They said computation was clean.
They said it was weightless.
They lied.

Artificial intelligence was supposed to be the future. It may be the last thing the grid can afford.

Keep Computing

How Light Solves Computing's Impossible Problem

Derek W Bailey

TRUE PHOTONIC, INC.

Santa Rosa, California

Published by True Photonic, Inc. 1260 N. Dutton Street Santa Rosa, CA 95401
www.truephotonic.com

ISBN: 978-0-692-98288-4
Printed in the United States of America
978-0-692-98288-4

For my daughters, Alexandra and Maxine, who have paid all the tolls on my journey road.

For my mother Ozie-Mae, and my brothers Orrin and Walter.

For Pattie Wexler - best friend- who keeps me balanced.

For Del—brother.

For Gary—brother in Christ.

And for everyone who was told it couldn't be done—go build it anyway.

Acknowledgments

A book like this does not happen alone. Neither does a company.

Del Wolverton has been my partner for twenty-seven years. Everything in these pages that works—the technology, the architecture, the vision—carries his fingerprints. Brother, thank you.

Dr. Gary Poovey trusted us with his life's work. That trust made everything else possible.

Former Congressman William Jefferson provided guidance and opened doors as we work to bring True Photonic to New Orleans. His belief in what photonic computing can mean for the Gulf South has been invaluable.

Roy Geddie, CEO of WolvertonBailey, has been a constant force moving this project forward. His steady hand keeps us grounded when the vision gets ahead of the execution.

Mark Newland leads our project management team with the discipline and precision the work demands. He was one of the first engineers at SanDisk; we are fortunate he chose to build something new with us.

Michael Miles, Steven Branch, Taylor Paul, Alex Hicks, Nathaniel Anderson, and Douglas Eze have worked across both Derek Automotive and True Photonic, supporting me at every step

Robert Switzer and the team at Astera Energy brought us the substrate of the future. The graphene-boron nitride-sapphire architecture that enables the Poovey Stack's thermal performance exists because of their work.

John and Andrea of Iron Horse Capital believed in the SkyLight vision early. Their investment in SkyLight Holdings gave us the runway to build infrastructure, not just technology.

Carlos Austin has been a steady advisor—the kind who tells you what you need to hear, not what you want to hear. "Your suggestions made this a better book."

Mike Rodriguez made the introduction that started everything—a handshake at a small church in Dunsmuir that changed the trajectory of three lives and, we hope, an industry.

Searcie Cassidine of Cassidine Consulting brought rigor to our IP valuation and insisted I write this book. She was right on both counts.

To the engineers, attorneys, advisors, and staff across True Photonic, WolvertonBailey, and Derek Automotive who do the work that makes the vision real—thank you. Your names may not all appear on these pages, but your contributions appear on every one.

And to my family, who have lived with "impossible" projects for decades: I know the cost. I do not forget it.

CONTENTS

For forty years, photonic computing was a solution in search of a problem. The problem has arrived.

Prologue

This Is Not My First Impossible System

Del met Dr. Gary Poovey at a small church in Dunsmuir, California.

This was 2022.

Mike Rodriguez, a friend and WolvertonBailey shareholder, made the introduction—the kind of handshake that happens a thousand times and leads nowhere. Del did not know Gary was a physicist. Gary did not know that Del had received a DARPA grant for work that made light do impossible things, or that he had been on the early military teams writing code that used artificial intelligence to read satellite imagery. They were just men in a congregation, exchanging pleasantries.

Mike mentioned that Dr. Poovey was looking for funding for some research he did not really understand, but he knew Del would. Del is the smartest guy Mike knows. Del is the smartest guy everyone knows. A Navy Senior Chief. Trained in engineering, radar, missile systems, and photonics. Del listened to Gary's quick story about his work. It registered immediately. He invited Dr. Poovey to our lab.

A few weeks later, Gary came to Santa Rosa. He brought data—not theories, not projections, but a validation report from Technion University in Israel, one of the world's leading materials science institutions. His team had tested optical switches built from a material he had developed, one that responds to the magnetic field of light. The measured switching speed was 150 to 200 femtoseconds. The fastest optical switching speed ever published in scientific literature was 600 femtoseconds. Gary's material was three to four times faster than anything that existed.

As he always does when he sees something I should see, Del called me. He described the encounter, the report, the background. Before he could finish, I asked: "Couldn't we make a transistor with that?"

"Exactly—I trained you well," Del said. "But say nothing to Dr. Poovey right now. He doesn't know what he really has. I'm going to recreate the Technion test in our lab. If I get the same result, we move. Fund it. Buy the tech."
"How far along is he?"

"He has a first-generation optical switch. But no logic. His research team was strong, but there were no computer guys on it. And the material composition is not optimal." Del paused. "I can fix all of it."

"Then let's do it," I said. "At Derek speed." Derek speed is an inside joke at the company. It means now.

This is how Del and I have always worked. Del sees things early. I trust his judgment completely. And unlike most people in technology or investing, we do not need the whole picture—because we choose projects where we are the ones who complete the picture. Most investors wait for proof. They want validation, traction, a clear path to returns. They want someone else to have taken the risk of being wrong. By the time they see opportunity, the opportunity has been priced accordingly.

Del and I invest while the idea is still on the napkin. That is what separates the daring from the prudent. We look for situations where the physics is sound but the engineering is incomplete. Where the insight exists but the system does not. Where someone has discovered something real but does not yet understand what it could become.

Gary Poovey had a switching material. He did not have a transistor. He did not have logic gates. He did not have an architecture, a manufacturing strategy, or a path to market. He had the first piece of something that could change computing—and no way to build the rest.

We had the way. Del and I had spent twenty-five years building things that were not supposed to work, developing exactly the capabilities Gary's discovery required.

CopperSnake—video over copper wire when the industry said copper was dead.

Popa Media—digital billboards installed on decommissioned payphone stands across urban centers, offering free phone calls through voice-over-IP. This was when people were still paying for minutes on their cell phone plans. They could not understand how we could give calls away for free. We understood. We were just early.

The Counterpoise engine—a combustion system that defied what the textbooks said was possible.

Each project demanded we solve problems no one else had solved because no one else had attempted them. When I looked at Gary's Technion report, I did not see a physics result. I saw an incomplete picture that we knew how to complete. Del would build the team and the technology. I would build the narrative and the business structure—which for an idea this ambitious, in an industry as entrenched as semiconductors, would have to be as novel as the technology itself.

That is what True Photonic is.

That is what True Photonic Ventures will become: a venture arm that seeds start-ups built on our photonic platform, companies that will locate their operations in our SkyLight Clean Compute Centers. Not a venture fund that waits for founders to figure everything out, but a venture partner that brings the capabilities to complete pictures others cannot. We will not just write checks. We will finish architectures.

Del Wolverton is a Navy man. A Senior Chief. A U.S. military technical advisor to foreign governments—the man who first took the Saudi Navy to sea. He does not speak carelessly. When he says he can do something, he has already done the math. When he says a technology is real, he has already seen how to build it.
In twenty-seven years of working together, I learned to trust his quiet confidence more than most people's enthusiasm.

What Del understood, in that first conversation about Gary's work, was something Gary did not yet fully grasp. A switching material is not a transistor. A transistor is not a processor. The gap between a laboratory phenomenon and a computational system requires logic gates, architecture, signal routing, thermal management, manufacturing strategy—the accumulated engineering knowledge that Del had spent his career developing.

Gary had the physics. Del had everything else. What I understood was simpler: this was a foundation I could build something incredible on.

We made Gary an offer. He did not accept immediately. He went home and

prayed with his family. Gary Poovey is a deeply religious man, and he was being asked to entrust his life's work to people he had met at church a few weeks earlier. The decision deserved more than a handshake.

When he called back, the answer was yes—with a condition. He wanted assurance that we would be a different kind of company. An ethical company. In his words: "Christian businessmen."

Dr. Gary Poovey merged his technology into the company. True Photonic was born.

What I have not yet described is the negotiation that preceded that birth.
Gary came to us with patents, with validation, with a material that worked. He did not come with a path to market. His research team in Israel had spent millions of dollars over years of work. They had proven the physics. They had not built a product, had not identified a customer, had not constructed the business architecture that translates laboratory results into commercial returns. Gary was approaching eighty. The funding sources that back early-stage technology do not typically write checks to octogenarians, regardless of how sound the science. The Technion report was not a beginning for Gary. It was, in a sense, an ending—the culmination of a research program that had nowhere left to go.

Del saw this clearly. So did I.

The conversation was direct. Del would take control of the technical development—not as a collaborator, but as the architect. The compositions would be optimized. The logic gates would be designed. The system architecture would be built from Del's specifications, informed by Gary's physics but not constrained by Gary's original approach. I would handle business development, capital formation, and narrative—the work of translating technology into enterprise.

This was, in effect, a takeover.

We were not offering to help Gary complete his vision. We were offering to absorb his breakthrough into ours. But it was a takeover with good intent, and Gary understood the alternative. His options were to partner with people who could complete the picture, or to watch his life's work remain forever incomplete—a validated phenomenon with no path to impact. He chose partnership. He chose us.

The trust that required still humbles me. I made a promise to Gary's wife: I would make their family name famous around the world. And so I named the light-activated light switch—the device at the heart of everything that follows—the Poovey Switch.

It is a promise I intend to keep.

Del started building the team. First was Mark Newland. Mark was one of the early engineers at SanDisk, part of the team that built the flash memory industry from laboratory curiosity to global infrastructure. His career had traced the arc of modern computing—a bachelor's degree in applied mathematics and computer science from California Polytechnic, an MBA in computer information systems from National University, decades of software development in C++, C#, and Linux environments. He had led cross-platform application engineering efforts in both industry and academia, directed the Information Network Technology Program at Pratt Community College, taught computer science at Emmaus Bible College and Barclay College, bridging theory and practice in ways that most engineers never attempt.

Mark understood what it takes to move technology from prototype to production. He had done it before, at scale, at a company whose products now lived in billions of devices worldwide. When Del called, Mark listened. When Del explained what the Poovey Switch could become, Mark saw what Del saw: not merely a faster switch, but the foundation of a new computational paradigm. -
Del has smart friends too.

A physicist, an engineer, and a dreamweaver walked into a lab. It sounds like the setup to a joke. The punchline is that it worked.

I have not received a paycheck—one that was not 100 percent commission—since I was nineteen years old. I am no longer a spring chicken.

That is not a boast. It is context.

When I tell you that impossible tasks and setbacks do not frighten me, I am not speaking from theory. I am speaking from more than four decades of operating without a net, without institutional backing, without the assumption that effort would be rewarded simply because it was applied.

Hospitality management came first. Then real estate. Then insurance. I became a registered investment representative, then left to sell home improvements. I owned a restaurant. I ran food trucks. I built businesses in Asia and Latin America, flirted with business in Russia—export and marketing operations that taught me how systems work differently in different places and how they fail for the same reasons everywhere.

The formula I learned: Have a vision. Build a narrative. Go sell that shit.

The breakthrough came in temporary staffing—a business that scaled, generated real capital, and still operates today. Most people, at that point, would have consolidated. Stayed in their lane. Protected what they had built.

The family has always hated the way I put us at risk. But I have never been good at staying in lanes.

As a kid, when adults asked what I wanted to be when I grew up, I said I wanted to own a company like GM. I am not sure I even knew GM made cars. I just remembered commercials that said something like "What's good for GM is good for America." That sounded right to me. I want a company like that someday, I would tell them.

People would laugh. This kid.

The staffing profits went into technology. Then into vehicles. Now into the problem that underlies everything: computation itself.

Del entered my life around 1998.
He came to work on a project most people considered impossible: transmitting full-motion, broadcast-quality video over copper wire. Fiber optic infrastructure was just beginning to develop. The industry consensus was that copper could not handle video. Too much impedance. Bandwidth too narrow. The physics did not support it.

We called the system CopperSnake. We built it anyway. It worked. Video transmission up to a mile over copper wire. That was my first lesson in the difference between "impossible" and "not yet attempted by the right team." Del made things work that were not supposed to work.

He had done it before—at companies he founded, Space Microwave and SpectraSwitch—and he would do it again, across the decades we worked together.

The phone calls between us are a constant.

One came at three in the morning. Del had woken from a dream about building the perfect engine. He had gone, as he put it, to God's place of all knowledge and seen how to make a better combustion engine. He got up, sketched the design on a napkin, and called me. "You have to come out to California. I have something to show you. A new engine design." I was half asleep. "Aren't you an electrical engineer? What do you know about engines? That's mechanical."

But I remembered a conversation from a few months earlier. We had been playing chess, talking about what project to take on next. I had just bought my first pickup truck, and the fuel costs were shocking. I told Del we should do something about that. We should build a better engine.
He laughed. "I can do everything except defy physics. Everything there is to do to an engine has been done. If something was left, don't you think one of the major engine companies would have done it by now?"

Four months later: the call.

I was on a 7:30 a.m. flight from Philadelphia. I knew nothing about engines.
But I knew Del.

We stopped working for each other and started working together.
WolvertonBailey, Inc. was formed to develop the Counterpoise engine—a generator system that proved to be 83 percent more efficient than the diesel cycle. It completes its rotation in 360 degrees instead of 720. It is a third the size for equivalent power. It uses half the fuel. Now patented.

That work led me to form Derek Automotive, where the goal was a parallel-hybrid, self-charging electric vehicle. A car that generates its own electricity using a small amount of fuel while operating. That project continues—painfully, through tariffs and supply chain disruptions. I have not abandoned it. I have deferred it. Because somewhere along the way, I realized I was solving the wrong problem first.

With a photonic computer, that self-charging vehicle could become fully autonomous. Real autonomy—not the assisted driving that companies market as self-driving—requires computation at speeds and efficiencies that electronic systems cannot deliver. The processors would be obsolete before the car reached production.

The foundation was wrong.

When you discover the foundation is wrong, you have two choices. Continue building and hope the structure holds. Or stop, go back, and fix the foundation first.

I chose the foundation.

Then Gary Poovey walked into our lab with a Technion report and a material that could switch light faster than anything in the scientific literature. Del looked at me and said: Fund it. Let's buy the tech.

The foundation, it turned out, was waiting to be built.

I am not an engineer. I am not a physicist. But I have learned so much from Del and from years of self-study that I can hold my own. The first white paper on the Counterpoise engine—I wrote it. Most of the published articles on the Poovey Switch—I wrote them. The corporate structure of True Photonic—I designed it. The narrative that holds this entire vision together—I built it.

What I am is someone who has spent four decades learning to recognize incomplete pictures and to judge whether I am capable of completing them. That judgment is not intuition. It is not luck. It is pattern recognition, earned through failure and success and repetition. Building things that were not supposed to work. Failing at things that were supposed to succeed. Learning the difference between a bad bet and bad timing.

CopperSnake was not supposed to work.
Popa Media was not supposed to work—free phone calls when everyone was paying by the minute.

The Counterpoise engine was not supposed to work. The self-charging vehicle is still considered impossible by people who have not seen the prototypes.
Photonic computing was not supposed to work either—until Gary's switches proved that it could, and Del's engineering showed us how to complete the picture.
The Poovey Switch became the Poovey Stack. A laboratory phenomenon became an architecture.

But I am moving too fast. I need to tell you about the day the switch became real.

Del had said he would recreate the Technion test in our Santa Rosa lab. If the results held, we would move. I waited. Days became weeks. Del does not provide progress reports; he provides conclusions.

The call came while I was in Savannah, Georgia, receiving a shipment of vehicles and electric bikes for Derek Automotive—the parallel-hybrid project I had not abandoned, only deferred. I was standing in a warehouse, signing manifests, sweating through my shirt in the coastal humidity, when my phone buzzed.

"It works," Del said. No preamble. No buildup. "The switching speed is real. The compositions I optimized are faster than Gary's originals."

I stepped outside, away from the forklifts and the delivery drivers. "How much faster?"

"We're measuring 150 femtoseconds. Consistent. Repeatable. This is not an artifact. This is not instrument error. The material does what Gary said it does—and with my optimizations, it does it better."

I leaned against a loading dock and tried to process what I was hearing. One hundred fifty femtoseconds. A quadrillionth of a second. A timescale so brief that light itself travels only the width of a human hair.

"So we move," I said.
"We move," Del confirmed. "At Derek speed."

The vehicles behind me—the electric bikes, the hybrid prototypes, the machinery of a project I had spent years building—suddenly felt like artifacts from a previous life. They were still important. The work would continue. But standing in that Savannah warehouse, phone pressed to my ear, I understood that the foundation had shifted beneath me.

The call lasted four-five more minutes. Del outlined next steps: logic gate designs, fabrication partners, patent filings. I half-listened, still absorbing the central fact. The switch was real. The physics worked. Everything else was execution.

The fuller implications came later, in a conversation with Del and Mark Newland at the Santa Rosa lab. I had been running numbers in my head for days—the way I do when a pattern starts to emerge, when the shape of something larger begins to resolve from scattered data points.

"Wait," I said. "If the switch is 5,000 times faster than electronic transistors—headed toward 10,000 times faster—does that mean we can divide the number of transistors on an Nvidia chip by 5,000 and maintain the same processing speed? While reducing water and electricity consumption per logic event by the same factor?"
There was a pause. Mark looked at Del. Del looked at me. The silence stretched long enough that I wondered if I had said something stupid.

"That's exactly what it means," Del said finally.

"Fuck," I said. "That's a patent worth—who the hell knows what that's worth."

Mark laughed. Del did not. He was already thinking about implementation, about the engineering that would translate this arithmetic into architecture. That is the difference between us. I see the opportunity and start calculating value. Del sees the opportunity and starts solving problems.
It was worth, as it turned out, the foundation of everything that followed. Fourteen patent filings. Over six hundred claims. An architecture that exists not on napkins or in slide decks but in fabricated wafers and validated test results.

The switch became a stack. The stack became a platform. The platform became the subject of this book.

That is what happens when you see early, trust your judgment, and choose projects where you are the one who completes the picture. We could not have afforded to invest in the Poovey Switch at a later stage of development. Our most recent funding was $30 million for 25 percent of a division of True Photonic. We got in early because early is where we live.

The computing crisis described in this book is not speculation. The data is public. The trends are measurable. The industry is hitting limits that cannot be solved by working harder within the current paradigm—only by changing the paradigm itself. Gary Poovey understood this. He had spent years in the semiconductor industry watching Moore's Law decelerate, watching thermal limits constrain what chips could achieve, watching an entire industry optimize within boundaries rather than question whether the boundaries had expired. His response was to ask a different question:

What if the switching medium itself could change?

The material he developed was the answer. The Technion validation was the proof. But proof is not a product. Physics is not a processor. Del brought the engineering—logic gates, optimized compositions, the system-level architecture that transforms a switching phenomenon into a computational platform. Everything he had done on fiber optics, on military systems, on every impossible project we had tackled together, had prepared him for this moment.

I brought the business model. The capital from decades of building companies. The tolerance for risk that comes from never having operated any other way. And the pattern recognition that told me this was not just interesting technology—this was a structural solution to a structural problem. An incomplete picture we were uniquely positioned to complete. We had been preparing for this partnership for twenty-seven years without knowing it.

I do not know if we will succeed. No one can know until the work is complete and the market delivers its verdict. What I know is that success is no longer theoretical. The Technion validation proved the switching speed. The fabricated wafers proved the manufacturability. Fourteen patent filings document an architecture that exists—**not one that is imagined.**

What I also know is that the current path is unsustainable. Power consumption is rising faster than generation. Water usage is straining communities that never agreed to become industrial cooling systems. Supply chains are concentrated in ways that create unacceptable risk. Something will change. Either better technology enables a different path—or demand for computation contracts to fit what current technology can supply. This book argues that light offers a better path than retreat.

One more thing.

I mentioned that Del met Gary Poovey at church. I mentioned that Gary prayed with his family before joining us. I want to be clear about what that means—and what it does not.

I do not claim certainty about outcomes. I do not believe faith substitutes for work, or that conviction guarantees success. What I believe is that work worth doing is worth doing regardless of whether success is guaranteed—and that the people who show up when outcomes are uncertain are different from those who show up when outcomes are assured. Many of the people who came together around True Photonic share that orientation. We are comfortable with ambiguity. We function under pressure. We have been underestimated before, and we have learned that underestimation clarifies teams. The people who leave when odds are long are not the people you want when odds improve.

Gary prayed before he committed. Del calculated before he spoke. I looked at an incomplete picture and saw one we could complete. Some call the convergence of those three things coincidence. I do not. But I am not asking you to share my interpretation. I am asking you to read what follows with an open mind, evaluate the evidence on its merits, and draw your own conclusions.

The physics is sound. The technology is validated. The need is urgent.

That is enough to justify the work ahead.

And it is enough, I hope, to justify the pages that follow.

— Derek W. Bailey
Atlanta, Georgia
January 2026

PART I

THE IMPOSSIBLE PROBLEM

The water did not come back.

The data center is the factory of the twenty-first century. And like every factory before it, it extracts.

CHAPTER 1

The Cost of a Question

You asked a question.

Perhaps it was idle curiosity—the capital of Burkina Faso, the lyrics to a half-remembered song. Perhaps it was consequential: a medical symptom that needed interpretation, a contract clause that required explanation, a line of code that refused to compile. Whatever the question, you typed it into a box on your screen, pressed enter, and within seconds received an answer that would have taken your parents a trip to the library, and your grandparents a consultation with an expert they may not have been able to afford.

The answer appeared as if by magic. Clean text on a white screen. No visible machinery. No evident cost.

But there was a cost. There is always a cost.

In the fraction of a second between your question and its answer, a cascade of physical events unfolded across thousands of miles. Your query traveled through fiber-optic cables to a data center—perhaps in Virginia, perhaps in Iowa, perhaps on three continents simultaneously. There, it was routed to a cluster of servers stacked in racks that stretch toward forty-foot ceilings, each server packed with processors running so hot they would melt without constant cooling. Those processors performed billions of calculations, drawing electricity from a grid that may have burned coal or natural gas to generate it. The heat they produced was absorbed by water—thousands of gallons of it—which carried the thermal energy away from the machines before evaporating into the atmosphere or cycling back through cooling towers the size of apartment buildings.

Your answer materialized. You read it, nodded, and moved on with your day.

The water did not come back.

Light does not heat
what it touches.

The infrastructure that answers our questions is, by design, invisible. This is a triumph of engineering and a catastrophe of perception. We have built the most energy-intensive industrial system in human history and hidden it so completely that most of its users do not know it exists.

The system is called "the cloud," a word chosen for its weightlessness, its ethereal connotation of something floating benignly overhead. The metaphor is a lie. The cloud is not floating. The cloud is a constellation of windowless buildings, many of them larger than aircraft hangars, consuming more electricity than small cities. The cloud requires its own power plants. The cloud drinks rivers.

A modern hyperscale data center—the kind operated by Google, Microsoft, Amazon, and Meta—covers between 500,000 and two million square feet. Inside, tens of thousands of servers operate continuously, processing the queries, storing the data, and running the applications that constitute our digital lives. A single facility can draw 100 megawatts of power or more. For context, that is enough electricity to power approximately 80,000 homes. It is the output of a small nuclear reactor, or a medium-sized natural gas plant, dedicated entirely to computation. And computation generates heat. The processors that answer your questions operate by pushing electrons through silicon pathways so small they are measured in billionths of a meter. This is violent work at the atomic scale. Resistance generates heat. Switching generates heat. The very act of thinking—if we may call it that—generates heat. Left unchecked, this heat would destroy the machines within minutes. The processors would throttle, then fail, then melt into useless slag.

So we cool them. We cool them constantly, at enormous expense, using systems that have themselves become marvels of engineering. Cold water circulates through pipes that wind among the server racks, absorbing heat and carrying it away. In evaporative cooling systems—still the most common approach—that water eventually becomes steam, rising through cooling towers and dispersing into the air. It is gone. A medium-sized data center can consume 300,000 gallons of water per day. A large one can drink over five million gallons daily—the water usage of a town of 50,000 people.

In 2024, Google's single largest data center, located in Council Bluffs, Iowa, consumed one billion gallons of water. That is enough to supply the residential water needs of the entire state of Iowa for five days. One building. One year.

One billion gallons. This is the machine that answers your questions.

This is the cost you do not see.

• • •

For two decades, data center energy consumption grew relatively slowly. Engineers achieved remarkable efficiency gains that offset the explosive growth in digital traffic. Between 2010 and 2018, global data center workloads tripled while energy consumption increased by only six percent. It was a quiet triumph of optimization—better chips, smarter cooling, more efficient software. The industry could credibly claim that the digital economy was decoupling from its physical footprint. Then came artificial intelligence.

On November 30, 2022, a San Francisco startup called OpenAI released a chatbot named ChatGPT to the public. Within five days, it had one million users. Within two months, it had one hundred million—the fastest adoption of any technology in human history. By 2025, ChatGPT was serving 700 million weekly active users and processing over 2.5 billion queries per day. And ChatGPT is only one product from one company in what has become a global arms race to deploy artificial intelligence. The computational demands of AI dwarf anything the industry had previously encountered. A traditional Google search consumes approximately 0.0003 kilowatt-hours of electricity. A ChatGPT query consumes roughly ten times that amount—and can consume fifty times more for complex queries with long inputs or outputs. The queries that once took fractions of a second now require sustained computation across thousands of specialized processors. The models that power these systems required months of training on clusters consuming megawatts of power continuously.

The efficiency gains that had masked two decades of growth evaporated almost overnight. According to the Lawrence Berkeley National Laboratory, U.S. data centers consumed approximately 176 terawatt-hours of electricity in 2023—roughly 4.4 percent of total national electricity consumption. By 2028, that figure is projected to reach 325 terawatt-hours at minimum, and could exceed 580 terawatt-hours under aggressive AI adoption scenarios. The upper estimate represents twelve percent of projected U.S. electricity consumption—more than the steel, aluminum, and chemical manufacturing industries combined.

The same explosion is happening with water. U.S. data centers directly consumed 17 billion gallons of water in 2023. By 2028, that figure could double—

or quadruple. And direct cooling water is only part of the story. The electricity that powers data centers must itself be generated, and thermoelectric power plants—which still provide the majority of U.S. electricity—consume enormous quantities of water for cooling. When you account for this indirect water consumption, the total footprint multiplies by a factor of twelve.

We are building an industry that drinks like agriculture and burns like manufacturing, and the costs are no longer invisible.

Data centers are not distributed evenly across the landscape. They cluster in particular places, for particular reasons, and their concentration creates particular problems. Northern Virginia is the capital of the cloud. Seventy percent of the world's internet traffic passes through servers in four counties surrounding Washington, D.C.—Fairfax, Loudoun, Prince William, and Fauquier. The region hosts more than 300 operational data centers, with dozens more under construction. The concentration is a historical accident that became self-reinforcing: the early internet backbone ran through the area, attracting the first data centers; their presence attracted more connectivity; more connectivity attracted more data centers. Today, Loudoun County alone contains more computing capacity than most nations.

The consequences of this concentration are now impossible to ignore. In 2023, all data centers in Northern Virginia collectively consumed nearly two billion gallons of water—a 63 percent increase from 2019. Loudoun Water, the county's water authority, has had to rely increasingly on potable water supplies to feed the data centers, diverting water that would otherwise serve residential and agricultural needs. The electrical grid strains under the load. Dominion Energy, the regional utility, is planning $5.2 billion in new transmission infrastructure, including hundreds of miles of new power lines, simply to keep pace with data center demand.

But Virginia is not unique. Similar concentrations are emerging wherever the economics of computation align: cheap electricity, available water, favorable tax treatment, and proximity to major fiber-optic routes. In Iowa, data centers cluster along the Missouri River, drawing both power from wind farms and water from the aquifers below. In Arizona, they spread across the Sonoran Desert, tapping the Colorado River system even as the American Southwest enters its worst drought in 1,200 years. In Oregon, they line the Columbia River

Gorge, competing for hydroelectric power with aluminum smelters and the endangered salmon runs that depend on managed river flows. Each location makes a certain sense in isolation. Collectively, they represent a pattern of resource arbitrage that is becoming increasingly difficult to sustain. The data centers go where the resources are cheapest. They consume those resources. The resources become less cheap. And the industry begins looking for the next place to extract what it needs. It is the logic of extraction, applied to electricity and water in the service of computation. It is the cloud touching down on Earth, and leaving a mark.

• • •

The communities that host data centers are beginning to calculate what they have traded, and whether the exchange was fair. The industry's promises were seductive: tax revenue, construction jobs, the prestige of hosting critical digital infrastructure. In exchange, localities offered tax incentives, expedited permitting, and access to resources. But the ledger is proving more complex than the initial pitch suggested. Data centers employ very few people relative to their footprint and resource consumption. A facility covering twenty acres and consuming fifty megawatts of power might employ fifty workers once construction is complete. Compare this to a manufacturing plant of similar size, which might employ several hundred. The jobs that remain are specialized—network engineers, security personnel, HVAC technicians—and often filled by workers who commute from outside the community.

