COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

COLUMBIA NANO INITIATIVE

Yuyang Wang, Ph.D.
Postdoctoral Research Scientist
Columbia Nano Initiative
Columbia University in the City of New York
530 West 120th Street, New York, NY 10027
+1 (781) 428-0408
yw3831@columbia.edu

January 3, 2024

Faculty Search Committee
Department of Computer Science and Engineering
Pennsylvania State University
W209 Westgate Building
University Park, PA 16802

Dear Members of the Faculty Search Committee,

Re: Application for the Tenure-Track Assistant Professor Position

I am writing in response to the advertised faculty position (**Tenure-Track Assistant Professor—Job # REQ_0000051058**) in the Department of Computer Science and Engineering at the Pennsylvania State University. I am currently a Postdoctoral Research Scientist at the Columbia Nano Initiative of the Columbia University under the mentorship of Prof. Keren Bergman. My Ph.D. in Electrical and Computer Engineering was awarded by the University of California, Santa Barbara in 2021, where I was co-advised by Prof. Kwang-Ting Cheng and Prof. John E. Bowers. My research portfolio, which is both diverse and synergistic, focuses on silicon photonics (SiPh) optical interconnects, with significant implications for computer systems and architecture, particularly bridging integrated silicon photonics and hardware-software co-design for AI/machine learning applications. I believe that my interdisciplinary research background positions me well to contribute to your department's ongoing excellence.

The prevalence of communication bottleneck in the era of data ubiquity—emerged in modern distributed computing infrastructures and exacerbated by the pervasiveness of data-intensive AI/machine learning applications—drives my research that seeks **transformative connectivity solutions** for future system scalability. I aim to cultivate the potential of CMOS-compatible silicon photonics to develop ultra high-bandwidth and energy-efficient optical interconnects. Building upon a unique link architecture that facilitates unprecedented channel parallelism, I envision system-level breakthroughs that boost the density, adaptability, and functionalities of optical interconnects amidst the ever-evolving data landscape.

My Ph.D. research laid the groundwork for this vision, **vertically integrating design enablement technologies** for integrated silicon photonics across device, circuit, and system levels, including modeling, simulation, and variation mitigation techniques to democratize the design of yield- and performance-optimized SiPh devices and circuits. The research resulted in publications at leading design automation conferences like DAC and ICCAD, esteemed photonics venues like OFC and JLT, and a book chapter in press with Springer, bridging electronics and photonics research communities.

Orthogonally expanding on my Ph.D. work, my postdoctoral work at the Columbia University focuses on the design and implementation of massively scalable SiPh chip I/Os, **horizontally integrating optimization techniques** across the entire design cycle, from chip layout to post-fabrication tuning and runtime reconfiguration. I spearheaded two generations of the SiPh transceiver chips—co-designed with both academia and industry partners for 3D-integrated electronic drivers and advanced packaging—integrating over 2,000 microresonators into an 8 mm × 8 mm footprint to achieve 16 Tbps chip I/O bandwidth with sub-pJ/b energy consumption. The featured link architecture played a pivotal role in securing a $35M SRC JUMP 2.0 grant involving 23 principal investigators, to which I have contributed through proposal writing and ongoing research efforts. An invited paper detailing the chip's design and characterization is

currently under review for CICC 2024, and a manuscript focusing on the end-to-end link demonstration is slated for an invited submission to *Nature Communications Physics*.

Moving forward, I anticipate an even more diverse and dynamic data landscape, shaped by the growing heterogeneity of computing applications. Emerging privacy-focused frameworks such as federated learning emphasize the exchange of models over data, which—coinciding with the expansion of large models like GPTs—further pushes the system bottleneck from computation to communication capabilities. This shift necessitates the next generation of **reconfigurable system connectivity** that can adapt on-demand to varying traffic needs while maintaining high bandwidth and energy efficiency. My foundation work in system reconfiguration techniques, built upon the scalable link architecture, will be the cornerstone of my ongoing research endeavor to equip future computing infrastructures for the ever-evolving data context.

I also foresee opportunities arising from a deeper integration of silicon photonics within computing sockets, a direction that my postdoctoral research is actively pursuing. Anticipating advances in dense 3D optical I/Os—a concept that I helped formulate for showcasing at the 2023 DARPA ERI Summit—I envision **innovative system architectures** that leverage the manifolded reach of on-board optical links. This approach could eliminate the reliance on silicon interposers, currently nearing their performance limits, and surpass the density constraints of conventional fiber optics, thus catalyzing new computational paradigms and interconnect functionalities with redefined chip-to-chip connectivity.

My research agenda is situated at the system level, deeply engaged with the interdisciplinary nexus of electronics/photonics, devices/systems, and design/applications. I am excited about the prospect of collaborating with the diverse faculty in the Department of Computer Science and Engineering, who are leading in areas such as *computer architectures and networked systems* (Chitaranjan Das, Mahmut Kandemir, Vijaykrishnan Narayanan, John (Jack) Sampson, Anand Sivasubramaniam), *systems for big data/AI* (George Kesidis, Daniel Kifer, Dong Xie), as well as areas in Electrical Engineering such as *integrated photonics and optoelectronics* (Iam-Choon Khoo, Xingjie Ni, Shizhuo Yin) and *high-performance integrated circuits* (Mehdi Kiani, Wooram Lee), to tackle the grand challenge of revolutionizing connectivity in future computing infrastructures across the full system stack.

My training in California and New York has enriched me with diverse and inclusive academic experiences, fueling my passion to contribute to the Pennsylvania State University's community through not only research, but also teaching, mentoring, and outreach. As a teaching assistant and guest lecturer, I have taught at both undergraduate and graduate levels, focusing my teaching philosophy on the balance of engagement, conveyance, and inspiration through visualization. I also value the role of mentoring in fostering students' intellectual growth beyond technical skills and have actively mentored students tailored to their research interests and career aspirations. My advocacy for diversity and inclusion is reflected in my experiences of working with—as well as my efforts to promote the accessibility of STEM education to—underrepresented groups in my endeavors.

I am enthusiastic about joining the Department of Computer Science and Engineering at the Pennsylvania State University to further its excellence in research and education. Enclosed are my application materials, and I am glad to provide additional information if required. I look forward to the possibility of discussing my qualifications and research vision with you. Thank you for considering my application, and I anticipate your response.

