) links, by introducing a new device “mo-detector” capable of both modulation and detection for hybrid plasmonics. The modulation function is integrated on the racetrack ring waveguide controlled by a  $2 \times 2$  plasmonic switch which is the blue stripe shown in **Fig. 9a**. By applying a voltage on the metal contact of the plasmonic switch, the refractive index of the active material (Indium Tin Oxide, ITO) can be changed and further changes the operation status from ‘OFF’ to ‘ON’ (**Fig. 9b**). The overall performance of this Mo-detector device makes it a good choice for NoCs with insertion loss as low as 0.08 dB, which means the transmission loss is almost negligible when bypassing this device. On the other hand, the On-Off-Ratio (the power ratio between modulating a ‘1’ and a ‘0’, also called extinction ratio) of this device is as high as 15.53 dB, which provides quite clear signals for light detection. We augmented a base mesh network with various options – photonics, HyPPI, and CHyPPI. Results from CHyPPI are still evolving. We evaluated different network

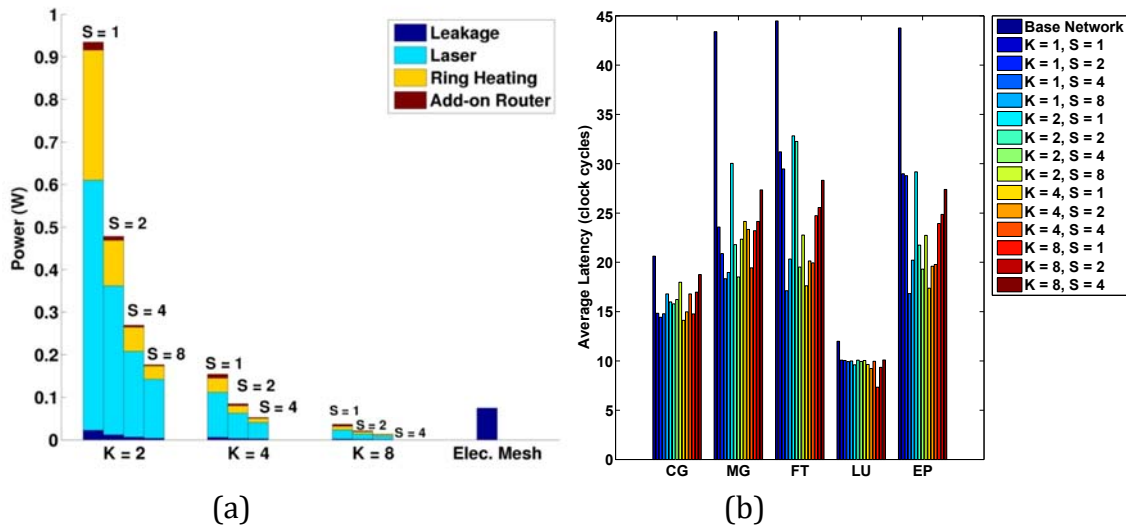


Figure 8. (a) Static power consumption for different number of snakes (K) and different stride (S). For example, K=2 indicates two snakes, and S=2 indicates that only every alternate router in the NoC connects to the snake, thus using a stride of 2. Conventional nanophotonics has higher power consumption over electronic mesh, but (b) demonstrates impressive latency gains for the NAS Parallel benchmarks. The network simulations were carried out using Booksim, with the benchmark traces derived from running the benchmarks on a Cray XE6m supercomputer [7].

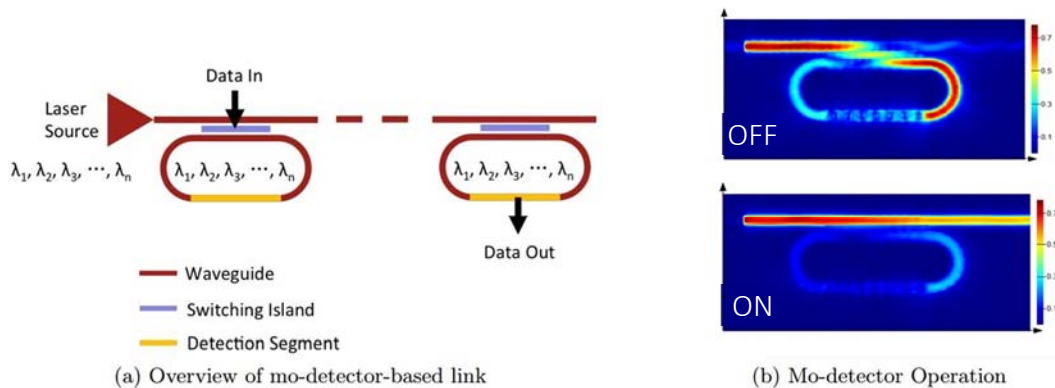
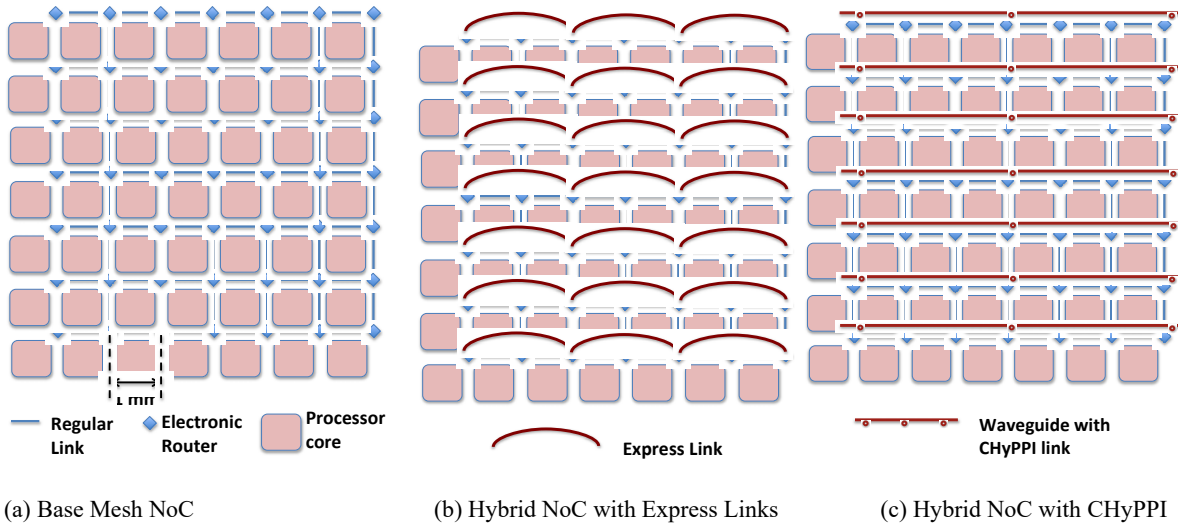


Figure 9. Plasmonic Mo-detector Device for Configurable HyPPI [15].

types to investigate dynamic reconfigurability to include a mesh augmented with fixed express link (Fig. 10b), and a mesh network augmented with CHyPPI links that can establish point-to-point connection between any two nodes connected to the waveguide (Fig. 10c), while comparing it to the base mesh NoC (Fig. 10a). Using the CLEAR FOM, we investigated the benefits of using the different hybridization options, namely, the technology used for the base network’s links, as well as the technology used for the augmented (overlaid) network. Initial investigations reveal the advantages of using a HyPPI base network (with electronic routers) augmented with CHyPPI (Fig 11). While we observe some performance improvements, reconfigurability does yet require



(a) Base Mesh NoC (b) Hybrid NoC with Express Links (c) Hybrid NoC with CHyPPI  
 Figure 10. Networks evaluated for different technology options. The small number of cores is for illustration only. Express Links shown are for Hops = 2. The illustrated CHyPPI has Stride = 3. All links are bidirectional (not shown).

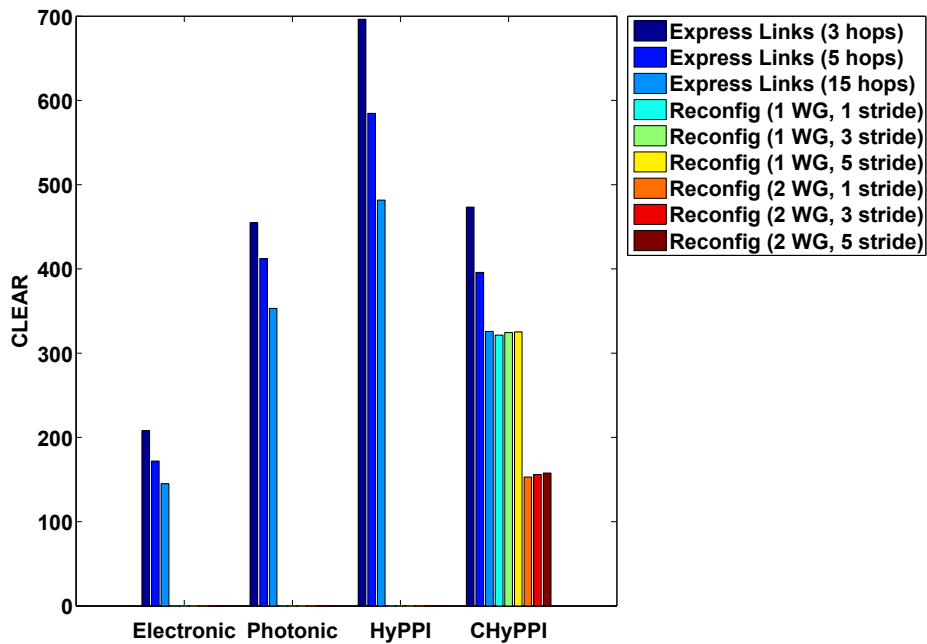


Figure 11. Comparing different hybridization options of Fig. 9. with the base network links (short links) using HyPPI. Base network+CHyPPI (the last group of bars) shows are good value of CLEAR.



more investigations to determine benefits. For instance, a natural limitation of these preliminary results are that the current design only allows one link to be configured at a time on a CHyPPI waveguide. As we allow for multi-level connectivity per link, we expect further improvements as we are working on improving the CHyPPI links in order to support multiple links at different segments of the same waveguide.