The tax revenue, too, has proven less generous than promised. Many states and localities offered aggressive incentives to attract the industry, exempting data centers from sales taxes on equipment, property taxes on facilities, and various other levies. In some jurisdictions, the net fiscal impact of data centers is now being questioned. "They don't really bring any jobs, the taxation levels were lowered to the point where they don't really bring a fiscal benefit to the city, and they require an enormous amount of power," the planning administrator of Chandler, Arizona, told journalists. "We struggle with being able to tell the story that this is a communal benefit."

Meanwhile, the costs are becoming harder to ignore. Residential electricity rates are rising in data center corridors as utilities invest in new generation and transmission capacity. Aquifers are declining. Noise from cooling equipment disturbs neighboring homes. Power lines march across landscapes that were once farmland or forest.

In some communities, data centers are consuming more than a quarter of the local water supply.

And so the backlash has begun.

According to a 2025 analysis by Data Center Watch, $64 billion in data center projects across the United States have been blocked or delayed by local opposition since May 2024. In Virginia alone, $900 million in projects have been killed outright, and $45.8 billion face significant delays. Across 28 states, researchers identified 142 activist groups organizing against data center development. The opposition crosses political lines. In Warrenton, Virginia, more than 500 residents—including the actor Robert Duvall—packed a town council meeting to oppose an Amazon data center. In Chandler, Arizona, the city council voted unanimously against a 422,000-square-foot facility despite personal lobbying from former Senator Kyrsten Sinema. In Peculiar, Missouri, a group called "Don't Dump Data in Peculiar" successfully pressured officials to amend zoning laws to prohibit data centers entirely. In Prince William County, Virginia, every town council member who supported a controversial data center project subsequently lost reelection.

A nationwide poll found that only 44 percent of Americans would welcome a data center near their home—making data centers less popular than gas plants, wind farms, or even nuclear facilities. When one Maryland developer proposed converting an abandoned mall into a data center, more than 20,000 people signed a petition against it. The industry that once flew under the radar is now a central issue in local politics across America. And the question it raises is simple but profound: Can the communities that host our digital infrastructure afford its physical costs?

This book is about that question—and about an answer that might seem, at first, impossible. The crisis described in these opening pages is real. The energy demands of artificial intelligence are growing faster than our ability to supply them sustainably. The water consumption of data centers threatens communities and ecosystems. The communities that host digital infrastructure are pushing back, and their grievances are legitimate. The current trajectory is unsustainable by any honest measure. And yet the demand for computation is not going to diminish. Artificial intelligence is not a fad that will pass. The applications being built on these systems—medical diagnosis, climate modeling, drug discovery,

autonomous vehicles, fraud detection, accessibility tools for the disabled—are too valuable to abandon.

The question is not whether we will keep computing.

The question is how.

The conventional answer is to build more: more data centers, more power plants, more transmission lines, more cooling towers. This is the answer embodied in proposals like OpenAI's Stargate project—multi-gigawatt AI factories that would each consume more electricity than some small countries. It is an answer that assumes the fundamental physics of computation cannot change, and that our only option is to feed the machine more resources than it currently consumes. But there is another answer. It requires changing the machine itself. For sixty years, computation has been performed by pushing electrons through silicon. This is not a law of nature. It is a technological choice—one that made sense when it was made, and one that has delivered extraordinary benefits. But it is a choice that carries inherent physical limitations, limitations that manifest as heat, as power consumption, as water usage, as all the costs described in this chapter. There is another substrate for computation. One that does not generate heat in the same way. One that does not require water for cooling. One that operates not at the nanosecond timescales of electron-based switching, but at the femtosecond timescales of light.

That substrate is photons. And the technology built upon it is called photonic computing.

This is the subject of this book. In the chapters that follow, we will examine the full scope of the impossible problem—the physical, political, and strategic dimensions of our current crisis. We will then explore the physics of light as a medium for computation, and the engineering advances that have made photonic computing practical. We will look at the infrastructure being built to deploy this technology, and the economic and geopolitical implications of a transition from electrons to photons. The story is ultimately hopeful. The crisis is real, but it is not insoluble. The demand for computation can be met—not by consuming ever more resources, but by changing the way we compute.

We can keep computing. We can do it sustainably. We can do it now.

But first, we must understand what we are up against.

The cost of a question is only the beginning.

CHAPTER 2

The End of Moore's Law

In 1965, Gordon Moore made an observation that would shape the next sixty years of human civilization.

Moore was the director of research and development at Fairchild Semiconductor, one of the pioneering companies of what would become Silicon Valley. He had been asked to write a brief article for Electronics magazine predicting where the semiconductor industry might be heading. Looking back at the short history of integrated circuits—the first had been demonstrated only seven years earlier—Moore noticed a pattern. The number of components that engineers could fit onto a single chip was doubling roughly every year. If that trend continued, he speculated, integrated circuits would eventually contain thousands of components, then tens of thousands, then more.

It was a modest prediction, offered with appropriate hedging. Moore was not claiming to have discovered a law of physics. He was describing an empirical trend in manufacturing capability, driven by improvements in lithography, materials science, and engineering ingenuity. He later revised his estimate to a doubling every two years. But the observation took on a life of its own. It became known as Moore's Law, and the semiconductor industry organized itself around it as if it were a commandment from on high.

For six decades, the industry delivered on that commandment with remarkable fidelity. In 1970, transistors measured approximately 12,000 nanometers across. By 2000, they had shrunk to 130 nanometers. By 2020, leading-edge chips were manufactured at 5 nanometers. Each generation packed more transistors onto the same area of silicon, enabling more computation at lower cost. The smartphone in your pocket contains more processing power than the computers that guided Apollo 11 to the moon—and costs a tiny fraction of what those room-sized machines did.

This exponential progress was not an accident. It was not even, strictly speaking, a prediction that happened to come true. It was a self-fulfilling prophecy, a roadmap that the entire industry followed because following it was profitable.

Chipmakers invested billions in research and development because they knew their competitors were doing the same. Equipment manufacturers built ever more sophisticated tools because they knew customers would pay for them. Software developers wrote ever more demanding applications because they knew faster hardware was coming. The expectation of continued progress created the conditions for continued progress.

Until it didn't.

There is a reason Moore's observation was never called Moore's Theory, or Moore's Principle. It was always Moore's Law—a term that carries the implication of something immutable, something beyond human control. But unlike the laws of thermodynamics or gravity, Moore's Law was never written into the fabric of the universe. It was an engineering achievement, and engineering achievements eventually run into physics. The physics that now constrains the semiconductor industry operates at scales almost too small to conceptualize. At these dimensions, the classical behavior of matter gives way to quantum mechanics, and quantum mechanics does not care about roadmaps. To understand why Moore's Law is ending, you must understand what happens inside a transistor at three nanometers.

A transistor is a switch. Current flows, or it does not. The gate that controls this flow is, on a modern chip, approximately ten atoms thick. Ten atoms. The transistor itself is approximately fifteen silicon atoms across. If you could somehow line up the gates of every transistor in a smartphone processor, the line would be shorter than a human hair is wide. At these dimensions, the rules change.

An electron is not a tiny ball that bounces predictably off barriers. An electron is a probability distribution—a smear of likelihood that extends beyond the physical boundaries of the materials that nominally contain it. When a gate is ten atoms thick, there is a meaningful probability that an electron will simply appear on the other side, having passed through the barrier without ever occupying the space within it. This is quantum tunneling. It is not a flaw in the engineering. It is not a problem that better manufacturing can solve. It is the physics of matter at scales approaching the atomic, and it sets a hard limit on how small transistors can become while still functioning as switches. The consequences are not abstract. A transistor that leaks current when it should be off wastes energy—energy that becomes heat. Heat that must be removed.

Water that must be evaporated. Power that must be generated. The smaller the transistors, the more of them fit on a chip; the more transistors, the more heat; the more heat, the more infrastructure required to prevent the chip from destroying itself. This is the physics that makes the crisis described in Chapter 1 inevitable. It is not a failure of execution. It is not a failure of investment. It is electrons doing what electrons do when you confine them in spaces measured in handfuls of atoms.

The only way out is to stop using electrons.

And quantum tunneling is not the only problem. As transistors shrink, they pack more tightly together, and the heat they generate concentrates in smaller areas. Modern processors can produce power densities exceeding 100 watts per square centimeter—comparable to the surface of a nuclear reactor. This heat must be removed, constantly, or the chip will destroy itself. But heat removal becomes harder as devices shrink, because the surface area available for cooling decreases faster than the power density increases. The industry has hit what engineers call the "power wall."

There is also the problem of wiring. At advanced nodes, more than 70 percent of a chip's area and a substantial portion of its power consumption are dedicated not to transistors but to the nanoscale wires connecting them. These wires have such high resistance at small scales that they act like tiny heaters, and the delay in signals traveling across them often becomes the primary performance limiter—not the transistor switching speed. The speed of light and the resistivity of metals now dictate performance more than the transistor itself. The industry consensus is converging on a date. Somewhere around 2030, classical silicon chips will reach the end of what physics allows. The road that Gordon Moore first described in 1965 is approaching a cliff.

If the physics of small were the only problem, the industry might find ways to work around it—at least for a while longer. But there is a second crisis unfolding in parallel, and it may prove even more constraining: the economics of extreme miniaturization are breaking down.

Moore's original insight was not just about transistor counts. It was about cost. The subtitle of his 1965 article was "Cramming More Components onto Integrated Circuits," but the point was that each component would become cheaper

as more were added. The cost per transistor would fall with each generation, making computation progressively more affordable. This economic dynamic—not just the technical achievement of miniaturization—is what made the digital revolution possible. That dynamic is now reversing. The cost per transistor has stopped falling and, by some measures, has begun to rise. The reason is visible in Veldhoven, a small city in the Netherlands that has become the most critical chokepoint in the global technology supply chain.

Veldhoven is home to ASML, a company that most people have never heard of but that every advanced technology depends upon. ASML manufactures lithography machines—the equipment that prints circuit patterns onto silicon wafers. And ASML has a monopoly on the machines required to produce the most advanced chips in the world.

The monopoly is not the result of anticompetitive behavior. It is the result of physics and decades of relentless investment. ASML's extreme ultraviolet lithography machines—EUV machines—use light with a wavelength of just 13.5 nanometers to etch patterns that are only a few atoms wide. To generate this light, the machines drop microscopic tin droplets into a vacuum chamber, then blast each droplet twice with a high-powered laser 50,000 times per second. The resulting plasma emits the EUV light that makes modern chips possible.

The engineering challenges are almost absurd. EUV light is absorbed by nearly everything, including air, so the entire optical path must be in a vacuum. The mirrors that guide the light must be polished to within a single atom of perfection—if one of ASML's mirrors were scaled to the size of Germany, its tallest bump would be less than a millimeter. The machines contain over 100,000 components sourced from more than 800 suppliers across the globe. They are the size of a bus, require multiple cargo planes to ship, and cost approximately $200 million each. The next-generation high-numerical-aperture EUV machines, now entering production, cost over $350 million. ASML spent thirty years and tens of billions of dollars developing this technology. Intel, Samsung, and TSMC—the three companies capable of manufacturing leading-edge chips—each invested billions more directly in ASML to ensure the technology would exist. No other company has the accumulated knowledge, the supplier relationships, or the financial resources to replicate what ASML has built. China has reportedly invested tens of billions attempting to develop domestic EUV capability. **Experts estimate they are at least a decade behind.**

The result is a concentration of technological power that has no parallel in modern industry. Every advanced processor—every smartphone chip, every AI accelerator, every data center server—passes through ASML's machines. There is no alternative supplier. There is no backup plan.

And the machines are only part of the cost. A leading-edge semiconductor fabrication facility—a "fab"—now costs upward of $20 billion to construct, and the figure is rising with each generation. Only three companies in the world can afford to compete at the cutting edge: TSMC in Taiwan, Samsung in South Korea, and Intel in the United States. TSMC alone manufactures more than 90 percent of the world's most advanced chips.

This concentration has consequences. When the cost of manufacturing leadership becomes prohibitive for all but a few players, the competitive dynamics that drove Moore's Law begin to break down. Investment in the next generation of technology becomes a bet-the-company decision rather than routine business. The pace of progress slows not because the physics is impossible but because the economics no longer support the pace of investment. The consequences of these twin crises—physical and economic—are already visible in the performance of new chips. In the golden era of Moore's Law, each new manufacturing node delivered roughly 40 percent more performance at the same power level, or the same performance at roughly half the power. This was Dennard scaling, the silent partner to Moore's Law that made transistor shrinking worthwhile. Smaller transistors meant faster switching and lower voltage, which meant exponential improvement in the amount of computation that could be performed per watt of electricity.

Dennard scaling ended around 2006. Since then, voltage has remained stubbornly fixed—it cannot be reduced further without causing transistors to become unreliable—and power density has risen with each new generation. The industry has compensated by shifting from faster clock speeds to more processor cores, spreading the work across multiple units operating in parallel. Your computer no longer runs faster than the one you bought five years ago; it runs more things at once. This shift to parallelism has masked the slowdown in fundamental progress, but it has limits of its own. Not all software can be parallelized efficiently. Coordinating multiple cores introduces overhead. And adding more cores does not help when the bottleneck is how quickly a single thread of computation can

execute—which is often the case for the AI workloads now driving industry demand. The performance gains from each new manufacturing node are now measured in single-digit percentages, not the 40 percent improvements of the past. Moving from 5 nanometers to 3 nanometers might yield a 10 percent speed improvement at the same power, or 25 percent power reduction at the same speed. These are valuable gains, but they are not the exponential leaps that built the digital economy.

Meanwhile, the node designations themselves have become largely meaningless. A "3-nanometer" transistor is not actually three nanometers in any meaningful dimension—the gate pitch, the fin pitch, and the metal pitch are all larger. The naming convention has become a marketing exercise, a way for manufacturers to signal continued progress even as the physical reality of that progress diminishes. Industry veterans now refer to the phenomenon with dark humor: the numbers on the roadmap bear less and less relationship to anything happening on the silicon. The industry has responded to these diminishing returns by going vertical—literally. Three-dimensional chip stacking allows manufacturers to pile multiple layers of transistors atop one another, connected by microscopic pillars called through-silicon vias. This approach sidesteps some of the limitations of planar scaling, but it introduces new challenges in heat dissipation, manufacturing yield, and design complexity. It is an engineering workaround, not a solution to the underlying physics.

The industry has also turned to specialized accelerators—chips designed for specific workloads rather than general-purpose computation. Graphics processing units, tensor processing units, and various AI accelerators perform their narrow functions far more efficiently than general-purpose processors. But this, too, is a workaround. It does not solve the fundamental problem; it merely postpones the reckoning.

• • •

In 2022, Jensen Huang, the CEO of Nvidia—the company whose graphics processors have become the workhorses of artificial intelligence—declared that Moore's Law was dead. A few days later, Pat Gelsinger, the CEO of Intel—the company that had been most closely identified with Moore's Law for half a century—countered that Moore's Law was alive and well. The exchange captured the industry's uneasy relationship with its own trajectory. The executives were talking past each other, each using "Moore's Law" to mean something different. But both understood the underlying reality: the exponential progress that had defined their industry for sixty years was coming to an end.

This matters because the demand for computation is not ending. The artificial intelligence systems described in the previous chapter require ever more processing power. The models are getting larger. The training runs are getting longer. The inference workloads are multiplying. And the chips that power these workloads are no longer getting exponentially better.

The industry's current strategy is to paper over the gap with brute force—to build more data centers, deploy more chips, consume more electricity. But brute force has limits, as we have seen. The power and water consumption of data centers are already straining communities and ecosystems. If the chips themselves stop improving at historical rates, the strain will only grow.

What is needed is not another incremental improvement in the existing paradigm. What is needed is a new paradigm altogether—a way of computing that does not rely on pushing electrons through ever-smaller silicon channels. The successor to the transistor will not be a smaller transistor. It will be something else entirely.

That something is the subject of Part II of this book. But before we can turn to solutions, we must complete our survey of the crisis. The problems of energy consumption and physical limits are only part of the story. There are also questions of infrastructure, geography, politics, and national security that must be understood.

The cloud has a weight.
The supply chain has vulnerabilities.

And the communities that bear the burden of our digital infrastructure are no longer willing to bear it silently.

The physics was never the problem. The timing was.

CHAPTER 3

The Weight of the Cloud

In January 2025, President Trump stood in the White House alongside the CEOs of OpenAI, SoftBank, and Oracle to announce the largest infrastructure project in American history. It was not a highway system. It was not a power grid. It was not a network of bridges or tunnels or dams or bullet trains.

It was a data center.

The project was called Stargate, and its ambitions matched its name. Five hundred billion dollars in investment over four years. A network of artificial intelligence factories spread across Texas, Ohio, Wisconsin, and eventually abroad. The first site, in Abilene, Texas, would be operational within months, consuming more than one gigawatt of electricity—enough to power 750,000 homes. By the time the project reached full scale, it would require more power than many small countries. The announcement was theatrical, as such things tend to be. The president compared Stargate to the Manhattan Project, invoking the national mobilization that had built the atomic bomb. OpenAI's CEO spoke of hundreds of thousands of jobs and American leadership in artificial intelligence. The numbers were almost too large to process: $100 billion in immediate deployment, nearly seven gigawatts of planned capacity, facilities spreading across multiple states and eventually multiple continents.

What the announcement did not mention was where all that electricity would come from. Or how the power lines would be built to deliver it. Or what would happen to the communities that would host these facilities. Or whether the physical infrastructure of the United States could support such ambitions at all. The cloud, it turns out, is very heavy. And the ground beneath it is beginning to buckle.

• • •

For fifteen years, American electricity demand was essentially flat. This was not supposed to happen. Utilities had built their business models on steady growth—population increases, economic expansion, the natural accumulation of appliances and devices. But beginning around 2008, efficiency gains offset new consumption. LED bulbs replaced incandescents. Factories optimized their

processes. The grid operators learned to predict demand with precision, and the system settled into a comfortable equilibrium.

That equilibrium is now shattering.

In 2023, American utilities submitted their five-year demand forecasts to regulators, as they do every year. The collective projection showed 38 gigawatts of new peak demand expected by 2028. One year later, in 2024, the same utilities submitted updated forecasts. The new projection: 128 gigawatts. In twelve months, the expected growth had more than tripled. The cause was not a mystery. Data centers, particularly those being built for artificial intelligence, were consuming electricity at rates that defied historical patterns.

A single hyperscale AI facility can draw 100 megawatts or more—the equivalent of a small city.

And the facilities were multiplying. According to research firm 451 Group, U.S. data center grid power reached approximately 62 gigawatts in 2025 and is projected to exceed 134 gigawatts by 2030. The Lawrence Berkeley National Laboratory, using more conservative assumptions, projected that data centers could consume between 325 and 580 terawatt-hours annually by 2030—representing between seven and twelve percent of total U.S. electricity consumption.

The higher estimates are staggering. Boston Consulting Group calculated that under aggressive AI adoption scenarios, U.S. data centers could consume more than 1,000 terawatt-hours annually by 2030—roughly a quarter of total American electricity generation in 2023. That scenario may be unlikely, but even the conservative projections represent a fundamental transformation of the American energy landscape. The utilities are scrambling to respond. American Electric Power, which serves eleven states across the Midwest and South, announced that it expects 24 gigawatts of new demand by 2030—of which 18 gigawatts will come from data centers alone. That is five times the company's current system capacity. Dominion Energy, which serves Virginia's data center corridor, has committed $50 billion in infrastructure investment between 2025 and 2029 simply to keep pace with growth.

But building new power plants and transmission lines takes time—far more time than building data centers. A natural gas plant requires five to seven years from

conception to operation. A nuclear plant takes a decade or more. Transmission lines face years of permitting, land acquisition, and environmental review.

The data centers are being built in months. The infrastructure to power them is measured in decades. The gap between demand and supply is opening like a chasm, and no one is quite sure how to bridge it.

• • •

Before a new power plant or data center can connect to the electrical grid, it must pass through what is known as an interconnection queue. This is the waiting room of American energy infrastructure—a bureaucratic process by which grid operators evaluate whether the system can handle new loads and what upgrades might be required to accommodate them.

The queue is broken.

According to Lawrence Berkeley National Laboratory, nearly 2,600 gigawatts of generation and storage capacity are now waiting in interconnection queues across the United States—more than twice the total installed capacity of the entire existing power plant fleet. The median wait time from initial application to commercial operation has grown from under two years in 2008 to nearly five years in 2023. In high-demand regions like Northern Virginia, delays can reach seven years or more. Over 70 percent of applications are eventually withdrawn, often because the projects are no longer economically viable by the time approval comes through. The problem is structural. The interconnection process was designed for a different era—one in which new projects trickled in gradually and could be evaluated one at a time. The system used a serial, first-come-first-served approach: each application was studied in sequence, and if any project ahead of you in the queue dropped out, the entire analysis had to be redone. Developers learned to game the system by submitting multiple speculative applications, hoping that at least one would make it through. The result was a queue clogged with phantom projects and a backlog that grows faster than it can be cleared.

In Texas, CenterPoint Energy reported a 700 percent increase in large load interconnection requests between late 2023 and late 2024, growing from one gigawatt to eight gigawatts. Utilities across the country are reporting more gigawatts of data center applications than their historical maximum peak demand. The Federal Energy Regulatory Commission has implemented reforms—shifting to a cluster-based approach, requiring higher deposits and proof

of project readiness—but the changes will take years to show results, and the queue continues to grow.

Meanwhile, the physical equipment needed to connect new loads to the grid is itself in short supply. Large power transformers—the devices that step voltage up for transmission and down for distribution—now have lead times of 120 weeks or more. Some orders stretch to 210 weeks, more than four years. The North American Electric Reliability Corporation has warned that transformer shortages pose a risk to grid reliability. Wood Mackenzie estimates that the United States faces a 30 percent supply shortfall for power transformers in 2025.

The shortage is not easily fixed. Transformer production depends on grain-oriented electrical steel, a specialized material with particular magnetic properties. Only one company in the United States—Cleveland-Cliffs—manufactures it domestically, at plants in Pennsylvania and Ohio. Roughly 80 percent of large transformers have historically been imported from Mexico, China, and Thailand. But global demand has surged, and the supply chains that served a stable grid cannot keep pace with explosive growth. The transformers that once cost $400,000 now cost $600,000 or more. Prices have risen four to six times since 2020. And even at those prices, the equipment may not arrive for years. Two data centers in Silicon Valley have been built but cannot begin operations: the transformers to supply them with electricity do not exist.

The costs of this infrastructure crisis are not being borne equally. They are landing on the communities least equipped to absorb them. In the PJM Interconnection, the regional grid operator serving much of the Mid-Atlantic and Midwest, capacity prices for the 2025-2026 planning year increased by $9.3 billion in a single auction. Residential customers in Maryland will see their bills rise by approximately $18 per month. In Ohio, the increase is projected at $16 per month. These are not one-time charges; they are permanent increases in the cost of electricity, driven largely by the infrastructure investments required to serve data centers.

A study by Carnegie Mellon University found that average American electricity bills could rise by 8 percent by 2030 due to data center growth. In Northern Virginia, the epicenter of the data center industry, the increase could reach 25 percent or more. The households paying these higher bills are not the ones benefiting from the AI revolution. They are simply the ones who happen to live

near the infrastructure that powers it. Some states are pushing back. Ohio has implemented new tariffs requiring large data centers to guarantee 85 percent minimum usage of their requested capacity—a response to speculative interconnection requests that tied up grid resources for facilities that were never built. The policy caused interconnection requests in one utility's territory to drop from 30 gigawatts to 13 gigawatts overnight, as developers were forced to demonstrate that their projects were real.

Other jurisdictions are going further. In 2024, PJM proposed rules that would require large data centers to bring their own generation—to build power plants alongside their server farms, rather than relying solely on the grid. The proposal reflects a fundamental shift in how regulators view these facilities: not as ordinary customers entitled to service, but as industrial operations that must bear the full cost of their extraordinary demands. The data center operators are responding in kind. Microsoft has contracted to restart a reactor at Three Mile Island, the nuclear plant famous for its 1979 partial meltdown, to secure dedicated power for its facilities. Amazon has signed agreements with nuclear plants in Pennsylvania and Illinois. A company called CloudBurst is building a natural gas pipeline capable of delivering 450,000 MMBtu per day—enough to generate 1.2 gigawatts of electricity—directly to its Texas data center campus. The cloud is becoming its own power system, separate from and competing with the grid that serves everyone else.

This privatization of energy infrastructure has its own consequences. When the largest electricity consumers exit the shared grid, the fixed costs of maintaining that grid are spread across fewer customers. Rates rise for those who remain. The grid becomes a service for those who cannot afford alternatives, while the wealthy build their own systems. It is a familiar pattern in American infrastructure—the decay of shared resources as the privileged opt out—now playing out at the scale of the national power system.

• • •

The constraints facing the data center industry are not merely regulatory or financial. They are physical. The materials required to build digital infrastructure are finite, and the supply chains that deliver them are fragile.

A hyperscale data center is, at its foundation, a building—and buildings require steel and concrete. The American Concrete Association projects that the United States will need approximately one million metric tons of cement for AI data

center construction by 2028. That is in addition to the concrete required for housing, roads, bridges, and all the other demands of a growing economy. The data centers are not replacing other construction; they are adding to it. Beyond the shell, data centers require specialized equipment with long lead times and limited suppliers. Uninterruptible power supply systems, power distribution units, generators, switchgear—each component passes through supply chains that were built for steady demand, not exponential growth. Some items have lead times exceeding six months. Fiber optic cable manufacturing is controlled by approximately eight suppliers globally, only two of which are in the United States.

The concentration is even more extreme for the servers themselves. The AI accelerators that power modern data centers are manufactured almost exclusively by Nvidia, using chips fabricated almost exclusively by TSMC in Taiwan. A single company's factories, located on an island 100 miles from mainland China, produce the silicon that powers the AI ambitions of every major technology company in the world. The supply chain is not merely concentrated; it is singular. And then there is land. The largest data center facilities require 500 to 800 acres of space, in locations with access to fiber optic networks, reliable power, water for cooling, and willing local governments. Such sites are not infinite. The industry is spreading from its traditional strongholds in Virginia and the Pacific Northwest to new territories: Texas, Arizona, Georgia, the Midwest. But each expansion encounters the same constraints—grid capacity, water availability, community resistance—that drove the industry away from its previous homes.

A 2025 study found that power constraints were extending data center construction timelines by 24 to 72 months. Projects that were once limited primarily by land availability and permitting are now constrained by whether utilities can deliver power at all. The bottleneck has shifted from what can be built to what can be powered. And the bottleneck is tightening.

• • •

Return, for a moment, to Stargate.

The project embodies every tension in the data center industry's trajectory. It promises transformative investment—hundreds of billions of dollars, hundreds of thousands of jobs, American leadership in the technologies that will define the

twenty-first century. And it requires transformative infrastructure—gigawatts of power generation, thousands of miles of transmission lines, fleets of transformers that do not yet exist, supply chains that cannot yet deliver at scale. Within days of the announcement, critics pointed out that the project's funding was uncertain, its timelines were aggressive, and its power requirements were essentially impossible to meet through conventional means. Elon Musk, himself no stranger to ambitious infrastructure projects, publicly questioned whether Stargate's backers actually had the money they claimed. The project's defenders countered that the skeptics underestimated American ingenuity, that the investment would flow as promised, that the infrastructure would somehow materialize.

Both sides may be right. Stargate may indeed attract hundreds of billions in investment. It may indeed create significant employment and advance American competitiveness in artificial intelligence. And it may also run headlong into the physical limits that this chapter has described—the interconnection queues that stretch for years, the transformers that take four years to deliver, the grid that cannot expand fast enough to serve it. The broader question is not whether Stargate will succeed or fail. It is whether the paradigm that Stargate represents—building ever larger electronic data centers, consuming ever more electricity, demanding ever more from infrastructure that is already strained—can continue at all. The physics described in the previous chapter suggests that the chips themselves are approaching fundamental limits. The infrastructure described in this chapter suggests that the systems supporting those chips are approaching limits of their own.

Something has to give. Either the demand for computation must be curtailed—an outcome that seems unlikely given the economic and strategic imperatives driving AI development—or the way we compute must change.

The problems are real.

The constraints are binding.

The current path is unsustainable.

But the problems are also, in a sense, clarifying. They define the specifications that any solution must meet. A sustainable computing infrastructure must consume dramatically less power. It must generate dramatically less heat.

It must operate within the physical constraints of an electrical grid that cannot expand indefinitely and a planet whose resources are finite. But the physical constraints described in this chapter are only part of the story. The infrastructure that powers computation is strained. The supply chain that builds it is something else entirely: fragile, concentrated, and vulnerable in ways that no amount of domestic investment can quickly resolve.

That vulnerability has a geography. It runs through a single Dutch company's factories in the Netherlands. It passes through fabrication plants on an island 100 miles from mainland China. It depends on materials and expertise concentrated in a handful of locations that could be disrupted by a single geopolitical crisis.

The cloud is heavy.

The stack beneath it is vulnerable.

And the communities that bear its weight are beginning to push back.

CHAPTER 4

The Vulnerable Stack

In Veldhoven, a small city in the southern Netherlands, a company called ASML manufactures machines that weigh 180 tons, cost up to $400 million each, and require three Boeing 747 cargo planes to transport. Each machine contains more than 100,000 components sourced from over 800 suppliers across the globe. The mirrors inside must be polished to atomic smoothness—deviation from perfection must be measured not in millimeters or micrometers but in individual atoms.