Yours sincerely,

Yuyang Wang

encl: Curriculum Vitae
Research Statement
Teaching Statement

Yuyang Wang

Columbia Nano Initiative, Columbia University
530 West 120th St, New York, NY 10027
+1 (781) 428-0408 yw3831@columbia.edu wangyy.phd

Research Interests

My research centers on enabling the design of system connectivity that is ultra high-bandwidth, energy-efficient, and adaptable in today's data-driven world. Tackling the significant communication bottlenecks found in modern distributed computing infrastructures, exacerbated by data-intensive AI and machine learning workloads, I am committed to designing—as well as democratizing the design process for—scalable architectures that maximize the capabilities of integrated silicon photonics, bridging communication and computation. With a deeper integration of photonics within computing sockets, I envision a paradigm shift in computing architectures that promises unparalleled interconnection bandwidth density, versatility, and functionalities.

Current Appointment

Columbia University in the City of New York — New York, NY, USA
Postdoctoral Research Scientist, Columbia Nano Initiative — 2021–Present

- Mentored by Prof. Keren Bergman

Education

University of California, Santa Barbara — Santa Barbara, CA, USA
Ph.D. in Electrical and Computer Engineering — 2018–2021

- Co-advised by Prof. Kwang-Ting Cheng and Prof. John E. Bowers

University of California, Santa Barbara — Santa Barbara, CA, USA
M.S. in Electrical and Computer Engineering — 2015–2018

Tsinghua University — Beijing, China
B.Eng. in Electronic Engineering — 2011–2015

Professional Experience

Semiconductor Research Corporation (SRC) Research Scholars Program
Research Scholar, Center for Ubiquitous Connectivity (CUbiC) under SRC JUMP 2.0 — 2023–Present

- Contributed to the writing and visualization of the proposal that led to the award of $35M JUMP 2.0 grant for the CUbiC Center
- Co-led the CUbiC Scholar Leadership Council to organize monthly workshops with industry liaisons

Hong Kong University of Science and Technology — Hong Kong SAR, China
Postgraduate Visiting Intern, Department of Electrical and Computer Engineering — Aug. 2019–Dec. 2019

Cadence Design Systems, Inc. — San Jose, CA, USA
Design Engineering Intern - Photonics, Custom IC & PCB Group — Jun. 2018–Sep. 2018

Rice University — Houston, TX, USA
Student Intern, Department of Electrical and Computer Engineering — Jul. 2014–Sep. 2014

Publications

Book Chapter

B1 **Y. Wang**, Z. Zhang, J. E. Bowers, and K.-T. Cheng, "Silicon photonics optical interconnects for data-centric artificial intelligence applications: A design automation perspective," in *Frontiers of Electronic Design (FED)*, A. Iranmanesh, Ed., in press, Cham: Springer International Publishing.

Refereed Journal Articles

J1 A. James, A. Rizzo, **Y. Wang**, A. Novick, S. Wang, R. Parsons, K. Jang, M. Hattink, and K. Bergman, "Process Variation-Aware Compact Model of Strip Waveguides for Photonic Circuit Simulation," *Journal of Lightwave Technology*, pp. 1–14, 2023. doi 10.1109/JLT.2023.3238847.

J2 A. Novick, A. James, L. Y. Dai, Z. Wu, A. Rizzo, S. Wang, **Y. Wang**, M. Hattink, V. Gopal, K. Jang, R. Parsons, and K. Bergman, "High-bandwidth density silicon photonic resonators for energy-efficient optical interconnects," *Applied Physics Reviews*, vol. 10, no. 4, p. 041 306, Nov. 2023. doi 10.1063/5.0160441.

J3 Z. Wu, L. Y. Dai, **Y. Wang**, S. Wang, and K. Bergman, "Flexible silicon photonic architecture for accelerating distributed deep learning," *Journal of Optical Communications and Networking*, 2023, to appear.

J4 **Y. Wang**, P. Sun, J. Hulme, M. A. Seyedi, M. Fiorentino, R. G. Beausoleil, and K.-T. Cheng, "Energy Efficiency and Yield Optimization for Optical Interconnects via Transceiver Grouping," *Journal of Lightwave Technology*, vol. 39, no. 6, pp. 1567–1578, Mar. 2021. doi 10.1109/JLT.2020.3039489.

J5 Z. Zhang, R. Wu, **Y. Wang**, C. Zhang, E. J. Stanton, C. L. Schow, K.-T. Cheng, and J. E. Bowers, "Compact Modeling for Silicon Photonic Heterogeneously Integrated Circuits," *Journal of Lightwave Technology*, vol. 35, no. 14, pp. 2973–2980, Jul. 2017. doi 10.1109/JLT.2017.2706721.

Refereed Conference Proceedings

C1 A. Novick, M. Hattink, A. Rizzo, **Y. Wang**, V. Gopal, S. Wang, R. Parsons, and K. Bergman, "Integrated photonic resonant modulator-based equalization and optimization for DWDM," in *Optical Fiber Communication Conference (OFC) 2024*, to appear, Optica Publishing Group, 2024.

C2 S. Wang, **Y. Wang**, X. Meng, K. Hosseini, T. T. Hoang, and K. Bergman, "Automated tuning of ring-assisted MZI–based interleaver for DWDM systems," in *Optical Fiber Communication Conference (OFC) 2024*, to appear, Optica Publishing Group, 2024.

C3 Z. Wu, R. Parsons, S. Wang, **Y. Wang**, and K. Bergman, "Wavelength reconfigurable transceiver for multi-interface compute accelerator networks," in *Optical Fiber Communication Conference (OFC) 2024*, to appear, Optica Publishing Group, 2024.

C4 G. Michelogiannakis, Y. Arafa, B. Cook, L. Y. Dai, A.-H. Hameed Badawy, M. Glick, **Y. Wang**, K. Bergman, and J. Shalf, "Efficient Intra-Rack Resource Disaggregation for HPC Using Co-Packaged DWDM Photonics," in *2023 IEEE International Conference on Cluster Computing (CLUSTER)*, Santa Fe, NM, USA: IEEE, Oct. 2023, pp. 158–172. doi 10.1109/CLUSTER52292.2023.00021.

C5 S. Wang, A. Novick, A. Rizzo, R. Parsons, S. Sanyal, K. J. McNulty, B. Y. Kim, Y. Okawachi, **Y. Wang**, A. Gaeta, M. Lipson, A. Gaeta, M. Lipson, and K. Bergman, "Integrated, Compact, and Tunable Band-Interleaving of a Kerr Comb Source," en, in *CLEO 2023*, San Jose, CA: Optica Publishing Group, 2023, STh3J.6. doi 10.1364/CLEO_SI.2023.STh3J.6.