## 2.2. Year 2

### 2.2.1. Fundamental Scaling Laws in Nanophotonics

The success of information technology has clearly demonstrated that miniaturization often leads to unprecedented performance, and unanticipated applications. This hypothesis of “smaller-is-better” has motivated optical engineers to build various nanophotonic devices, although an understanding leading to fundamental scaling behavior for this new class of devices is missing. Here we analyze scaling laws for optoelectronic devices operating at micro and nanometer length-scale. We show that optoelectronic device performance scales non-monotonically with device length due to the various device tradeoffs, and analyze how both optical and electrical constraints influence device power consumption and operating speed. Specifically, we investigate the direct influence of scaling on the performance of four classes of photonic devices, namely laser sources,

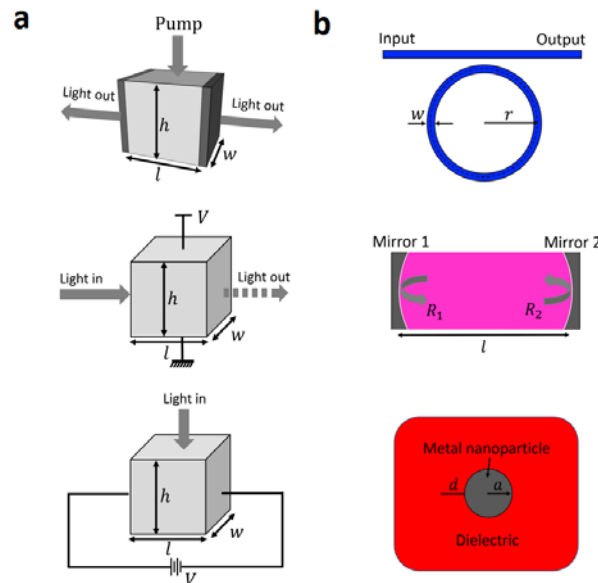


Figure 12. Schematic structures of devices and cavities. (a) Here we investigate performance scaling of four photonic devices, namely a laser source, an electro-optic modulator, a photodetector, and an all-optical nonlinearity-based switch (the latter is not shown). The physical device volume is given by the device geometry. (b) Here we utilize three device-underlying cavity types; namely a ring resonator (RR) cavity with the waveguide width,  $w$ , and the ring radius,  $r$ ; a Fabry-Pérot (FP) cavity comprised of a dielectric material sandwiched by a pair of highly reflecting metal mirrors with the reflectivity of  $R_1$  and  $R_2$ ; and a plasmon cavity formed by metal nanoparticle (MNP) embedded in a dielectric, and  $a$  is the radius of metal nanoparticle.  $d$  represents the normal distance for the dipole position from the metal particle surface as is equal to 10 nm. The scaling parameters are  $r$ , for the RR,  $l$  for the FP, and  $a$  for the MNP cavity, respectively [7].

electro-optic modulators, photodetectors, and all-optical switches based on three types of optical resonators; microring, Fabry-Perot cavity, and plasmonic metal nanoparticle (Fig. 12). Results show that while microrings and Fabry-Perot cavities can outperform plasmonic cavities at larger length-scales, they stop working when the device length drops below 100 nanometers, due to insufficient functionality such as feedback (laser), index-modulation (modulator), absorption (detector) or field density (optical switch). And our results provide a detailed understanding of the limits of nanophotonics, towards establishing an opto-electronics roadmap, akin to the International Technology Roadmap for Semiconductors.

For our scaling law analysis, we define the critical length for the three underlying cavities as the radius for the RR and the MNP, and the physical distance between two mirrors for the FP. We derive analytical expressions for both the cavity quality factor  $Q$  and the optical mode volume  $V_m$  for the RR and FP cavities and estimate the Purcell factor (Fig. 13), defined as  $F_p =$

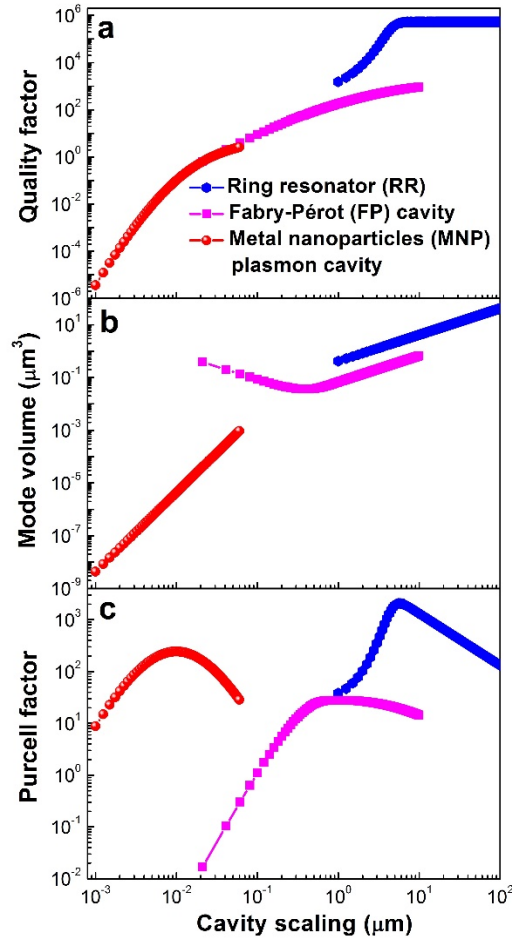


Figure 13. Cavity performance as a function of scaling. Scaling of (a) quality factor  $Q$ , (b) mode volume,  $V_m$  and (c) Purcell factor,  $F_p$ , for all three cavity configurations. While the general trend shows a reduced  $Q$  upon scaling, significant differences between the three cavity types exist. Nominal parameters in this study are as follows; the propagation loss of a diffraction limited beam,  $\alpha_p=1.0$  dB/cm such as used in the RR; Silver metal mirrors,  $\tilde{n}_{Ag} = 0.41+10.05i$ , the dielectric refractive index is taken to be  $\tilde{n}_d = 3.0- i0.001$ , the Silver conductivity  $\sigma_{Ag}=6.3\times 10^7$  mho/m, the dipole distance from the MNP to be  $d=10$  nm, and the damping rate for the MNP to be  $\gamma_d=2.0\times 10^{15}$  rad/s. Clear maxima upon scaling are observed for all three cavities [7].

$\frac{3}{4\pi^2} \left(\frac{\lambda_R}{n}\right)^3 \left(\frac{Q}{V_m}\right)$ , where  $\lambda_R$  is the resonant wavelength of the cavity, and  $n$  is the cavity material refractive index. The discontinuity of the displacement current across the metal particle in the plasmonic cavity requires a different approach; we find the ratio of the effective density of the surface plasmon modes,  $\rho_{SP}$ , relative to that of the radiation continuum,  $\rho_{rad}$ , which directly gives  $F_p$ . The mode volume  $V_m$  is obtained by a permittivity-modified geometric volume, from which we can estimate  $Q$ .

Regarding cavity performance,  $Q$  as a function of the critical length is a key metric since it relates the ability to spectrally store optical energy relative to its loss (**Fig. 13a**). For the RR, at larger length,  $Q$  is almost independent of length, since the increased propagation loss with the circumference of RR is cancelled by the increased round-trip time. However, at small radius ( $<10 \mu\text{m}$ ), bending losses overcome the propagation losses resulting in a sharp drop in  $Q$ . For the FP and MNP,  $Q$  decreases with scaling, mainly due to an increased overlap of the electromagnetic field with the lossy metal. We therefore choose a lowest-volume approach by selecting the single-mode condition for the ring waveguide cross-section dimension and scale only the ring radius for the RR, whereas the MNP scales cubically,  $V_m \sim a^3$  (**Fig. 13b**). For the FP, the mode-volume has a non-monotonic variation with the critical length; upon scaling and approaching the diffraction limit ( $\sim 200 \text{ nm}$ ),  $V_m$  decreases. Based on the above analysis, we investigate the scaling behavior of the Purcell factor for each cavity (**Fig. 13c**). For all the three cavities, we observe a non-monotonic dependence of  $F_p$  with scaling: for longer lengths an increased  $V_m$  reduces the Purcell factor, whereas at short lengths the reduction in  $Q$  is detrimental to  $F_p$ . However, the optimal cavity length and the maximum Purcell factor for each cavity are vastly different.

### 2.2.2. CLEAR+: Predicting and Post-dicting the Computing System/Unit Evolution

In order to determine the applicability time period of the respective FOM we take the deviation point where the data deviates away from a single exponential growth rate. After comparing all four FOMs, we find that when a FOM incorporates more relevant factors its starting point of deviation from its original trend occurs later. Unlike the CLEAR, however, which shows an accurate fit to the data. Thus, we empirically find, that computing performance grows consistently at a fixed rate at  $2\times/\text{year}$  and is independent of technology. This assumption can be proved by the scientific theory that the result is observable and can be subjected to minor adaptations in FOM factors. Testing emerging compute engine performance such as those augmented with photonics as predicted by IBM for examples, we find that such technologies indeed appear to continue the evolutionary  $2\times/\text{year}$  development trend.

In addition, these multi-factor FOMs are able to classify computing systems by their relative deviation from the  $2\times/\text{year}$  trend line. For instance, the additional overhead on (i.e. physical size, parallelism, cooling, low economy-of-scale, and manufacturing costs) of supercomputers show their inferior CLEAR values relative to all other computer types such as laptops and mobile devices, despite their higher performance (dashed circles, **Fig. 12c, 12d**). The high parallelism of multi-core technologies used in supercomputers is challenged by computing-to-energy returns described by Amdahl's law. We observe that while supercomputers deliver petaflop performance, the entire infrastructure resembles that of computers 5-30 years back thus questioning future scale-up.

In addition to applying CLEAR to the system-level for technology evolution prediction, it can be amended to guide the development of the device- and network-level alike, provided minor

modifications are applied; for the device-level, the signal distance becomes the device length and hence cancels. Thus the area reduces to the device length. The data capacity and latency from the link become the device operating speed and response time, respectively. Thus, at the device-level CLEAR becomes Capability-to-Length-Energy-Area-Ratio, which breaks-down as follows: i) the device operating frequency is the capability (C); ii) the scaling efficiency which is the reciprocal of the critical scaling length (L) of the device describes the interaction length to provide functionality; iii) the energy consumption (E) of the energy 'cost' per bit is the reciprocal of the energy efficiency; iv) the on-chip footprint, or area (A), and v) the economic resistance (R) in units of dollars (\$) is the reciprocal of the device cost efficiency. Here the critical scaling length in the denominator does not conflict with the area factor but indicates the scaling level or ability of the device to deliver functionality given its length. For instance, the critical scaling length of the CMOS transistor is the length of its logic gate, which controls the ON/OFF states. For photonic and plasmonic devices, it can be regarded as the laser or modulator (linear) length, or micro-ring diameter when cavities are utilized.

Next, we demonstrate how we compare the performance of a) and electronic transistor; b) a microdisk ring modulator; c) a plasmonic electro-optic modulator and d) a hybrid plasmon polariton modulator among different technologies based on CLEAR FOM. We represent the device-CLEAR results as five merit factors in a radar plot (Fig. 14). Note, each factor is represented in such a way that the larger the colored area in Fig. 3 the higher the CLEAR FOM of the device technology. Moreover, some of the factors of the device-CLEAR have physical constraints that fundamentally limit further growth independent of chosen technology. For example, the energy efficiency of the device is ultimately limited by the Landauer's principle ( $E_{\min} \geq k_B T \ln(2)$ ), which restricts the minimum energy consumption to erase a bit of information to 2.87 zeptojoule at room temperature ( $T = 300\text{K}$ ). Given this device energy limit, the Margolus–Levitin theorem set a cap for the maximum operating frequency of the device. Based on the fundamental limit of quantum computing, a device with the amount of  $E$  energy requires at least a time in units of  $h/4E$  to transfer \*what\* from one state to the other resulting in a switching bandwidth of about 16 THz for energy levels at the Landauer's

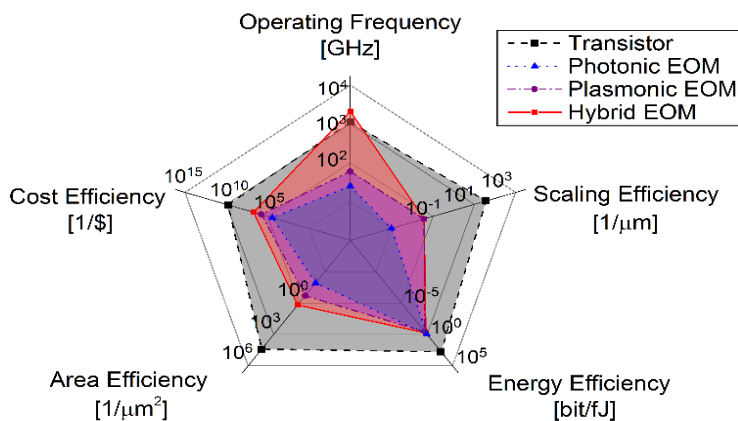


Figure 14. The CLEAR comparison at device level. Each axis of the radar plot represents one factor of the device-CLEAR and is scaled to the actual physical limit of each factor. Four devices compared from different technology options are: 1) the conventional CMOS transistor at 14 nm process; 2) the photonic microdisk silicon modulator; 3) the MOS field effect plasmonic modulator; and 4) the photonic plasmonic hybrid ITO modulator. The colored area of each device also demonstrates the relative CLEAR value of each device.

limit. When approaching the quantum limit for data communication, the device’s critical length would be scaled down to the dimension of about 1.5 nm based on the Heisenberg uncertainly Principle ( $x_{min} \geq \hbar/\Delta p = \hbar/\sqrt{2mE_{min}} = \hbar/\sqrt{2mk_B T \ln 2} = 1.5nm$ ).