These machines create light that exists nowhere else on Earth. A laser fires 50,000 times per second at droplets of molten tin, vaporizing them into plasma hotter than the surface of the sun. That plasma emits extreme ultraviolet light with a wavelength of just 13.5 nanometers—light that must be captured and directed by mirrors operating in a vacuum, because air would absorb it completely. The light passes through a precise pattern and etches circuits onto silicon wafers, drawing lines 6,000 times thinner than a human hair.

ASML is the only company on Earth that makes these machines. There are no alternatives. Every advanced semiconductor fabricated anywhere in the world—the chips powering AI systems, smartphones, data centers, military systems—requires ASML's extreme ultraviolet lithography equipment. The company holds 100 percent of the EUV market and approximately 90 percent of the broader lithography market. Thirty years of research, tens of billions of euros in development, and an ecosystem of 5,000 specialized suppliers have produced what may be the most consequential monopoly in modern industry.

Chris Miller's Chip War, published in 2022, documented how this monopoly emerged and why it matters. Miller traced the decades of investment, the geopolitical maneuvering, and the technical achievements that produced ASML's dominance—and the strategic vulnerabilities that dominance creates. His book became essential reading in boardrooms and government offices precisely because it revealed how thoroughly the global economy depends on a single company, in a single city, building machines that no one else can replicate.

Since Chip War was published, the concentration has only intensified. Export controls have tightened. Capacity constraints have worsened. The question Miller's book raised but could not answer—what happens when the chokepoint fails?—is no longer hypothetical. It is the operating environment of every company that depends on advanced computation. This is not a story about corporate dominance. **It is a story about fragility.**

The semiconductor supply chain—the foundation upon which the entire digital economy rests—is concentrated in ways that would have seemed reckless if anyone had designed it deliberately. The machines come from one company in the Netherlands. The most advanced chips are fabricated overwhelmingly by one company in Taiwan. The rare gases and specialized materials flow through chokepoints that can be disrupted by wars, earthquakes, or political decisions. A single failure at any of these points could bring the global technology industry to its knees.

The production of a single advanced semiconductor involves more than 1,000 processing steps and can cross international borders 70 or more times before reaching a customer. Silicon ingots are grown in Japan, sliced into wafers in South Korea, fabricated in Taiwan, packaged in Malaysia, and assembled into devices in China. The process takes four to six months, requires uninterrupted clean-room operations lasting weeks, and depends on specialized knowledge concentrated among fewer than 50,000 engineers worldwide. At the center of this global network sits Taiwan Semiconductor Manufacturing Company. TSMC fabricates approximately 64 percent of the world's contracted semiconductors and 92 percent of the most advanced logic chips—those produced at process nodes of 7 nanometers and below. The company's customers include Apple, Nvidia, AMD, Qualcomm, and virtually every other major technology firm that designs chips but does not manufacture them. In 2024, Taiwan's total semiconductor output reached $165 billion, a 22 percent increase from the previous year, and the island's share of global advanced chip production continued to grow.

Taiwan is an island approximately the size of Maryland, located 100 miles from mainland China. The Chinese government claims Taiwan as its territory and has not renounced the use of force to achieve what it calls reunification. The Taiwan Strait, the body of water separating the island from the mainland, carries half of the world's container ships and 80 percent of the largest vessels

by tonnage. Military exercises, blockade scenarios, and the possibility of invasion are not abstract concerns; they are the subject of detailed planning by governments on both sides of the Pacific. The economic consequences of disruption would be staggering. Bloomberg Economics estimates that a Chinese blockade of Taiwan would reduce global economic output by 2.8 percent in the first year alone—approximately $2.7 trillion in lost production. Economists at Johns Hopkins calculate that a blockade scenario would immediately disrupt 90 percent of advanced chip production. It would cost an estimated $350 billion and take three years to replace Taiwan's semiconductor capacity, assuming such replacement were even possible given the concentration of expertise and equipment.

The risks are not limited to military scenarios. In April 2024, a 7.4-magnitude earthquake struck Taiwan, forcing TSMC to evacuate facilities and halt production at plants responsible for manufacturing the world's most advanced chips. The disruption lasted only hours—the company estimated losses of approximately $92 million after insurance—but it served as a reminder of how concentrated and vulnerable the supply chain has become. An earthquake, a typhoon, a pandemic, or a political crisis could interrupt the flow of chips upon which the global economy depends.

• • •

The physical infrastructure of chip manufacturing depends on materials whose supply chains are equally precarious. Consider neon gas, which is used in the lasers that perform photolithography—the process by which circuits are etched onto silicon wafers. Semiconductor-grade neon must be refined to extraordinary purity, and before 2022, Ukraine supplied approximately 90 percent of the highly purified neon imported by the United States. The gas is a byproduct of steel production, specifically from older mills equipped with air separation technology, and the former Soviet Union's industrial legacy made Ukraine the world's dominant supplier. When Russia invaded Ukraine in February 2022, two of the three major Ukrainian neon producers—Ingas in Mariupol and Cryoin in Odesa—were forced to halt operations. Neon prices in China quadrupled. Krypton prices in Japan rose by similar amounts. The semiconductor industry's dependence on Ukrainian neon, largely invisible before the invasion, became front-page news. Companies scrambled to diversify suppliers, invest in recycling technology, and build stockpiles. The immediate crisis was averted, but the vulnerability it revealed persists.

The situation with Chinese materials is more deliberate and more dangerous. China produces 98 percent of the world's primary gallium, a metal essential for compound semiconductors and advanced electronics. It controls approximately 68 percent of germanium production and dominates the processing of rare earth elements critical for everything from electric vehicles to defense systems. In July 2023,

China imposed export controls on gallium and germanium, requiring licenses for shipments and detailed end-use documentation. In December 2024, those controls escalated to an effective embargo on exports to the United States. The export controls were a direct response to American restrictions on semiconductor technology sales to China—part of an escalating trade and technology conflict between the world's two largest economies. But the underlying vulnerability is structural. China spent decades building dominant positions in critical materials, often through industrial policy and subsidies that made Chinese production cheaper than alternatives. Western countries, comfortable with cheap imports, allowed their domestic industries to atrophy. By the time policymakers recognized the strategic implications, the dependencies were deeply entrenched. By mid-2025, gallium prices in Rotterdam had risen more than 150 percent from pre-control levels.

The U.S. Geological Survey estimated that a complete gallium and germanium export ban could reduce American GDP by $3.4 billion, with nearly half the impact falling on the semiconductor industry. China temporarily lifted some restrictions following diplomatic negotiations, but the underlying message was clear: the materials upon which the technology industry depends can be weaponized at any time.

The United States and its allies have not ignored these vulnerabilities. The CHIPS and Science Act, signed into law in August 2022, authorized approximately $280 billion in new funding to strengthen domestic semiconductor manufacturing and research. The law included $52.7 billion in direct support for chip production, with $39 billion in manufacturing subsidies and 25 percent investment tax credits for fabrication equipment. The European Union committed over $100 billion to similar initiatives. South Korea announced plans to invest more than $470 billion through 2047. Japan launched aggressive programs to rebuild its chipmaking capacity.

The money has begun to flow. TSMC is building six new fabrication plants in Arizona, with the first entering high-volume production in late 2025. Intel received $7.86 billion in CHIPS Act funding for facilities in Arizona, New Mexico, Ohio, and Oregon. Samsung is constructing a major fab in Taylor, Texas. Micron secured $6.2 billion for memory chip production in Idaho and New York. By late 2025, the Semiconductor Industry Association counted more than $630 billion in announced investments across 140 projects in 28 states. But the projects have encountered difficulties that reveal the challenges of rebuilding an industrial base that took decades to offshore. TSMC's Arizona facilities faced delays attributed to labor disputes, cultural clashes between Taiwanese managers and American workers, and the simple reality that building semiconductor fabs is extraordinarily difficult. Samsung's Texas project, originally scheduled to begin production in late 2024, was pushed back repeatedly. Intel's ambitious Ohio facility, once expected to open in 2025, is now scheduled for 2030. The delays are measured not in months but in years.

More fundamentally, the reshoring effort addresses geography without solving the underlying physics. A fab in Arizona still requires ASML's lithography machines—there is no American alternative. It still requires the specialized gases, chemicals, and materials that flow through fragile global supply chains. It still depends on expertise that exists primarily in Taiwan and South Korea. The CHIPS Act may eventually produce a more distributed manufacturing base, but it cannot eliminate the chokepoints that define the industry's structure. The Relaxed Lithography approach described later in this book offers a different path. If photonic devices can be fabricated on mature process nodes—180 nanometers, 90 nanometers, geometries that dozens of foundries worldwide can produce—then the ASML chokepoint becomes irrelevant. The strategic calculus changes. The supply chain constraints that currently limit who can build advanced computing systems would no longer apply. This is not a theoretical advantage. It is a structural one—the kind of advantage that matters when export restrictions tighten, when geopolitical tensions escalate, when the single point of failure that Miller documented finally fails.

Industry analysts estimate that even with aggressive investment, the United States might produce approximately 10 percent of the world's advanced memory chips by 2035, up from less than 2 percent today. That is progress, but it is not independence. The dependency on Taiwan for the most advanced logic chips—the processors that power AI systems—will persist for years.

The dependency on ASML will continue indefinitely. The materials controlled by China will remain strategic vulnerabilities unless entirely new supply chains can be built. Taiwan's dominance of semiconductor manufacturing is sometimes described as a "silicon shield"—the idea that the island's irreplaceable role in chip production deters Chinese aggression because disrupting Taiwan would devastate China's own economy and technological ambitions. The theory has a certain logic. China imports approximately 60 percent of its chips from Taiwan, and Chinese companies depend on TSMC for the advanced semiconductors they cannot produce domestically.

But the shield may be weakening. As Western nations work to reduce their dependence on Taiwan, Taiwan's bargaining position erodes. Every fab that TSMC builds in Arizona or Japan is a fab that does not strengthen the island's strategic value. Taiwanese law prohibits the overseas production of the company's most advanced chips—currently the 2-nanometer process—but each generation of technology eventually becomes available abroad. The shield depends on irreplaceability, and irreplaceability is being systematically undermined. Meanwhile, China is investing heavily in domestic semiconductor capabilities. Chinese equipment manufacturers like Naura and AMEC have captured growing shares of certain market segments, though they remain far behind in the most advanced technologies. In December 2025, reports emerged that Chinese researchers had secretly built a prototype EUV lithography machine using knowledge from former ASML engineers, though experts estimated that producing working chips would require several more years of development. The technological gap is real, but it is not permanent.

The geopolitical trajectory is troubling. Trade restrictions escalate. Export controls multiply. Supply chains fragment into competing blocs. A technology that once represented globalization's triumph—the collaborative production of humanity's most complex manufactured object—is becoming a weapon in great power competition. The integration that made semiconductors cheap and abundant also made them vulnerable. The diversification that would make them resilient also makes them expensive and scarce.

The vulnerability of the semiconductor supply chain is not an accident.

It is the accumulated consequence of decades of decisions that optimized for efficiency over resilience, cost over security, globalization over strategic

autonomy. No single actor designed the system to be fragile; it emerged from countless choices that each made sense individually but collectively created a structure that a single earthquake, a single war, a single policy decision could disrupt catastrophically. The previous chapters described problems of physics and infrastructure—the heat that chips generate, the limits of shrinking transistors, the grid that strains to power them. This chapter has described a different kind of problem: the concentration of irreplaceable capabilities in locations that are geographically, politically, and strategically exposed. The computing infrastructure of the twenty-first century rests on a foundation that is simultaneously essential and precarious.

The industry's response to these vulnerabilities has been to build more of the same: more fabs, more supply chains, more redundancy within the existing paradigm. That response makes sense as far as it goes. But it does not address the fundamental question of whether the paradigm itself—electronic computation at ever-smaller scales, requiring ever-more-exotic materials and ever-more-concentrated expertise—can deliver the computing capacity that the world will demand. **The problems are compounding.**

The physics limits described in Chapter 2 constrain what transistors can do. The infrastructure limits described in Chapter 3 constrain how many can be powered. The supply chain limits described in this chapter constrain where they can be built and by whom. At every level of the stack, the current trajectory encounters obstacles that money and effort alone cannot overcome. And yet, within these constraints, one company has built an empire.

Any honest assessment of the computing landscape must reckon with Jensen Huang and Nvidia. The CEO of Nvidia recognized, earlier than almost anyone, that parallel processing would become the foundation of artificial intelligence. His leather jacket has become a symbol—part tech-founder uniform, part rock-star affect—and his keynote presentations draw audiences that rival product launches from Apple. Nvidia's developer conferences are not merely commercial events; they are gatherings of the faithful. Engineers get Nvidia tattoos. They speak of CUDA—the proprietary programming framework that locks developers into Nvidia's ecosystem—with the reverence that earlier generations reserved for Bell Labs or Xerox PARC. Nvidia has built the most valuable semiconductor company in history by mastering the current paradigm—by pushing electronic computation to its limits within the constraints this chapter has described.

The company's GPUs power the AI revolution. Its market capitalization exceeds two trillion dollars. Its developer ecosystem numbers in the millions. Jensen Huang is not waiting passively for the future; he is actively shaping it, investing in efficiency improvements, interconnect technologies, and responses to the power and cooling constraints that define the current crisis.

This is the incumbent that any new paradigm must displace. Not a complacent monopolist resting on legacy advantages, but a brilliant, aggressive, technically sophisticated competitor with resources measured in hundreds of billions and relationships with every major technology company on Earth. The question is whether any transition can outrun an incumbent this capable—or whether the physics advantages of alternative approaches are so fundamental that even Nvidia's resources cannot close the gap. The answer to that question is the subject of Part II.

But the stack has one more layer to examine. The data centers, the power plants, the fabrication facilities—they must be located somewhere. And the communities that host them are increasingly unwilling to bear the costs. The backlash against digital infrastructure is growing, and it threatens to constrain the industry's expansion in ways that even its leaders have begun to acknowledge.

CHAPTER 5

The Backlash

On the evening of February 14, 2023, more than five hundred residents of Warrenton, Virginia, gathered at Fauquier High School for what would become a Valentine's Day massacre of a different sort. The town council had scheduled a public hearing on Amazon's proposal to build a 220,000-square-foot data center at the eastern gateway to their community. Among those who rose to speak was Robert Duvall, the Oscar-winning actor who had made his home in the surrounding horse country. Nearly 130 speakers followed, each allocated three minutes to make their case. The hearing stretched past two o'clock in the morning. All but three opposed the project. The council voted 4-3 to approve it anyway.

What happened next defied the conventional wisdom that communities rarely sustain opposition to development once approvals are secured.

Yard signs reading "**Stop Amazon, Stop The Towers**" remained in place for months, then years. Residents formed advocacy groups. They filed lawsuits, requested public records, and began organizing for the next election cycle. When town officials refused to release correspondence between staff and Amazon's representatives, citizens' groups sued—and won at the Virginia Court of Appeals, which ruled that the town must disclose more than 3,000 emails and other documents related to the decision. By November 2024, voters had systematically removed every council member who had supported the data center.

"All of the people who voted for Amazon are gone," one of the newly elected council members observed. "This election was a referendum on Amazon."

The newly elected officials issued a joint statement declaring that "no data center has any place in a small town like Warrenton" and pledging to pursue "whatever course of action is available" to reverse the decision. The statement also promised transparency and accountability—explicit repudiations of the closed-door process that had produced the original approval. Slowly and methodically, the statement noted, "all four Warrenton Town Council members who sold out their constituents to Amazon" had been ushered out of office.

Warrenton is not an outlier. It is a harbinger.

The backlash against data centers has emerged with a speed and intensity that has caught the technology industry off guard. According to Data Center Watch, a research project that tracks opposition across twenty-eight states, community resistance blocked or delayed an estimated sixty-four billion dollars in proposed data center construction between May 2024 and March 2025. By the second quarter of 2025, that figure had grown to approximately one hundred billion dollars in a single three-month period—more than the cumulative total of every quarter since 2023. Twenty projects fell through in ninety days. The pace of opposition was accelerating faster than the pace of construction.

The geography of resistance spans the political spectrum and the national map. In Peculiar, Missouri, population six thousand, the Board of Aldermen amended the city's zoning ordinance to exclude data centers entirely after residents organized against a proposed development. In Goodyear, Arizona, a fourteen-billion-dollar project by developer Tract was withdrawn after activists pressured local authorities to block required rezoning. In Chesterton, Indiana, public forums highlighted community concerns until the developer pulled its application. In Prince George's County, Maryland, officials imposed a 180-day pause on new data center proposals after a petition garnered more than twenty thousand signatures to oppose a project that would have converted an abandoned mall into server farms.

The pattern has become sufficiently pronounced that Data Center Watch now identifies 142 advocacy groups mobilizing across twenty-four states. On Change.org, the petition platform reported that data center-related petitions surged from a single filing in all of 2024 to 113 in 2025, accumulating roughly fifty thousand signatures. Senator Bernie Sanders of Vermont announced he would push for a national moratorium on AI data center construction to "give democracy a chance to catch up" to an expansion that has proceeded largely without regulation. More than 230 environmental and community organizations signed a letter to Congress requesting a pause on new construction until federal standards are established.

For an industry that had grown accustomed to operating in obscurity, the shift has been disorienting. "Before, local opposition was more of an anecdotal possibility," observed Miquel Vila, an analyst at Data Center Watch. "Now, it's

becoming a core feature of development, in the same way issues like land, energy, and water are taken into account." The complaints against data centers cluster around a constellation of grievances, each reinforcing the others. Water consumption ranks among the most visceral, particularly in regions where drought has made every gallon precious. On May 17, 2024, the City Council of Mesa, Arizona, approved an eight-hundred-million-dollar data center development that would require up to 1.25 million gallons of water daily to keep its servers from overheating. Vice Mayor Jenn Duff cast the lone dissenting vote. "This has been the driest twelve months in 126 years," she said, citing data from the National Oceanic and Atmospheric Administration. "We are on red alert, and I think data centers are an irresponsible use of our water."

The typical data center uses between three and five million gallons of water per day, according to researchers at Texas Tech University's Water Resources Center—roughly equivalent to the consumption of a city of thirty thousand to fifty thousand people. The water is needed primarily for evaporative cooling, a process in which liquid absorbs heat from servers and dissipates into the atmosphere. Unlike industrial processes that discharge wastewater for treatment and reuse, much of this consumption simply vanishes. A study by the Houston Advanced Research Center projected that data centers in Texas alone would consume forty-nine billion gallons of water in 2025, potentially rising to 399 billion gallons by 2030. That final figure would be equivalent to drawing down Lake Mead, the largest reservoir in the United States, by more than sixteen feet in a single year.

The geography of data center expansion has proceeded with apparent indifference to hydrology. According to a Bloomberg analysis, roughly two-thirds of data centers built since 2022 have been located in water-stressed regions. Phoenix, where Maricopa County hosts the nation's second-largest concentration of data centers, sits in a desert where the state has already limited home construction to preserve groundwater. The irony is not lost on residents watching their wells run dry while gleaming server farms consume aquifers that took millennia to fill. In Spain, Amazon's three planned data centers in the Aragon region are licensed to use 755,720 cubic meters of water annually—enough to irrigate more than 570 acres of farmland. "These data centers use water that comes from northern Aragon, where I am," farmer Chechu Sánchez told The Guardian. "They consume water. Where do they take it from? They take it from you, of course." The opacity surrounding water usage has itself become a point of contention.

In Oregon, the city of The Dalles attempted to block a newspaper's request for Google's water consumption data, triggering a thirteen-month legal battle before releasing the records. When the data finally emerged, it showed that 29 percent of the city's water was flowing to a single Google facility. Lawmakers in at least eight states introduced legislation in 2025 requiring data centers to report their water use. Democratic governors in both California and New Jersey vetoed such bills, citing concerns about imposing rigid requirements on an industry of global importance. The industry's resistance to transparency has not improved its standing with communities that feel they are being asked to sacrifice resources without even knowing how much.

In Tucson, Arizona, the city council voted unanimously in August 2025 to oppose Project Blue, the code name for a proposed $3.6 billion Amazon data center campus on the city's southern edge. The decision came after months of sustained community opposition driven by fears about the project's estimated water consumption and concern over how the cost of intense energy needs might get passed on to local ratepayers. "Is the increase in tax revenue and the relatively paltry number of jobs worth the water?" asked Kathryn Sorensen, a professor at Arizona State University and former water director for Mesa. The question has no easy answer, and increasingly, communities are deciding they would rather not find out.

• • •

Carlos Yanes believes he can identify the moments when global internet activity spikes each night. That is when the sounds from the Amazon data center six hundred feet from his Prince William County, Virginia, home grow most pronounced: the revving of machinery, the whir of exhaust fans cycling to cool servers processing the world's queries. On the morning of February 17, 2025, Yanes and his neighbors in the Great Oak community awoke to a new sound—dozens of Amazon emergency diesel generators blaring after a power outage. The noise, he said, was like "a hair dryer right on your ear."

Data center noise emanates primarily from two sources. Cooling systems, which operate continuously to prevent equipment from overheating, generate sound levels ranging from 55 to 85 decibels. Backup generators, which can be the size of tractor-trailers, produce noise comparable to city traffic or jet engines when they activate. For context, the American Speech-Language-Hearing Association identifies 70 decibels as the threshold for safe prolonged exposure; above that level, hearing damage becomes a risk.

The constant drone of a data center may not reach that threshold, but it operates around the clock, replacing the silence of night with an industrial hum that never stops. "It's not necessarily that data center cooling systems have incredibly high decibel ratings—though the generators do," explained Michael, a former engineer leading noise study efforts for a Northern Virginia community group. "It's that they emit noise nearly constantly." The distinction matters. Traditional noise ordinances focus on peak levels, not duration. A data center that technically complies with decibel limits can still transform a quiet neighborhood into something that feels like living beside a highway. The steady background roar accumulates into a quality-of-life assault that residents describe in terms of lost sleep, elevated stress, and the inability to enjoy their own backyards.

Many local noise ordinances do not account for cooling systems mounted on rooftops, which are standard designs for large-scale data centers. Community activists in Northern Virginia have taken to conducting their own noise studies with professional equipment, documenting that facilities regularly exceed county limits—particularly on hot days and nights when cooling demands peak. In Fairfax County, Virginia, regulators responded by requiring data centers to maintain at least 200-foot setbacks from residential properties and mandating noise studies before and after construction. Prince William County is considering new limits of 52 decibels during the day and 47 at night—levels that some existing facilities may find difficult to meet.

The health implications remain understudied, but the concerns are real. A 2025 paper in the journal Environmental Science & Policy called for empirical research on long-term health outcomes in data center host communities, noting that "the public health implications of data center expansion remain largely overlooked" despite the rapid growth of the industry. Data centers generate significant noise pollution primarily from diesel generators and heating, ventilation, and air conditioning systems—the same equipment that also produces air pollutants harmful to nearby communities. The combination of constant low-frequency noise and intermittent generator exhaust creates an environment that some residents describe as industrial in character, incongruous with the residential neighborhoods where many facilities have been permitted.

In Memphis, Tennessee, the consequences have been more acute. In 2024, Elon Musk's artificial intelligence company xAI opened a facility called Colossus in South Memphis to power its Grok chatbot. The site is located in Boxtown,

a predominantly Black community founded by formerly enslaved people who built homes from discarded railroad boxcars. The neighborhood's cancer risk from industrial sources was already 4.1 times the national average, according to a ProPublica analysis. Shelby County receives a failing grade from the American Lung Association for ozone pollution and has the highest rates of children hospitalized for asthma in Tennessee. Into this environment, xAI installed thirty-five methane gas turbines capable of generating 420 megawatts of electricity—equivalent to a medium-sized power plant—without obtaining the required air pollution permits. The Southern Environmental Law Center estimates these turbines emit between 1,200 and 2,000 tons of nitrogen oxides annually, making xAI one of the largest sources of smog-forming pollution in the county. Researchers at the University of Tennessee found that peak nitrogen dioxide concentrations in areas immediately surrounding the facility increased by 79 percent after operations began.

Five days after Grok 4 became one of the most powerful language models in the world, Alexis Humphreys, a twenty-eight-year-old Boxtown resident, experienced her first severe asthma attack in fifteen years. "It felt like my chest was caving in," she told Time magazine. Other residents describe chronic coughs, a persistent smell of gas, and a fear of going outside. "I don't go outside no more," one community member said in a video documenting conditions. "I'm scared to breathe." The NAACP and Southern Environmental Law Center filed notice of intent to sue xAI under the Clean Air Act. More than two thousand public comments were submitted opposing the company's belated permit application, with nearly every comment in opposition. "What's happening in Memphis is a human rights violation," said KeShaun Pearson, executive director of Memphis Community Against Pollution. "Elon Musk and xAI are violating our human right to clean air and a clean, healthy environment."

If water and noise are local grievances, electricity costs have become a regional crisis. The infrastructure required to power data centers must be paid for somehow, and increasingly, that cost is showing up on residential utility bills. In the PJM Interconnection, the regional grid operator serving thirteen states from Illinois to New Jersey, wholesale electricity costs for the 2025-2026 planning year increased by more than 500 percent compared to three years earlier. The amount PJM agreed to pay to secure capacity from power plants—ensuring they are available when electricity use rises—soared from $2.2 billion in late 2022 to $14.7 billion in 2024, then climbed another 9 percent to $16.1 billion in 2025.

The independent watchdog that monitors PJM identified the primary culprit: "The current conditions in the capacity market are almost entirely the result of large load additions from data centers, both actual historical and forecast." Those capacity prices are ultimately passed down to household electricity bills. In Maryland, residential customers are seeing increases of approximately $18 per month. In Virginia, where Dominion Energy serves the densest concentration of data centers on Earth, prices rose 13 percent in 2025 compared to the previous year. In New Jersey, increases reached 21 percent. In Ohio, the increase is projected at $16 per month. These are not temporary spikes; they are structural increases in the cost of electricity, driven by infrastructure investments that would not be necessary but for the data center explosion.

A Harvard Law School paper found that Virginia and Maryland ratepayers are on the hook for billions of dollars in transmission infrastructure being built to accommodate data center growth—not residential customers. "Without reforms," the researchers wrote, "consumers will be paying billions of dollars for regional infrastructure that is designed to address the needs of just a few of the world's wealthiest corporations. Residential ratepayers are not causing PJM to plan new transmission." The cost of transmission upgrades in the PJM region reached $4.3 billion in 2024 alone, including $1.3 billion for thirty-seven projects in Ohio. These costs are passed on to electricity customers through higher rates.

The political consequences have arrived with remarkable speed. In November 2025, Abigail Spanberger won Virginia's gubernatorial race after promising to make data centers "pay their own way and their fair share" of rising electricity costs. In New Jersey, governor-elect Mikie Sherrill pledged to declare a state of emergency over electric bills on her first day in office. In Georgia, Democrats ousted two Republican incumbents from the state's utility regulatory commission, with one candidate explicitly tying price increases to data center expansion. Georgia Power had raised monthly bills six times in two years, to an average of $175; the company's proposal to spend $15 billion expanding capacity—primarily to serve data centers—had become a political liability.

An Associated Press poll from October 2025 found that electricity bills are a "major" source of stress for 36 percent of American adults. In data center corridors, that stress has found a target.

The industry's defenders point to the tax revenue that data centers generate—

and they are not wrong that the facilities pay substantial sums. But the ledger has grown more complicated as communities examine what they traded to attract that revenue. According to Good Jobs First, a nonprofit watchdog that tracks economic development incentives, at least thirty-six states have crafted subsidies specifically for data center projects. Texas will forego over one billion dollars in tax revenue in 2025 through sales tax exemptions, including exemptions on the electricity data centers consume. The state's Comptroller now estimates the program will cost $1.7 billion in 2030 alone and $9 billion between 2025 and 2030. Virginia, home to the largest data center market in the world, abates nearly one billion dollars annually in state and local sales and use taxes—without disclosing which companies benefit or how much they receive.

Over the past decade, data centers have been exempt from paying $2.7 billion in Virginia alone. Illinois lost $370 million in 2024, up from $10 million just four years earlier, an increase of 3,600 percent. The combined cost across states with substantial programs now exceeds several billion dollars annually, and the figures are rising rapidly. "We know of no other form of state spending that is so out of control," the Good Jobs First researchers concluded. The programs typically have no caps on how much any facility or company can benefit, and no limits on how much any state can lose in a given year. Virginia's data center tax exemption constitutes 80 percent of all economic incentive spending in the state. A 2024 study by the Virginia Joint Legislative Audit and Review Commission found that "like most economic development incentives, the data center exemption does not pay for itself" when considering lost revenue.

The jobs promised have often failed to materialize in numbers that match the facilities' footprints. A 2017 U.S. Chamber of Commerce report found that during construction, data centers average 1,688 workers. Once operational, they employ only 157 permanent staff. A 250,000-square-foot data center might employ twenty-five direct workers and another twenty-five contractors. "That building gets finished," observed Terry Rephann, a regional economist at the University of Virginia, "and then what you have is a data center that really doesn't, per square foot, employ a lot of people."