C6 **Y. Wang**, S. Wang, A. Novick, A. James, R. Parsons, A. Rizzo, and K. Bergman, "Dispersion-Engineered and Fabrication-Robust SOI Waveguides for Ultra-Broadband DWDM," en, in *Optical Fiber Communication Conference (OFC) 2023*, San Diego California: Optica Publishing Group, 2023, Th3A.4. doi 10.1364/OFC.2023.Th3A.4.

C7 A. James, **Y. Wang**, A. Rizzo, and K. Bergman, "Flexible, Process-Aware Compact Model of Effective Index in Silicon Waveguides for Commercial Foundries," in *2022 International Conference on Numerical Simulation of Optoelectronic Devices (NUSOD)*, Turin, Italy: IEEE, Sep. 2022, pp. 173–174. doi 10.1109/NUSOD54938.2022.9894784.

C8 **Y. Wang** and K.-T. Cheng, "Traffic-Adaptive Power Reconfiguration for Energy-Efficient and Energy-Proportional Optical Interconnects," in *2021 IEEE/ACM International Conference On Computer Aided Design (ICCAD)*, Munich, Germany: IEEE, Nov. 2021, pp. 1–9. doi 10.1109/ICCAD51958.2021.9643475.

C9 **Y. Wang**, J. Hulme, P. Sun, M. Jain, M. A. Seyedi, M. Fiorentino, R. G. Beausoleil, and K.-T. Cheng, "Characterization and Applications of Spatial Variation Models for Silicon Microring-Based Optical Transceivers," in *2020 57th ACM/IEEE Design Automation Conference (DAC)*, San Francisco, CA, USA: IEEE, Jul. 2020, pp. 1–6. doi 10.1109/DAC18072.2020.9218608.

C10 **Y. Wang** and K.-T. Cheng, "Task Mapping-Assisted Laser Power Scaling for Optical Network-on-Chips," in *2019 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, Westminster, CO, USA: IEEE, Nov. 2019, pp. 1–6. doi 10.1109/ICCAD45719.2019.8942146.

C11 **Y. Wang**, M. A. Seyedi, J. Hulme, M. Fiorentino, R. G. Beausoleil, and K.-T. Cheng, "Bidirectional tuning of microring-based silicon photonic transceivers for optimal energy efficiency," in *Proceedings of the 24th Asia and South Pacific Design Automation Conference*, Tokyo Japan: ACM, Jan. 2019, pp. 370–375. doi 10.1145/3287624.3287649.

C12 **Y. Wang**, M. A. Seyedi, R. Wu, J. Hulme, M. Fiorentino, R. G. Beausoleil, and K.-T. Cheng, "Energy-efficient channel alignment of DWDM silicon photonic transceivers," in *2018 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, Dresden, Germany: IEEE, Mar. 2018, pp. 601–604. doi 10.23919/DATE.2018.8342079.

C13 R. Wu, M. A. Seyedi, **Y. Wang**, J. Hulme, M. Fiorentino, R. G. Beausoleil, and K.-T. Cheng, "Pairing of microring-based silicon photonic transceivers for tuning power optimization," in *2018 23rd Asia and South Pacific Design Automation Conference (ASP-DAC)*, Jeju: IEEE, Jan. 2018, pp. 135–140. doi 10.1109/ASPDAC.2018.8297295.

C14 R. Wu, **Y. Wang**, Z. Zhang, C. Zhang, C. L. Schow, J. E. Bowers, and K.-T. Cheng, "Compact modeling and circuit-level simulation of silicon nanophotonic interconnects," in *Design, Automation & Test in Europe Conference & Exhibition (DATE), 2017*, Lausanne, Switzerland: IEEE, Mar. 2017, pp. 602–605. doi 10.23919/DATE.2017.7927057.

C15 A. Ghofrani, M. A. Lastras-Montaño, **Y. Wang**, and K.-T. Cheng, "In-place Repair for Resistive Memories Utilizing Complementary Resistive Switches," in *Proceedings of the 2016 International Symposium on Low Power Electronics and Design*, San Francisco Airport CA USA: ACM, Aug. 2016, pp. 350–355. doi 10.1145/2934583.2934590.

C16 C. Xu, F. X. Lin, **Y. Wang**, and L. Zhong, "Automated OS-level Device Runtime Power Management," in *Proceedings of the Twentieth International Conference on Architectural Support for Programming Languages and Operating Systems*, Istanbul Turkey: ACM, Mar. 2015, pp. 239–252. doi 10.1145/2694344.2694360.

Invited Conference Papers

I1 **Y. Wang**, A. Novick, R. Parsons, S. Wang, K. Jang, A. James, M. Hattink, V. Gopal, A. Rizzo, C.-P. Chiu, K. Hosseini, T. T. Hoang, and K. Bergman, "Scalable architecture for sub-pJ/b multi-Tbps comb-driven DWDM silicon photonic transceiver," in *Next-Generation Optical Communication: Components, Sub-Systems, and Systems XII*, G. Li, K. Nakajima, and A. K. Srivastava, Eds., San Francisco, United States: SPIE, Mar. 2023, p. 55. doi 10.1117/12.2649506.

I2 **Y. Wang**, L. Shao, M. A. Lastras-Montano, and K.-T. Cheng, "Taming Emerging Devices' Variation and Reliability Challenges with Architectural and System Solutions [Invited]," in *2019 IEEE 32nd International Conference on Microelectronic Test Structures (ICMTS)*, Kita-Kyushu City, Fukuoka, Japan: IEEE, Mar. 2019, pp. 90–95. doi 10.1109/ICMTS.2019.8730924.

Under Review and In Preparation

P1 **Y. Wang**, S. Wang, R. Parsons, A. Novick, V. Gopal, K. Jang, A. Rizzo, C.-P. Chiu, K. Hosseini, T. T. Hoang, S. Shumarayev, and K. Bergman, "Silicon photonics chip I/O for ultra high-bandwidth and energy-efficient die-to-die connectivity," in *2024 IEEE Custom Integrated Circuits Conference (CICC)*, invited, under review, IEEE, 2024.