### 2.3 Non-Blocking WDM Hybrid Photonic-Plasmonic Router for HyPPI

The requirements for data communication rates become more demanding driven by power and thermal budget constraints driven by to the simultaneous increase of data transfer rate and multicore technology, where the latter caused drawbacks to parallelism challenged by ‘dark silicon’. With the success of long-haul optical networks, optical interconnects at the board- or even at the chip-level have become of interest to mitigate the processing/communication gap. In this work we find that designing a router using a hybrid plasmonic-photon approach and emerging unity-high index tuning materials simultaneously improves all three factors. The enabling technological insights are based on the strong index tunability of the underlying optical plasmonic hybrid mode enabling short 2×2 switches based on voltage-controlled directional-couplers. Cascading a network of these nanoplasmonics 2×2 switches we can design a compact optical router since the switching length scales inversely with index-tuning capability. In addition, given the low quality-factor of the 2×2 switches due to the lossy plasmonic mode, this leads to an advantageous property of being broadband, hence being not being wavelength limited enabling a wide spectral WDM operation without thermal tuning thus saving energy consumption during operation. In this paper, we use the terminology ‘all-optical router’ to describe the lack-of O-E-O conversion inside the router, but note that signal routing requires electrical decision-making from the control circuit.

The fundamental building block of the optical router is a 2×2 optical switch, for example, the voltage-controlled directional coupler whose performance directly impacts the overall performance of the router. Recently, photonic 2×2 switches with microring resonators (MRRs) or Mach-Zehnder Interferometers (MZIs) have been applied to perform this routing function in variety of optical networks. With high spectral sensitivity (< 5 nm free spectral range) and low insertion loss (< 1 dB per ring), photonic MRR-based switch is still suffering from the ring-tuning (dynamic) power and dense packaging since the ring radius is usually chosen to be 10µm or larger in order to have higher quality factor (Q factor) and low bending loss.

The total number of 2×2 switches needed for a non-blocking router scales with 2(N-1), where N is the number of ports of that router. Thus, as a router for an optical mesh network of a NoC, requires 4 ports to connect to the north, south, east and west neighbors, and 1 additional port for connection to the local processing core. Thus, eight 2×2 hybrid routers are needed in total to

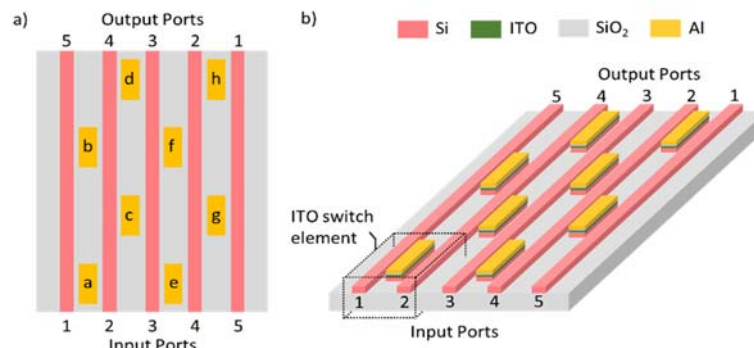


Figure 15. The top view and the schematic plot of the 5x5 Port non-blocking optical router. 8 individual 2x2 ITO switches are placed with certain pattern in order to achieve non-blocking routing function. The length of the ITO switches is not to scale for clarity [15].

achieve the non-blocking routing functionality that requires assigning a random input port to a random output port anytime during operation without disturbing other data streams (**Fig. 15**). We note that other input ports are still able to maintain connections with the remainders of the output ports without affecting the initially set switches. Moreover, self-communication (communication between same input and output port number, resulting in a U-turn) is forbidden because: 1) it can be done with higher energy- and latency- efficiency with other local (electrical) interconnect links, and 2) avoiding self-communication can simplify the router from  $N^2$  number of switches required for all-to-all connection down to only  $2(N-1)$ , which can also reduce the average loss of the router.

In summary, we have shown for the first time a hybrid photonic-plasmonic non-blocking broadband router with fast response time (2 ps) and high-energy efficiency (82 fJ/bit) enabled by hybridizing plasmonics with a photonic device. By comparison MRR and MZI based photonic routers offer microseconds-to-nanoseconds and picojoule levels, respectively. Integration of the ITO plasmonic switches scales the device on-chip area down to  $250 \mu\text{m}^2$ , which gives  $10^2\sim 10^3$  times area-efficiency improvement. This router operates over a broadband 3-dB signal discrimination bandwidth over 200 nm allowing for 76~89 Tbps theoretical noisy Shannon channel capacity. The high performance and scalability of this hybrid router are promising towards future large-scale multi-core optical networks requiring all-optical routing.

#### 2.2.4. HyPPI NoC: Bringing Hybrid Plasmonics to an Opto-Electronic Network-on-Chip

While HyPPI had been studied in the first year, its implications at the onchip network level had not been explored, and is thus the focus of the investigations in this year. HyPPI is an excellent candidate as a point-to-point link, to replace electronic links in a network-on-chip (NoC). However, due to reliance on the electronic routers for directing flits across the NoC, there are a lot of optical to electrical (O-E) and electrical to optical (E-O) conversions that occur as a result. For instance, consider the Mesh NoC shown in **Fig. 16a**. Each one of the 'Regular Link' can be optical, however, a node communicating from the left end to the right end will incur several O-E-O conversions. One possible approach to address this issue is through the use of express links. An example with 2 hops express links in the horizontal direction is shown in **Fig. 16b**. Since additional links demand a larger number of ports from the participating routers, we consider express links only in the horizontal direction.

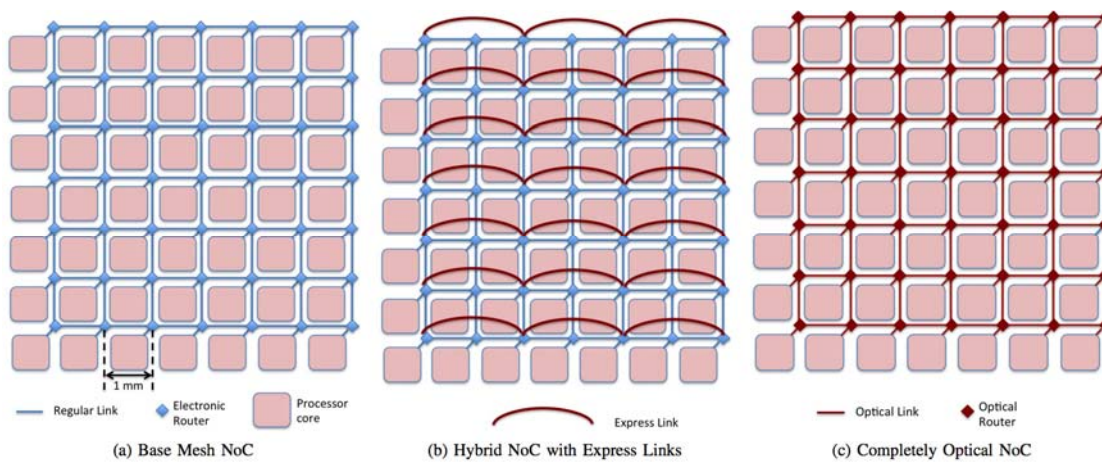


Figure 16. Networks evaluated for different technology options. The small number of cores is for illustration only [11].



The other option is to use an all-optical NoC, see **Fig. 16c**. However, in our opinion, completely optical NoCs are not yet fully mature for migration from contemporary electronic networks. We thus believe that it is better to deploy photonic links only for long-range traffic and for nodes that communicate heavily. Furthermore, with the lack of memory storage in optics (no flip flops or registers or buffers), an all-optical network will require a suitable infrastructure for arbitration and/or routing, with proposed approaches using token-based arbitration or a parallel electronic path for channel setup. Thus, we prefer to adopt the cheaper and well-understood and easily routable electronics for short distances. Furthermore, due to additional clock cycles overhead in opto-electric conversions, optical links become inferior for short distance traffic between, for instance, neighboring core routers.

In order to help design hybrid networks incorporating express links, we adopted a unified metric called CLEAR, which is defined in earlier in this report.

We demonstrated results for link and network evaluations using this metric. The results of these simulations along with latency, power, and area simulations are depicted in **Fig. 17**. These evaluations demonstrated that electronic NoCs augmented with HyPPI provided a  $1.8\times$  improvement in CLEAR over a base electronic mesh. These results indicated up to  $1.64\times$  latency improvement over a base electronic mesh, with negligible energy overheads due to the HyPPI express links. Finally, we carried out performance projections for all-optical NoCs. The projections indicate that all-HyPPI as well as all-photonic NoCs would be significantly more energy efficient than electronic NoCs ( $255\times$ ), although electronic route setup requirements may diminish this result. Furthermore, an all-HyPPI NoC would be two orders of magnitude smaller in area compared with an all-photonic.

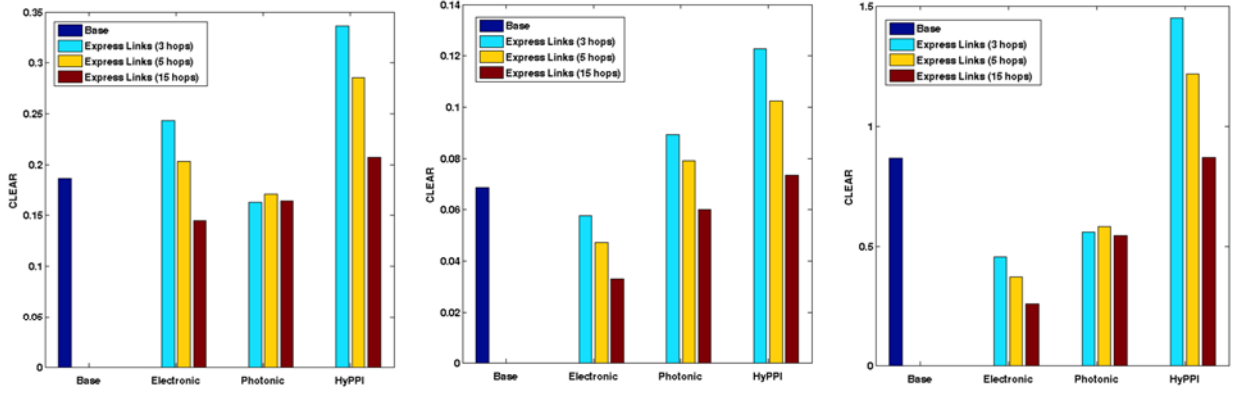
### **2.2.5. D<sup>3</sup>NoC: A Dynamic Data Driven Network on Chip**

So far our proposed HyPPI NoCs have been static and any reconfiguration envisioned has been realized at design time. Now, we investigate realization of an adaptive dynamic reconfigurable NoC. Our NoC adapts to changes in environment by taking measurements of environment and react to the measurements by augmenting the topology by an adaptable optical express bus. In fact, we also enable the dynamic adaptation of our measurement system in response to behavior changes of environment. To authors knowledge the latter approach had not been investigated prior to this study in the context of NoCs. The primary objective of this line of research is to show the potential of adaptive dynamic measurement in addition to conventional reconfiguration techniques. Our design is motivated by the Dynamic Data Driven Application System (DDDAS) paradigm. In DDDAS, computations and measurements form a dynamic closed-loop feedback in which they tune one another in response to changes in the environment. We expect the practice of applying DDDAS concept in NoC design would improve both performance and power efficiency. The idea is to augment the base electronic mesh NoC with an optical HyPPI bus as shown in **Fig. 18**. The bus is dynamically allocated to source-destination nodes that are projected to have heavy communication during the next time interval.