Georgia's legislature passed a bill in 2024 to pause the state's data center incentive program for review. Governor Brian Kemp vetoed it. In Ohio, the legislature voted to eliminate the state sales tax break—an early sign of data centers' emerging political vulnerability—to help pay for a broader income tax

cut. Governor Mike DeWine preserved it, writing in a veto message that Ohio needs the exemption to compete "with other states for technology jobs and capital investment." The race to the bottom continues, even as communities question whether they are winning or losing.

The backlash against data centers has scrambled traditional political alliances in ways that have surprised observers on both sides. Republican officials, traditionally friendly to business development, have raised concerns about tax incentives that benefit the world's wealthiest corporations while straining local resources. Democratic officials, typically focused on environmental and labor issues, have found themselves vetoing transparency requirements and defending subsidies for Big Tech. The opposition crosses party lines because the grievances are fundamentally local: rising utility bills affect Republican and Democratic households equally; noise pollution does not respect voting patterns; aquifer depletion is bipartisan.

In Warrenton, the voters who removed pro-Amazon council members included conservatives concerned about property values and progressives worried about environmental impacts. In Peculiar, Missouri, the group "Don't Dump Data in Peculiar" united residents across the political spectrum. In Prince William County, the lawsuits challenging the PW Digital Gateway project—a $24.7 billion data center complex that remains tied up in court—were filed by Citizens for Fauquier County, a group whose membership spans the ideological range of rural Virginia. The industry has begun to respond. A relatively new trade group, the National Artificial Intelligence Association, has distributed talking points to members of Congress and organized local data center field trips to better pitch voters on their value. Meta has taken out advertising campaigns to sell communities on the economic benefits of hosting facilities. Amazon released a study claiming its data centers actually drove down electricity prices by introducing a massive new customer that helps spread the cost of maintaining existing infrastructure. The argument has not convinced residents watching their monthly bills rise.

The moratoriums are spreading. In Georgia, eight counties and cities passed temporary bans on data center development in September 2025 alone. DeKalb County extended its moratorium while commissioners debated regulations that would restrict facilities to industrial areas. St. Charles, Missouri, imposed a one-year ban after residents strongly opposed a project later linked to Google.

St. Louis is considering its own pause while updating zoning codes that predate the digital age. In San Marcos, Texas, city council members moved to pause a $1.5 billion proposal after residents voiced concern that the development would create too much demand on energy in an area already vulnerable to brownouts. Atlanta prohibited data centers within a twenty-two-mile radius of the Belt-line overlay district. More than one hundred counties and cities have passed temporary moratorium, zoning limits, or new environmental rules since 2023.

The industry's political response has been mixed. Public Citizen, a consumer advocacy organization, published a report recommending that governments consider temporary moratoriums on new data center development until baseline protections are implemented. The report also called for federal regulation under bulk power market reliability standards, mandatory load flexibility requirements, and a prohibition on federal preemption that would allow states and localities to maintain their own zoning regulations. The Trump administration, by contrast, has moved in the opposite direction—reportedly preparing an executive order that would block states from creating their own AI rules and could be used to override local zoning restrictions on data center construction.

A nationwide poll found that only 44 percent of Americans would welcome a data center near their home—making the facilities less popular than gas plants, wind farms, or even nuclear reactors. The cloud has descended from abstraction to concrete reality, and the reality is unwelcome.

The phenomenon is not uniquely American. In central Mexico, Microsoft opened a large data center in 2024; residents in the area subsequently complained of more frequent power cuts and water outages lasting days or weeks. In the Netherlands, where data centers already consume an estimated 4 percent of national electricity, the government has imposed restrictions on new construction in certain regions. In Ireland, which hosts a disproportionate share of European data centers, regulators have begun refusing grid connections in areas where power demand threatens stability. The global expansion of digital infrastructure is encountering local resistance everywhere it touches down.

I have watched this backlash unfold with particular attention, because I am trying to build something different in the middle of it. In March 2024, Meta announced plans for a hyperscale data center in Richland Parish, Louisiana—a facility that will eventually consume more than 100 megawatts of power.

The company has deep relationships with state government, virtually unlimited capital, and a playbook refined across dozens of similar projects worldwide. When I mentioned to advisors that True Photonic was exploring Louisiana for our own facilities, the response was blunt: "Meta is a crusher. They will crush you."

Perhaps. But I have also done the math.

A photonic **Clean Compute Center** occupying two to four floors of a commercial building in New Orleans could match the computational output of Meta's hundred-acre hyperscale facility—at a fraction of the power consumption, a fraction of the water usage, a fraction of the grid impact. No cooling towers evaporating millions of gallons. No diesel generators rattling windows at three in the morning. No transmission lines marching across wetlands that are already disappearing into the Gulf.

If we can demonstrate that capability—if we can show that the computational power of a hyperscale facility can fit inside an office building, powered by a fraction of the grid capacity, requiring no special water allocation—the message to communities like Warrenton and Mesa and Boxtown would be unmistakable: there is another way. The era of hyperscale extraction is not inevitable. The photonic railroad is open, and everyone currently riding the silicon rail is invited to switch tracks. That is why the backlash described in this chapter matters to me not just as context but as opportunity. The communities rejecting data centers are not opposed to computation. They are opposed to a particular mode of computation—one that takes without giving, that externalizes costs onto people who never agreed to bear them, that treats water and electricity and silence as inputs to be consumed rather than resources to be shared.

Photonic computing offers a different bargain. The physics that makes it possible—light instead of electrons, femtoseconds instead of nanoseconds, transparency instead of resistance—also makes it clean. A technology that does not generate the heat does not need the water to dissipate it. A technology that does not strain the grid does not raise the bills of families who never asked to subsidize it. A technology that does not require rare earths cannot be weaponized by export controls from Beijing. The January 2024 rare earth restrictions that sent gallium prices soaring would have been irrelevant to photonic switches. The materials in the Poovey Stack are globally abundant.

The manufacturing processes use mature equipment available from dozens of suppliers. The strategic vulnerabilities described in the previous chapter simply do not apply.

This is not idealism. It is physics, translated into business strategy.

The communities that have said no to hyperscale data centers might say yes to something that solves their problems rather than creating them. The backlash is not an obstacle to photonic computing. It is the market signal that photonic computing was designed to answer.

The crisis is now complete.

In the preceding chapters, we have traced the impossible problem from its origins to its current manifestations. We have seen how the demand for artificial intelligence is driving energy consumption to levels that strain electrical grids and deplete water resources. We have seen how the physics of shrinking transistors is approaching fundamental limits, ending the exponential progress that sustained the digital revolution for sixty years. We have seen how the infrastructure required to power computation—the transformers, the transmission lines, the generating capacity—cannot expand fast enough to meet exploding demand. We have seen how the supply chain that builds the chips themselves is concentrated in locations that are geographically precarious and politically vulnerable. And now we have seen how the communities that host this infrastructure are pushing back. The sixty-four billion dollars in blocked or delayed projects is not merely a business problem for developers. It is a signal that the current trajectory is unsustainable not just physically, not just economically, but politically. The social license to build data centers—the implicit consent of communities to host the physical infrastructure of the digital economy—is being withdrawn.

The activists who have organized against data centers are not Luddites opposed to technology. Many use the very services that data centers enable. They shop online, stream video, query chatbots, store their photographs in the cloud. Their opposition is not to computing but to a particular mode of computing—one that externalizes costs onto communities that derive little benefit, that consumes resources without transparency, that offers tax revenue in exchange for social disruption and then extracts the tax incentives that would have made the exchange worthwhile. Their question is simple: **Is there a better way?**

The industry's response has been to build more of the same, faster. More data centers. More power plants. More transmission lines. More cooling towers. The Stargate project promises five hundred billion dollars in investment; other hyperscalers are committing similar sums. The assumption underlying these commitments is that the fundamental architecture of computation cannot change, and that the only path forward is to feed the machine more resources than it currently consumes.

But what if the assumption is wrong?

What if there is another way to compute—one that does not generate the heat that requires water to dissipate, that does not consume the electricity that strains grids and raises bills, that does not demand the exotic materials concentrated in vulnerable supply chains? What if the sixty years of electronic computation that brought us to this crisis were not the end of the story, but a prologue to something different? The crisis has defined the specifications that any solution must meet. A sustainable computing infrastructure must consume dramatically less power per operation. It must generate dramatically less heat per calculation. It must operate within the physical constraints of an electrical grid that cannot expand indefinitely and communities that will not accept indefinite burdens. It must be buildable with materials and equipment that are not concentrated in a single earthquake zone or controlled by a single geopolitical rival.

Those specifications sound impossible. For sixty years, they would have been impossible. The path of computation ran through smaller transistors, more electrons, more heat, more power, more of everything.
But there is another path. It runs not through electrons but through photons—through light itself. And the physics of that path is different in ways that matter profoundly for everything we have discussed.

Part I of this book has described the impossible problem. Part II will introduce the nature of light as a medium for computation.

The transition from crisis to solution begins now.

What if the data center that powers artificial intelligence required no cooling towers, no river, and one-tenth the electricity?

PART II

THE NATURE OF LIGHT

In a femtosecond, light travels three hundred nanometers. In that same moment, an electron has barely begun to move.

CHAPTER 6

What Light Can Do

In 1905, a twenty-six-year-old patent clerk in Bern, Switzerland, published four papers that would reshape humanity's understanding of the physical world. The most famous of these introduced special relativity and the equation $E = mc^2$. But another paper, published in March of that year, addressed a puzzle that had confounded physicists for decades: the photoelectric effect.

The problem was simple to describe and impossible to explain. When light struck certain metals, it ejected electrons from the surface. This much was expected—light carried energy, and energy could dislodge particles. But the details were wrong. Classical physics predicted that brighter light should eject faster electrons, since more energy was hitting the surface. Instead, experiments showed that the speed of ejected electrons depended only on the color of the light, not its intensity. Red light, no matter how bright, could not eject electrons from certain metals at all. Violet light, even when dim, sent them flying.

Albert Einstein's explanation was radical. Light, he proposed, was not a continuous wave but a stream of discrete packets—particles of energy that would later be called photons. Each photon carried a specific quantity of energy determined by its frequency. A single violet photon packed enough energy to knock an electron loose; a trillion red photons did not, because each individual packet was too weak. The intensity of light determined how many photons arrived, but the frequency determined what each photon could do.

The Nobel Committee awarded Einstein the 1921 prize not for relativity but for this insight into the nature of light. It was, in many ways, the more revolutionary idea. Relativity told us that space and time were stranger than we had imagined. The photoelectric effect told us that light itself—the phenomenon humans had used to see, to signal, to measure—was something we had fundamentally misunderstood.

A century later, that misunderstanding has consequences for everything described in Part I of this book. The data centers consuming gigawatts of

power, the cooling towers evaporating billions of gallons of water, the transformers straining under unprecedented loads, the communities blocking construction of facilities they cannot afford to host—all of these crises trace back to a choice made in the middle of the twentieth century. We chose to compute with electrons. We could have chosen photons. The choice made sense when it was made. But the physics of electrons and photons are profoundly different, and those differences matter more with each passing year. To understand why light might solve computing's impossible problem, we must first understand what light is—and what it can do that electrons cannot.

• • •

The physics in this chapter took me three years to learn. I am not a physicist. I am not an engineer. I am a businessman who has spent four decades learning to recognize incomplete pictures—and learning that you cannot complete a picture you do not understand. When Del first explained what the Poovey Switch could become, I nodded in the right places but grasped perhaps a tenth of what he was saying. The terms were unfamiliar. The concepts were abstract. I understood that 150 femtoseconds was fast, but I did not understand why it mattered. I understood that photons generated less heat than electrons, but I could not have explained the physics to anyone else.

Del taught me. Patiently, repeatedly, with diagrams on napkins and whiteboards and, eventually, actual technical papers that I read with a dictionary of physics terms open beside me. Mark Newland filled in the engineering implications. Gary Poovey, walked me through the materials science that made his switching material possible.

I learn fast - though the education will never be complete. But I learned enough to hold my own in conversations with researchers who had spent careers in optics, enough to write the white papers and articles that have introduced the Poovey Stack to investors and partners, enough to explain to you, in the pages that follow, what light can do and why it matters for the future of computation. What I am about to describe is not speculation. It is physics—the same physics Einstein elucidated in 1905, extended and refined by a century of research. The applications are new. The principles are as old as light itself.

Begin with the electron.

An electron is a fundamental particle with a property called charge. When electrons move through a conductor, their collective motion constitutes electrical current. This current can be switched on and off, directed through gates, manipulated to represent the ones and zeros of binary logic. The transistor—the device that makes modern computing possible—is essentially a valve that controls the flow of electrons through a semiconductor material.

But electrons do not move through conductors the way water flows through a pipe. They interact constantly with the atomic lattice of the material. An electron drifting through copper encounters nuclei, other electrons, vibrating atoms. Each encounter deflects its path, absorbs some of its momentum, converts some of its kinetic energy into thermal energy. This is electrical resistance, and it is not a design flaw that better engineering can eliminate. It is a fundamental property of charged particles moving through matter.

The consequences are inescapable. Every time an electron passes through a transistor, some of its energy becomes heat. Every time a signal travels along a wire, some of its energy becomes heat. Every computation performed by an electronic processor—every multiplication, every comparison, every bit flipped in memory—releases thermal energy into the surrounding environment. The faster the processor runs, the more operations it performs, the more heat it generates. This is not a limitation of current technology that future innovations will overcome. **It is physics.**

In 1961, the physicist Rolf Landauer of IBM quantified the minimum possible energy cost of computation. Even in a theoretically perfect computer operating at room temperature, he showed, erasing a single bit of information must release at least 2. - 75 × "10-21" joules of heat—a quantity now known as the Landauer limit. This is the thermodynamic floor below which computation cannot go, regardless of how clever the engineering becomes. Modern processors operate roughly a million times above this limit, meaning enormous room exists for improvement in principle. But the practical improvements have slowed, and the gap between theoretical minimum and actual performance has stubbornly refused to close.The reason is not that engineers have run out of ideas. The reason is that electrons, by their nature, convert energy to heat through every material they traverse. The resistance can be reduced but not eliminated. The switching can be optimized but not made lossless. **The heat will always come.**

Now consider the photon.

A photon is a quantum of electromagnetic radiation—a discrete packet of oscillating electric and magnetic fields propagating through space at approximately 299,792,458 meters per second. Unlike the electron, the photon carries no charge. It does not interact with atomic nuclei through electromagnetic forces in the same way. It does not scatter off every atom it encounters. In transparent materials, photons can travel extraordinary distances with minimal loss.

When light passes through a fiber-optic cable, for instance, it can travel hundreds of kilometers before the signal degrades appreciably. The loss in modern optical fiber is approximately 0.2 decibels per kilometer at certain wavelengths—meaning that 95 percent of the light entering the fiber emerges after traveling ten kilometers. The photons are not bouncing off atoms, losing energy with each collision. They are propagating through the material's electromagnetic structure, interacting weakly with the medium, converting almost none of their energy to heat. This property is why the global telecommunications network runs on light rather than electricity. The transatlantic cables that carry internet traffic between continents are not copper wires conducting electrical current; they are glass fibers guiding photons. The choice was made decades ago for simple economic reasons: transmitting information as light requires vastly less energy than transmitting it as electricity, and the infrastructure to amplify and regenerate signals can be spaced much farther apart. The world's data already moves as photons. It is only processed as electrons. The distinction matters because processing—computation—is where the energy crisis originates.

A Google search consumes most of its electricity not in transmitting your query to the data center or sending the results back to your screen, but in the calculations performed on servers once the query arrives. An AI model drawing 100 megawatts is not spending that power moving data through fiber-optic cables; it is spending it running electrons through transistors, generating the heat that must be carried away by cooling systems consuming millions of gallons of water.

What if the processing happened in light? The physics of photonic computation is not speculative.

The fundamental operations required for computing—switching signals on and off, combining inputs to produce outputs, storing information for later

retrieval—can all be performed with light. The question has never been whether photons can compute, but whether they can compute practically. Consider switching. A transistor in an electronic processor operates by controlling the flow of electrons through a channel. Apply a voltage to the gate, and current flows; remove the voltage, and current stops. The transition between states—from off to on, from zero to one—takes time. In a modern high-performance processor, that transition occurs in approximately one nanosecond: one billionth of a second, a duration in which light travels roughly thirty centimeters.

Optical switching can occur in femtoseconds. A femtosecond is one quadrillionth of a second—one millionth of a nanosecond. In a femtosecond, light travels approximately 0.3 micrometers, about the wavelength of ultraviolet radiation. The switching happens so fast that the term itself requires recalibration of intuition. If a nanosecond were stretched to one second, a femtosecond would be one millionth of that second—one microsecond in the analogy, a duration so brief it would be imperceptible even in the slowed-down timescale.

The speed differential is not incremental. It is not a matter of photons being somewhat faster than electrons. The physics permits operations occurring one thousand to ten thousand times faster than electronic switching allows. This is the difference between a car traveling at highway speed and a spacecraft traveling at orbital velocity. They are not in the same category.

Speed alone would be valuable. But the more consequential difference involves energy. An electronic transistor switching states must charge and discharge capacitances—tiny stores of electrical energy built into the device structure. Every transition requires pushing electrons, and pushing electrons requires overcoming resistance and dissipating power. A photonic switch manipulates light, and light passing through transparent materials generates negligible heat. The energy required to perform an optical switching operation can be orders of magnitude lower than the energy required for an equivalent electronic operation. The numbers vary depending on implementation, but the theoretical potential is staggering.

Analyses of photonic computing architectures suggest power reductions of 90 to 95 percent compared to equivalent electronic systems—not through incremental improvements, but through the fundamental physics of using a medium that does not convert energy to heat in the same way.

Ninety percent power reduction means a data center consuming 100 megawatts could perform the same computation while drawing 10 megawatts. It means the cooling infrastructure that dominates operating costs could shrink proportionally. It means the water currently evaporating through cooling towers could remain in aquifers and reservoirs. It means the electrical grid currently straining under data center load could accommodate far more computation without the infrastructure investments now projected to cost hundreds of billions of dollars.

The mathematics are simple.

The physics is established.

The question has always been whether the engineering could make it real.

• • •

The properties that make photons attractive for computation also introduce challenges that electrons do not face. Electrons interact with each other. This interaction is what makes electronic logic possible. A transistor works because electrons flowing through one channel can be influenced by electrons accumulated in a nearby gate. The charged particles affect each other's behavior, and this mutual influence enables computation.

Photons, in their ordinary behavior, do not interact with each other.

Two beams of light can cross paths without affecting each other at all—this is why you can see objects illuminated by multiple light sources without the beams interfering in complicated ways. The property that makes photons energy-efficient—their weak interaction with matter and each other—also makes them difficult to use for logic. How do you build an AND gate, which outputs a signal only when both inputs are present, if the inputs cannot influence each other? This is not an insurmountable problem, but solving it requires care. Certain materials and configurations can make photons interact indirectly—through the medium they traverse, through nonlinear optical effects, through coupling with electronic systems at specific points. The solutions vary in their elegance and their practicality. Some approaches require exotic materials that are difficult to manufacture. Some require intense laser light that reintroduces energy consumption concerns. Some work in laboratory demonstrations but prove fragile or inconsistent in real-world conditions.

The history of photonic computing is substantially a history of attempts to solve this interaction problem—to find ways of making light do logic without sacrificing the advantages that made light attractive in the first place. That history is the subject of the next chapter. The attempts span decades, involve some of the most accomplished physicists and engineers of the twentieth century, and resulted in repeated frustration. The fact that we are returning to the subject now—after so many false starts—reflects not a triumph of hope over experience, but a genuine change in what is possible.

Before turning to that history, though, it is worth understanding more precisely what photons can do when properly harnessed. The physics is not merely different from electronics. It is different in ways that align remarkably well with the specific demands of artificial intelligence—the application that has pushed electronic computing to its breaking point. The computational demands of artificial intelligence are unlike the demands of traditional software. When you run a word processor or a spreadsheet, the computer executes a series of sequential operations: load this value, add it to that value, store the result, check this condition, branch to that instruction. The operations are complex in aggregate but simple individually, and they follow a logical sequence that mirrors how humans think about the problem being solved.

Neural networks—the computational structures underlying modern AI—work differently. A neural network consists of layers of artificial neurons, each connected to neurons in adjacent layers by weighted connections. When an input enters the network, every neuron in the first layer processes it simultaneously, producing outputs that flow to the second layer, where every neuron processes all incoming signals simultaneously, and so on through potentially hundreds of layers. The fundamental operation is not sequential logic but parallel matrix multiplication: taking large arrays of numbers and combining them according to mathematical rules, over and over, billions of times per inference.

Electronic processors can perform matrix multiplication, but they were not designed for it. A traditional CPU processes one or a few operations at a time, cycling through calculations sequentially even when the underlying mathematics is inherently parallel. Graphics processing units—GPUs—were developed for rendering images, a task that involves similar parallel calculations, and were adapted for AI workloads because their architecture aligned better with neural network requirements. Nvidia's dominance in AI hardware stems from this

architectural match: their GPUs can perform many calculations simultaneously, accelerating the matrix operations that neural networks require. But even GPUs are electronic devices, subject to all the limitations described above. The electrons still flow through silicon, still encounter resistance, still generate heat. The massive clusters of GPUs that train large language models consume megawatts of power not because matrix multiplication is inherently expensive, but because electronic implementation of matrix multiplication is expensive.

Photons offer a different approach. Light has a property that electrons lack: it can perform certain mathematical operations as a natural consequence of how it propagates through optical systems. When light passes through a lens, the lens performs a Fourier transform on the light—a mathematical operation fundamental to signal processing. When multiple beams of light combine at a detector, their intensities add, performing summation. When light passes through materials with varying optical properties, the pattern of transmission performs multiplication. The operations that dominate AI workloads—the matrix multiplications and accumulations at the heart of neural network inference—can be executed by light passing through appropriately designed optical elements.

This is not a theoretical curiosity. It is a physical reality that has been demonstrated repeatedly in laboratory settings. A beam of light encoded with input values, passing through an optical matrix of controlled transmission elements, emerges with output values that represent the matrix-vector product—the core operation of neural network computation. The calculation happens at the speed of light, limited only by how fast the optical elements can be reconfigured. And because light passing through transparent materials generates minimal heat, the operation consumes a fraction of the energy required by electronic equivalents.

The alignment between photonic physics and AI requirements is not coincidental. Matrix operations are among the few computational tasks for which analog approaches—computing with continuous physical quantities rather than discrete digital bits—can outperform digital electronics. Light, as an electromagnetic wave with continuous amplitude and phase, is naturally suited to analog computation. The resurgence of interest in photonic computing tracks almost exactly with the resurgence of neural networks: both emerged from long periods of relative dormancy within the past decade, and both are flourishing because the computational demands of modern AI reward their particular strengths.

The energy implications deserve concrete illustration.

Training GPT-4, the large language model that powers ChatGPT, reportedly required approximately 50 gigawatt-hours of electricity over several months—enough to power 4,500 American homes for a year. The training run consumed electricity continuously, around the clock, with thousands of GPUs running at near-maximum capacity. The heat generated required constant cooling; the cooling required water and electricity of its own. The total carbon footprint has been estimated at thousands of tons of CO_2, depending on the power sources feeding the data centers where training occurred.

Inference—the process of actually using a trained model to answer questions—consumes less energy per query than training consumed per example. But inference happens billions of times per day across millions of users. The cumulative energy consumption of inference is now believed to exceed training, and the gap widens as AI deployment accelerates. Every query to ChatGPT, every image generated by DALL-E, every video analyzed by automated systems, draws electricity that ultimately becomes heat.

A photonic system performing the same matrix operations at 90 percent lower energy consumption would not merely reduce electricity bills. It would transform the economics and physical footprint of AI deployment.

A facility that currently requires 100 megawatts could deliver equivalent computation with 10 megawatts. The cooling infrastructure sized for 100 megawatts of heat dissipation could serve facilities performing ten times the computation. The water consumption that has made data centers unwelcome in drought-stricken regions could fall to levels that communities might accept. The numbers are not precise—real-world implementations involve compromises, and the exact efficiency gains depend on architectural details that vary between systems. But the order of magnitude is consistent across analyses. Photonic computing does not promise marginal improvements. It promises improvements commensurate with the crisis: reductions in energy and heat sufficient to change whether AI development is sustainable. There is a temptation, when confronting a crisis as severe as the one described in Part I, to grasp at any solution that promises relief. The history of technology is littered with approaches that seemed revolutionary in the laboratory but proved impractical in the field. Readers who have followed claims of technological transformation before—

cold fusion, superconducting computers, quantum computing's imminent arrival—may reasonably suspect that photonic computing belongs in the same category: promising in principle, perpetually five years away in practice.

The skepticism is warranted, and the next chapter will take it seriously. Photonic computing has a history, and that history includes decades of disappointment. Researchers in the 1980s and 1990s believed they were on the verge of optical computers that would supplant electronics. Companies invested hundreds of millions of dollars. Governments funded major research programs. The revolution did not arrive. The electronics industry continued its relentless improvement, and photonics remained confined to communications—moving data between electronic processors but never replacing them.

Understanding why previous attempts failed is essential to understanding why the current moment is different. The failures were not random; they stemmed from specific technical barriers that could not be overcome with the materials, devices, and fabrication techniques available at the time. Those barriers have not all disappeared, but some have, and others have been circumvented through approaches that were not available to earlier researchers.

The physics described in this chapter has not changed. Light still propagates faster than electrons. Photons still generate less heat than charge carriers traversing resistive materials. Optical systems still perform certain mathematical operations with extraordinary efficiency. What has changed is the ability to build practical systems that exploit these properties—systems that can be manufactured at scale, integrated with existing infrastructure, and deployed in the real world where data centers currently consume rivers and burn coal.

Part I of this book described an impossible problem. The energy demands of artificial intelligence are growing faster than our ability to supply them sustainably. The physics of electronic computation imposes fundamental limits on efficiency. The infrastructure required to power current approaches cannot be built fast enough, and the communities asked to host it are refusing. Something must change, or the AI revolution will consume itself.

Part II introduces the nature of the change that is possible. Light can compute. Light can compute fast—a thousand to ten thousand times faster than electronics. Light can compute efficiently—consuming a fraction of the energy and

generating a fraction of the heat. Light can compute in ways particularly suited to the matrix operations that dominate AI workloads.

The question is not whether light can solve the problem. The question is whether we can build machines that harness what light can do.

That question has been asked before. The answers have been disappointing. But the asking has never stopped, and the conditions for success have never been better aligned.

The history of those attempts—and what has changed—is where we turn next.

The question was never whether light could compute. The question was whether anyone could build it.

CHAPTER 7

The History of Photonic Computing

In 1990, the cover of Scientific American featured an image that captured the technological optimism of its era: a beam of light passing through a prism, refracting into a spectrum of colors, above the headline "The All-Optical Computer." The accompanying article, written by researchers from AT&T Bell Laboratories, described a future in which computation would occur entirely in photons, without the heat and limitations of electronic circuits. The authors were not speculating about distant possibilities.

They were reporting on working devices in their laboratory—optical logic gates, optical memory, optical switches—and predicting that practical optical computers were perhaps a decade away. That decade passed. Then another. The all-optical computer did not arrive.

The history of photonic computing is a history of brilliant ideas, substantial investment, and repeated disappointment. Understanding that history is essential to evaluating claims made today, because the physics that makes light attractive for computation has not changed since 1990. What has changed—and what makes the current moment genuinely different—is a combination of factors that previous generations of researchers could not have anticipated. To understand why this time might be different, we must first understand why all the previous times were not.

The dream of optical computing predates the electronic computer itself. In 1819, the French physicist Augustin-Jean Fresnel demonstrated that light behaves as a wave, establishing principles that would eventually enable the manipulation of optical signals with precision. By the late nineteenth century, Lord Rayleigh had shown that optical systems could perform Fourier transforms—mathematical operations fundamental to signal processing—simply by passing light through appropriate lenses. The physicist Ernst Abbe developed diffraction theory that explained how optical elements could manipulate wavefronts. The mathematics of light manipulation was established more than a century before the first electronic computer was built.

The first practical optical computing systems emerged in the 1960s, driven by military applications. The United States Air Force needed to process radar signals faster than electronic systems could manage, and researchers discovered that optical correlators—devices that compared incoming signals against stored patterns using the physics of light interference—could perform the necessary calculations at extraordinary speed. By 1964, the Perkin-Elmer Corporation had built optical processing systems for the National Reconnaissance Office, analyzing satellite imagery using coherent light rather than electronic circuits. The systems were specialized, expensive, and classified, but they demonstrated that light could perform useful computation.

The invention of the laser in 1960 accelerated research dramatically. Lasers provided coherent light—waves oscillating in perfect synchronization—which was essential for precise optical manipulation. Through the 1970s, research groups at Bell Labs, IBM, Stanford, and universities across the world explored optical approaches to computing. The work was largely theoretical and experimental, building understanding of what might be possible without yet producing practical devices.