P2 **Y. Wang**, S. Wang, R. Parsons, S. Sanyal, A. Novick, A. Rizzo, K. Jang, V. Gopal, K. J. McNulty, B. Y. Kim, Y. Okawachi, C.-P. Chiu, K. Hosseini, T. T. Hoang, S. Shumarayev, M. Lipson, A. Gaeta, and K. Bergman, "Scalable co-packaged dwdm silicon photonics chip i/o driven by microresonator Kerr frequency combs," *Nature Communications Physics*, 2024, invited, in preparation.

Talks and Presentations

Invited talk SPIE Photonics West, San Francisco, CA, USA — Jan. 2023

Scalable Architecture for Sub-pJ/b Multi-Tbps Comb-Driven DWDM Silicon Photonic Transceiver

Poster Ph.D. Forum, ACM/IEEE Design Automation Conference (DAC), online virtual event — Jun. 2020

Design and Optimization of Variation-Aware Runtime-Reconfigurable Optical Interconnects

Invited talk Optical/Photonic Interconnects for Computing Systems (OPTICS) workshop, Dresden, Germany — Mar. 2018

Optimal Pairing and Non-Uniform Channel Alignment of Microring-based Transceivers for Comb Laser-Driven DWDM Silicon Photonics

Invited talk ECE Departmental Seminar, Hong Kong University of Science and Technology, Hong Kong SAR, China — Jan. 2018

Variation-Aware Modeling and Design of Silicon Photonic Systems

Poster Optical/Photonic Interconnects for Computing Systems (OPTICS) workshop, Lausanne, Switzerland — Mar. 2017

Variation-Aware Modeling and Design of Nanophotonic Interconnects

Leadership

Center for Ubiquitous Connectivity (CUbiC) — SRC JUMP 2.0

- Co-led the CUbiC Scholar Leadership Council to organize monthly workshops with industry liaisons — 2023–Present
- Co-led the Socket-to-Socket Distributed AI/ML/HPC Fabric Platform (SoSFab) research task for system testbed development — 2023–Present

Columbia University in the City of New York — New York, NY, USA

- Led the photonic integrated circuit (PIC) team, supervised by Prof. Keren Bergman, under the DARPA CHIPS program — 2021-Present
- Co-led the design aggregation of a custom 300 mm full-wafer run with AIM Photonics involving multiple internal/external riders — 2023

Professional Service

Journal Reviewer

- Nature Nanotechnology
- IEEE Journal on Selected Areas in Communications
- IEEE Transactions on Computers
- IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems
- IEEE Transactions on Very Large Scale Integration (VLSI) Systems
- IEEE Access

Conference External Reviewer

– IEEE International Symposium on High-Performance Computer Architecture (HPCA) 2020

Textbook Translation

T1 C. Hawkins, J. Segura, and P. Zarkesh-Ha, *CMOS Digital Integrated Circuits: A First Course (Chinese Edition)*, trans. by **Y. Wang** and Y. Yin. China Machine Press, 2016, original work published by the Institution of Engineering and Technology (IET) in 2013.

T2 S. Kundu and A. Sreedhar, *Nanoscale CMOS VLSI Circuits: Design for Manufacturability (Chinese Edition)*, trans. by **Y. Wang** and W. Xie. China Science Publishing, 2014, original work published by McGraw-Hill Education in 2010.

Teaching Experience

Columbia University in the City of New York New York, NY, USA

Guest Lecturer Spring 2023 ELEN 9403: Seminar in Photonics (Graduate-level)

University of California, Santa Barbara Santa Barbara, CA, USA

Teaching Assistant Winter 2019 ECE 153B: Sensor & Peripheral Interface Design (Undergraduate-level)

Student Mentoring

Songli Wang
Ph.D. Student at Columbia University
Scalable link architectures, automated post-fabrication tuning
Resulting joint publication(s): [C2], [P1], [P2], [I1]

Zhenguo Wu
Ph.D. Student at Columbia University
Reconfigurable architecture for optically connected systems
Resulting joint publication(s): [C3], [J3]

Robert Parsons
Ph.D. Student at Columbia University
Compact modeling of silicon photonic devices and circuits
Resulting joint publication(s): [P1], [I1]

Aneek E. James
Ph.D. Student at Columbia University, now with Draper Laboratory
Wafer-scale process variation extraction and characterization
Resulting joint publication(s): [J1], [C7]

Kaylx Jang
Ph.D. Student at Columbia University
Dispersion-engineered and fabrication-robust (de-)interleavers
Resulting joint publication(s): [C6]

Max Haimowitz
Ph.D. Student at Columbia University
Scripted and automated large-scale silicon photonics chip layout
> 80 mm²/reticle fully-scripted layout on a 300 mm full-wafer run

Awards and Honors

Graduate Fellowship, Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA, USA 2015

Outstanding Thesis Award, Department of Electronic Engineering, Tsinghua University, Beijing, China 2015

Scholarship for Sports Excellence, Department of Electronic Engineering, Tsinghua University, Beijing, China 2014

References

Ⓐ: from academia; Ⓘ: from industry

Ⓐ **Keren Bergman**
Charles Batchelor Professor of Electrical Engineering
Columbia University in the City of New York
bergman@ee.columbia.edu

Ⓐ **Kwang-Ting Cheng**
Vice-President for Research and Development, Chair Professor
Hong Kong University of Science and Technology
timcheng@ust.hk

Ⓐ **John E. Bowers**
Director of Institute for Energy Efficiency, Distinguished Professor
University of California, Santa Barbara
jbowers@ucsb.edu

Ⓘ **Raymond G. Beausoleil**
Senior Fellow and Senior Vice President
Hewlett Packard Enterprise
ray.beausoleil@hpe.com

Ⓐ Yuan Xie
Chair Professor
Hong Kong University of Science and Technology
yuanxie@ust.hk

Ⓘ Kaveh Hosseini
Technical Lead, Photonics Co-Packaging Architect
Intel Corporation
kaveh.hosseini@intel.com

Ⓘ M. Ashkan Seyedi
Principal, Silicon Photonics Products
Nvidia Corporation
mseyedi@nvidia.com

Research Statement

Yuyang Wang

January 3, 2024

In the dynamic landscape of distributed computing, the exponential growth in traffic demands within data centers and high-performance computing systems has been distinctive, fueled by a deluge of data-intensive workloads. This trend is prominently exemplified by the rapid expansion of machine learning, big data analytics, and most notably, deep learning (DL)–driven artificial intelligence (AI) applications. The recent advent of large language models, which has revolutionized natural language processing and creative content generation, is propelling the broad adoption of ever-larger DL models and datasets, marking a significant milestone toward the era of data ubiquity. The continued scaling of these applications has pushed the limits of computational hardware, notably via increased parallelism and specialization. Yet, this rapid progress has outpaced the evolution of the underlying communication infrastructure, rendering chip-to-chip data movement a formidable barrier impeding performance and energy efficiency. This communication bottleneck has become the grand challenge to the quest of upscaling the computing systems toward exascale.