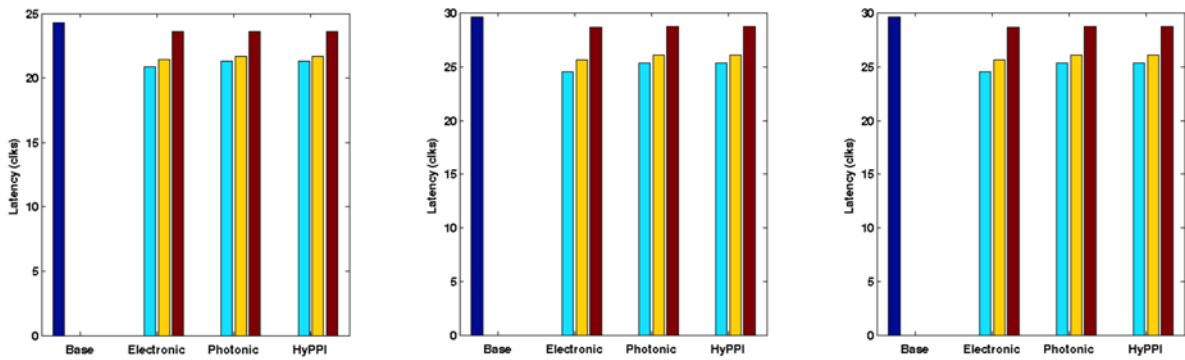
## **2.3. Year 3**

### **2.3.1. DDDAS Adaptive Routing Algorithm**

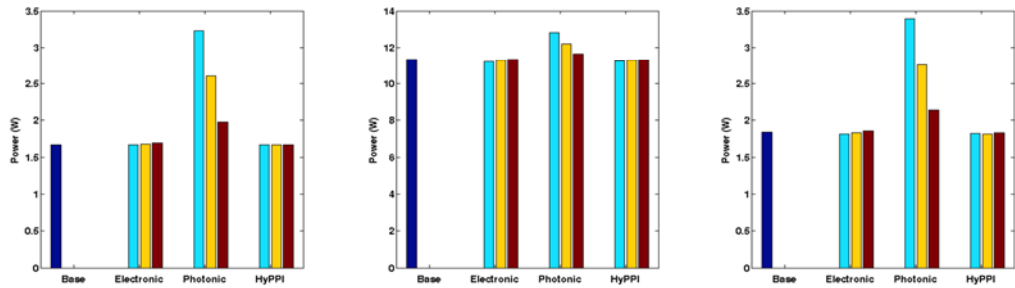
Based on the DDDAS concept, a novel adaptive routing algorithm named Weighted Recent Communication Driven (WReCD) adaptive routing algorithm is developed. Instead of setting up



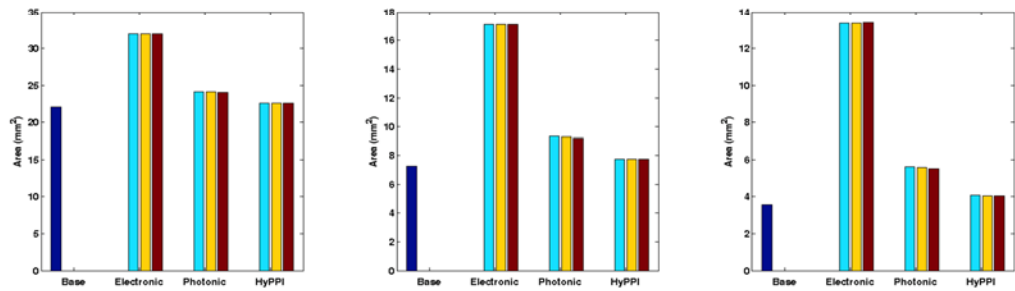
(a) CLEAR: Base E-Mesh + Express (b) CLEAR: Photonic Mesh + Express (c) CLEAR: HyPPI Base Mesh + Express



(d) Latency: Base E-Mesh + Express (e) Latency: Photonic Mesh + Express (f) Latency: HyPPI Base Mesh + Express



(g) Power: Base E-Mesh + Express (h) Power: Photonic Mesh + Express (i) Power: HyPPI Base Mesh + Express



(j) Area: Base E-Mesh + Express (k) Area: Photonic Mesh + Express (l) Area: HyPPI Base Mesh + Express

Figure 17. Comparing different flavors of hybrid NoCs (injection rate = 0.1) [11]



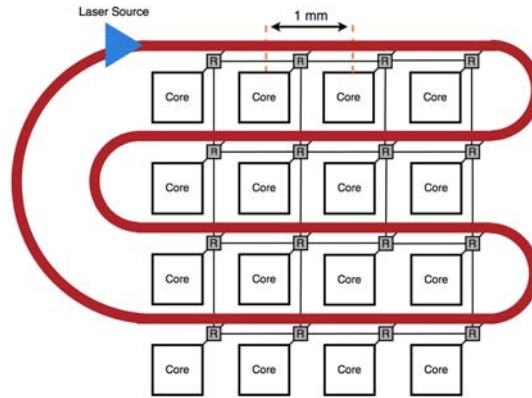


Figure 18. serpentine-like optical bus connects all the cores in a NoC. For simplicity we plotted a 4x4 NoC [16]

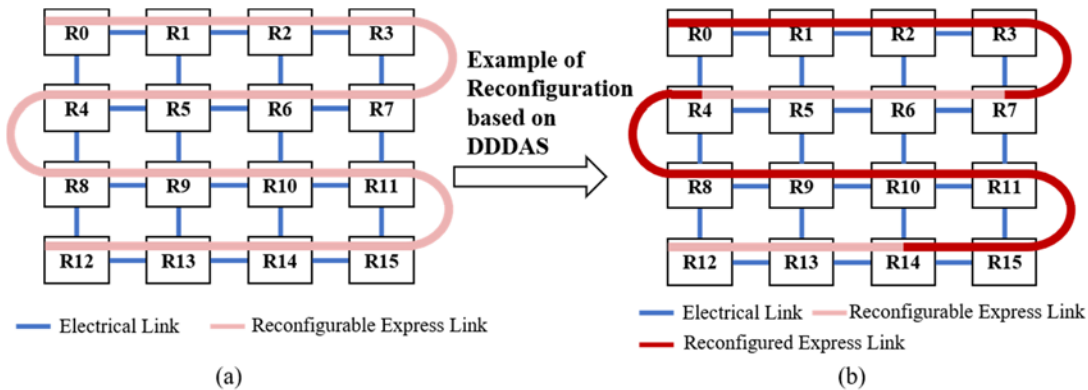


Figure 19. (a) A snake-like HyPPI link on top of a basic electrical 4x4 mesh network. (b) An example of reconfigured HyPPI link setup based on DDDAS concept

a static express link only, we proposed the use of Hybrid Photonic-Plasmonic Interconnect (HyPPI) links that can be setup dynamically. HyPPI links is built on top of a regular electrical mesh network, as depicted in **Fig. 19**. To enable the possibility of all-to-all communication, the HyPPI express links are placed as a snake. Based on the traffic pattern, the HyPPI links might be reconfigured to different segments, which allows packets to be transmitted in an express path. For example, **Fig. 19** shows one of the possible reconfigurations of the HyPPI links. R0 connects to R7 with HyPPI, while R4 is connected with R14 directly. It means that packets traversing at the beginning of the HyPPI links (R0 or R4), may use the express HyPPI links if destinations of the packets are close to the end of the express links (R7 or R14).

Our system has two execution windows; namely, the operating window and the reconfiguration window. By switching the window, the whole system will either run the application or reconfigure the HyPPI links. **Fig. 20(a)** depicts the details of both the operating window and the reconfiguration window. In the operating window, the average destination of each router will be collected. A router runs the application normally as long as the difference between the current average destination and old average destination does not reach a preset parameter (the *drift threshold*). Once the *drift threshold* is exceeded, a router will send a signal to the central control unit, requesting to reconfigure the HyPPI link. This procedure is considered as a *vote*. Once the total number of *votes* reaches a certain percentage of the NoC, *vote threshold*, the central control

unit will change the window to the reconfiguration state (**Fig. 20b**). To avoid the situation in which a configured long HyPPI link is not used anymore, a parameter, *decay type*, is proposed. This parameter is used to eliminate idle HyPPI links during the next configuration.

In the reconfiguration window, first the network uses a distributed sorting algorithm<sup>1</sup> to sort the collected data from each router along the snake like HyPPI link that spans all the routers in the NoC. The sorted data is then used to segment the snake HyPPI link into shorter non-overlapping segments. The details of the different tasks in the operating and reconfiguration windows are detailed in the following subsections.

**Data Collection.** Each router collects the destination from incoming packets, and the destination would be decoded as X and Y coordinates in a 2D mesh network. The destinations of all packets within the current operating window are tracked. We compute the average destination using following equation.

$$X_{avg} = \left\lfloor \frac{\sum_i 2^i x_i}{2^n} \right\rfloor, Y_{avg} = \left\lfloor \frac{\sum_i 2^i y_i}{2^n} \right\rfloor$$

where  $x_i$  and  $y_i$  represent the coordinates of destination of incoming packet.  $X_{avg}$  and  $Y_{avg}$  represent coordinates of the weighted average destination for  $n$  packets the arrives the current router. This equation represents a weighted average of the destination that gives higher weight to more recent destination, hence the name of our algorithm, Weighted Recent Communication Driven (WRcCD) adaptive routing. For example, assume that there are three packets with destination (2, 2), (4, 8) and (10, 20) respectively arrive the same router in this sequence. The default  $X_{avg}$  and  $Y_{avg}$  are (0, 0). The weighted average destination will be (2, 4) after the first two packets arrived, and finally, it will be (6, 12) after all these packets arrived. For the same case, the regular average destination, which simply average destinations for all packets, will be (5, 10). Our method focuses more on the recent traffic pattern, resulting in the weighted average destination will be closer to the destination of most recent incoming packet, while comparing to the regular average method. It intends the next packet arrive at the same router will go close to the previous packet.

Moreover, the hardware implementation of such a computation only involves additions and shifts. The result is stored in two registers, named  $X_{avg}$  and  $Y_{avg}$ . Based on this average, an express HyPPI link might be setup between the current node and the destination during the reconfiguration window.