The breakthrough moment appeared to arrive in 1985, when researchers at Bell Labs demonstrated the first optical bistable device small enough to be integrated onto a chip. A bistable device can exist in two stable states—the optical equivalent of a transistor's on and off positions—and can switch between them when triggered by an incoming light signal. The Bell Labs device, called a Self Electro-optic Effect Device or SEED, was fabricated using semiconductor manufacturing techniques similar to those used for electronic chips. It seemed to offer a path from laboratory curiosity to manufacturable technology. The enthusiasm that followed was intense and, in retrospect, premature. DARPA, the Defense Advanced Research Projects Agency, launched major funding programs for optical computing research. AT&T, IBM, and other technology giants invested heavily. Academic conferences on optical computing drew hundreds of researchers. The 1990 Scientific American article was not an outlier; it reflected a genuine consensus among experts that optical computing's moment had arrived.

The consensus was wrong.

The failures of optical computing in the 1990s were not failures of physics.

The physics worked exactly as the researchers understood it. The failures were failures of engineering, economics, and timing—a convergence of practical obstacles that proved far more difficult to overcome than the underlying science had suggested.

The first obstacle was materials. The optical devices that performed logic operations required exotic semiconductor compounds—gallium arsenide, indium phosphide, and more complex alloys—that were expensive to produce and difficult to manufacture at scale. Silicon, the material that had enabled the explosive growth of electronics, was poorly suited to optical applications because it does not naturally emit or efficiently manipulate light. The photonics industry was attempting to build a new manufacturing infrastructure from scratch, at enormous cost, while the electronics industry was refining processes it had perfected over decades.

The second obstacle was integration. An optical computer requires not just optical logic gates but optical memory, optical interconnections, and interfaces to the electronic systems that would provide inputs and receive outputs. Each component presented its own challenges. Optical memory, in particular, proved extraordinarily difficult. The devices that could store information in optical form required constant power to maintain their state, consumed substantial energy, and could not match the density or reliability of electronic memory. A computer without adequate memory is not a computer at all, merely a demonstration.

The third obstacle was the relentless improvement of electronics. Gordon Moore's observation continued to hold throughout the 1990s and into the 2000s. Every year, electronic transistors became smaller, faster, and cheaper. The performance targets that optical computing was aiming for kept receding as electronics achieved them first. An optical device that would have been competitive in 1990 was hopelessly outclassed by 1995; a device competitive in 1995 was obsolete by 2000. The optical computing community was running a race in which the finish line moved faster than they could approach it.

The fourth obstacle was perhaps the most fundamental: optical computing was attempting to replace electronics in applications where electronics already worked well. The computers of the 1990s performed word processing, spreadsheet calculations, database queries, and simple graphics rendering. These tasks were sequential in nature, processing one instruction after another,

and electronic processors handled them with increasing efficiency as Moore's Law delivered its annual improvements. Optical computing offered potential advantages in speed, but those advantages were irrelevant if the applications did not require them. A word processor does not need to perform calculations in femtoseconds; it needs to respond quickly enough that the human user does not notice delay. Electronic processors exceeded that threshold decades ago.

By the late 1990s, the major corporate research programs had largely shut down. Bell Labs, restructured and diminished after the breakup of AT&T, no longer had the resources for long-term fundamental research. IBM shifted its optical efforts toward telecommunications rather than computing. DARPA funding dried up as the promised breakthroughs failed to materialize. The researchers who had devoted careers to optical computing dispersed to other fields or retired. The dream did not die—academic research continued at reduced levels—but it faded from the technological mainstream.

The all-optical computer, it seemed, had been a mirage.

• • •

I think about those researchers sometimes—the ones who spent the 1980s and 1990s chasing the optical computing dream, only to watch it evaporate. They were not wrong about the physics. They were not wrong about the potential. They were simply early, in a field where being early and being wrong look identical until history decides otherwise. **Gary Poovey was one of them.**

When Gary developed his switching material, he was working within a research tradition that the broader industry had abandoned. He spent years refining compositions, testing optical properties, pursuing validation—all while the conventional wisdom held that optical computing was a dead end. His team in Israel invested millions of dollars. They proved the physics worked. They achieved switching speeds that the scientific literature said were impossible. And then they hit the same wall that had stopped everyone else.

The physics was validated, but there was no path to a product. No architecture to plug the switch into. No manufacturing strategy. No business model. No one willing to fund an octogenarian physicist pursuing technology that the industry had written off decades earlier.

When Gary walked into our Santa Rosa lab with his Technion report, he was carrying not just his own work but the accumulated weight of a field's unfulfilled promise. The researchers who had given up, the programs that had been shuttered, the careers that had been redirected—all of that history was present in his data.

What Del saw, and what I came to understand, was that Gary had solved the hardest part of a problem that everyone else had abandoned. The switching speed was real. The physics was validated. What remained was everything else: the logic gates, the architecture, the manufacturing strategy, the business model, the narrative that could attract capital and talent to finish what Gary had started.

The history in this chapter is not abstract to me. It is the context in which True Photonic exists. We are not the first to attempt optical computing. We are the latest in a long line, and we are building on foundations laid by researchers who never got to see their work completed.

If we succeed, it will be because we learned from their failures. If we fail, we will join them in the history of promising approaches that could not cross the gap from laboratory to market.

My team intends not to fail.

What emerged from the wreckage of optical computing research was something more modest but commercially viable: optical communications. The physics that made photons attractive for computation—their ability to travel long distances without losing energy, their immunity to electromagnetic interference, their capacity to carry multiple signals simultaneously through wavelength multiplexing—made them even more attractive for moving data between locations.

The telecommunications industry adopted fiber optics enthusiastically, replacing copper cables with glass fibers that could carry thousands of times more information over thousands of times greater distances. By 2000, the backbone of the global internet was optical. The transatlantic cables, the transcontinental links, the connections between major data centers—all ran on light. But the light stopped at the edge of the computer. Signals arrived as photons, were converted to electrons for processing, and were converted back to photons for transmission to the next destination. The conversion happened billions of times per second

across the global network, at every router, every server, every switch. Each conversion consumed energy and introduced delay. The optical network was a series of islands connected by light, with electronic computation happening on each island.

This architecture reflected a practical compromise.

The conversion between optical and electronic signals required devices called transceivers, and transceivers were expensive—but they were less expensive than replacing all the electronic computing equipment with optical alternatives that did not yet exist. The compromise worked well enough that it became entrenched. The telecommunications industry optimized for optical transmission; the computing industry optimized for electronic processing; the interfaces between them were accepted as a necessary cost.

Through the 2000s and into the 2010s, this division of labor held stable. Data centers grew larger, but their internal architecture remained electronic. The servers, the switches, the storage systems—all processed information as electrons. The optical connections between data centers grew faster and more numerous, but the computation that justified those connections remained stubbornly electronic.

Photonics had found its niche, and that niche was **communication** rather than **computation.**

CHAPTER 8

The Breakthrough

The history of technology is populated by ideas that worked in principle but not in practice, by demonstrations that dazzled in laboratories but withered in the field, by breakthroughs announced with fanfare that faded into footnotes. The skepticism warranted by this history applies with particular force to photonic computing, a field whose previous chapter described decades of disappointment and promises unfulfilled. Any claim to have solved the problems that defeated Bell Labs, IBM, and DARPA-funded research programs deserves the most rigorous scrutiny.

This chapter examines one such claim.

It introduces an approach to photonic computing developed by a company called True Photonic, founded on technology invented by Dr. Gary Poovey, Del Wolverton, and Derek Bailey (myself). The examination will be specific where specificity is possible: naming the architecture, describing its principles, identifying where it has been validated and where validation remains incomplete. The tone will remain analytical rather than promotional, because the reader who has followed the crisis described in Part I deserves honest assessment rather than marketing language.

The approach is called the Poovey Stack. Its central claim is direct: it performs logic operations in light, addressing the photon interaction problem that stymied previous generations of researchers through an architecture that requires neither cryogenic temperatures nor cutting-edge fabrication equipment. If the claim is valid, the Poovey Stack represents the breakthrough that photonic computing has sought for forty years. If it is not, it represents another entry in a long catalog of false starts. The stakes justify careful attention. The crisis is real. The demand for a solution is urgent. And the specific characteristics of this approach—if they hold—address not merely the physics of computation but the manufacturing, infrastructure, and geopolitical constraints described in earlier chapters.

Recall the central obstacle that previous chapters identified: photons, unlike electrons, do not naturally interact with each other.

This property is what makes light attractive for computation—**photons traversing transparent materials generate minimal heat, carry information at extraordinary speed, and consume far less energy than electrons encountering resistance in silicon**.

But the same property that enables efficiency creates a fundamental challenge for logic. **Computation requires signals to influence each other.** An AND gate must produce an output only when both inputs are present. A transistor must allow current to flow or block it based on the state of a control signal.

Without interaction, there can be no logic, and without logic, there can be no computation.

The challenge runs deeper than it might first appear. In electronic circuits, the solution is elegant: electrons carry charge, charge creates electric fields, and electric fields influence the behavior of other electrons. The transistor exploits this chain directly—electrons accumulated in a gate electrode create a field that either permits or blocks current flow through a nearby channel. The interaction is inherent in the physics of charged particles. No special materials or conditions are required; the effect is present whenever electrons are present.

Photons have no charge. They do not create static electric fields. Two beams of light can cross paths in empty space without affecting each other at all—this is why you can see objects illuminated by multiple light sources without the beams interfering in complicated ways. The property is not a limitation that clever engineering can work around; it is a fundamental consequence of the quantum electrodynamics that governs electromagnetic radiation. Making photons influence each other requires introducing something that can mediate the interaction—a material, a structure, a mechanism that can translate the presence of one light beam into a change that affects another.

The photonic computing efforts of the 1980s and 1990s addressed this problem through various approaches, each with limitations. Some used nonlinear optical materials—substances in which the optical properties change depending on the intensity of light passing through them. In these materials, a strong control beam could alter the refractive index enough to affect the propagation of a weaker signal beam. The approach worked in principle, but the effects were weak, requiring high-intensity light that reintroduced energy consumption problems and demanding materials that were expensive and difficult to integrate into practical

systems. Others used hybrid approaches in which light carried data but electronic components performed logic—converting optical signals to electrical ones at each computational node, performing the logical operation with transistors, then converting back to light for transmission to the next stage. These systems captured some benefits of optical communication while retaining electronic computation, but the constant conversion consumed much of the efficiency advantage that light offered. Each optical-to-electrical and electrical-to-optical conversion required energy and introduced delay. The approach reduced to using light as better wire rather than as a medium for computation itself.

Still others achieved optical bistability in exotic semiconductor materials—gallium arsenide, indium phosphide, and more complex III-V compounds that could change their optical properties in response to optical inputs. The Bell Labs SEED devices mentioned in the previous chapter were examples of this approach. They worked as laboratory demonstrations, but the materials were expensive, difficult to manufacture, and fundamentally incompatible with the silicon fabrication infrastructure that had driven electronic computing's cost reductions over decades. Building a commercial industry on these materials would have required creating an entirely new manufacturing ecosystem—billions of dollars in investment for facilities that would produce components at far higher cost than their silicon equivalents.

The approaches failed not because the underlying physics was wrong, but because the specific implementations could not deliver practical systems. The physics of light still permits computation. The question was always whether an implementation could be found that worked at scale, at reasonable cost, at room temperature, using manufacturable materials.

The Poovey Stack represents a different approach to the interaction problem—one that does not attempt to make photons interact with each other directly, but instead uses carefully designed structures that allow light to control light through intermediate mechanisms. The distinction is subtle but consequential. Rather than forcing photon-photon interactions that physics resists, the architecture creates conditions in which photons can influence the behavior of optical pathways that subsequent photons traverse. The effect is logical switching without requiring the direct interactions that nature prohibits.

An analogy may help. Consider a railroad switching yard, where tracks can

be reconfigured to direct trains to different destinations. The trains themselves do not interact directly—one locomotive does not push another onto a different track. But a control mechanism can observe an incoming train and set the switches appropriately, so that the next train takes a different path.

The "interaction" between trains is mediated by the switching mechanism, not by direct contact between the trains themselves.

The Poovey Stack operates by a conceptually similar principle, though the physics is obviously different. The "switching mechanism" consists of carefully engineered material structures whose optical properties can be altered by light under specific conditions. When a control signal arrives, it triggers a change in the structure that persists long enough for a subsequent data signal to encounter the altered path. The data signal is then routed or modified according to the state that the control signal established. The photons in the control signal and the photons in the data signal never interact directly; they are mediated by the structure they traverse.

The details of how this mechanism operates—the specific materials, the layer structure, the physical phenomena exploited—are the subject of the next chapter. Here, the essential point is conceptual: the Poovey Stack sidesteps the photon interaction problem rather than solving it head-on. It achieves logical switching by engineering the environment through which photons travel, rather than by engineering interactions between photons themselves. This approach opens design space that was not available to researchers attempting direct photon-photon interactions.

• • •

Del Wolverton is not YET a name that appears in the canonical histories of semiconductor technology. He did not work at Bell Labs during its golden era or lead research programs at IBM or Intel. But his career in fiber-based systems began earlier than most—as founder of Space Microwave, he contributed to the early development of fiber optic switching technologies that would eventually transform telecommunications. His path ran through military radar electronics and guidance systems, through decades of work on high-frequency signal processing, to the specific problem of making light perform logic.

The Poovey Stack emerged from that persistence, from years of attention to problems that larger organizations had abandoned.

The collaboration that produced True Photonic combined complementary capabilities. Dr. Gary Poovey brought the theoretical foundation and breakthrough materials science—understanding the physics of how specific material compositions could enable optical switching at femtosecond timescales. Wolverton contributed device-level insight from decades of work in fiber optic switching and high-frequency systems—understanding how light could be made to do logic through structures that could actually be built. Bailey brought systems-level thinking: how such devices would integrate with existing infrastructure, what manufacturing approaches would scale, what markets would justify the development investment. Robert Switzer's contribution of graphene-sapphire substrate technology provided the final essential element—a foundation on which the Poovey Stack's switching structures could be reliably fabricated with the thermal and optical properties the architecture requires.

The partnership reflected a lesson from the history described in the previous chapter. Many photonic computing efforts failed not because the physics was wrong but because the translation from physics to product was never adequately addressed. Brilliant device concepts remained stuck in laboratories because no one had thought through how they would be manufactured, what systems they would plug into, or who would pay for them. The Poovey Stack was developed with these questions in mind from the beginning—not as afterthoughts but as constraints that shaped the technical choices.

The name "Poovey Stack" refers to the layered architecture of the switching structure, in which different physical phenomena are orchestrated across vertical layers to achieve the logical operations that computation requires. The term "stack" carries deliberate resonance with the technology industry's usage—the "software stack" that layers applications atop operating systems atop hardware, the "network stack" that layers protocols atop transport mechanisms. The Poovey Stack is a physical stack, a three-dimensional arrangement of materials and structures that enables light to perform logic. The technical details of how the stack achieves switching will be examined in the following chapter. Here, the focus is on what the architecture accomplishes rather than how—on the specifications that would matter to someone evaluating whether this approach could address the crisis described in Part I.

The first specification concerns speed. Electronic transistors in modern processors switch states in approximately one nanosecond—one billionth of a second.

This has been sufficient for most applications; a nanosecond is far shorter than any duration humans can perceive, and billions of nanosecond operations per second have powered the digital revolution. **But the limits are real.** The clock speeds of processors have plateaued as heat generation prevents further acceleration. The data movement between processors and memory, which happens at nanosecond timescales, has become a bottleneck limiting system performance regardless of how fast the processors themselves can operate.

The Poovey Stack has been documented at **switch speeds of 150 to 200 femtoseconds.** A femtosecond is one quadrillionth of a second—one thousandth of a trillionth of a second, a million times shorter than a nanosecond. In a femtosecond, light travels approximately 300 nanometers, about half the wavelength of visible red light. The switching time is so brief that measuring it requires specialized metrology at the edge of what current instruments can resolve.

The improvement factor is between 5,000 and 10,000 times—not incremental progress but a categorical change in what operations are possible.

To place this in context: if electronic switching were a car traveling at highway speed, photonic switching at these timescales would be a spacecraft traveling at a significant fraction of the speed of light. The comparison is not merely rhetorical; it reflects the actual difference in the physics. We are not discussing a somewhat faster version of the same phenomenon but a fundamentally different regime of operation.

Speed alone would not solve the crisis described in earlier chapters. A processor that operates faster but consumes proportionally more power would simply generate more heat. The value of femtosecond switching lies not merely in completing operations more quickly but in enabling computation to occur through a different physical mechanism—one that does not generate heat in the same way electronic switching does. The speed is a consequence of the physics, and the physics is what enables the efficiency.

The second specification concerns power consumption. Electronic processors dissipate energy primarily through resistance—the collisions between moving electrons and the atomic lattice of the materials they traverse. This is not a flaw that better engineering can eliminate; it is a consequence of how charge moves through matter. Every transistor switch charges and discharges capacitances,

and every charge movement through resistive materials converts electrical energy to heat. The Landauer limit mentioned in Chapter 6 sets a theoretical floor, but practical processors operate roughly a million times above that floor, and the gap has proven extraordinarily difficult to close.

Photons moving through transparent materials encounter no such resistance. Light in a fiber-optic cable can travel hundreds of kilometers with minimal attenuation. The photons are not colliding with atoms, not losing momentum with each interaction, not converting their energy to thermal motion of the medium. This is why telecommunications adopted optical transmission decades ago: the energy required to move information as light is a small fraction of the energy required to move it as electrical current.

The Poovey Stack achieves power reductions of 90 to 95 percent compared to equivalent electronic systems. The precise figure depends on the specific comparison—which electronic architecture, which workload, which operating conditions—but the order of magnitude is consistent across analyses. A computation that requires 100 watts in electronic implementation requires 5 to 10 watts in the Poovey Stack. The reduction is not marginal; it is transformative. The implications for the infrastructure crisis are direct and substantial. A data center consuming 100 megawatts to power electronic processors could deliver equivalent computational output at 10 megawatts with photonic systems. The cooling infrastructure sized for 100 megawatts of heat dissipation could serve facilities performing ten times the computation. The water that currently evaporates through cooling towers—the 300,000 gallons per day that a medium-sized data center consumes, the billion gallons per year that Google's Council Bluffs facility requires—could be reduced proportionally. The transmission lines and transformers straining under data center load, the utilities projecting hundreds of billions in infrastructure investment, the communities experiencing rising electricity rates—all would see relief proportional to the efficiency gain.

The 90 percent figure is not a marketing claim selected for impressiveness. It emerges from the fundamental physics: photons in transparent media simply do not dissipate energy through resistance the way electrons in silicon do. The engineering question is whether a practical system can capture this physical advantage—whether the control structures, interconnections, and auxiliary electronics required to make a photonic processor work do not reintroduce the energy consumption that photonic switching avoids.

The Poovey Stack architecture addresses this question through designs that minimize the electronic components required for control and read-out, maintaining the efficiency advantage through the complete system rather than losing it at the interfaces.

The third specification may be the most consequential for practical deployment: Manufacturing.

Chapter 4 described the concentration of advanced semiconductor fabrication in a handful of companies and locations, the $400 million lithography machines that only ASML can build, the $20 billion fabrication facilities that only three companies in the world can afford. The leading edge of electronic chip manufacturing is an extraordinary achievement of human engineering—and an extraordinary vulnerability, a chokepoint that could be disrupted by a single geopolitical crisis, a single earthquake, a single export restriction.

The Poovey Stack is not envisioned to require leading-edge fabrication.

It can be manufactured using what the industry calls "mature" processes—the 90-nanometer to 180-nanometer technology nodes that were cutting-edge fifteen to twenty years ago but have since been superseded by the relentless progression of Moore's Law. These mature processes are widely available, operated by dozens of fabrication facilities around the world, and do not require the extreme ultraviolet lithography machines that represent the critical choke-point in the semiconductor supply chain.

The insight that enables this is what True Photonic calls "**Relaxed Lithography.**"

It reflects a fundamental difference in how photonic and electronic devices operate. Electronic transistors must be small because their speed depends on how quickly electrons can traverse the channel—and the smaller the channel, the shorter the transit time. The race to smaller nodes in electronics is driven by this relationship: speed requires miniaturization, which requires ever more exotic fabrication technology, which requires ever more concentrated supply chains and ever higher capital investment.

Photonic devices operate by different physics entirely. Light does not speed up when the device shrinks. The relevant dimension for photonic operation is the

wavelength of light itself—hundreds of nanometers for visible light, roughly 1,550 nanometers for the infrared light commonly used in telecommunications. Structures much smaller than these wavelengths do not improve photonic performance; they simply make fabrication more difficult without commensurate benefit. The Poovey Stack is designed to operate at scales that match the physics rather than fight it—scales that happen to align with mature, widely available fabrication technology.

The strategic implications are profound. A photonic computing technology that requires access to TSMC's most advanced facilities in Taiwan would be subject to all the vulnerabilities described in Chapter 4—the geopolitical tensions, the single points of failure, the capacity constraints, the years-long queues for manufacturing slots. A technology that can be fabricated at dozens of facilities worldwide, using equipment that is commercially available without restriction, using processes that have been refined over decades of industrial experience, represents a different kind of strategic asset entirely. Such a technology can be manufactured domestically without requiring the massive investments that the CHIPS Act is attempting to mobilize for leading-edge fabrication. It can be sourced from multiple suppliers, eliminating single points of failure. It does not depend on the continued stability of any single region or the cooperation of any single company. The strategic advantage of photonic computing is not merely technical but geopolitical: it offers a path to computational sovereignty that does not require winning the race to the most advanced fabrication technology.

The counterargument would be that mature processes cannot achieve the density or performance of leading-edge nodes—that eschewing advanced fabrication means accepting inferior capabilities. This argument applies to electronic devices, where density and performance are intimately linked through the physics of electron transport. It does not apply to photonic devices operating on different physical principles.

The Poovey Stack's performance specifications—the femtosecond switching, the 90 percent power reduction—are achieved at mature process nodes. The relaxation of lithography requirements does not compromise performance; it is a consequence of designing for the physics of light rather than the physics of electrons.

• • •

Before proceeding to validation, a clarification is necessary. The previous chapter distinguished between photonic transport and photonic logic—between companies that use light to move data between electronic processors and companies that use light to perform computation itself. This distinction is essential for understanding where the Poovey Stack fits in the current landscape **and why its claims, once valid, are significant.**

Companies like **Ayar Labs** (**valuation $1B+**) and **Lightmatter (valuation $4.4B)** have achieved genuine technical accomplishments in photonic interconnects.

They have developed technology that moves data between chips using light rather than electrical signals, reducing the bandwidth bottlenecks that limit system performance. Their products are being adopted by major technology companies and represent real progress in data center efficiency.

But their technology **does not perform** *computation* **in light**; it performs *transport* in light. The processors receiving and sending data through these optical interconnects remain electronic. The *heat* they generate, the *power* they consume, the *water* required to cool them—none of these are reduced by faster optical connections between chips.

The Poovey Stack makes a different claim: **that** *computation* **itself—**the logical operations that constitute processing**—occurs in photons.**

This is the claim that previous photonic computing efforts failed to deliver. **It is the claim that, if validated, would address the fundamental energy crisis rather than merely improving data movement.**

The distinction is not between good and better versions of the same thing; it is between addressing a symptom and addressing a cause. The transport companies and True Photonic are not competitors in the usual sense. Their technologies address different parts of the system and could potentially be complementary. But for the reader trying to understand whether photonic computing can solve the crisis described in Part I, the distinction matters enormously. Faster data movement does not reduce the power consumption of computation.

Faster computation—computation that generates less heat by operating through different physics**—addresses the problem at its** *source***.**

The gap between *transport* and *logic* is not unbridgeable.

The NIBLE BUS—documented in U.S. Patent Publication 20240313860, with Del Wolverton as inventor—represents an architecture that crosses this divide: an all-optical system for managing data transfer through photonic switches, steering components, and buffer memories. Unlike transport systems that merely move data between electronic processors, the NIBLE BUS brokers signals using only light signals, performing the data management functions that computing requires without converting to electronic form.

A companion publication, U.S. Patent Publication 20240411203, documents a purely optical NAND logic gate—the universal building block from which all Boolean operations can be constructed. These are not theoretical proposals but documented implementations with sufficient specificity to receive USPTO publication, the standard that indicates an invention has been reduced to practice. **The publications establish that all-optical data handling and all-optical logic have progressed from laboratory curiosity to formal intellectual property**—searchable in public databases, available for examination by anyone who wishes to verify the claims.

Claims require validation. The history of technology is replete with demonstrations that worked in idealized conditions but failed in practical deployment, with specifications measured in laboratories that could not be replicated in production environments. A responsible assessment of the Poovey Stack must ask: What has been demonstrated? What remains theoretical? Where has the technology been validated, and where does validation remain incomplete?

The switching speed—the 150 to 200 femtosecond figure that anchors the performance claims—*has been validated independently.* Technion University in Israel, one of the world's leading materials science institutions, tested optical switches built from Gary Poovey's material compositions using pump-probe spectroscopy, the standard technique for measuring ultrafast optical phenomena. The Technion report documented switching times consistent with the claimed specifications. This is not a measurement performed by True Photonic and reported by True Photonic; it is an independent validation by a respected research institution with no financial stake in the outcome.

The significance of independent validation cannot be overstated.

The history of technology includes too many cases where claimed specifications were measured by interested parties using favorable conditions, only to prove unreproducible when tested independently. The Technion validation establishes that the femtosecond switching speed is not an artifact of measurement methodology or optimistic interpretation; it is a physical reality that external researchers have confirmed.

I remember the day Del called with those results.

I was still processing what we had—still running the numbers in my head, still not quite believing that the physics Gary had described was actually real. The Technion report changed that. It was no longer our claim. It was documented. Validated. Independently confirmed by researchers who had no reason to tell us what we wanted to hear. That was the moment the company became real. Not when we incorporated, not when we filed our first patent, not when we closed our first investment. The day the independent validation came back—that was the day I knew we were not chasing a mirage. **The physics worked. Everything else was execution.**

As of this writing, True Photonic has fabricated working optical switches on silicon wafers using commercial foundry processes. The switches exist as physical devices, not merely as simulations or designs. They have been produced through the same manufacturing infrastructure that produces electronic chips—demonstrating that the Relaxed Lithography approach is not merely theoretical but practically achievable. The devices switch at the claimed timescales and consume power at the claimed levels. The significance of working wafer-level devices should not be understated. Many promising technologies never progress beyond theoretical analysis or small-scale laboratory demonstrations using custom-built apparatus. The transition to wafer-level fabrication using commercial foundry processes is a critical milestone because it demonstrates compatibility with the manufacturing infrastructure that would be required for commercial scale. A technology that works only with hand-built laboratory equipment is interesting; a technology that works on commercial wafers is potentially deployable.

What has not yet been demonstrated is a complete photonic processor— an integrated system that performs useful computation rather than individual switching operations. The gap between working switches and working

processors is significant. Integration introduces challenges that do not appear at the device level: thermal management across large arrays of switches operating simultaneously, signal routing between components without introducing losses or crosstalk, interfaces to electronic systems for input and output that do not negate the efficiency advantages, yield rates in manufacturing that make production economically viable at scale. These challenges are not unique to photonic computing; they are the standard obstacles that any new semiconductor technology must overcome. The history of electronics includes many device technologies that demonstrated excellent individual performance but failed during integration. Magnetic bubble memory, superconducting Josephson junctions, and resonant tunneling diodes all showed promising device characteristics but never achieved the system-level integration necessary for commercial success.

The Poovey Stack must navigate the same transition, and history provides no guarantees about the outcome.

The validation trajectory follows a recognizable pattern from semiconductor development: individual devices first, then small circuits, then larger integrated systems, then commercial products. True Photonic is progressing along this trajectory, with working devices established and integration underway. The timeline from current status to commercial deployment involves years of engineering, not decades. But it also involves the resolution of challenges that have defeated other photonic computing efforts in the past. Success is plausible; it is not assured.

The honest position is this: the Poovey Stack represents a technically credible approach to photonic logic that has demonstrated individual device functionality but has not yet been proven at system scale. The physics is sound. The manufacturing approach is validated. The performance specifications are supported by device-level measurements. What remains is the integration and scaling work that separates laboratory demonstration from commercial product. The company's claims are testable, and the next few years will determine whether they can be substantiated across the full development arc.

• • •

The skeptic who has followed photonic computing's history of disappointment deserves an answer to the obvious question:

Why should this time be different?

The previous chapter identified five factors that distinguish the current moment from earlier eras of photonic computing enthusiasm: the rise of AI workloads suited to optical processing, the maturation of silicon photonics manufacturing, the end of Moore's Law scaling, the emergence of efficiency as a primary metric, and the regulatory and social pressure for sustainability. The Poovey Stack is not merely an alternative technology arising at a fortunate moment; it is a technology specifically designed to exploit these changed conditions.

The alignment with AI workloads is direct and fundamental. The matrix multiplications that dominate neural network computation are precisely the operations that photonic systems perform most efficiently. The Poovey Stack architecture is optimized for these operations—not because it was designed to ride a trend, but because the inventors recognized early that AI's computational demands would eventually outstrip what electronic systems could deliver sustainably. The technology addresses the workloads that are driving the infrastructure crisis described in Part I, not the general-purpose computing tasks that electronics handle adequately. The mature-process approach directly addresses the manufacturing barriers that defeated earlier efforts.