My research endeavors to find **transformative connectivity solutions**, maximally harnessing the potential of integrated silicon photonics (SiPh). In this pursuit, I have devised a dual-thrust research agenda for my independent career. The first thrust focuses on **reconfigurable system connectivity**. It aims to develop optical interconnects that not only provide unprecedented bandwidth but also adapt in real-time to the ever-evolving demands of emerging applications. The second thrust looks into **innovative system architectures**. It targets redefining chip-to-chip communication with groundbreaking optical I/O technologies, thereby pioneering new computational paradigms and interconnect functionalities. The synergy of these thrusts introduces unique design challenges, which I am equipped to tackle with my interdisciplinary research experience, ensuring the readiness of essential design tools and methodologies for these advanced connectivity solutions.

My research agenda is situated at the system level, squarely fitting into the interdisciplinary nexus of computer systems and architecture that bridges silicon photonics, networked systems, computer architecture, design automation, and AI/machine learning. I look forward to the possibility of collaborating with the esteemed colleagues in the Department of Computer Science and Engineering at the Pennsylvania State University to confront the grand challenge of data movement in future computing infrastructures across the full system stack.

1 Research Accomplishments and Skills

My doctoral and postdoctoral work have established a solid foundation for my anticipated research (Fig. 1) and equipped me with the skills necessary to address the upcoming research challenges. I was among the first to integrate accurate **compact models and simulation methodologies** of silicon photonic devices into widely used electronic design automation (EDA) platforms like Cadence Virtuoso [1, 2]. This integration is crucial for the accurate co-simulation of electronic and photonic components, enabling the efficient development of complex photonic integrated circuits (PICs) with reliable performance estimation. My expertise in **process variation characterization, mitigation, and tolerance** ensures the robustness and energy efficiency of fabricated designs [3–5]. This is particularly important for advanced technologies that often rely on emerging fabrication processes and require post-fabrication tuning. My work in these areas has been recognized in leading design automation conferences like DAC and ICCAD, respected photonics venues like OFC and JLT, and a forthcoming book chapter with Springer, effectively connecting the electronics and photonics research communities.

These design enabling techniques have been practically utilized in creating two generations of SiPh transceiver chips, featuring a **scalable link architecture** that facilitates unprecedented channel parallelism and delivers a chip I/O bandwidth of over 16 Tbps with energy consumption below 1 pJ/b [6, 7]. Fabricated in partnership with AIM Photonics through two full-wafer runs, each chip, measuring ~70 mm², densely integrates over 2,000 microresonators. The chip layout process was fully scripted and automated, showcasing not only significant technological advancements but also remarkable design efficiency. The highlighted link architecture was instrumental in securing a $35M SRC JUMP 2.0 grant with 23 principal investigators, a program to which I have contributed through proposal writing and ongoing research efforts. This work has also resulted in invited papers and presentations at both photonics and electronics design conferences (Photonics West and CICC), and an invited journal submission to *Nature Communications Physics*.

These accomplishments have advanced my research into exploring **traffic adaptability** for optical interconnects in distributed computing systems, grounding them in credible performance models and hardware validation. Notably, I have delved into runtime adjustments of parameters such as laser power and link bandwidth, aiming at accelerating distributed machine learning applications with reduced energy consumption [8–10]. A promising off-chip prototype, designed to redistribute wavelength channels across multiple ports at application runtime, is currently under review for publication. Additionally, an on-chip implementation is being fabricated at AIM Photonics and slated for testing in May 2024. These investigations underscore the significance of integrating architectural innovations and optimization strategies at the system level, a process which—without meticulous execution—could inadvertently counteract the advancements achieved at both device and link levels. This realization is a key driver behind my future research directions.

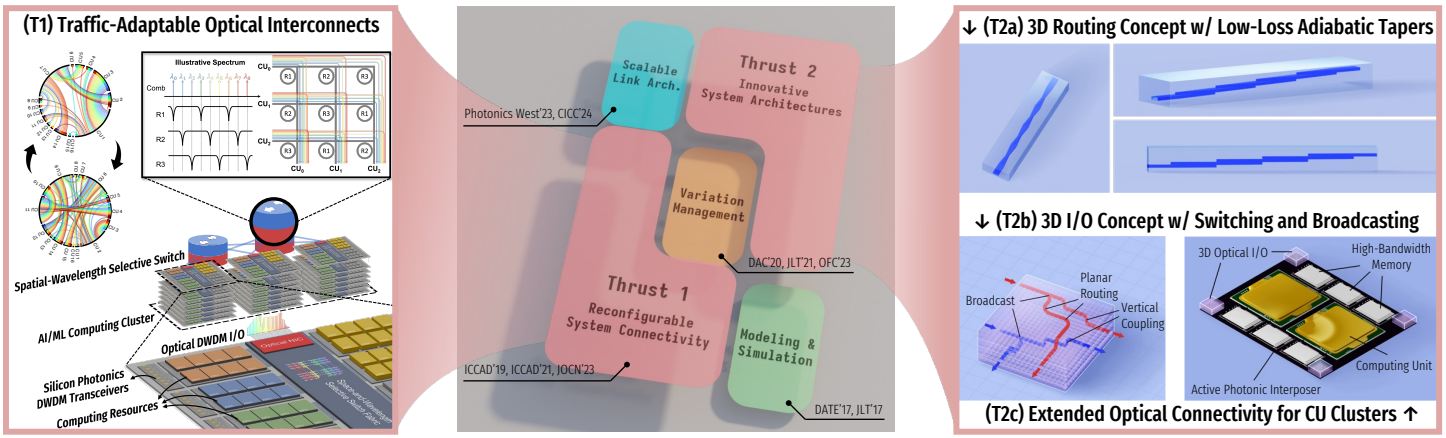

Figure 1: Overview of my research accomplishments and proposed research directions.