Once the difference between the current weighted average destination and the weighted average destination from the previous operating cycle is larger than or equal the *drift threshold*, the router

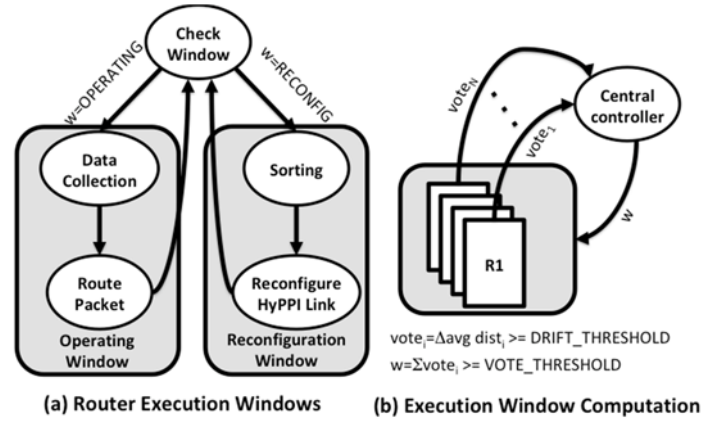


Figure 20. Weighted Recently Communication Driven Adaptive Routing Algorithm with Reconfigurable Topology

<sup>1</sup> Lang, H.W., Schimmler, M., Schmeck, H. and Schroder, H., 1985. Systolic sorting on a mesh-connected network. *IEEE Transactions on Computers*, (7), pp.652-658.

sends a *vote* signal to the central unit for requesting reconfiguration. This *vote* does not guarantee that the system will enter reconfiguration as shown in **Fig. 20b**. For short communication, e.g. neighbor communication, there is no need to use the express links. Hence, the *drift threshold* is set as 3, which means if the Manhattan distance between the current average and the previous average is smaller than 3, the router will not request reconfiguration. The local communication path (electrical mesh), in this case, should be enough. As the communication patterns change, the computed average will change following the new pattern, and thus, the system will adapt to the application needs.

The current average can also change when a router is *idle* for a period greater than or equal the *decay threshold*. In this case, the average destination is decreased in order to decrease the priority of this node at the sorting stage. This is because it is desired to build HyPPI links that serve the most recent traffic pattern instead of the previous traffic pattern. Three decay types are considered, including DECREASE, RESET and NONE. Both DECREASE and RESET monitor *idle* time. As it reaches the *decay threshold*, the *weighted average destination* changes accordingly. It will be reset as zero in RESET decay type, while it will be decreased by *decrease value* in DECREASE decay type. The NONE decay type disables this feature.

**Next Hop Selection.** Considering the HyPPI link as an express lane, and the electrical mesh connection as local routes, packets take the express path if the destination is long enough. Packets will utilize express HyPPI link if the Manhattan distances between the current router and destination is longer than the distance between the end of HyPPI link and the destination. The routing algorithm is designed to avoid deadlocks by ensuring that a packet will always be getting closer to its destination. Alternatively, packets follow XY routing.

In case of using a HyPPI link, the information stored inside the packets will be encoded to optical signal by switching the modulators, which is considered as an electrical-optical (E-O) conversion. Using this simple comparison, the area and power overhead should be relatively low. Similarly, the destination node contains the photodetector, in addition to an optical-electrical (O-E) conversion component, which will transfer the optical signal it receives back to electrical signal.

**Reconfiguration.** The first stage of reconfiguration is using a distributed sorting algorithm that sorts the data collected in the previous operating window along the optical snake. For a  $n \times n$  mesh network, this type of sorting has a time complexity of  $O(n)$ . It is important to use a very efficient sorting algorithm as the sorting time represents the majority of the reconfiguration cycle overhead.

After the sorting is completed, each router stores one set of possible HyPPI link connection (including source and destination), starting from the longest Manhattan distance. Only the values stored in the first two rows of the mesh is going to be used to configure non-overlapping segments on the optical snake. The decision to only consider the possible HyPPI link sets stored in the first two rows is taken because we are interested in generating the longest possible optical segment and in order to save reconfiguration time as it takes one extra cycle per considered node. Once the reconfiguration is done, the central unit sets the execution window back to *operating*.

### 2.3.2 Design and Implementation of Router on FPGA

To adapt the HyPPI link with the electrical mesh routing, a new router is designed. D.U Becker implemented a router<sup>2</sup>, with efficient control logic for high-performance computing of a NoC

---

<sup>2</sup> Becker, D.U., 2012. *Efficient microarchitecture for network-on-chip routers* (Doctoral dissertation, Stanford University).

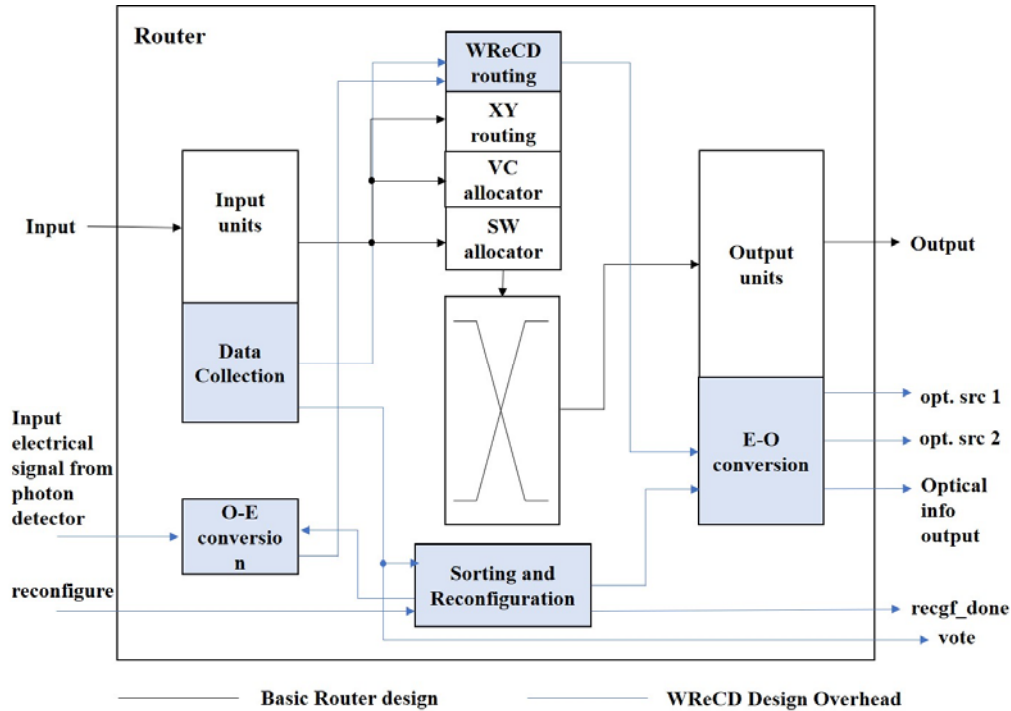


Figure 21. Router connection with both electrical component and HyPPI link.

system. Based on this RTL design, additional hardware is required to implement the proposed WReCD adaptive routing algorithm. The microarchitecture of the regular router design and WReCD router design are shown in **Fig. 21**. The blue boxes and lines illustrate the additional hardware requirements, while the black parts illustrate the original hardware design of D.U Becker's work. First, the data collection components are added. This includes two registers,  $X_{avg}$  and  $Y_{avg}$ , another set of registers are required to store the previous value of  $X_{avg}$  and  $Y_{avg}$ . In addition, a register that stores the *idle time* is built in the data collection module. The WReCD routing unit is necessary for next hop routing, in addition to the regular XY-routing unit. It is embedded in the WReCD routing module. To make the HyPPI link reconfigurable, then a sorting and reconfiguration unit is implemented as well. This unit enables sorting the Manhattan distance (including source and destination nodes) into a snake-like indexing scheme. The reconfiguration unit sets up all the essential control signals in order to activate/deactivate the Optical-Electrical (O-E) conversion or the Electrical-Optical (E-O) conversion. Sorting and reconfiguration are done using software routines that are executed on all nodes during the reconfiguration window.

Now the WReCD router is available to connect to other WReCD routers and the HyPPI-related optical components (**Fig. 22**). As a basic router in a mesh topology, it is connected to other four routers, as well as to the node connected to itself, resulting in ten flit-size ports, including input and output, on each router in total. There are two optical sources for both downstream and upstream communication since the light could propagate in one direction only. Also, there is a control signal, named as  $MoD_{enable}$ , to control the hybrid photonic-plasmonic device termed MoDetector (MoD). As a symmetrical device, MoD is able to provide either a modulation function or a light detection function by using electrical bias with bi-directional communication capability. To be more specific, a three-waveguide based switching mechanism has been put next to the main bus in order to provide off-bus modulation or signal bypassing. Once the  $MoD_{enable}$  signal is

enabled, the switching island is active. The light goes cross the waveguide will be routed into the ring. Moreover, when the light is coupled into the racetrack ring, the detection segment (in orange) is able to convert the optical signal into the electrical domain with high efficiency. Note, avoiding the O-E-O conversion when node bypassing is needed is the key to maintain good optical link performance and high energy efficiency.

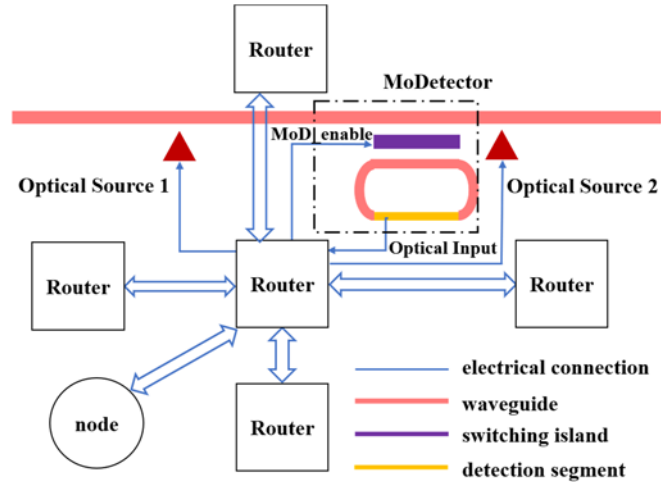


Figure 22. Router connection with both electrical component and HyPPI link.

In addition, there is a control signal to enable the O-E converter inside the router. With the exception of the destination nodes on HyPPI link, other nodes will not expect any signal from the photodetector; however, as long as the nodes are on the active HyPPI link, optical signal might be detected. Hence, this control signal is designed to avoid any unexpected optical input. If the O-E converter is enabled, the packets from the detection segment (yellow) are considered as one of the input signals from other routers/self-node and follows the WReCD routing algorithm as other normal packets.

We implemented the revised router on Xilinx Virtex UltraScale+ FPGA VCU1525. By adding the extra modules shown in Fig. 4, we built a DDDAS router with small overhead in terms of area and power consumption. Only 7.26% logic power and 1.82% static power are required as extra overhead compared to the original router. Only 2.15% of LUT as logic and 4.42% number of registers are considered as additional area on FPGA.

### 2.3.3 Evaluation on Benchmarks

To test our performance, we implemented a full network simulator. Both synthetic benchmarks and several NAS (NASA Advanced Supercomputing) parallel Benchmarks (NPB) on the simulator. To get the best performance, the parameters mentioned in Section 2.3.1 are swept, including the drift value, drift threshold, vote threshold, decay type, decay threshold and decrease value. According to our simulation results, RESET type works best with small decay threshold (equals to 3) for long communication traffic pattern.