Previous photonic computing technologies required exotic materials—gallium arsenide, indium phosphide, complex compound semiconductors—that could not be produced at the scale and cost that commercial deployment required. Building an industry on those materials would have required creating an entirely new fabrication infrastructure from scratch, competing for investment against the established silicon ecosystem that was still delivering Moore's Law improvements. The Poovey Stack uses silicon and standard semiconductor materials, fabricated at process nodes where manufacturing has been refined over decades. The technology does not require building a new fabrication infrastructure; it leverages the existing one.

The timing with respect to electronic competition is favorable in ways that were not true in the 1990s. Then, electronic processors were improving exponentially year over year; any photonic alternative had to hit a moving target that kept receding faster than the technology could advance. Now, as Chapter 2 described, that improvement has slowed dramatically. The process shrinks that once delivered 40 percent generational improvements now deliver single-digit

percentages. A photonic technology that is competitive today may remain competitive for years, rather than being leapfrogged by the next electronic generation. The target has stabilized. And the pressure for sustainable alternatives has never been greater. The communities blocking data center construction, the utilities struggling to meet demand, the water authorities watching aquifers decline, the grid operators projecting infrastructure investments in the hundreds of billions—these are not abstract concerns but immediate constraints on the industry's expansion. A technology that can deliver equivalent computation at a fraction of the power and water consumption is not merely an improvement; it may be a prerequisite for continued growth. The market is not just ready for photonic computing; it may be demanding it.

The claims made in this chapter have been deliberately general, describing what the Poovey Stack accomplishes without explaining how. This is appropriate for an introductory treatment but insufficient for serious evaluation. The physics of how photons achieve logical switching, the specific structures that constitute the "stack," the mechanisms that enable control and read-out without reintroducing electronic inefficiency—these details matter, and they will be examined in the chapters that follow.

Part III of this book is titled "**The Architecture**." Chapter 9 will examine the Poovey Stack in technical detail, explaining the layer structure and the physical phenomena that enable switching.

Chapter 10 will address how computation actually occurs—how logical operations are performed and combined, how data flows through the system, how the architecture maps to the AI workloads that are driving demand. It will also address how the Poovey Stack relates to the von Neumann architecture that has defined computing since the 1940s—whether this is a departure from that model or an implementation of it through different physics. Chapter 11 will examine manufacturing: the specific processes used, the yields achieved, the path from current capabilities to commercial scale.

The examination will be rigorous. The claims will be tested against physical principles, against manufacturing realities, against the specific requirements of the applications they purport to serve. Photonic computing has disappointed before, and healthy skepticism serves the reader better than uncritical enthusiasm.

The technology either works or it does not.

The efficiency either materializes in real systems or it does not.

The manufacturing either scales or it does not.

But the examination will also be fair.

The crisis described in Part I is real. The demand for computation that does not consume rivers and burn coal is genuine. If the Poovey Stack represents a path forward—a way to keep computing without the costs that have made computing increasingly controversial—that possibility deserves serious engagement rather than reflexive dismissal.

The all-optical computer promised by Scientific American in 1990 did not arrive. What may be arriving instead is something both more modest and more practical: a photonic architecture that performs specific operations—the matrix multiplications of AI—with extraordinary efficiency, while interfacing with electronic systems for other functions. Not the replacement of electronics but a strategic complement to it. Not a revolution in all computation but a transformation in the computation that matters most for the crisis at hand.

The breakthrough, if it is a breakthrough, lies not in achieving what physics always permitted but in making what physics permits practically achievable. That is the claim. The architecture is the evidence. And the evidence deserves detailed examination.

PART III
THE ARCHITECTURE

NOT
was trivial in electronics.
In photonics, it defeated
a generation of researchers.
Until now.

CHAPTER 9

The Poovey Stack

The architecture is the evidence.

The previous chapter introduced the Poovey Stack's specifications—femtosecond switching speeds, ninety percent power reductions, compatibility with mature fabrication processes. It described what the technology accomplishes without explaining how. For the skeptical reader, that gap is significant. Claims are cheap. Specifications announced in press releases have a long history of evaporating when engineers attempt to build actual products. The all-optical computer promised in 1990 foundered not on its specifications but on the physical mechanisms that were supposed to deliver them. This chapter examines the mechanism. It describes the physical structure of the Poovey Stack—the specific layers, their compositions, their functions—and explains how those structures enable photons to perform logic. The explanation will be technical but not gratuitously so. The goal is not to replicate a physics journal but to provide enough detail that a serious reader can evaluate whether the approach is plausible, where the uncertainties lie, and what distinguishes this architecture from the approaches that failed in previous decades.

The explanation begins at the foundation—literally. Every semiconductor device rests on a substrate, the material that supports the active structures and provides the physical and thermal environment in which they operate. The choice of substrate is not incidental; it shapes what can be built above it. The Poovey Stack's substrate is unconventional, and its unconventionality is essential to the architecture's function.

• • •

A conventional electronic chip rests on silicon—a thin wafer of crystalline silicon dioxide that has been purified to extraordinary levels and processed through dozens of steps to create the transistors, interconnects, and other structures that constitute the circuit. Silicon has dominated semiconductor manufacturing for six decades because it combines useful electrical properties with manufacturability at scale. It forms a stable oxide that serves as an excellent insulator. It can be doped with impurities to create the p-type and n-type regions that transistors require. Its crystal structure is well understood, and the processes for growing,

cutting, and polishing silicon wafers have been refined over generations of industrial experience. For photonic devices, silicon presents difficulties. Silicon does not emit light efficiently—a fundamental limitation arising from its indirect band gap that cannot be engineered away without altering the material's basic nature. Silicon absorbs light at the wavelengths most useful for photonics, particularly in the visible and near-infrared range, introducing losses that accumulate as signals propagate through the material. The telecommunications industry chose 1,550 nanometer wavelength partly because silicon's absorption is lower there, but even at this wavelength the material remains far from ideal. And silicon's thermal conductivity, while adequate for electronic devices with their distributed heat generation, becomes a constraint when the goal is to build dense arrays of optical switches that must maintain precise alignment and timing.

Previous photonic computing efforts struggled with silicon's limitations in different ways. Some accepted silicon as the substrate and built switching structures from other materials deposited on top—gallium arsenide, indium phosphide, lithium niobate—creating hybrid devices that inherited silicon's thermal and mechanical properties while attempting to overcome its optical limitations. Others abandoned silicon entirely, fabricating devices on substrates of the III-V semiconductors themselves. Both approaches encountered the manufacturing problem described in Chapter 7: the processes for working with these materials were expensive, low-yield, and incompatible with the economies of scale that had made electronic computing affordable.

The Poovey Stack uses a different foundation: a composite substrate of graphene, hexagonal boron nitride, and sapphire. Each material contributes specific properties that the architecture requires, and the combination creates a platform fundamentally suited to optical rather than electronic operation. Sapphire—crystalline aluminum oxide—forms the structural base. It is mechanically robust, optically transparent across a wide range of wavelengths from ultraviolet through mid-infrared, and thermally stable at temperatures far exceeding any that the devices will encounter in operation. The material's thermal expansion coefficient is well characterized and predictable, allowing engineers to design structures that maintain dimensional stability across operating temperature ranges.

Sapphire substrates are already used in certain specialized semiconductor

applications, including LEDs and high-power electronics, and can be produced at the scales and quality levels that commercial manufacturing requires. The material provides a rigid, stable foundation that does not expand, contract, or deform in ways that would disrupt the precisely engineered structures built upon it.

Hexagonal boron nitride—sometimes called "white graphene" for its structural similarity to graphene—serves as an intermediate layer and as an encapsulation material for critical structures. Boron nitride is an exceptional electrical insulator, with a band gap exceeding five electron volts, while being thermally conductive—a rare combination that allows it to channel heat away from active regions without providing pathways for electrical interference. When deposited properly, it forms atomically flat surfaces with roughness measured in fractions of a nanometer, creating the smooth interfaces that the single-atom-thick graphene layers require for optimal performance. And it is transparent to the infrared wavelengths at which the Poovey Stack operates, introducing minimal optical loss even in multilayer configurations.

Graphene itself is the most exotic component and the most critical. A single layer of carbon atoms arranged in a hexagonal lattice, graphene possesses properties that seem almost designed for photonic switching. Its electrons behave as massless Dirac fermions, a consequence of the material's unique band structure, responding to electromagnetic fields with extraordinary speed and minimal energy dissipation. Unlike bulk semiconductors where electrons must navigate complex band structures and scatter frequently against lattice defects, electrons in graphene move through a two-dimensional landscape where scattering is minimized and transit times are measured in femtoseconds. Graphene can absorb and release photons across a remarkably broad spectrum of wavelengths, from terahertz through visible and into ultraviolet, making it useful for the infrared light that telecommunications and data systems commonly employ.

A single graphene layer absorbs approximately 2.3 percent of incident light regardless of wavelength—a seemingly modest figure that becomes significant when multiplied across multiple layers and interaction zones. Its optical properties can be tuned by applying electric fields—a characteristic called electro-optic modulation—allowing external signals to control how light interacts with the material without requiring physical reconfiguration. And when properly

integrated with surrounding structures, graphene can serve as the active switching element that enables photons to influence the paths of other photons. The thermal properties of this substrate stack deserve particular attention. The crisis described in Part I of this book is fundamentally a thermal crisis: electronic processors generate heat that must be removed at enormous expense in energy, water, and infrastructure. Any technology claiming to address that crisis must demonstrate not only that it generates less heat but that what heat it does generate can be managed effectively.

The graphene-boron nitride-sapphire substrate achieves thermal conductivity exceeding 2,000 watts per meter-kelvin in the plane of the graphene layers. For comparison, copper—the standard material for heat spreaders in electronic systems—conducts heat at approximately 400 watts per meter-kelvin. Aluminum, widely used in heat sinks, manages about 200 watts per meter-kelvin. Silicon itself conducts at roughly 150 watts per meter-kelvin. Diamond, often cited as the ultimate thermal conductor and used in specialized high-power applications, reaches roughly 2,000 watts per meter-kelvin.

The substrate matches diamond's thermal performance using materials that can be deposited through industrial processes rather than requiring the extreme pressures and temperatures of diamond synthesis. This thermal conductivity means that whatever heat the devices do generate spreads rapidly across the substrate rather than accumulating in hot spots. The Poovey Stack generates far less heat than electronic equivalents—a consequence of the switching physics described below—and what heat it does generate dissipates efficiently through a substrate designed for thermal management. The combination addresses both sides of the thermal equation: producing less heat and removing it more effectively.

• • •

The substrate would be useless without the ability to manufacture it reliably at scale. This is where many advanced materials technologies have failed: laboratory demonstrations of remarkable properties that cannot be reproduced in production environments, or can be reproduced only at costs that make commercial deployment impossible. Graphene in particular has suffered from a gap between its theoretical potential and its practical availability. First isolated in 2004 by Andre Geim and Konstantin Novoselov at the University of Manchester—work that earned them the 2010 Nobel Prize in Physics—graphene spent the following

decade as a material of extraordinary promise and limited commercial impact.

The original method for producing graphene involved mechanical exfoliation—essentially peeling layers from graphite using adhesive tape. This approach yielded high-quality material suitable for research but was obviously unsuitable for manufacturing. Chemical vapor deposition on copper foils emerged as a scalable alternative, but it required transferring the graphene from the copper growth substrate to the target substrate, a process that introduced defects, contamination, and yield losses. The graphene community spent years developing transfer techniques, each iteration reducing defects incrementally but never eliminating them entirely.

Robert Switzer's contribution to True Photonic addresses this challenge directly. Switzer, through Astera Energy, developed a process for growing graphene directly on sapphire substrates using microwave plasma-enhanced chemical vapor deposition—a technique abbreviated MW-PECVD.

The process eliminates the transfer step entirely: graphene grows where it will be used, avoiding the defects and contamination that transfer introduces. And it does so quickly—depositing high-quality graphene in under three minutes, a timeframe compatible with high-volume manufacturing. Previous approaches to graphene synthesis required hours of processing time, exotic precursor gases, or the problematic transfer steps. MW-PECVD eliminates these bottlenecks.

The significance of a three-minute growth cycle extends beyond simple throughput calculations. Short process times mean that manufacturing equipment can produce more substrates per day, reducing capital costs per unit. They mean that process variations—the inevitable fluctuations in temperature, gas flow, plasma power, and chamber conditions that occur in any industrial process—have less time to accumulate into defects. A small variation that causes no problem over three minutes might produce significant non-uniformity over three hours. And short cycles mean that the feedback loop between process engineers and production results is fast enough to enable rapid optimization.

A process that takes hours to complete can be run only a few times per day; a process that takes minutes can be iterated dozens of times, accelerating the learning curve that moves technology from laboratory to factory.

The graphene produced through this process exhibits the properties that the Poovey Stack requires: uniform coverage across the substrate area, minimal defects in the atomic lattice, and consistent optical and electronic characteristics from batch to batch. Raman spectroscopy—the standard technique for characterizing graphene quality—confirms single-layer coverage with the characteristic G and 2D peaks that indicate proper crystal structure.

The manufacturing process has been validated through production of actual substrates used in device fabrication—not merely through theoretical analysis or small-scale demonstrations.

• • •

Above the substrate, the Poovey Stack builds a vertical structure of functional layers—hence the name "stack." The specific number and composition of these layers varies depending on the device being constructed, but the general architecture supports configurations ranging from five layers for simple switching elements to more than one hundred fifty layers for complex integrated circuits. This scalability is essential: a technology that works only for isolated devices would be a laboratory curiosity, while a technology that can scale to large integrated systems has commercial potential. The layers serve distinct functions within the architecture. Some contain the active switching structures—regions where light can be controlled by modifying the optical properties of the material.

These active layers incorporate graphene or related two-dimensional materials in configurations designed to maximize the switching effect while minimizing parasitic losses. Others serve as waveguides, channels that confine and direct light between active regions. The waveguide layers must be engineered to support the optical modes that carry information, with dimensions and refractive indices calculated to ensure that light propagates efficiently without leaking into surrounding structures or coupling inadvertently between adjacent channels.

Still other layers provide isolation, preventing optical or electrical crosstalk between adjacent structures. In a dense array of switching nodes, signals must remain confined to their intended paths; stray light that couples from one channel to another would corrupt data and introduce errors. The isolation layers use materials with carefully chosen refractive indices to create barriers that light cannot penetrate, while remaining compatible with the manufacturing processes used for the rest of the stack. The arrangement is three-dimensional in a way that conventional planar semiconductor designs are not. Electronic chips are

essentially two-dimensional structures; the transistors and interconnects lie in a thin layer near the surface of the silicon wafer, with the substrate itself serving merely as mechanical support and heat dissipation. The Poovey Stack exploits the vertical dimension to achieve density and functionality that would be impossible in a single plane. Light can propagate upward, downward, and laterally through the stack, with switching nodes positioned at strategic intersections to control the flow.

The individual layers are thin—measured in nanometers for some critical structures—but the total stack can extend to substantial thickness. This is a departure from the trend in electronic semiconductors, where the active regions have become progressively thinner as feature sizes have shrunk. In photonic devices operating at infrared wavelengths, the relevant dimension is hundreds of nanometers to match the wavelength of the light being manipulated, and useful structures can span micrometers without penalty. The Poovey Stack takes advantage of this design freedom, building upward in ways that electronic designs cannot. The vertical architecture also enables a manufacturing approach that differs from conventional semiconductor processes. Rather than building the entire circuit in a single wafer and then dicing it into individual chips, the Poovey Stack can be fabricated in layers that are deposited and patterned sequentially. Each layer can be inspected and characterized before the next layer is added, allowing defects to be identified early rather than discovered only after the entire fabrication sequence is complete. This layer-by-layer approach reduces waste and improves yields—critical factors for commercial viability in an industry where profitability depends on the percentage of manufactured devices that actually work.

• • •

The switching mechanism is the heart of the architecture, the element that distinguishes the Poovey Stack from optical interconnect technologies that merely transport light between electronic processors. Recall the challenge outlined in Chapter 8: photons do not naturally interact with each other. Two beams of light can cross in space without either beam affecting the other—a property exploited every time you see multiple light sources illuminating the same scene without interference. A technology that performs computation in light must find a way to make the presence or absence of one light beam affect the behavior of another—without the direct photon-photon interactions that physics does not permit at practical power levels. The Poovey Stack achieves this through engineered structures that the inventors call Delbert Tracks.

Each track consists of a region where the optical properties of the material can be altered by a control signal, changing how a subsequent data signal propagates through the structure. The control signal is itself optical—a pulse of light at a specific wavelength and intensity—but it does not interact directly with the data signal. Instead, it interacts with the switching structure, changing the structure's state in ways that affect subsequent optical signals passing through.

The mechanism exploits phenomena that occur when light interacts with the piezoelectric and two-dimensional materials in the switching structure. When photons of appropriate energy strike the active region, they induce changes in the material's optical properties—its refractive index, its absorption coefficient, its transmission characteristics—for a brief period. During this window of altered properties, a data signal passing through encounters different optical conditions than it would if the control signal had not arrived.

The effect is binary: light turns off light, creating the logical switching that computation requires. The physics is more nuanced than a simple on-off switch. The switching structure uses an optical channel whose dimensions can be altered by light-induced strain in piezoelectric materials. When the control signal arrives, it triggers a nanometer-scale contraction that changes the channel's cut-off frequency. A data signal at a wavelength near this cutoff—1,600 nanometers in the validated configuration—finds itself suddenly unable to propagate through a channel that was transparent moments before. The effect persists only while the control signal maintains the altered state, then relaxes as the piezoelectric material returns to equilibrium.

The duration of this switching window is measured in femtoseconds. The piezoelectric response and the optical channel dynamics operate on timescales far faster than in conventional semiconductors where electronic switching involves charging and discharging capacitances through resistive pathways—processes that take nanoseconds or longer. **This is the origin of the Poovey Stack's femtosecond switching speed.** The speed is not achieved through clever engineering of conventional materials; it is a consequence of using materials and mechanisms whose fundamental physics operates on these timescales.

The switching energy is correspondingly low.

The optical power required to trigger the switching event is measured in **milliwatts** at the source, reduced by the optical path to approximately 0.1

milliwatts at the switching structure itself. The control pulses that flip switching states consume microjoules or less, compared to the millijoules or joules required for electronic switching operations.

The ratio is roughly a million to one, which translates directly to the ninety percent or greater power reduction that the specifications claim.

A single switching node is not a computer. It is not even a logic gate, in the sense that a transistor is not a logic gate but rather a component from which logic gates can be constructed. The translation from switching elements to computational capability requires additional architecture: ways to combine multiple switches to perform logical operations, ways to route signals between operations, ways to synchronize the timing of inputs and outputs so that data flows correctly through the system.

This translation is the subject of the next chapter, which examines how the Poovey Stack implements the Boolean logic operations that computation requires. Here, the focus remains on the physical mechanism—on establishing that the switching phenomenon is real, that it occurs at the claimed speeds and energies, and that it can be manufactured using processes compatible with commercial production. The mechanism has been validated through independent testing at Technion University in Israel, one of the world's leading research institutions in photonics and materials science. The validation, conducted by Dr. Gary Poovey with equipment setup and calibration provided by Dr. Alex Bekker, employed pump-probe spectroscopy using a cross-correlation machine—the standard methodology for characterizing ultrafast optical phenomena. The testing used a femtosecond erbium fiber laser, a lock-in amplifier for signal extraction, and a high-speed indium gallium arsenide detector capable of resolving the infrared wavelengths at which the switch operates.

The experimental configuration matched the intended operating conditions of the Poovey Stack. A control signal at 1,550 nanometers—the standard wavelength for telecommunications fiber optics—triggered the switching event. A data signal at 1,600 nanometers passed through the switching structure, its transmission monitored as the timing between control and data pulses was varied. When the two pulses arrived at the structure simultaneously, the data signal was switched off. When they arrived at different times, the data signal passed through unimpeded.

The switching behavior was exactly as the theory predicted. The measured switching speed was 150 to 200 femtoseconds—faster than the theoretical prediction of 283 femtoseconds, and dramatically faster than any previously reported optical switching result. The fastest switching speed in the published scientific literature prior to this validation was 600 femtoseconds, achieved through different mechanisms and materials. The Poovey Switch operates three to four times faster than the previous record. For context, conventional optical networks today are switched by electronic transistors operating at approximately 1,000 picoseconds—one nanosecond. The Poovey Switch is five thousand times faster than the electronic switches currently deployed in telecommunications infrastructure.

The validation confirmed several critical aspects of the switching behavior. The signal appeared at the correct position in the cross-correlation measurement, indicating that the switching was occurring through the intended mechanism rather than some artifact of the experimental setup. The signal shape matched theoretical predictions, demonstrating that the switching dynamics followed the expected physics. And the signal magnitude was consistent with the calculated switching contrast, confirming that the effect was strong enough for practical use in logical operations. The testing identified areas for future optimization as well. The signal exhibited noise that could potentially be reduced through narrower wavelength filtering. The equipment constraints limited exploration of even shorter pulse durations that might reveal faster switching components. These are engineering refinements rather than fundamental limitations—the kind of incremental improvements that characterize the progression from laboratory validation to commercial product.

• • •

The characterization process itself deserves mention, because it illustrates both the capability and the current limitations of the technology. Measuring events that occur in 150 femtoseconds requires instrumentation operating at the frontier of what metrology can achieve. The standard oscilloscopes and logic analyzers used to characterize electronic circuits cannot resolve events shorter than a few hundred picoseconds—roughly a thousand times too slow. Even the fastest photodetectors available commercially have response times measured in tens of picoseconds, far too slow to capture the dynamics of femtosecond switching. The cross-correlation technique used at Technion circumvents these limitations through an elegant physical principle. Rather than trying to detect the switching

event directly with a fast detector, the method uses the speed of light itself as the timing reference. Two pulses derived from the same laser source travel different path lengths before reaching the switching structure. By varying one path length with a precision translation stage, researchers can control the relative arrival time of the pulses with femtosecond resolution. A change in path length of 0.03 millimeters corresponds to a time difference of 200 femtoseconds—the stage movement during which the switching signal was observed.

This measurement approach confirms that individual switching nodes operate as designed. What it cannot confirm—what no measurement technique can confirm until the devices are built—is that large arrays of switching nodes will work together as an integrated system. The gap between individual device performance and system-level functionality is where many promising technologies have failed. The stage that True Photonic is currently in is doing the engineering work that addresses this gap, moving from validated switches to validated circuits and ultimately to validated processors.
True Photonic has estimated an 18-month time-line.

• • •

The honest assessment is this: the Poovey Stack represents a physically plausible approach to photonic logic with validated device-level performance and a manufacturing pathway that does not require exotic equipment or processes. The switching mechanism operates on sound physical principles using materials whose properties are well characterized. The substrate technology addresses both thermal management and manufacturability. The architecture scales from individual devices to complex systems through a layer structure that can be fabricated with existing industrial capabilities. And the switching speed has been validated at a world-class research institution using standard ultrafast characterization methods. **What remains unvalidated is integration.**

The challenge of building working circuits from working switches, working systems from working circuits, and working products from working systems has defeated photonic computing efforts before. The physical principles are not in question; the engineering is. The manufacturing yields that make commercial production viable, the interconnection schemes that route signals without prohibitive losses, the interfaces to electronic systems that must provide inputs and receive outputs—these are the challenges that separate laboratory demonstration from commercial deployment.

The approach to integration that True Photonic has adopted reflects lessons from the semiconductor industry's history. Rather than attempting to build a complete photonic processor in a single leap—an approach that failed spectacularly for several prominent computing startups in recent years—the development proceeds through stages: first simple circuits of a few switches, then more complex circuits, then subsystems, then systems. Each stage validates not only the technical performance but the manufacturing processes, the test methods, and the design tools that production will require. The complete timeline is measured in years, not months, and the uncertainties are real.

But the physics is real as well.

The switching speeds are not projections or targets; they are measurements taken at Technion University and documented in validation reports. The power consumption is not an estimate; it is a calculation from measured switching energies. The substrate thermal conductivity is not a theoretical prediction; it is a characterization of fabricated materials. The foundation of the Poovey Stack is not aspirational; it is demonstrated. The question is whether what has been demonstrated can be extended, integrated, and scaled to create useful systems. That question can only be answered through engineering—through the painstaking work of building, testing, debugging, and refining that transforms validated devices into validated products.

The work is underway.

The outcome is not guaranteed.

• • •

For the reader who has followed photonic computing's history of disappointment, the details in this chapter may provoke skepticism rather than enthusiasm. The materials are different from previous efforts, but materials have been different before. The speeds are faster, but speed claims are easy to make and difficult to verify independently. The manufacturing approach is more practical, but manufacturing approaches always seem practical until they encounter the realities of production at scale.

This skepticism is warranted. The appropriate response to the Poovey Stack is neither acceptance nor rejection but observation. The technology makes testable claims. The device-level performance has been verified at Technion using equipment and methods available to any properly equipped ultrafast optics

laboratory. The integration stages that follow will succeed or fail in ways that will become evident. The market will not accept photonic computing on the basis of promises; it will accept it on the basis of working products that deliver the claimed advantages in real deployments.

What the Technion validation demonstrates is that light can turn off light at speeds three to four times faster than any previously published result, and five thousand times faster than the electronic switches in current optical networks. This is not a theoretical projection or a simulation result; it is a measurement. The switching structures were fabricated, tested, and characterized using the same methods that the photonics research community has refined over decades. The results can be replicated by anyone with access to appropriate equipment and the technical expertise to use it. (Note: Materials have been changed and improved since the Technion testing.)

What the architecture demonstrates is that photonic logic is physically possible in a form compatible with commercial manufacturing. This is not a small achievement. The previous chapter reviewed decades of efforts that could not make this demonstration—that required exotic materials available only from specialized suppliers, cryogenic temperatures maintained by expensive refrigeration systems, or fabrication processes that existed only in a handful of research facilities worldwide. The Poovey Stack uses materials that can be produced at scale, operates at room temperature, and has been fabricated through standard foundry processes. These characteristics do not guarantee success, but they remove barriers that blocked previous efforts.

The next chapter examines how computation emerges from the architecture—how switching nodes combine to form logic gates, how logic gates combine to perform useful operations, and how those operations map to the AI workloads that are driving the infrastructure crisis. The physics of individual switches is necessary but not sufficient; what matters for the crisis described in Part I is whether those switches can be organized into systems that deliver meaningful computation at meaningful scale.

The architecture provides the foundation. The logic provides the capability. And the capability is what the world desperately needs—computation that does not require rivers of water, forests of cooling towers, and the output of power plants dedicated to answering questions.

Whether the Poovey Stack delivers on that need remains to be proven. That it offers a credible path—validated at femtosecond timescales in the laboratories of one of the world's leading technical universities—**is what this chapter establishes.**

CHAPTER 10

Logic-in-Light

A switch is not a thought. The validated Poovey Switch— operating at 150 to 200 femtoseconds, consuming microjoules of energy, fabricated through standard foundry processes—represents a physical capability, not a computational one. The previous chapter established that light can turn off light at speeds and energies that dwarf electronic alternatives. This chapter addresses the more consequential question: Can that capability be organized into something that actually computes?

The question is not rhetorical.

The history of computing includes numerous switching technologies that worked beautifully in isolation but could not be assembled into useful systems. Magnetic bubble memory switched reliably; it never became a computer. Josephson junctions switched faster than any transistor of their era; they powered no commercial products. The gap between a working switch and a working computer is vast, and many technologies have fallen into it. What follows is an examination of how the Poovey Stack bridges that gap—how individual switching nodes combine to form logic gates, how those gates combine to perform useful operations, and how those operations map to the AI workloads that are driving the infrastructure crisis described in Part I. The examination will be technical, because the details matter. It will also be honest about what has been demonstrated and what remains theoretical, because readers deserve to understand where validated physics ends and engineering projection begins.

•••

Every digital computer ever built—from the room-sized ENIAC of 1945 to the smartphone in your pocket—performs its operations through Boolean logic. Named for George Boole, the nineteenth-century English mathematician who formalized the algebra of logical operations, Boolean logic reduces all computation to combinations of simple binary decisions. A value is either true or false, one or zero, on or off. Complex calculations emerge from the systematic combination of these elementary choices. The fundamental operations of Boolean logic are surprisingly few. NOT inverts a value: if the input is true, the output

is false, and vice versa. AND produces true only when both inputs are true. OR produces true when either input is true. From these three operations, all others can be constructed. XOR—exclusive or—produces true when exactly one input is true. NAND—not and—inverts the result of AND. NOR—not or—inverts the result of OR. XNOR—exclusive nor— produces true when both inputs are the same. Seven operations in total, and from them every calculation a computer performs can be derived. This is not a simplification for pedagogical purposes; it is a mathematical fact with profound practical implications. A technology that can implement the complete Boolean logic family can, in principle, perform any computation. A technology that cannot implement even one of the fundamental operations is fundamentally limited, useful perhaps for specialized applications but incapable of general- purpose computation. The completeness of the logic family is the threshold that separates a switching technology from a computing technology.