2 Research Vision and Agenda

My research agenda is set to continue at the system level, leveraging the latest breakthroughs in device designs and link architectures, while simultaneously informing their future advancements from a system application perspective. In light of the evolving data landscape, I plan to focus on two complementary and synergistic research thrusts in pursuit of groundbreaking connectivity solutions.

2.1 Thrust 1: Reconfigurable System Connectivity

With the advent of augmented reality (AR), virtual reality (VR), and Metaverse applications, distributed machine learning frameworks are seeing an increase in data privacy concerns that were previously confined to sectors with sensitive information, such as banking and healthcare. These sectors typically handle smaller volumes of data with more flexible latency requirements. In response, decentralized learning frameworks like federated learning have received growing popularity, as they allow the exchange of model parameters over raw data. Yet, certain applications still prioritize data parallelism to meet stringent requirements on model accuracy. Consequently, the data landscape in distributed computing is evolving toward both larger volumes and greater heterogeneity. This evolution, coinciding with the expansion of large models like GPTs, necessitates the next generation of optical interconnects to further excel in traffic adaptability, in addition to bandwidth and energy efficiency.

In this research thrust, my objective is to significantly enhance traffic adaptability by co-designing reconfigurable link architectures along with dynamic reconfiguration strategies (Fig. 1-T1). Building upon the SiPh transceiver developed during my postdoctoral research [6, 7]—which stands out for its leading bandwidth capacity and energy efficiency among state-of-the-art solutions—I aim to incorporate greater reconfigurability into its design. My prior work, namely on runtime laser power scaling and link bandwidth reconfiguration [8–10], serves as a proof-of-concept for the effectiveness of traffic-adaptable tuning knobs in improving both the performance and the energy efficiency of optically connected computing systems. Moving forward, I anticipate the success of this research thrust to be contingent on the following critical tasks:

1. Profiling and characterizing the traffic patterns of a diverse range of distributed computing applications, expected to exhibit greater heterogeneity and temporal dynamics compared to the collective communications typically observed in current computing clusters, as referenced in [10].
2. Introducing additional reconfigurable parameters beyond laser power and link bandwidth, such as wavelength allocation and switching/routing, and developing runtime reconfiguration strategies tailored to these characterized traffic patterns.
3. Conducting system-level simulations to assess the energy and performance impacts of the proposed reconfigurability, supported by credible performance models that accurately reflect real link designs.
4. Designing and integrating reconfiguration modules with state-of-the-art SiPh transceiver implementations, and validating the enhanced interconnect architecture with reconfiguration strategies on a hardware testbed driven by realistic/production network traces.

Throughout this endeavor, I also anticipate deriving valuable insights from a system application perspective. These insights will be instrumental in informing the design of SiPh devices and circuits, focusing on essential aspects such as tuning range and reconfiguration speed, to meet key performance metrics at the system level. This collaborative synergy across multiple design hierarchies is essential to maintain cutting-edge system connectivity in an ever-changing data landscape.

2.2 Thrust 2: Innovative System Architectures

Complementary to the first research thrust aimed at advancing chip-to-chip connectivity, the second thrust strives to address the notable gap between on-chip and off-chip communication bandwidths. This gap is particularly pronounced in accelerator systems comprising clusters of computing units (CUs) that frequently access data from both on-chip memory banks and off-chip memory pools. Expanding the number of on-chip high-bandwidth memory (HBM) stacks is increasingly impractical as the bandwidth capacity of electronic interposers approaches its

limits. Conventional approaches using optical fibers to connect CU clusters and memory pools are also constrained by the size and pitch of fiber arrays. Nonetheless, the emerging concept of 3D optical I/Os, benefiting from dense waveguide routing across multiple layers, presents a promising avenue to scale up CU clusters with optical connectivity that stays on-board with extended reach (Fig. 1-T2a–c). My contribution to assisting the formulation of this concept, which was successfully showcased at the 2023 DARPA ERI Summit, has inspired me to further explore this cutting-edge area. The key challenges I plan to address in this research thrust include:

1. Formulating the 3D routing problem with objectives such as maximized density and minimized loss, and developing efficient routing algorithms that draw from traditional EDA expertise and the latest in machine learning techniques.
2. Informing the design of 3D routing elements with performance and area constraints, and optimizing their physical design employing recent advances in areas such as photonic inverse design and topology optimization.
3. Conducting system-level design space explorations for computing architectures with transformed memory connectivity to delineate optimal system configurations, such as the ideal size of CU clusters that benefit from the expanded reach of on-board optical connectivity, and the optimal balance between on-chip and off-chip memory capacities.

In addition to eliminating the bandwidth taper at chip boundaries and allowing for continued upscaling of CU clusters, this research thrust also promises to expand the role of optical interconnects beyond traditional data communication. For instance, certain computational tasks, such as matrix multiplication, can be offloaded to the optical domain, for which existing explorations have been limited by the vast difference in physical dimensions of electronic and photonic implementations. This limitation can be significantly alleviated by the manifolded density of optical components enabled by 3D routing. This thrust, therefore, not only addresses current technological limitations but also fosters the development of new computing paradigms, where optical interconnects assume a more dynamic and integral role in future computing system architectures.

3 Research Collaborations and Initiatives

My research experience has been deeply rooted in multidisciplinary collaboration, a skill I mastered during my postdoctoral training at the Columbia University. There, I led research initiatives within our group, guided by my supervisor's mentorship and backed by funding from agencies like DARPA, SRC, and ARPA-E. These initiatives required seamless teamwork with colleagues from academia, industry, and governmental bodies. In addition, I have a proven track record in assisting both my doctoral advisor and postdoctoral supervisor with fundraising activities. My responsibilities also encompassed preparing and compiling reports and materials, as well as participating in presentations at quarterly reviews to fulfill the requirements of our funded projects.

Given the interdisciplinary essence of my research agenda, I am enthusiastic about the opportunity to collaborate with the diverse faculty in the Department of Computer Science and Engineering. The expertise present in areas such as *computer architectures and networked systems* (Chitaranjan Das, Mahmut Kandemir, Vijaykrishnan Narayanan, John (Jack) Sampson, Anand Sivasubramaniam), *systems for big data/AI* (George Kesidis, Daniel Kifer, Dong Xie), as well as areas in Electrical Engineering such as *integrated photonics and optoelectronics* (Iam-Choon Khoo, Xingjie Ni, Shizhuo Yin) and *high-performance integrated circuits* (Mehdi Kiani, Wooram Lee) is particularly compelling. I am excited to contribute my experience and enthusiasm to your esteemed institution, keenly anticipating the chance to work with a community that resonates my commitment to innovation and making a meaningful impact on the future of technology.