**Synthetic Benchmarks.** To validate the simulator, first we tested a synthetic benchmark, named bigX. In this benchmark, the corners of the mesh network generate diagonally traces. We tested different configurations to see how important to the performance for each parameter. Fig.23 depicts all the changes in terms of latency. As the number of packets increase, the traffic pattern intends to be more stable. In all simulations, as long as the traffic patterns are stable enough, it gains benefits from our design, up to 88%. In all the parameter we tested, the *drift vote* plays the most important role in terms of latency. It is because *drift vote* represents how often the system should be reconfigured. In the reconfiguration process, sorting covers a lot of time. The less reconfiguration encountered, the less time consumed. Though the algorithm would like to adapt

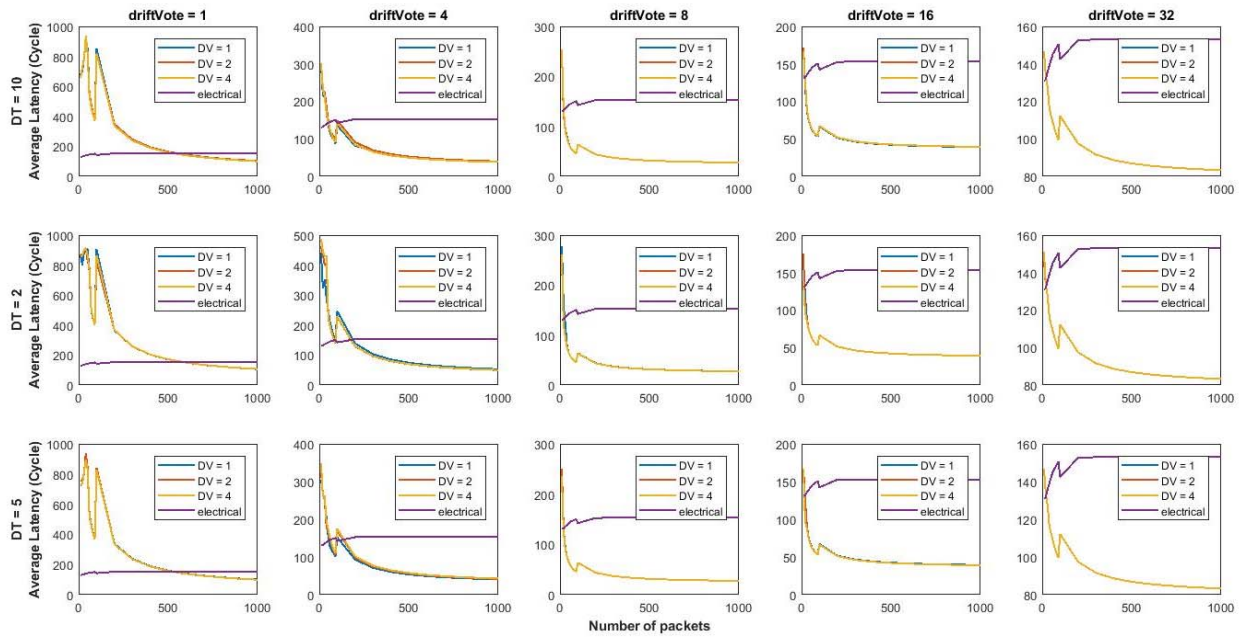


Figure 23. Simulation Results of Running Synthetic Benchmark Big-X with Different Configurations

the traffic pattern during runtime, it is inefficient to reconfigure the HyPPI links if only one router requests.

### 2.3.2 Performance of NPB Benchmarks

A trace-driven simulator is built to test the traffic of NAS parallel benchmark. Several parameters are defined for performance testing. The network change is monitored while the parameters are swept, including *drift value*, *vote threshold*, *reconfiguration threshold*, *decay type*, *decrement value*, *decrease threshold*, and *reset threshold*.

As the *vote threshold* increases, the average number of hops for each traffic trace increases as well, getting close to the XY-routing algorithm. On the other hand, a smaller *vote percentage* reduces the number of hops. By using the optimal parameters setting, our design saves the number of hops for all the benchmarks (**Fig. 24**).

An average of 31% of the total number of hops is saved. The savings reach more than 50% for the FT benchmark, however, a negligible percentage of saving can be seen in LU. This is because the communication patterns in LU are very short, thus we cannot achieve much savings.

The savings of the number of hops represents the upper bound to how much we can save in latency. In fact, the difference between the savings in the number of hops and latency is dominated by the reconfiguration overhead (the number of reconfigurations multiplied by one reconfiguration time). The reconfiguration time is mostly spent in sorting. The smallest time complexity of a snake-like  $N \times N$  mesh is  $O(N)$ .

Here, we report the best time saving for each benchmark (**Fig.25**). We consider two different implementations for our express snake-like link: the HyPPI link and the electrical implementation. We can see that the HyPPI link is about 5% faster than the electrical implementation. However, the electrical link consumes more power due to capacitive effects.



The saving in latency is more than 30% for CG and FT, and around 17% for MG and SP. LU does not gain the benefit from WReCD routing since most of the communication require one hop only,

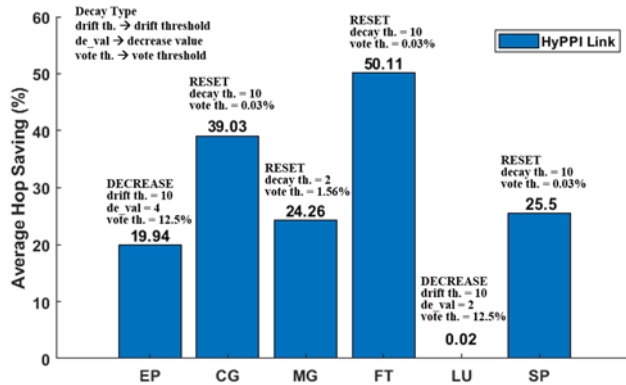


Figure 24. Average Saving for Number of Hops for Testing Benchmarks, as the *Drift Threshold* is 3 on RESET Type with *Reset Threshold* is 2. The *drift vote* is 0.03%

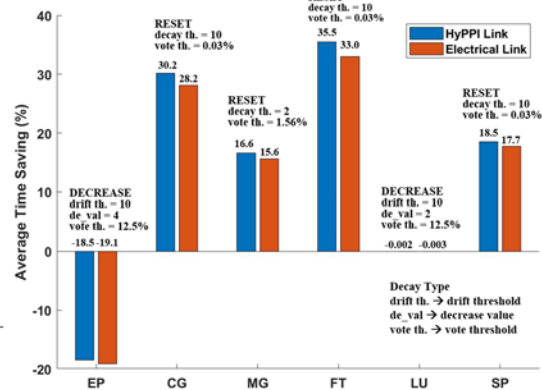


Figure 25. Optimal Average Time Saving for Testing Benchmarks.

similar to the hop saving. WReCD algorithm harms the time saving of EP. By using our algorithm, EP runs 18.5% slower than XY routing. Although it saves 47% number of hops, the reconfiguration overhead is too high because the communication patterns dynamics are too fast which requires a large number of reconfigurations.

In general, WReCD adaptive routing algorithm is suitable for irregular and long communication patterns. In such cases, the reconfiguration time of HyPPI links is more valuable. It provides the express highway links for more frequent long communications. Hence, the time saving increases.

In addition, the time saving will improve as the size of the network increases. Our experimental results show that for the same benchmark going from a 4×4 network to an 8×8 network, we achieve 20% more savings. We even achieve more savings as we move to a 16×16 network. This implies that for future processors with even larger network sizes, our algorithm will achieve even better savings.

## 2.3.4 Hybrid Photonic-Plasmonic Devices Design for NoC

### 2.3.4.1 HPP 5×5 Non-Blocking Broadband Router

With the success of long-haul optical networks, optical interconnects at the board, and even at the chip-level, have become of interest in order to mitigate the processing-to-communication gap. However, the majority of optical network-on-chip (NoC) routers perform their role not exclusively in the photonic domain but often in capacitive-limiting electronics. The later also requires an overhead-heavy optic-electric-optic (O-E-O) conversion. On the other hand, one can perform routing entirely in the electronics. Yet, the known performance bottlenecks of electronic devices, namely mainly delay and power dissipation, and clamping performance. While photonic routers based on microring resonators have been proposed and demonstrated, the high sensitivity (i.e. spectral and amplitude) require dynamic tunability which is both power hungry and relatively slow if high Q-factor rings are used. Hence taken together, optical routing is a) technologically cumbersome, b) latency- and energy-prone mainly due to O-E-O conversion, and c) suffers from

high energy overhead due to signal error correction at the detectors TIA and laser stages, and from thermal tuning in rings-based routers.

In contrast, here we demonstrate an optical router design using a hybrid plasmonic-photon approach and emerging unity-high index tuning materials simultaneously to improve photonic integrated routing performance in all three factors. Cascading a network of these plasmonic  $2\times 2$  switches we can design a compact optical router since the switching length scales inversely with index-change per voltage. In addition, given that the  $2\times 2$  switches are non-resonant devices due to the lossy plasmonic mode, this optical router allows for spectrally broadband operation for WDM applicability. Furthermore, unlike microrings, thermal tuning is not required, thus saving energy consumption.

The fundamental building block of the optical router is a  $2\times 2$  optical switch, namely a voltage-controlled directional coupler whose performance directly impacts the overall performance of the router. To overcome the fundamental and practical drawbacks such as high tuning energy and large on-chip footprint, routing switches utilizing emerging materials beyond silicon, such as ITO, has been studied and carrier-based Drude tail modulation demonstrated. While a physical demonstration of the actual index tuning speed-potential in ITO is still outstanding, we estimate the carrier drift time to be sub-ps given a mobility of  $15\text{ cm}^2/\text{Vs}$  for 10-20 nm thin ITO films. We note that this estimation does not violate physical fundamentals, as the corresponding drift velocity is about a third of ITOs Fermi-velocity. However, based on our previous ITO experimental result, the observed index change was an averaged value for an ITO thickness of 10 nm; meaning the actual index change is higher at near the interface, and lower further away from it. That is, we double the thickness of the ITO layer (20 nm) while biasing it simultaneously from both the top and the bottom with opposite-sign voltages to achieve two accumulation layers at each ITO-insulator surface, which is beneficial for reducing the physical switch length thus enhancing the coupling efficiency discussed below. The selection for ITO as the switching material is based on its unity-strong index tunability and possible CMOS compatibility.

Utilizing hybrid plasmon polaritons (HPPs), we added a tunable ITO layer within the metal-oxide-semiconductor (MOS) structure in order to form an electrical capacitor towards changing the optical mode's index via voltage control (**Fig. 26**). The switch structure includes two bus waveguides, one on each side as the input (port 1 and port 4) and the output (port 2 and port 3) ports of the switch. The center island is the actively index-tunable location of the switch. The active material is "sandwiched" between two oxide layers structure to achieve dual bias operation. The fundamental operation principle of this device is to use the index-tunable active layer (ITO layer) to switch between the CROSS state (light travels from one side of the first bus to the second bus on the other side when bias voltage  $V_{\text{bias}}$  is  $V_0 = 0\text{V}$ ) and the BAR state (light stays within the bus on the same side when bias voltage is  $V_{\text{dd}}$ ) by changing the carrier concentration of the ITO layer, thus further affecting the effective index of the supermodes governing this device; three lowest-order TM modes are spread across the cross-section of this 3-waveguide structure and can be regarded as the supermodes TM1, TM2, and TM3 of the device (**Fig. 26a**). Our final optimized design and resulting performance parameters of the  $2\times 2$  hybrid plasmonic-photon switch are summarized in **Table 1**.