The Poovey Stack implements the complete Boolean logic family. This claim requires substantiation, because previous photonic computing efforts often stumbled on exactly this point. Optical AND gates were demonstrated in the 1990s; optical NOT gates proved far more difficult. Without NOT, there is no NAND; without NAND, there is no functional completeness; without functional completeness, there is no general-purpose computer. The mechanisms by which the Poovey Stack achieves each logical operation can be described at the architectural level without disclosing the specific material compositions and geometries that constitute True Photonic's intellectual property.

•••

The NOT gate is the foundation. In electronic circuits, NOT is trivially implemented: a transistor inverts its input signal as a natural consequence of its switching behavior. In photonics, inversion is far from trivial. Light does not naturally negate other light. Two photons passing through the same space do not cancel each other; they simply coexist. Creating optical NOT required engineering a structure where the presence of an input signal causes the absence of an output signal—where light, in effect, turns off light. The Poovey Stack achieves optical NOT through the switching mechanism described in the previous chapter. A continuous probe beam passes through an optical channel—a waveguide engineered to transmit light at a specific wavelength. When no control signal is present, the probe beam propagates through the channel and emerges as the output: input absent, output present. When a control signal arrives at the switching

structure, it induces the piezoelectric contraction that shifts the channel's cutoff frequency, blocking the probe beam: input present, output absent. The truth table is precisely that of a NOT gate, achieved entirely through optical means. The probe beam is not the input being negated; it is a carrier signal that the input modulates. This distinction matters for understanding how the architecture scales. The energy required to switch the channel is independent of the energy in the probe beam, allowing weak control signals to gate strong data signals—an optical analog to the amplification that transistors provide in electronic circuits. The architecture avoids the fan-out limitations that plagued earlier photonic logic proposals, where each stage necessarily attenuated the signal until it became unusable.

The AND gate combines two input signals through sequential switching structures. Each input controls a separate switching node, and the nodes are arranged in series along the optical path. The output signal must traverse both switches to reach the output port. If either input is absent—if either switch remains open—the signal is blocked. Only when both inputs are present, triggering both switches to their transmissive state, does the output emerge. The physical arrangement directly implements the logical requirement: output true if and only if both inputs are true.

The OR gate uses a parallel rather than serial arrangement. Two optical paths, each controlled by one input, converge at a common output. If either input is present, its corresponding path transmits, and the output receives a signal. Only when both inputs are absent—both paths blocked—does the output remain dark. The spatial arrangement of the waveguides implements the logical function without requiring any additional control circuitry. The compound gates—NAND, NOR, XOR, XNOR—emerge from combinations of these primitives. NAND cascades an AND structure with a NOT structure, inverting the AND output. NOR similarly inverts OR. The cascade adds delay— two switching events rather than one—but femtosecond operations cascade into hundreds of femtoseconds, still thousands of times faster than electronic equivalents.

XOR requires particular attention because it is computationally significant: XOR is the fundamental operation of addition in binary arithmetic, the gate that determines whether two bits sum to zero or one before considering the carry. The implementation in the Poovey Stack uses parallel paths with cross-coupled switching. Each input controls two switching nodes: one that enables its own

output path, and one that disables the parallel path controlled by the other input. When input A alone is present, it enables path A and disables path B; the output emerges through path A. When input B alone is present, the mirror situation obtains. When both are present, each disables the other's path; no output emerges. When neither is present, neither path is enabled; again, no output. The truth table of XOR emerges from the geometry without requiring sequential evaluation. XNOR—the complement of XOR—uses similar geometry with the addition of a NOT stage at the output. The gate produces true when both inputs match: both present or both absent. In cryptographic applications and error-detection circuits, XNOR serves as a bitwise equality comparator, a function that electronic implementations achieve through multiple transistors. The photonic implementation requires no more switching nodes than XOR itself, just different routing of the output signal.

The spatial nature of these implementations deserves emphasis. In electronic circuits, logic gates are switching devices whose inputs and outputs are voltage levels on wires. The physical size of the gate has no direct relationship to its logical function; a smaller transistor performs the same logical operation as a larger one, just faster and with less power. In the Poovey Stack, the logic is embedded in the geometry—in the arrangement of waveguides, the positioning of switching nodes, the routing of optical paths. The architecture is inherently parallel because light propagates through all paths simultaneously, reaching resolution in the time required for a single switching event regardless of the complexity of the logical function being computed.

•••

Functional completeness is a precise mathematical concept with a surprisingly simple definition: a set of logic gates is functionally complete if every possible Boolean function can be expressed using only gates from that set. The NAND gate alone is functionally complete—any Boolean function whatsoever can be implemented using only NAND gates. The same is true of NOR. The set containing AND, OR, and NOT is also functionally complete. The Poovey Stack, implementing all seven standard gates, exceeds the minimum requirement by a comfortable margin.

Why does this matter? Because functional completeness is the boundary between a technology that can perform specific operations and a technology that can perform any operation. A system limited to AND and OR gates—

lacking NOT—can compute only monotonic functions: functions whose output can only increase as inputs increase. Such a system cannot implement subtraction, cannot perform comparison, cannot execute conditional branching. It can do arithmetic, but only addition. It can store data, but cannot search it efficiently. The limitations compound until the system, however fast its individual operations, proves inadequate for practical computation.

The history of photonic computing includes multiple demonstrations of partial logic families—optical AND gates in 1987, optical OR gates using different mechanisms in 1992, optical XOR for specific telecommunications applications in 2001. What these demonstrations shared was an inability to achieve inversion efficiently. The NOT gate, trivial in electronics, required either electronic conversion at some stage—sacrificing the speed advantage of optical processing—or optical mechanisms that introduced unacceptable losses, power consumption, or complexity. The research community produced hundreds of papers describing optical gates that worked in isolation but could not be combined into complete systems.

The Poovey Stack's approach to NOT—using the channel cutoff mechanism to block a probe beam—solves the inversion problem without electronic conversion and without the power penalties of earlier optical approaches. The switching energy is the same for NOT as for AND or OR: approximately 0.1 milliwatts at the switching structure. The speed is the same: 150 to 200 femtoseconds. And because NOT is achieved, all compound gates follow automatically. NAND becomes AND followed by NOT, operating at twice the single-gate delay but still completing in under 400 femtoseconds. The architecture achieves what previous approaches could not: a complete optical logic family operating at consistent speed and power across all operations.

This achievement is now documented in the public record. U.S. Patent Publication 20240411203, with Del Wolverton and Kevin Joseph Bylow as co-inventors, describes a purely optical NAND logic gate—the universal building block from which all Boolean operations can be constructed. The publication establishes that the core logical element required for general-purpose computation has been reduced to practice in photonic form, with sufficient specificity to receive USPTO publication. This is not a theoretical proposal but a documented implementation, searchable in patent databases and available for examination by anyone who wishes to verify the claims.

The optical NAND represents the culmination of what decades of photonic computing research sought but failed to achieve: a manufacturable, all-optical gate capable of implementing any Boolean function.

I remember asking Del a naive question early in our work together. "If we can make light turn off light," I said, "why can't we just build a computer out of that?" Del looked at me the way he does when I have said something that is either very stupid or accidentally profound. Then he walked me through what functional completeness actually requires—why AND alone is not enough, why OR alone is not enough, why the NOT gate that electronics achieves trivially had defeated optical computing for decades.

"The problem," he said, **"was never making light do something. The problem was making light do everything."**

That conversation lasted a couple hours. By the end, I understood why the history described in Chapter 7 was a history of partial solutions—why researchers could demonstrate individual gates but could not build complete systems. The Poovey Stack was designed from the beginning to achieve functional completeness, not as an afterthought but as the primary engineering constraint. Every material choice, every layer configuration, every switching geometry was evaluated against the question:

Does this get us to a complete logic family? The **optical NAND** documented in that patent publication is the proof that the answer is **yes.**

•••

The reader familiar with computer architecture may be wondering how the Poovey Stack relates to the von Neumann model that has dominated computing since 1945. Named for John von Neumann, who formalized it in a famous 1945 report, the von Neumann architecture defines the structure of virtually every general-purpose computer: a processing unit that performs arithmetic and logical operations, a memory unit that stores both data and instructions, and a control unit that fetches instructions from memory and directs the processor to execute them. The architecture's defining characteristic is the stored-program concept—the insight that instructions themselves can be represented as data, stored in the same memory as the data they manipulate, and modified by the programs they constitute.

The von Neumann architecture is both a triumph and a bottleneck. Its flexibility enabled the software revolution: the same hardware can run word processors, video games, weather simulations, and artificial intelligence simply by loading different programs. But its structure creates a fundamental limitation. The processor and memory are separate, connected by a communication channel of finite bandwidth. Every instruction must be fetched from memory before it can be executed; every data element must be retrieved before it can be processed. The bandwidth of this channel—the rate at which information can flow between processor and memory—bounds the speed of the entire system.

This limitation is called the von Neumann bottleneck, and it has grown more severe as processors have become faster. A modern CPU can execute billions of operations per second, but it cannot fetch billions of instructions per second from main memory. The disparity is managed through elaborate hierarchies of cache memory—small, fast memories that store frequently accessed data close to the processor—but the fundamental mismatch remains. The processor spends substantial time waiting for data, its computational capacity idled by the bandwidth limitation that von Neumann's architecture imposes.

The question for photonic computing is whether it represents a departure from the von Neumann model or an optical implementation of it.

The answer is both, depending on how the technology is deployed.

At the system level, the Poovey Stack can implement a von Neumann architecture in optical form. The logic gates described above can be combined to create arithmetic units, control units, and memory interfaces, just as electronic gates are combined in conventional processors. An optical von Neumann machine would retain the stored-program flexibility while operating at femtosecond switching speeds and dramatically lower power consumption.

This configuration represents a direct replacement for electronic computing: **same architecture, different physics, better efficiency.**

But the Poovey Stack also enables architectures that depart from von Neumann's model in ways that exploit the unique properties of optical processing. Light propagates simultaneously through multiple paths; it does not queue for a shared bus.

A photonic processor can be designed with massive parallelism built into its geometry, performing thousands of operations in the time required for a single instruction fetch in a conventional architecture. For workloads that can exploit such parallelism—and the AI workloads driving the current infrastructure crisis are precisely such workloads—the non-von Neumann configuration may prove more valuable than the traditional one. The architecture also addresses the von Neumann bottleneck through a capability the inventors call Storage-in-Light, or SIL. Rather than shuttling data between separate processor and memory components, SIL enables optical storage within the processing fabric itself. Data persists in optical resonant structures that can be read and written at femtosecond timescales, eliminating the round-trip latency of conventional memory access. The processor does not wait for memory because memory is not elsewhere; it is distributed throughout the computational structure, accessible at the speed of light.

The honest assessment is that SIL remains earlier in its development arc than the switching technology. The Poovey Switch has been validated at Technion with measured performance matching or exceeding theoretical predictions. SIL has been designed and analyzed but not yet subjected to equivalent independent validation. The physics is sound— optical bistability and resonant cavity effects are well- characterized phenomena—but the gap between theoretical design and validated performance must be acknowledged. True Photonic's development roadmap positions SIL as a near-term milestone rather than a demonstrated capability.

•••

The infrastructure crisis described in Part I is not caused by computing in general; it is caused by a specific type of computing that has exploded in demand since 2022. Understanding how photonic logic addresses that crisis requires understanding what AI systems actually do when they process your queries—what operations consume the electricity, generate the heat, and drink the water. The dominant operation in modern AI is matrix multiplication. A neural network processes information by multiplying input vectors against weight matrices, applying activation functions, and passing the results to subsequent layers.

A single query to a large language model like GPT-4 involves trillions of multiply-accumulate operations— multiplying pairs of numbers and adding the results— distributed across hundreds of matrix operations involving matrices with billions of elements. The energy consumption that Part I documented, the water

usage that has communities blocking data center construction, the grid strain that has utilities scrambling to build new power plants —all of it traces back to this single mathematical operation performed at unprecedented scale. Consider what happens when you ask a large language model to summarize a document. The model encodes your query as a vector of numbers—typically 4,096 or more dimensions in current models. That vector is multiplied against a weight matrix with billions of parameters to produce a transformed representation.

The process repeats through dozens of transformer layers, each involving multiple matrix multiplications: the attention mechanism alone requires computing query, key, and value matrices, then multiplying queries against keys to compute attention scores, then multiplying scores against values to produce the output. A single forward pass through GPT-4 involves approximately 1.8 trillion multiply-accumulate operations. A conversation with dozens of exchanges multiplies that figure accordingly.

Matrix multiplication has a property that makes it exceptionally well-suited to parallel processing: every element of the output matrix can be computed independently. The entry in row i, column j of the product matrix depends only on row i of the first matrix and column j of the second matrix, not on any other entries in either matrix or any other entries in the output. A processor with sufficient parallel resources can compute all output entries simultaneously, completing the entire multiplication in the time required for a single element. Electronic processors exploit this parallelism through architectures like GPUs, which pack thousands of simple processing units onto a single chip. But even GPUs face limits: the units must communicate through shared memory, they must be powered and cooled, and the parallelism is bounded by the physical number of units that can be fabricated.

Photonic matrix multiplication exploits a different kind of parallelism. Light passing through an optical structure interacts with all elements of that structure simultaneously— not sequentially, not through time-shared resources, but in a single pass at the speed of light. A matrix encoded in the optical properties of a device—the transmissivities of a grid of switching nodes, for example—multiplies an input vector encoded as light intensities in a single optical transit. The operation that requires thousands of clock cycles in a GPU completes in femtoseconds in a photonic processor. The Poovey Stack implements matrix multiplication through arrays of switching nodes arranged in a crossbar configuration.

The rows encode one matrix; the columns encode the other. Input signals propagate along the rows, interacting with the switching nodes whose states have been programmed to represent matrix weights. The products accumulate along the columns, emerging as output intensities that represent the matrix product. The geometry directly implements the mathematics, with the parallelism inherent in optical propagation replacing the sequential accumulation that electronic processors require. The energy advantage is dramatic. A matrix multiplication that consumes millijoules in an electronic GPU consumes **microjoules in a photonic implementation—three orders of magnitude difference.**

The thermal advantage follows directly: less energy dissipated means less heat generated, which means less cooling required. A data center performing the same AI computations photonically rather than electronically would consume a small fraction of the power and virtually none of the water. The crisis described in Part I is not inevitable; it is a consequence of solving the wrong problem—*optimizing electronic multiplication instead of replacing it.*

The scaling behavior differs fundamentally between electronic and photonic approaches. Electronic matrix multiplication scales linearly with matrix size—doubling the matrix requires roughly double the energy and time. Photonic matrix multiplication, exploiting the inherent parallelism of optical propagation, can scale sublinearly: larger matrices require more optical elements but not proportionally more time, because all elements resolve simultaneously. For the massive matrices that characterize modern AI—weight matrices with billions of parameters, attention matrices spanning thousands of tokens—this scaling difference compounds into substantial advantages.

The mapping extends beyond matrix multiplication to the other operations that AI workloads require. Activation functions—the nonlinear transformations that give neural networks their expressive power—can be implemented through optical nonlinearities in the switching structures. The ReLU activation function, which outputs zero for negative inputs and passes positive inputs unchanged, maps naturally to optical thresholding behaviors. The softmax function, which converts raw scores into probability distributions, involves exponentials and divisions that require more complex optical arrangements but remain achievable within the architecture.

Attention mechanisms, the innovation that enabled transformer models like

GPT, involve matrix operations that the same crossbar architecture can perform. The convolutional operations used in image processing map naturally to the spatial arrangement of optical elements— convolution is, mathematically, a sliding dot product, and dot products are what optical crossbars compute. Pooling operations, which reduce spatial dimensions by selecting maximum or average values from regions, can be implemented through optical combining elements. The workloads that are driving the infrastructure crisis are precisely the workloads that photonic processing handles most efficiently.

•••

The skeptical reader should be asking what has not been demonstrated. The answer is significant: integrated systems. The individual switching nodes have been validated at Technion. The logic gate configurations follow from the switching behavior through well-understood optical principles. The matrix multiplication geometry exploits mathematical properties that are beyond dispute. But a chip that performs useful computation—a photonic processor that accepts real AI workloads and returns correct results—does not yet exist.

The gap is not unique to photonic computing. It is the standard gap that every new computing technology must cross: from validated components to integrated circuits, from integrated circuits to functional processors, from functional processors to commercial products. Intel's first microprocessor, the 4004, required years of engineering between the initial transistor technology and the working chip. The same progression lies ahead for the Poovey Stack, and the uncertainties are real. Integration introduces challenges that do not appear at the component level. Thousands of switching nodes operating simultaneously will generate thermal gradients that could affect timing and signal integrity. Although the Poovey Stack generates far less heat than electronic alternatives, it generates some heat, and that heat must be managed across structures measured in micrometers. The graphene-boron nitride substrate's exceptional thermal conductivity— exceeding 2,000 watts per meter-kelvin—provides the foundation for thermal management, but translating that material property into system-level thermal stability requires careful engineering.

Optical signals propagating through complex geometries will experience losses and crosstalk that must be managed through careful design. A signal traversing multiple waveguides, passing through multiple switching nodes, interacting with multiple control structures, will lose intensity at each stage.

The cumulative loss must remain within bounds that preserve signal integrity—the output must be strong enough to register correctly and distinct enough from noise to avoid errors. Previous photonic computing efforts encountered loss budgets that proved unworkable at scale; the Poovey Stack's design addresses loss through shorter optical paths and higher switching contrast, but integration will test whether those theoretical advantages survive practical implementation. The interfaces between photonic processing fabric and electronic input/output systems must be engineered to preserve the efficiency advantages. Data enters the system as electronic signals from memory or networks; it must be converted to optical form for processing. Results emerge from optical computation; they must be converted back to electronic form for storage or transmission.

These interfaces —optical-to-electrical and electrical-to-optical converters— consume power and introduce latency. If the conversion overhead exceeds the processing savings, the system fails economically regardless of how efficiently the photonic core operates.

Manufacturing yields—the percentage of fabricated devices that actually work—must reach levels that make production economically viable. Semiconductor manufacturing achieves yields exceeding ninety percent for mature processes; new technologies typically start much lower and improve through iterative refinement. A yield of fifty percent means half the wafers produced are scrap—acceptable for research but ruinous for commerce. The Poovey Stack's compatibility with mature fabrication processes provides a starting point higher than exotic technologies would achieve, but the specific yields that commercial production will reach remain to be determined.

These challenges are not insurmountable, but neither are they trivial. The history of semiconductor development includes many technologies that failed during integration despite excellent component performance. True Photonic's approach—progressing through stages from simple circuits to complex systems, validating at each stage before advancing—follows the methodology that successful semiconductor companies have used for decades. But following the right methodology does not guarantee the right outcome. What can be said with confidence is that the physics supports the claims. Light can perform Boolean logic at femtosecond speeds using the mechanisms described in this chapter. Matrix multiplication can be implemented optically with the energy efficiency that the specifications claim. The operations that AI requires are precisely the

operations that optical processing performs most naturally. The question is not whether photonic logic is possible—it demonstrably is— but whether it can be engineered into practical systems at commercial scale.

The implications of photonic logic, if it can be scaled, extend beyond efficiency improvements to structural changes in how computing systems are built. Electronic data centers are designed around the thermal problem: dense racks of servers, elaborate cooling systems, power distribution networks capable of handling megawatts per building. The architecture reflects the physics of electronic computation— the unavoidable heat generation that accompanies electron flow through resistive pathways. Photonic computing changes the physics and therefore the architecture. A processor that generates a fraction of the heat does not require the same cooling infrastructure. A facility that consumes a fraction of the power does not require the same electrical capacity. The windowless warehouses and cooling towers that characterize today's data centers may prove to be artifacts of a particular technological era— necessary for electronic computation, unnecessary for optical.

Chapter 12 will explore what this could mean for the physical infrastructure of computing—how photonic data centers might be configured, where they might be located, how they might integrate into urban environments that currently exclude computation as too resource-intensive. But that discussion presupposes that photonic processors can be manufactured, and manufacturing is itself a technical challenge with specific requirements and constraints.

The next chapter examines the manufacturing pathway: the specific processes by which Poovey Stack devices are fabricated, the compatibility with existing semiconductor infrastructure that enables commercial production, and the scaling path from current capabilities to volume manufacturing. The logic described in this chapter is meaningless without the ability to build it; the building described in the next chapter validates whether this particular logic can enter the world as physical products.

What this chapter establishes is that the Poovey Stack provides a technically credible pathway from optical switching to optical computing. The complete Boolean logic family is implementable. The von Neumann architecture can be replicated optically, or departed from in ways that exploit photonic parallelism. The AI workloads driving the infrastructure crisis map directly to operations

that optical processors perform with extraordinary efficiency. The pieces fit together in ways that previous photonic computing approaches could not achieve.

Whether those pieces can be assembled into working systems —whether the integration challenges can be solved, the manufacturing yields achieved, the commercial products delivered—remains to be demonstrated.

The physics is proven.

The engineering is underway.

The outcome is uncertain.

But for a world facing an infrastructure crisis driven by the thermal consequences of electronic computation, the possibility of a different approach— one grounded in light rather than electrons, in femtoseconds rather than nanoseconds, in microwatts rather than watts— deserves the serious attention that this book attempts to provide.

CHAPTER 11

Building the Machine

The semiconductor industry has a saying: the lab is not the fab. A technology that works in a research environment—with hand-tuned equipment, expert operators, unlimited time per device, and no requirement for profitability—may fail completely when transferred to manufacturing conditions. The history of advanced materials is littered with laboratory demonstrations that never became products: superconducting electronics that required helium cooling no factory could maintain, compound semiconductors that demanded precursor gases too toxic for industrial use, nanoscale structures that formed perfectly once but could not be reproduced reliably at the thousand-wafer-per-week pace that commercial production requires.

The previous two chapters described what the Poovey Stack can do: switching at 150 to 200 femtoseconds, complete Boolean logic, efficient mapping to AI workloads. This chapter addresses whether it can be built—not once, in a laboratory, by inventors who understand every nuance of the technology, but repeatedly, in factories, by equipment and processes that must work the same way every time. The question is not whether photonic logic is possible; Chapter 10 established that it is. The question is whether photonic logic is manufacturable at the scale and cost that commercial deployment requires. The answer depends on a strategic choice that distinguishes True Photonic's approach from previous photonic computing efforts. That choice is called Relaxed Lithography, and understanding it requires understanding why semiconductor manufacturing has become so expensive, so concentrated, and so difficult for new technologies to enter.

• • •

Modern electronic chips are manufactured through photolithography—a process that uses light to transfer patterns onto silicon wafers. The principle is straightforward: coat a wafer with light-sensitive material, shine light through a patterned mask, develop the exposed areas to create a pattern, then etch or deposit material through that pattern. Repeat dozens of times with different masks, and the accumulated layers form transistors, interconnects, and all the structures that constitute an integrated circuit. **The challenge is resolution.**

The smallest features that photolithography can produce are limited by the wavelength of light used—shorter wavelengths enable smaller features. For decades, the industry used deep ultraviolet light at 193 nanometers, achieving feature sizes down to approximately 40 nanometers through increasingly elaborate optical tricks: immersion lithography that placed liquid between the lens and wafer, multiple patterning that exposed each layer several times with offset masks, computational lithography that pre-distorted masks to compensate for optical aberrations.

By the early 2010s, these tricks had reached their limits. The industry needed a fundamentally shorter wavelength to continue shrinking features. The solution was extreme ultraviolet lithography—EUV—using light at 13.5 nanometers, more than fourteen times shorter than the previous generation. EUV enables features below 10 nanometers, the dimensions required for the most advanced chips in production today. The cost of this transition has been staggering. An EUV lithography machine costs approximately 350 million dollars. A single machine. A modern semiconductor fabrication facility—a fab—requires multiple EUV machines along with hundreds of other tools, cleanroom infrastructure, and support systems. The total cost to build a leading-edge fab now exceeds twenty billion dollars. Only three companies in the world—TSMC, Samsung, and Intel—operate fabs capable of producing chips at the most advanced nodes. Only one company in the world—ASML, based in the Netherlands—manufactures EUV lithography equipment. This concentration creates a bottleneck for any technology that requires leading-edge manufacturing. The capacity is limited, the cost is enormous, and the queue is long. TSMC's advanced nodes are booked years in advance by Apple, Nvidia, AMD, and other high-volume customers willing to pay premium prices. A startup with a novel technology—however promising—cannot simply purchase access to EUV manufacturing. The economics do not work at low volumes, and the capacity is not available regardless.

The concentration has only intensified as Moore's Law has slowed. When each process generation delivered forty percent improvement, companies invested in their own fabs to capture competitive advantage. Now that improvements come in single-digit percentages at costs measured in billions, most chip designers have become fabless—designing chips but outsourcing manufacturing to the few companies that can afford the capital investment. This model works for established products with predictable volumes, but creates barriers for new technologies that cannot guarantee the demand foundries require to justify their attention.

Previous photonic computing efforts largely accepted this constraint and attempted to work within it. They designed devices for leading-edge nodes, assuming that photonic processors would need the smallest possible features to achieve competitive density and performance. This assumption shaped their architectures, their manufacturing partnerships, and ultimately their fates. When EUV capacity proved unavailable or unaffordable, the technologies stalled—not because the physics failed, but because the manufacturing pathway was blocked.

• • •

Relaxed Lithography rejects the assumption that photonic computing requires leading-edge processes. The Poovey Stack is designed for fabrication at mature process nodes—90 to 180 nanometers—using equipment and techniques that the semiconductor industry perfected fifteen to twenty years ago. No EUV required. No immersion lithography required. No multiple patterning required. The technology uses the manufacturing infrastructure that exists abundantly, operates reliably, and costs a fraction of leading-edge alternatives. The strategy seems counterintuitive. If smaller features enable faster, denser, more efficient electronics, why would a technology claiming superior performance deliberately choose larger features? The answer lies in the fundamental difference between electronic and photonic devices.

Electronic transistors switch by controlling the flow of electrons through channels measured in nanometers. Smaller channels mean shorter distances for electrons to travel, faster switching, lower capacitance, reduced power consumption. The relentless drive to smaller features in electronics reflects the physics of electron transport: smaller genuinely is better, up to the quantum mechanical limits now being approached.

Photonic devices operate through different physics. Light at the wavelengths used in the Poovey Stack—1,550 to 1,600 nanometers in the telecommunications infrared band—has wavelengths measured in micrometers, not nanometers. The waveguides that confine and route optical signals must be sized to match these wavelengths; structures smaller than the wavelength cannot guide light effectively. The switching nodes that perform logical operations interact with light through phenomena that occur over hundreds of nanometers. The relevant dimensions for photonic devices are inherently larger than the leading-edge dimensions of electronic devices—not because of manufacturing limitations, but because of optical physics.

This mismatch between electronic and photonic length scales creates an opportunity. A photonic device fabricated at a 90-nanometer process node is not compromised by the "large" feature size; the features are still far smaller than the optical wavelengths being manipulated. The device can achieve its full theoretical performance—femtosecond switching, microjoule energy consumption, complete logic functionality—without requiring the extreme resolution that electronic devices demand. The manufacturing simplicity is a benefit without a corresponding cost.

The mature-node strategy also avoids the yield penalties that plague leading-edge manufacturing. At 3-nanometer nodes, defect densities are measured in defects per square centimeter; at 90-nanometer nodes, they are measured in defects per square meter—a thousand-fold difference that translates directly into manufacturing yield. A technology designed for mature nodes inherits the yield advantages that decades of process refinement have achieved, rather than struggling with the yield challenges that leading-edge processes continuously confront. The economic implications are substantial. A 90-nanometer fab costs hundreds of millions of dollars to build, not tens of billions. Dozens of such fabs operate worldwide, many with available capacity as their original customers have migrated to smaller nodes. The equipment is proven, the processes are characterized, the yields are high, the costs are known. A technology designed for these nodes can access manufacturing capacity that leading-edge technologies cannot—not in years, but in months.

• • •

The manufacturing process for Poovey Stack devices begins with the substrate described in Chapter 9: sapphire wafers with graphene grown directly through microwave plasma-enhanced chemical vapor deposition. This step occurs outside the traditional semiconductor fab, in equipment designed for the specific requirements of graphene synthesis. The MW-PECVD process deposits high-quality graphene in under three minutes per wafer—a timeframe compatible with high-volume production and a dramatic improvement over the hours-long processes that characterized earlier graphene manufacturing approaches. The significance of the three-minute cycle time deserves emphasis. Semiconductor manufacturing economics depend critically on throughput—the number of wafers processed per hour, per day, per year. Equipment that processes one wafer per hour, however excellent its results, cannot support volume production. Equipment that processes twenty wafers per hour can. The MW-PECVD

process, developed by Astera Energy and adapted for True Photonic's requirements, achieves the throughput that commercial manufacturing demands without sacrificing the material quality that device performance requires.
The graphene-on-sapphire substrates then enter a standard semiconductor fabrication flow, modified to accommodate the different materials but using the same fundamental process categories that the industry has refined over decades. Thin film deposition adds the boron nitride isolation layers and the piezoelectric materials that enable switching. Photolithography patterns each layer according to the device design. Etching removes material from areas not protected by the photoresist. The sequence repeats, layer by layer, building the three-dimensional stack structure that the architecture requires.