References

[1] R. Wu, Y. Wang, Z. Zhang, C. Zhang, C. L. Schow, J. E. Bowers, and K.-T. Cheng, "Compact modeling and circuit-level simulation of silicon nanophotonic interconnects," in *Design, Automation & Test in Europe Conference & Exhibition (DATE), 2017*. Lausanne, Switzerland: IEEE, Mar. 2017, pp. 602–605.

[2] Z. Zhang, R. Wu, Y. Wang, C. Zhang, E. J. Stanton, C. L. Schow, K.-T. Cheng, and J. E. Bowers, "Compact Modeling for Silicon Photonic Heterogeneously Integrated Circuits," *Journal of Lightwave Technology*, vol. 35, no. 14, pp. 2973–2980, Jul. 2017.

[3] Y. Wang, J. Hulme, P. Sun, M. Jain, M. A. Seyedi, M. Fiorentino, R. G. Beausoleil, and K.-T. Cheng, "Characterization and Applications of Spatial Variation Models for Silicon Microring-Based Optical Transceivers," in *2020 57th ACM/IEEE Design Automation Conference (DAC)*. IEEE, Jul. 2020, pp. 1–6.

[4] Y. Wang, P. Sun, J. Hulme, M. A. Seyedi, M. Fiorentino, R. G. Beausoleil, and K.-T. Cheng, "Energy Efficiency and Yield Optimization for Optical Interconnects via Transceiver Grouping," *Journal of Lightwave Technology*, vol. 39, no. 6, pp. 1567–1578, Mar. 2021.

[5] Y. Wang, S. Wang, A. Novick, A. James, R. Parsons, A. Rizzo, and K. Bergman, "Dispersion-Engineered and Fabrication-Robust SOI Waveguides for Ultra-Broadband DWDM," in *Optical Fiber Communication Conference (OFC) 2023*. Optica Publishing Group, 2023, p. Th3A.4.

[6] Y. Wang, A. Novick, R. Parsons, S. Wang, K. Jang, A. James, M. Hattink, V. Gopal, A. Rizzo, C.-P. Chiu, K. Hosseini, T. T. Hoang, and K. Bergman, "Scalable architecture for sub-pJ/b multi-Tbps comb-driven DWDM silicon photonic transceiver," in *Next-Generation Optical Communication: Components, Sub-Systems, and Systems XII*, G. Li, K. Nakajima, and A. K. Srivastava, Eds. SPIE, Mar. 2023, p. 55.

[7] Y. Wang, S. Wang, R. Parsons, A. Novick, V. Gopal, K. Jang, A. Rizzo, C.-P. Chiu, K. Hosseini, T. T. Hoang, S. Shumarayev, and K. Bergman, "Silicon photonics chip I/O for ultra high-bandwidth and energy-efficient die-to-die connectivity," in *2024 IEEE Custom Integrated Circuits Conference (CICC)*. IEEE, 2024, under review.

[8] Y. Wang and K.-T. Cheng, "Task Mapping-Assisted Laser Power Scaling for Optical Network-on-Chips," in *2019 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*. Westminster, CO, USA: IEEE, Nov. 2019, pp. 1–6.

[9] Y. Wang and K.-T. Cheng, "Traffic-Adaptive Power Reconfiguration for Energy-Efficient and Energy-Proportional Optical Interconnects," in *2021 IEEE/ACM International Conference On Computer Aided Design (ICCAD)*. Munich, Germany: IEEE, Nov. 2021, pp. 1–9.

[10] Z. Wu, L. Y. Dai, Y. Wang, S. Wang, and K. Bergman, "Flexible silicon photonic architecture for accelerating distributed deep learning," *Journal of Optical Communications and Networking*, 2023, to appear.

Teaching Statement

Yuyang Wang

January 3, 2024

My educational journey, profoundly shaped by the dedication and expertise of several remarkable teachers and mentors, have led me to a firm belief in the transformative power of good teaching on a student's life path. Over the years, I have contemplated and reflected on the quintessential pedagogy that balances the act of engaging, conveying, and inspiring. Specifically, in developing my course materials, I am guided by a series of introspective questions that aims at ensuring that the content not only is informative but also stimulates critical thinking.

I continually ask myself five key questions:

1. What insights will students gather while actively engaging with both my spoken words and the slide content?
2. What insights will students gather while independently examining the slide content at my lecture pace?
3. What are the takeaways for students when they revisit the slides on their own time?
4. How can I guarantee consistency in the key takeaways gathered across these varied learning contexts?
5. How do I structure the slide content to convey key messages to the students without them directly reading from it?

These practices have shaped my teaching philosophy, which emphasizes the effective use of visualization.

1 Teaching Philosophy

Visualization is central to my instructional approach, leveraging the human brain's rapid image processing ability, which reportedly surpasses text interpretation by 6x–600x[1] [1]. This capability enables students to extract information from visual aids alongside verbal explanations far more efficiently than text alone, allowing for profound engagement in class. In the realm of STEM education, where abstract theories and complex equations can be overwhelming, visualization serves as a vital bridge, translating intricate ideas into comprehensible and memorable images. It also elegantly addresses the pedagogical challenge of conveying essential concepts without resorting to simply reading from the slides—a practice that hinders critical thinking. Beyond the immediate classroom benefits, the ability to visualize data and concepts is an indispensable skill for students, one that is increasingly critical in both their academic pursuits and future research careers. Building on this philosophy, I address the challenge of ensuring consistent takeaways from the course materials in different learning contexts—whether inside or outside the classroom—by integrating concise bullet points alongside visuals, ensuring key messages being conveyed.

2 Pitfalls and Lessons Learned

Designing visuals that effectively encode information demands thoughtful consideration and attention to detail, ensuring accessibility for all learners. Informed by my personal experience with minor color vision deficiency, I am acutely aware of key pitfalls in visual information delivery, such as solely relying on color contrast to embed information. For instance, Fig. 1 demonstrates how using color as the sole differentiator between two spectra can make the data difficult to interpret for those with color vision impairments. The prevalence of color blindness, affecting approximately 8% of males and 0.5% of females [2], might surprise many. However, this statistic virtually guarantees that in any moderately sized classroom, at least one individual is likely to have a form of color vision deficiency. Acknowledging this, I am committed to employing multiple modes of differentiation in my teaching materials, such as patterns, textures, and annotations, to ensure that all students, regardless of visual ability, have equal access to the information presented.