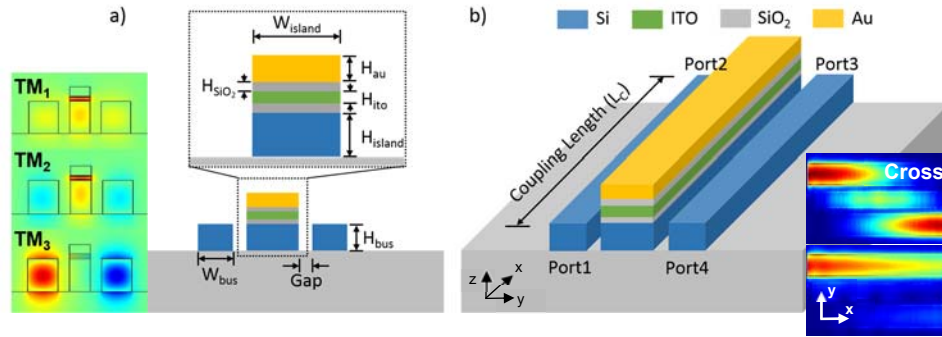


Figure 26. Schematic design of the  $2 \times 2$  hybrid photonic-plasmonic switch using ITO as the active material. The coupling length of the switch is equal to the CROSS state coupling length  $L_c$ . The insets are a) the  $TM_1$ ,  $TM_2$  and  $TM_3$  supermodes of the  $2 \times 2$  ITO switch and b) the electric field results of the device at BAR and CROSS states at 1550 nm wavelength. The length of the ITO switch ( $8.9 \mu\text{m}$ ) in the x-direction is not to scale.  $\lambda = 1550 \text{ nm}$  [14].

The elemental  $2 \times 2$  switches are interconnected with optical waveguides forming a switching fabric such as an  $N \times N$  spatial routing switch or "matrix switch" where  $N$  is the number of input ports, as well as the number of output ports. For such an  $N \times N$  switching network router, there are several practical architectures or layouts (Benes, Clos, etc). Here we have chosen to build the non-blocking router known as the permutation matrix. Generally speaking, the permutation matrix has the advantage that no waveguide crossings (intersections) are used throughout in the matrix, but the matrix has the disadvantage that the overall insertion loss between an input- $i$  and an output- $j$  depends upon the length of the optical path traversed between the two inputs, a length that varies depending upon the specific selected  $i$  and  $j$  pair. In other words, the IL is path dependent.

Table 1 Critical design parameters and performance list of two design cases. The energy consumption is calculated based on capacitor charging energy  $\frac{1}{2} CV^2$ , and the switching time is based on device RC delay.

Parameter	Values
Bus Diameter	400 nm $\times$ 340 nm
Switch Diameter	275 nm $\times$ 340 nm
Gap	150 nm
ITO Height	20 nm
Oxide Height	16 nm
Coupling Length	$8.9 \mu\text{m}$
Capacitance	1.63 fF
Resistance	500 $\Omega$
Bias Voltage	4 Volt
Energy per Switching	13.1 fJ
Switching Time	5.1 ps
BAR Insertion Loss	2.1 dB
CROSS Insertion Loss	0.4 dB
BAR Extinction Ratio	24.2
CROSS Extinction Ratio	9.3

The total number of  $2 \times 2$  switches needed for a non-blocking router scales with  $(N-1)^2/2$ , where  $N$  is an odd number of ports of that router. Thus, as a router for an optical mesh network of a NoC requires 4 ports to connect to the north, south, east and west neighbors, and 1 additional port for connection to the local processing core. This results in, eight  $2 \times 2$  hybrid switches needed to achieve  $5 \times 5$  non-blocking routing functionality that assumes assigning a random input port to a

random output port without disturbing other data streams (**Fig. 27**). We note that other input ports are still able to maintain connections with the remainders of the output ports without affecting the initially set switches. Moreover, self-communication (communication between same input and output port number, resulting in a U-turn) is forbidden because i) it can be achieved with higher energy- and latency- efficiency with other local (electrical) interconnect links, and ii) avoiding self-communication can simplify the router from  $N^2$  number of switches required for all-to-all connection down to only  $(N-1)^2/2$ , which can also reduce the average loss of the router.

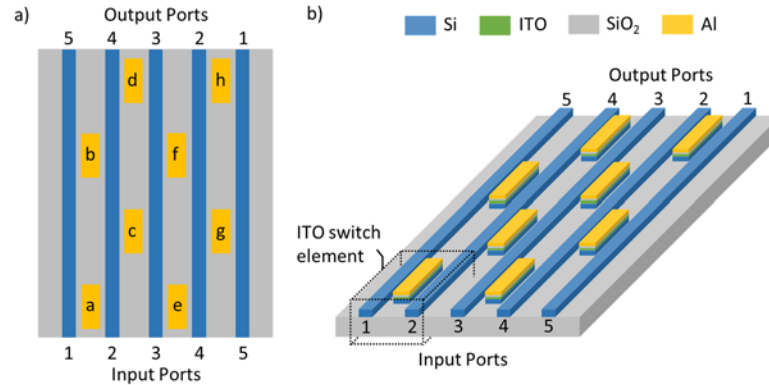


Figure 27. The top view and the schematic plot of the 5×5 Port non-blocking optical router. 8 individual 2×2 ITO switches are placed with certain pattern in order to achieve non-blocking routing function. The length of the ITO switches is not to scale [14].

The operational spectrum results for each output port with respect to cross-coupling from other routing paths are key parameters for signal quality and to assess the WDM ability (**Fig. 28**). For example, configuring the router to establish the following paths: 1 to 2, 2 to 3, 3 to 4, 4 to 5 and 5 to 1, and injecting a unity laser power ( $P_{\text{laser}} = 100\%$  a.u.) from port 1, results in the majority of the signal to be routed to port 2, as designed while the leakage is delivered to the remaining four output ports. The 3 dB spectral (not temporal) bandwidth, i.e. routed signal dropping to -3dB from maximum, is 106 nm wide on average for all 20 different routing paths (130 nm from 1.49 to 1.62  $\mu\text{m}$ ). The broad bandwidth with an average signal-to-noise ratio (SNR) of 123 resulting in an

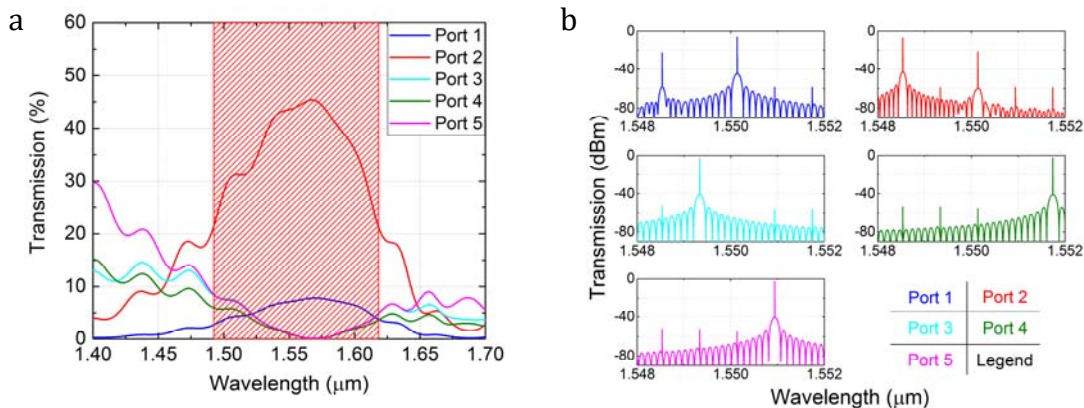


Figure 28. Router performance simulation. The router is configured to route the signal from each port to the next one (i.e. port 1 to port 2, port 2 to port 3, etc.). a) Single-wavelength-single-input from port 1 for operation spectrum testing; b) five-wavelength-five-input with each input port assigned to a wavelength for WDM testing with 0.8 nm wavelength spacing. The shaded area in a) represents the 3dB bandwidth which covers from 1.49  $\mu\text{m}$  to 1.62  $\mu\text{m}$  wavelength range [14].

average channel capacity of  $10 \times 5$  Gbps ( $10 \times 6$  Gbps in Fig. 4a due to above average bandwidth) per routing path based on CWDM standard across the S, C and L bands with 20 nm wavelength spacing (**Fig. 28a**) and 200 Gbps in total if all five ports are used. Here, the SNR is defined as the power ratio between the signal and the light leakage to the other ports. Furthermore, the data capacity can be improved by using DWDM in C band (1530~1560 nm wavelength) with 0.8 nm wavelength spacing which supports 40 wavelengths and results in 400 Gbps data capacity per channel (**Fig. 28b**). Note, this data capacity is calculated based on the standard of the 10 Gigabit Ethernet with 10 Gbps data rate per wavelength. However, the ideal Shannon data capacity based on the device 3 dB bandwidth and average SNR is about 92 Tbps based on Shannon Theory, which shows the maximum capacity of a single routing path with advanced coding strategies such as PAM, QAM, and PWM, etc. We note that this router is WDM capability in that it supports multiple wavelengths per light path. While individual wavelength routing is not possible, multiple pre-multiplexed wavelength channels could be routed jointly, and post-routing demultiplexed. Doing so increases the data capacity of this particular circuit-switched path by a factor equal to the number of wavelengths used (e.g. 100). This could be exploited in applications such as optical residue computing or optical reduction operations.

The port-to-port crosstalk is tested by injecting five light sources in five different wavelengths and we find that the port-to-port crosstalk is at least -13 dB higher than the signal power received by other ports (**Fig. 28b**). Interestingly, different from ring-based WDM optical routers that only support one wavelength at a given time window, the WDM ability of this router allows for multiple wavelengths to be supported simultaneously with no thermal resonance tuning needed.

#### 2.3.4.2 Dual-Functional On-Chip Modulator Detector (MOD)

The goal of MOD is to separate the light modulation and detection from the main bus in order to avoid the unnecessary conversions between electrical and optical domains which leads to extra losses. Since there is no conflict in separating either the modulator from the bus or the detector from the bus, we find positive synergies when both functionalities are combined into a single device as discussed here. Moreover, for network topologies like mesh, ring, and bus, some of the cores require bi-directional communication from both directions of the bus, which requires the MOD design to be symmetric. Based on these requirements, we consider a racetrack ring-based MOD structure that integrates an ‘expanded’ germanium photodetector on the ring via a  $2 \times 2$  hybrid plasmonic 3-waveguide switch to provide modulation functionality in **Fig. 29a**. The  $2 \times 2$  switch consists of a central switching island containing a highly optical index changeable material (indium tin oxide (ITO)) ‘sandwiched’ between two gate oxide layers ( $\text{SiO}_2$ ) to form a metal-oxide-ITO-oxide-semiconductor capacitive heterostructure; whereas the detector has a germanium block on top of the racetrack waveguide with that part of the silicon etched down to 100 nm for better light mode overlap with the high absorption region in **Fig. 29b**.

In order to obtain the bi-functionality of an optical transceiver (i.e. encoding, detection), a bias voltage is applied to MOD at the switch, and detector depending on the desired function; configuring the switch in the Bar states encodes a logic ‘1’ onto the downstream bus for an unmodulated light beam arriving at MOD represented in **Fig. 30d**. At the same time, a modulated signal arriving at MOD can be captured at the detector when the switch is in the Cross state in **Fig. 30(e)**. In this way, each MOD node in the NoC can act as either as a transmitter or receiver depending on the system’s demands.

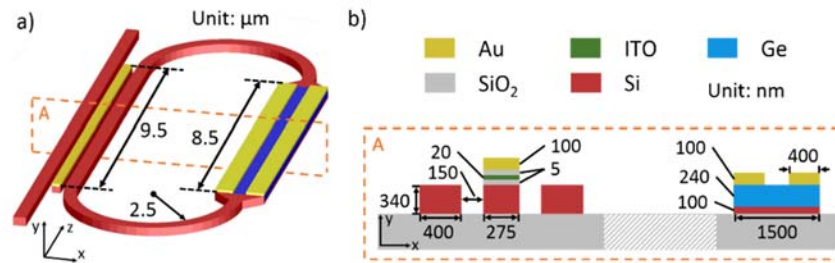


Figure 29. Schematic of the MODetector concept. a) 3D overview of MOD with the ITO hybrid switch on the left and Ge photodetector on the right. b) The cross-section of MOD at A plane. Both a) and b) are color-coded and sharing the same legend on the top-right. All the parameters are optimized for the highest coupling efficiency [15].

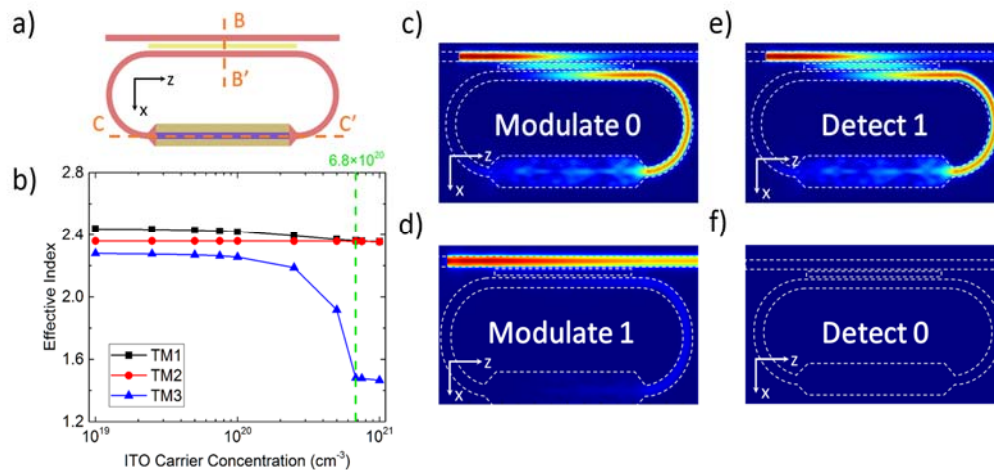


Figure 30. The switch analysis at the Cross and the Bar states. a) The top view of the MOD with the same color coding as Fig. 26. b) Fundamental TM mode effective indices change of the 3-waveguide switch at the cross-section (BB') based on ITO carrier concentrations. c)-f) The FDTD simulations of all four functionalities at different switch and detector state combinations: c) switch OFF, detector OFF; d) switch ON, detector OFF; e) switch OFF, detector ON; f) switch OFF, detector ON. All simulations are based on 1550 nm light source. The ITO refractive indices are calculated based on the Drude model.  $V_{bias}=V_{dd}=4V$ . Note, the MODetector is simulated in 3D using Lumerical FDTD software as a complete device [15].

For regular encoding operation, this dual functionality is used no time-concurrently. Interestingly, a time concurrent operation of MOD creates a copy of the data signal, which may have relevance for cyber security applications, yet, this case is not considered in this work. Placing the switch into the Cross state (0V bias voltage) allows dropping the optical signal from the bus to the racetrack ring, which enables three operation modes, namely modulate '0', detect '0', and detect '1' shown in Figs. 30(c), 30(e), and 30(f). For both detection modes, the detector must always be ON in order to generate photocurrent. On the other hand, for both modulation modes, the detector needs to be OFF in order to avoid any false photocurrent at the modulate '0' state as well as saving the energy. Note, the light will still be absorbed by the detector in this case. For regular operation, independent biasing is required, which eliminates the need for coordination logic-circuitry.

By integrating a hybrid photonic-plasmonic switch with a Germanium-based photodetector into one single device, we design a dual-function modulator-detector. This integrated device is able to detect optical signals up to 28 GHz and generate on-off keying signals up to 100 GHz. Based on

the symmetric design, it enables bi-directional all-to-all communication between multiple communication cores with only one bus waveguide, which significantly reduces the area for inter-chip connections. The performance shows over 10 dB extinction ratio and 0.7 A/W responsivity for the modulator and detector, respectively. This dual-functional device acts an optical transceiver capable of both sending and receiving optical data signals in optical networks and communications and could potentially be used as a reconfigurable optical element in analog photonic-optical compute engines and accelerators.

### 3. List of peer-reviewed Publications from this project (Journals and Proceedings)

- 1) S. Sun and V. J. Sorger, "Photonic-Plasmonic Hybrid Interconnects: a Low-latency Energy and Footprint Efficient Link," in *Advanced Photonics 2015*, OSA Technical Digest (Optical Society of America, 2015), paper IW2A.1.
- 2) S. Sun, A. Badawy, T. El-Ghazawi, V. J. Sorger, "The Case for Hybrid Photonic Plasmonic Interconnects (HyPPIs): Low-Latency Energy-and-Area-Efficient On-Chip Interconnects", *IEEE Photonics Journal*, 7, 6 (2015).
- 3) S. Sun, A. Badawy, V. Narayana, T. El-Ghazawi, and V. J. Sorger, "Bit Flow Density (BFD): An Effective Performance FOM for Optical On-chip Interconnects," in *Advanced Photonics 2016 (IPR, NOMA, Sensors, Networks, SPPCom, SOF)*, OSA technical Digest (Optical Society of America, 2016), paper ITu2B.6.
- 4) S. Sun, A. Badawy, V. Narayana, T. El-Ghazawi, and V. J. Sorger, "Bit Flow Density (BFD): An Effective Performance FOM for Optical On-chip Interconnects," in *Conference on Lasers and Electro-Optics*, OSA Technical Digest (2016) (Optical Society of America, 2016), paper JW2A.135.
- 5) S. Sun, A.-H. A. Badawy, V. Narayana, T. El-Ghazawi, V. J. Sorger, "Low latency, area, and energy efficient Hybrid Photonic Plasmonic on-chip Interconnects (HyPPI)" *Proc. SPIE 9753*, Optical Interconnects XVI, 97530A (2016).
- 6) Liu, K., Sun, S., Majumdar, A. and Sorger, V.J., 2016. Fundamental scaling laws in nanophotonics. *Scientific reports*, 6, p.37419.
- 7) Vikram K. Narayana, Shuai Sun, Abdel-Hameed A. Badawy, Volker J. Sorger, and Tarek El-Ghazawi. "MorphoNoC: Exploring the design space of a configurable hybrid NoC using nanophotonics." *Microprocessors and Microsystems* 50 (2017): 113-126.
- 8) Sun, S., Narayana, V.K., El-Ghazawi, T. and Sorger, V.J., 2017, May. Chasing Moore's law with CLEAR. In *CLEO: QELS\_Fundamental Science* (pp. JW2A-138). Optical Society of America.
- 9) Sun, S., Narayana, V., El-Ghazawi, T. and Sorger, V.J., 2017, July. CLEAR: A Holistic Figure-of-Merit for Electronic, Photonic, Plasmonic and Hybrid Photonic-Plasmonic Compute System Comparison. In *Optical Sensors* (pp. JTu4A-8). Optical Society of America.
- 10) Sun, S., Narayana, V., El-Ghazawi, T. and Sorger, V.J., 2017, July. High Performance Photonic-Plasmonic Optical Router: A Non-blocking WDM Routing Device for Optical Networks. In *Photonics in Switching* (pp. PM2D-3). Optical Society of America.
- 11) Narayana, V.K., Sun, S., Mehrabian, A., Sorger, V.J. and El-Ghazawi, T., 2017, August. HyPPI NoC: Bringing Hybrid Plasmonics to an Opto-Electronic Network-on-Chip. In *Parallel Processing (ICPP)*, 2017 46th International Conference on(pp. 131-140). IEEE.
- 12) Sun, S., Zhang, R., Peng, J., Narayana, V., Tarek, E.G. and Sorger, V.J., 2017, September. Hybrid Photonic-Plasmonic Directional Coupler Enabled Optical Transceiver. In *Laser Science* (pp. JW4A-55). Optical Society of America.

- 13) Sun, S., Narayana, V., Mehrabian, A., Zhang, R., Tarek, E.G. and Sorger, V.J., 2017, September. Holistic Performance-Cost Metric for Post Moore Era. In *Frontiers in Optics* (pp. JT2A-24). Optical Society of America.
- 14) Sun, Shuai, Vikram K. Narayana, Ibrahim Sarpkaya, Joseph Crandall, Richard A. Soref, Hamed Dalir, Tarek El-Ghazawi, and Volker J. Sorger. "Hybrid photonic-plasmonic nonblocking broadband  $5 \times 5$  router for optical networks." *IEEE*
- 15) Sun, Shuai, Ruoyu Zhang, Jiaxin Peng, Vikram K. Narayana, Hamed Dalir, Tarek El-Ghazawi, and Volker J. Sorger. "MO detector (MOD): a dual-function optical modulator-detector for on-chip communication." *Optics express* 26, no. 7 (2018): 8252-8259.
- 16) Mehrabian, Armin, Shuai Sun, Vikram K. Narayana, Jeff Anderson, Jiaxin Peng, Volker Sorger, and Tarek El-Ghazawi. "D 3 NoC: a dynamic data-driven hybrid photonic plasmonic NoC." In *Proceedings of the 15th ACM International Conference on Computing Frontiers*, pp. 220-223. ACM, 2018.
- 17) Sun, Shuai, Ruoyu Zhang, Jiaxin Peng, Vikram K. Narayana, Hamed Dalir, Tarek El-Ghazawi, and Volker J. Sorger. "An On-Chip Integrated Dual-Functional Modulator-Detector for Optical Communication." In *CLEO: QELS Fundamental Science*, pp. JW2A-3. Optical Society of America, 2018.
- 18) Sun, Shuai, Ruoyu Zhang, Jiaxin Peng, Hamed Dalir, Tarek El-Ghazawi, and Volker J. Sorger. "Dual-Functional Integrated Modulator-Detector for Optical Communication On-Chip." In *Laser Science*, pp. JT2A-115. Optical Society of America, 2018.
- 19) Jiaxin Peng, Yousra Alkabani, Erwan Favry, Armin Mehrabian, Shuai Sun, Sorger J. Volker, and Tarek El-Ghazawi. "Adaptive Routing for Hybrid Photonic-Plasmonic (HyPPI) using DDDAS on the Chip." (to be submitted)

[ ]  
SEP

#### 4. Patents

- 1) Full (#15/194,119): Hybrid Photonic Plasmonic Interconnects with Intrinsic and Extrinsic Modulation Option.
- 2) Full (#15/888,862) Hybrid Photonic Plasmonic Non-blocking Wide Spectrum WDM On-chip Router.
- 3) Provisional (#62/633,382) Dual Functional Broadband On-Chip Optical Modulator Detector (MODetector)

#### 5. Talks/Presentations/Colloquia Delivered

- El-Ghazawi presented at the Office of Science in the DoE; gave multiple related talks including two keynotes at IEEE International Conferences, IEEE CPSCOM June 2017 in Exeter UK, IEEE HPCC December 2016 in Sydney Australia, SOCC 2018, and HPCC 2018; is selected as IEEE Computer Society distinguished speaker giving talks about this work as one of the topics in IEEE Chapters Worldwide.
- Sorger gave many related talks in many visible venues including OSA Advanced Photonics Congress, New Orleans, July 2017, and SPIE Optics and Photonics, San Diego, California, September 2016.

- Armin Mehrabian presented at InfoSymbiotics/DDDAS Conference 2017, AFOSR Mid-Atlantic Student Research Day 2017 and SOCC 2018.
- Shuai Sun gave 8 talks at different venues.