3 Inclusive Teaching

My commitment to inclusivity extends beyond color vision awareness to embrace all facets of diversity and accessibility in education. I recognize that students come with a broad spectrum of cultural backgrounds, personal histories, and educational experiences, all of which influence their learning needs. In light of this, I will strive to create a classroom environment that is not only physically accessible but also cognitively and culturally welcoming. This entails the use of language that is inclusive and bias-free, as well as the incorporation of diverse examples in my teaching materials, practices that I will regularly reflect on to ensure adherence.

[1] Despite the discrepancy with the unfounded internet meme claiming 60,000x, as called out in *The 60,000 Fallacy* (https://policyviz.com/2015/09/17/the-60000-fallacy/), this is still substantial enough to warrant extra attention to the use of visualization in teaching.

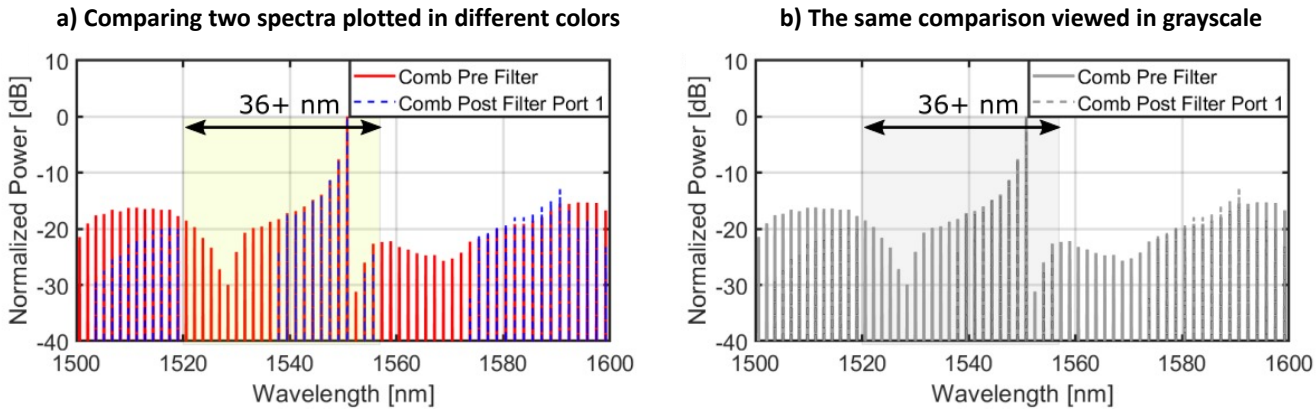

Figure 1: Example of non-robust information encoding solely with color contrast. **a)** Two spectra distinguishable by people with normal color vision. **b)** Illustrative rendering of the same two spectra possibly perceived by people with color vision deficiency.

4 Teaching Plans

In harmony with the esteemed curriculum of the Department of Computer Science and Engineering at the Pennsylvania State University, I am eager to contribute by teaching a range of courses that intersect with my expertise, such as

- Digital Design: Theory and Practice (CMPEN 270),
- Computer Organization And Design (CMPEN 331), and
- Fundamentals of Computer Architecture (CSE 530).

Believing that the preparation for teaching materials deepens my own understanding of the subject matter, I am also open to teaching courses beyond my immediate expertise, such as

- Communication Networks (CMPEN 362), and
- Parallel Processors and Processing (CSE 531).

Moreover, I am enthusiastic about the prospect of designing new courses, currently not offered at the Pennsylvania State University, such as

- Electronic Photonic Design Automation, and
- Optical Interconnects for Digital Systems,

which could expand and enrich the department's already distinguished curriculum.

5 Mentoring

My mentoring philosophy extends beyond knowledge sharing and intellectual guidance. I am dedicated to providing holistic support for the academic and professional development of the students. I will strive to create a supportive and collaborative research environment, inspiring students to excel in their studies, uphold professionalism, and explore their own research interests until they emerge as independent researchers. My time as a postdoctoral researcher at the Columbia University allowed me to partially implement this philosophy, guiding several Ph.D. students to significant milestones, including publishing their first papers as lead authors at premier conferences like OFC and CLEO, and securing their initial industry positions. I eagerly anticipate the opportunity to extend and refine my mentorship practices at the Pennsylvania State University.

6 Post–COVID-19 Considerations

In the post–COVID-19 era, I have adapted to the emerging norms of hybrid teaching and research settings. I am prepared to further refine my teaching and mentoring approaches to accommodate these evolving challenges. For example, I plan to enhance my slides with additional annotations and essential animations to offset the lack of real-time interaction that a physical whiteboard provides. Additionally, I plan to introduce regular coffee hours and/or lunch meetings within the group to facilitate in-person discussion whenever feasible and in compliance with the university guidelines. Such adaptability is critical for navigating through the evolving educational landscape, and I am committed to the continuous innovation in my teaching and mentoring practices to provide all students with accessible and effective education.

References

[1] "Research: Is A Picture Worth 1,000 Words Or 60,000 Words in Marketing? | Emailaudience." [Online]. Available: https://www.emailaudience.com/research-picture-worth-1000-words-marketing/

[2] "Types of Colour Blindness." [Online]. Available: https://www.colourblindawareness.org/colour-blindness/types-of-colour-blindness/

Keren Bergman
bergman@ee.columbia.edu
Charles Batchelor Professor of Electrical Engineering
Columbia University
530 W 120th St.
806 CEPSR, Department of Electrical Engineering
Columbia University
New York, NY 10027
+1 212 853 1657

Kwang-Ting Cheng
timcheng@ust.hk
Vice-President for Research and Development, Chair Professor
Hong Kong University of Science and Technology
Room 6359, Academic Building
The Hong Kong University of Science and Technology
Clear Water Bay, Kowloon, Hong Kong
(852) 2358 6161

John E. Bowers
jbowers@ucsb.edu
Director of Institute for Energy Efficiency, Distinguished Professor
University of California, Santa Barbara
3005 Henley Hall
University of California, Santa Barbara
Santa Barbara, CA 93106
(805) 893-8447

Raymond G. Beausoleil
ray.beausoleil@hpe.com
Senior Fellow and Senior Vice President
Hewlett Packard Enterprise
3000 Hanover Street
Palo Alto, CA 94304
(650) 258 1837