The layer count varies by device complexity. Simple switching elements require five to ten layers. Logic circuits incorporating multiple gates require more. Complex integrated systems—the photonic processors that would eventually replace electronic chips for AI workloads—could require one hundred fifty layers or more. The Poovey Stack architecture supports this range, with the layer-by-layer fabrication approach allowing inspection and validation at intermediate stages rather than only at final test.Several aspects of the process differ from conventional silicon fabrication. The thermal budgets must accommodate graphene's properties—high temperatures that would be routine for silicon processing could damage the two-dimensional carbon lattice. The etch chemistries must selectively remove some materials while preserving others in a materials stack more diverse than typical silicon devices. The alignment tolerances, while relaxed compared to leading-edge electronics, must still maintain the precision that optical structures require—misalignment of waveguide layers by even tens of nanometers could introduce unacceptable losses.

The deposition processes for the various layers require careful optimization. Boron nitride must be deposited in its hexagonal crystalline form to achieve the desired electrical and thermal properties; amorphous or randomly oriented material would compromise device performance. The piezoelectric layers must achieve the crystalline orientation that maximizes electromechanical coupling. These requirements add process constraints but do not require exotic equipment—the deposition techniques are variants of processes used throughout t he semiconductor industry, adapted for the specific materials the Poovey Stack employs.

These differences require process development—the systematic work of establishing recipes, characterizing variations, and qualifying the manufacturing sequence for production. True Photonic has completed this development for the initial device structures, working with foundry partners to demonstrate that the processes produce working devices with acceptable yields. The work is not complete; more complex structures require additional development, and the transition from engineering samples to volume production involves further qualification. But the fundamental compatibility of the Poovey Stack with semiconductor manufacturing has been established through actual fabrication, not merely theoretical analysis.

• • •

Manufacturing yield—the percentage of devices that work correctly—determines whether a technology is commercially viable. A process with fifty percent yield wastes half its materials, equipment time, and labor producing scrap. A process with ninety-five percent yield converts nearly all inputs into saleable products. The difference between these scenarios can be the difference between profit and bankruptcy. Yield depends on defect density: the number of flaws per unit area that cause device failure. Defects arise from particles in the cleanroom, impurities in materials, variations in process conditions, equipment malfunctions, and dozens of other sources that semiconductor fabs spend billions of dollars controlling. Mature processes at established nodes achieve defect densities low enough that large chips can be manufactured economically. Leading-edge processes, still being refined, typically suffer higher defect densities that improve over time as engineers identify and eliminate sources of variation.

The Poovey Stack benefits from the maturity of the processes it employs. A 90-nanometer process in 2025 has been running in production for nearly two decades. The equipment is understood, the failure modes are characterized, the defect sources have been systematically eliminated. Yields at these nodes routinely exceed ninety percent for conventional devices. The introduction of new materials—graphene, boron nitride, the piezoelectric layers—creates new potential defect sources, but the baseline process stability provides a foundation that leading-edge manufacturing cannot offer.

The layer-by-layer fabrication approach also enables yield management strategies unavailable in conventional chip manufacturing. Each layer can be inspected before subsequent layers are added. Defective regions can be identified early,

before additional processing wastes resources on devices that will ultimately fail. In principle, defective sections could be repaired or bypassed through redundancy designed into the architecture—approaches that the vertical structure of the Poovey Stack accommodates more readily than the planar structure of conventional chips. The economic model for photonic computing differs from electronics in another important respect: the value per device can be substantially higher. An electronic processor for AI applications—an Nvidia H100 GPU, for example—sells for approximately 25,000 to 40,000 dollars. If a photonic processor delivers equivalent computational capability at a fraction of the power consumption, the value proposition supports similar or higher pricing even at lower manufacturing volumes. The economics do not require the millions of units per year that justify leading-edge electronic fabs; they require thousands of units at prices that reflect the operational savings the technology delivers.

Consider the total cost of ownership for a data center operator. The purchase price of processors is significant but not dominant; electricity costs over a multi-year deployment often exceed hardware costs. A photonic processor priced at fifty thousand dollars that reduces electricity consumption by ninety percent could deliver payback in under a year for operators currently spending millions annually on power for AI computation. The economics favor photonic computing even at price points that would be untenable for consumer electronics—because the customers are not consumers, but enterprises whose operating costs dwarf their capital costs.

• • •

The path from current capabilities to commercial products follows a progression familiar from semiconductor history: device validation, circuit integration, subsystem demonstration, system qualification, and volume production. True Photonic has completed the first stage—individual switching devices have been fabricated and validated, including the independent testing at Technion described in Chapter 9. The subsequent stages involve progressively more complex integration, with each stage validating both the technology and the manufacturing processes before advancing.

The immediate next step is underway: circuit integration - combining multiple switching nodes into functional logic circuits that perform defined operations. A circuit containing tens of switches, implementing basic arithmetic or control functions, validates that the switches work together as designed—that the timing aligns, the signal levels are compatible, the thermal interactions are manage-

able. This stage is currently underway, with engineering samples in fabrication and testing.

The circuit integration stage reveals challenges invisible at the device level. Individual switches, tested in isolation, experience uniform optical conditions—controlled input power, defined wavelengths, characterized thermal environment. Switches integrated into circuits must operate amid the optical signals from neighboring switches, the thermal gradients from simultaneous switching events, and the accumulated losses from propagation through multiple waveguide segments. A switch that performs flawlessly in isolation may exhibit degraded performance in context. The circuit integration stage identifies these effects and develops mitigation strategies before they compound at larger scales. Subsystem demonstration follows: larger aggregations of circuits that perform useful computational tasks. A photonic matrix multiplier, for example, would combine thousands of switching nodes in the crossbar configuration described in Chapter 10, accepting input vectors and producing output vectors that represent the matrix product. This stage validates not only the technology but its application—demonstrating that photonic computation actually delivers the efficiency advantages that the physics predicts.

The subsystem stage also establishes the design methodology for larger systems. What software tools support circuit design at this scale? What simulation models predict performance accurately? What test procedures verify functionality efficiently? These questions have well-established answers for electronic design; photonic design requires developing equivalent capabilities. The subsystem stage produces not only working hardware but the design infrastructure that subsequent development requires.

System qualification integrates photonic subsystems with the electronic infrastructure required for practical deployment: input/output interfaces, control systems, packaging, and cooling. Even photonic processors that generate minimal heat require thermal management for stability; even optical systems require electronic interfaces for communication with the broader computing environment. This stage produces complete products that can be evaluated by customers in realistic operating conditions.

The packaging challenge deserves particular attention. Electronic chips are packaged in standardized formats—ball grid arrays, land grid arrays, and similar

configurations that connect to printed circuit boards through well-established interfaces. Photonic devices require optical connections as well as electrical ones: fiber-optic interfaces that couple light into and out of the chip, alignment mechanisms that maintain optical pathways, protective enclosures that shield sensitive structures from contamination. These requirements add complexity but are not unprecedented; the telecommunications industry has developed photonic packaging for decades, and that experience informs the approaches True Photonic is adapting for computing applications.

Volume production scales the qualified processes to manufacturing rates that support commercial deployment. The transition from engineering volumes—tens or hundreds of devices—to production volumes—thousands or tens of thousands—requires process optimization for throughput, supply chain development for materials and components, quality systems for consistency, and all the infrastructure that distinguishes a laboratory from a factory.

The volume production stage also establishes the supply chain resilience that commercial customers require. A technology dependent on single-source materials or sole-provider processes creates business risk that enterprise customers will not accept. True Photonic's strategy emphasizes multiple qualified suppliers for critical materials, multiple qualified foundries for device fabrication, and standard interfaces that avoid lock-in to particular equipment or process variants. The Relaxed Lithography approach supports this strategy by accessing manufacturing capacity that exists across multiple geographies and multiple providers.

True Photonic's timeline projects eighteen months from current status to initial commercial products. This estimate reflects the device validation already completed, the circuit integration currently underway, and realistic assessments of the engineering work remaining. The timeline could extend if unexpected challenges emerge during integration; it could compress if the work proceeds without significant obstacles. What the timeline does not include is the decade-long development cycles that characterized previous photonic computing efforts—cycles driven largely by manufacturing challenges that Relaxed Lithography avoids.

• • •

The geographic distribution of semiconductor manufacturing has become a matter of national security concern. Leading-edge fabs are concentrated in Taiwan and South Korea, with limited capacity in the United States, Europe,

and Japan. This concentration creates supply chain vulnerabilities that governments worldwide are now attempting to address through industrial policy, subsidies, and strategic investments. The CHIPS Act in the United States, the European Chips Act, and similar initiatives in Japan and elsewhere all aim to reshore semiconductor manufacturing capacity. These initiatives focus primarily on leading-edge manufacturing—the EUV-based processes that produce the most advanced chips. The assumption is that leading-edge capacity is strategically critical, and that mature-node capacity is adequately distributed and less significant. For conventional electronics, this assumption is reasonable. For photonic computing, it creates an opportunity.

Mature-node manufacturing capacity exists abundantly in the United States and allied nations. Fabs that once produced leading-edge chips have transitioned to mature nodes as their equipment aged; new fabs serving automotive, industrial, and analog markets operate at these nodes by design. A technology that requires 90 to 180 nanometer processes—not 3 to 7 nanometer processes—can be manufactured domestically without the multi-billion-dollar investments that leading-edge reshoring requires.

The strategic implications extend beyond simple supply chain security. A domestically manufactured photonic computing technology could reduce dependence on the concentrated leading-edge capacity that currently constrains AI development. Data centers powered by photonic processors would not compete for the limited TSMC capacity that every AI company currently requires. The bottleneck that Chapter 2 described—the shortage of advanced chips that limits AI deployment—could be bypassed rather than merely alleviated.

The policy implications are significant. Governments investing billions to reshore leading-edge semiconductor manufacturing face timelines measured in decades and outcomes measured in uncertainty. A photonic computing technology manufacturable at mature nodes could deliver strategic capability sooner, at lower cost, using infrastructure that already exists. This is not an a rgument against leading-edge investments—electronic computing will continue to require advanced manufacturing for the foreseeable future—but rather an observation that photonic computing offers a complementary pathway with different risk and timeline characteristics. This is not to suggest that photonic computing will immediately replace electronic computing across all applications. Electronic processors will continue to dominate general-purpose computing,

mobile devices, and applications where the existing infrastructure represents sufficient capability. But for the specific workloads driving the infrastructure crisis—the AI inference that consumes megawatts per data center, the training runs that require dedicated power plants—photonic alternatives manufactured outside the leading-edge bottleneck could provide both performance and strategic advantages.

• • •

The honest assessment of manufacturing readiness must distinguish between what has been demonstrated and what remains to be proven. True Photonic has fabricated working devices at commercial foundries using the Relaxed Lithography approach. The switches work. The processes are compatible with standard semiconductor manufacturing. The substrate technology produces material suitable for device fabrication. These are not projections; they are accomplishments. What has not been demonstrated is volume production of complex integrated systems. The engineering from validated switches to validated circuits is underway but incomplete. The further engineering from circuits to subsystems to products remains ahead. Each transition introduces challenges that cannot be fully anticipated until encountered. The eighteen-month timeline represents management's best estimate, not a guaranteed delivery date.

The manufacturing yields that will be achieved in production are unknown. Process development has established that working devices can be fabricated, but the yield percentages that determine commercial viability will emerge only from production experience. True Photonic's projections assume yields consistent with mature-node manufacturing of novel device structures—reasonable but unvalidated assumptions that production will eventually confirm or refute.

The cost structures that will characterize commercial products remain uncertain. Device costs depend on yields, throughput, materials costs, and manufacturing overhead—all of which will be refined through production experience. The pricing assumptions in True Photonic's business model are based on engineering estimates and market analysis, not production accounting. What can be said with confidence is that the manufacturing pathway is realistic. Unlike previous photonic computing efforts that required exotic materials, cryogenic temperatures, or processes available only in specialized research facilities, the Poovey Stack can be built using existing semiconductor infrastructure at accessible cost points. The pathway is not guaranteed to succeed, but it is not blocked by fundamental manufacturing barriers.

The lab may not be the fab, but the gap between them has been deliberately minimized.

• • •

Part III of this book has examined the architecture of photonic computing: the physical structures, the logical capabilities, and the manufacturing pathways that could enable a new approach to computation. The Poovey Stack provides validated switching at femtosecond speeds. The logic family achieves functional completeness through mechanisms compatible with commercial production. The Relaxed Lithography strategy accesses manufacturing capacity that leading-edge electronic technologies cannot reach.

Part IV turns from how the technology works to what it could mean—for the infrastructure crisis described in Part I, for the communities affected by data center expansion, for the industries that depend on continued growth in computational capability. The examination will be prospective rather than retrospective, exploring possibilities rather than documenting accomplishments. The uncertainties acknowledged throughout Part III apply with greater force to the projections that follow. Chapter 12 examines Clean Compute: the possibility of data centers that operate without the resource consumption that has made them controversial. If photonic processors deliver the efficiency the physics predicts, what does that mean for facility design, for site selection, for community relations? The infrastructure of computing could be fundamentally reconfigured.

Chapter 13 examines the new stack: the applications and industries that photonic computing could enable. From medical diagnostics to financial modeling to autonomous systems, computational capability has historically unlocked new possibilities. Photonic computing at photonic efficiency could unlock possibilities currently foreclosed by power and thermal constraints.

Chapter 14 examines the light ahead: the trajectory from current status to widespread deployment, the obstacles that remain, and the reasons for believing that 2026 might indeed become the year of light. The conclusion will be neither triumphalist nor defeatist, but realistic—acknowledging both the genuine promise and the genuine uncertainty that characterize any emerging technology.

The architecture has been described. The manufacturing has been explained. What remains is to explore what could be built—and why it matters.

PART IV

THE FUTURE

Computation that does not require rivers of water, forests of cooling towers, and the output of power plants dedicated to answering questions.

CHAPTER 12

Clean Compute

The data center that dominates the public imagination is a monument to industrial necessity: a windowless fortress covering acres of land, surrounded by cooling towers that exhale clouds of steam into the atmosphere, humming with the constant drone of fans and chillers, drawing power through transmission lines thick as a man's arm from substations that could serve a small city. This image is accurate because it reflects the physical requirements of electronic computation at scale. The machines generate heat. The heat must be removed. The removal requires infrastructure that cannot be hidden.

The communities that host these facilities experience them as industrial installations—because that is what they are. The tax revenue may be welcome, the jobs may be appreciated, but the facility itself is not something anyone would choose to live beside. The noise, the traffic during construction, the visual intrusion of industrial architecture on landscapes that were recently farms or forests—these are costs that communities accept in exchange for economic benefits, or resist when the benefits seem insufficient.

Part I of this book documented the backlash that has emerged as these costs have become clearer: the sixty-four billion dollars in blocked or delayed projects, the activist groups organizing in twenty-eight states, the local officials losing elections over data center approvals. The opposition crosses political lines and geographic boundaries. In Warrenton, Virginia, five hundred residents packed a town council meeting to oppose Amazon. In Chandler, Arizona, the city council voted unanimously against a facility despite lobbying from a former United States Senator. The backlash is not irrational; it reflects a genuine assessment that the costs and benefits of current data center development are not fairly distributed. But what if the data center itself could change? What if the ninety percent reduction in power consumption described in Part III translated into facilities that looked fundamentally different from the industrial installations that communities have learned to oppose? This chapter examines that possibility—not as speculation about distant futures, but as an exploration of what photonic computing's efficiency could mean for data center design, site selection, and community relations if the technology delivers what the physics predicts.

The physics of the transformation is straightforward, even if the implications are far-reaching. A data center's physical footprint is determined primarily by its cooling requirements. The servers themselves are relatively compact—racks of circuit boards and processors that pack enormous computational density into modest volumes. But those processors generate heat, and heat removal dominates facility design. Consider a hypothetical one-hundred-megawatt data center, representative of the hyperscale facilities operated by major cloud providers. One hundred megawatts of electrical power, converted almost entirely to heat by the electronic processors, must be continuously removed from the building. The cooling infrastructure required for this task is substantial: massive chillers that could fill a basketball court, cooling towers rising forty or fifty feet into the sky, pumping systems that circulate hundreds of thousands of gallons of water daily, air handling units the size of shipping containers pushing millions of cubic feet of air per hour. The cooling plant can occupy as much floor space as the computing equipment it serves. The cooling towers alone may cover an acre or more.

The numbers are not abstractions. Google's facility in Council Bluffs, Iowa—one of the largest in the world—consumes approximately one billion gallons of water annually. That is enough water to supply the residential needs of a city of fifty thousand people. Microsoft's data centers in Arizona draw water from the Sonoran Desert's already-stressed aquifer systems. Northern Virginia's data center cluster consumed nearly two billion gallons in 2023, a sixty-three percent increase from just four years earlier. These facilities are not merely using water; they are consuming it, sending it into the atmosphere through evaporative cooling where it disperses beyond recovery.

Now apply the ninety percent power reduction that photonic computing offers. The same computational output—the same queries answered, the same models trained, the same data processed—requires ten megawatts instead of one hundred. Ten megawatts of heat to remove instead of one hundred. The cooling infrastructure scales accordingly, not quite linearly but close enough for planning purposes. The cooling towers that dominated the site become unnecessary. The chiller plant that occupied a significant portion of the building shrinks to a fraction of its former size. The pumping systems, the air handlers, the elaborate ductwork—all reduce in proportion to the reduced thermal load. The transformation is not merely quantitative. Below certain thermal thresholds, the nature of cooling changes entirely. A facility generating one hundred megawatts of heat

requires active cooling—mechanical systems that consume significant energy themselves, typically adding fifteen to thirty percent to the facility's total power draw. A facility generating ten megawatts can potentially be cooled passively, or with minimal mechanical assistance. The heat can be dissipated through conventional building HVAC systems, through carefully designed natural ventilation, through thermal mass and building envelope design that would be utterly inadequate for electronic computing.

The metric that data center operators use to measure cooling efficiency is called Power Usage Effectiveness, or PUE. A PUE of 2.0 means that for every watt powering computing equipment, another watt powers cooling and other overhead. A PUE of 1.5 means that overhead consumes half as much as computation. The best hyperscale facilities achieve PUEs around 1.1 or 1.2 through extraordinary engineering investment. A photonic facility with ninety percent less heat to remove could achieve PUEs approaching 1.0—essentially no cooling overhead—because the heat load would fall below the threshold requiring dedicated cooling infrastructure. The water consumption that has made data centers controversial in drought-prone regions follows the same logic. Evaporative cooling—the cooling towers that consume millions of gallons annually—exists because it is the most efficient way to reject large amounts of heat. When the heat load drops by ninety percent, evaporative cooling may no longer be necessary. Closed-loop systems that reject heat without consuming water become practical. Air-cooled systems that would be inadequate for electronic data centers can handle photonic thermal loads. The facility that once drank a river needs only modest municipal water service.

• • •

The visual transformation would be equally dramatic. The cooling towers that announce a data center's presence from miles away—those massive concrete or steel structures venting plumes of water vapor into the sky—would simply not exist. The backup generators that line facility perimeters, ready to maintain operations during power outages, could be smaller or fewer. The electrical substations that can cover acres would shrink proportionally. The facility would look like what it is: a building containing computing equipment.

Not a power plant. Not a chemical factory. A building.

The architectural possibilities that emerge from this transformation are significant. Current data centers are designed around cooling constraints: floor layouts

optimized for hot-aisle and cold-aisle airflow patterns, ceiling heights of fifteen or twenty feet to accommodate air handling requirements, building orientations chosen for cooling tower placement and prevailing wind patterns. The buildings are utilitarian because utility is all that matters when thermal management dominates design. Remove the thermal constraints, and data centers could be designed like other commercial buildings—for efficiency of space utilization, for aesthetic compatibility with surrounding development, for integration into urban fabric rather than isolation from it. The ceiling heights could drop to standard commercial dimensions. The floor plates could be configured for human occupancy rather than optimized airflow. Windows become possible when the building no longer needs to be a hermetically sealed thermal management system. The architecture could express something other than industrial necessity.

The footprint reduction compounds these possibilities. A photonic data center delivering equivalent computational output to a current hyperscale facility might require one-tenth the land area—not because the computing equipment is smaller, but because the cooling infrastructure that dominates current facilities would largely disappear. A facility that once required forty acres might fit on four. A facility that demanded a rural site because of its scale and visual impact might be appropriate for suburban or even urban locations.

This is not to suggest that every data center could become an urban high-rise. Large-scale computing will continue to require significant infrastructure: robust electrical service with multiple redundant feeds, diverse network connectivity from multiple carriers, physical security appropriate for critical infrastructure, and access for maintenance and equipment replacement. But the constraints that have pushed data centers to remote locations—primarily the need for space and resources to support massive cooling operations—would be fundamentally relaxed. The site selection calculus would change.

• • •

SkyLight Holdings Inc. was created to pursue exactly this opportunity—but with a focus that goes beyond building new facilities. The insight driving SkyLight is that photonic computing does not merely enable a new kind of data center; it unlocks an entirely new category of real estate for computational use. Specifically, it transforms stranded commercial towers from depreciating liabilities into appreciating infrastructure assets.

The commercial real estate crisis that followed the pandemic has left cities worldwide with a specific and growing problem: office towers that cannot sustain their original purpose. Remote work has permanently reduced demand for traditional office space. Occupancy rates in major markets remain twenty to thirty percent below pre-pandemic levels, with no clear path to recovery. The buildings exist—substantial structures with robust electrical service, elevator systems, floor loading designed for commercial use—but the tenants who once filled them are not coming back. The obvious solution, residential conversion, has proven economically challenging in most cases. Office floor plates are typically too deep for residential use, with interior spaces too far from windows for livable apartments. The plumbing infrastructure is wrong—offices need restrooms clustered around cores, while residences need kitchens and bathrooms distributed throughout. The conversion costs frequently exceed the value of the resulting residential units, particularly in markets where residential rents cannot support the investment required.

The result is a class of stranded assets measured in the hundreds of billions of dollars globally. Real estate investment trusts hold these buildings on their balance sheets at values that bear diminishing relationship to economic reality. Pension funds that invested in commercial real estate face write-downs that affect millions of beneficiaries. Municipal tax bases built on commercial property assessments are eroding as values decline. The problem is structural, not cyclical, and traditional real estate solutions have not solved it.

Photonic computing offers a different kind of conversion—one that works with the characteristics of office towers rather than against them. Deep floor plates that frustrate residential conversion are ideal for computing equipment, which does not need window access. Robust electrical service designed for commercial loads can power computational infrastructure. Elevator systems built for office worker traffic can handle equipment installation and maintenance. The buildings were designed for exactly the kind of use that photonic data centers require, minus the thermal constraints that have historically made such use impossible. The thermal constraints are the key. A traditional data center cannot occupy an office tower because the cooling requirements are incompatible with building systems designed for human comfort. An office building's HVAC system might handle ten or twenty watts per square foot of heat load; a data center generates one hundred watts per square foot or more.

The mismatch is an order of magnitude. No amount of retrofit engineering can make an office tower's cooling infrastructure adequate for electronic computing at scale.

But photonic computing changes the equation. A facility generating ninety percent less heat falls within the range that commercial building systems can manage. Ten watts per square foot is not dramatically different from the heat load of a densely occupied office floor with its computers, lighting, and occupants. The building's existing HVAC, perhaps modestly upgraded, can handle the load. The transformation from stranded office tower to operational data center becomes a matter of installing computing equipment and network connectivity, not reconstructing the building's mechanical systems from scratch.

• • •

The strategic implications of this insight extend beyond individual buildings to the structure of the technology industry's support. Data centers have historically created opposition: communities that resist their construction, ratepayers who bear higher electricity costs, environmental groups that challenge their resource consumption. The political economy of data center development is one of concentrated benefits—to operators and their customers—and distributed costs borne by everyone else.

Tower conversion creates a different political economy. Real estate owners holding stranded assets become advocates for the technology that makes those assets valuable again. Institutional investors—the pension funds, insurance companies, and sovereign wealth funds that hold commercial real estate—develop financial interest in photonic computing's success. Municipal governments facing eroding tax bases see a path to restoring property values without the burdens that traditional data centers impose. The stakeholder alignment shifts from opposition to support. Consider the magnitude of the aligned interests. The commercial real estate market in the United States alone represents trillions of dollars in asset value, with office properties comprising a substantial share. The institutional owners of these assets—Blackstone, Brookfield, pension systems like CalPERS and the Teachers' Retirement System of Texas—manage capital measured in hundreds of billions. If photonic computing can transform even a fraction of stranded office inventory into productive infrastructure, these institutions become natural advocates for the technology's adoption and success.

SkyLight Holdings, as a subsidiary of True Photonic—seventy-five percent owned by the parent company—is positioned to capture this opportunity while creating these aligned stakeholders. The corporate structure reflects a strategic assessment: True Photonic's core business is developing the photonic computing technology itself, while building and operating data centers requires different capabilities, different capital structures, and different relationships. SkyLight was established to pursue the infrastructure opportunity, to build the partnerships with real estate owners that tower conversion requires, and to demonstrate that photonic computing creates value beyond the technology itself.

The name carries deliberate significance. "SkyLight" evokes both the light-based computation that enables the facilities and the architectural transformation that becomes possible—towers with their original windows and natural light preserved, not hermetically sealed industrial installations. But the name also signals the strategic focus: not greenfield development on rural sites, but vertical deployment in the towers that define urban skylines worldwide.

The towers serve purposes beyond computation alone. Their height and urban locations make them natural nodes in communications networks. True Photonic's patent portfolio includes innovations in optical signal distribution that could transform converted towers into infrastructure serving multiple functions: computation, certainly, but also network connectivity that extends the reach of photonic systems beyond the buildings themselves. The details of this network architecture—including the use of existing structures like billboards as signal repeaters—belong to Chapter 14's discussion of the broader infrastructure vision. Here, the essential point is that tower conversion is not merely about repurposing real estate; it is about creating an integrated infrastructure that compounds value across multiple dimensions.

• • •

The urban deployment that tower conversion enables deserves examination on its own terms. Current hyperscale facilities are located where they are—in rural Iowa, in suburban Virginia, in the Arizona desert—because urban locations cannot accommodate their requirements. The land costs would be prohibitive for facilities covering twenty or forty acres. The power requirements would strain urban grids already serving dense populations. The cooling infrastructure would be impossible to site amid surrounding development.

Photonic computing in converted towers inverts this logic entirely. The buildings already exist; there are no land costs for new construction. The electrical service is already in place, designed for commercial loads that approximate what photonic computing requires. The cooling can be handled by building systems designed for human occupancy. The infrastructure barriers that have excluded computation from urban cores dissolve. The advantages of urban deployment are substantial for certain applications. Latency—the delay between a user's request and the system's response—decreases with proximity. Light travels through fiber optic cable at roughly two hundred thousand kilometers per second, but even at light speed, distance introduces measurable delay. A data center in Manhattan serves New York users faster than a data center in Iowa by perhaps twenty or thirty milliseconds. For most web browsing, that difference is imperceptible. For financial trading, it is the difference between executing a trade and missing it. For real-time collaboration, it is the difference between natural conversation and awkward delay. For autonomous vehicles receiving updates from cloud systems, it could be the difference between safe operation and dangerous lag.

Current edge computing strategies attempt to address this by deploying smaller facilities closer to users. Companies like Equinix and Digital Realty operate urban data centers in major metropolitan areas, accepting higher costs and smaller scales to achieve proximity benefits. But edge facilities face the same thermal constraints as hyperscale centers, merely at smaller scale. They still require cooling, still generate noise, still present siting challenges. A five-megawatt edge facility needs proportionally the same cooling infrastructure as a fifty-megawatt hyperscale center. The heat per server is identical regardless of facility size.

Photonic edge facilities in converted towers would face these challenges in attenuated form—potentially mild enough to allow deployment on multiple floors of a single building, or across a portfolio of buildings in a metropolitan area. The facility would be invisible to the building's other occupants if any remain—no different in impact from a large IT installation or trading floor server room. The computation that currently happens in distant data centers could happen in the same building where users work, or in the tower across the street, or distributed across dozens of towers throughout a city.

The community relations transformation follows from these changes in physical impact and stakeholder alignment. The data center opposition documented in Part I emerges from a straightforward assessment: the facilities consume

resources, create nuisances, and provide limited local benefit while enriching distant corporations. Photonic computing in converted towers changes every element of this equation. The resource consumption drops by ninety percent. The nuisances—noise, visual impact, water consumption—largely disappear when cooling requirements fall below the threshold requiring industrial infrastructure. The local benefit expands when the facility preserves and enhances an existing building rather than replacing productive land with industrial installation. The enrichment flows in part to local stakeholders—the building owners, the municipal tax base, the workers who operate and maintain the facility—rather than exclusively to distant technology companies.

SkyLight's approach to community benefit goes beyond merely eliminating nuisances. Even a facility with ninety percent reduced heat generation produces some thermal output, and that output can be captured rather than wasted. District heating systems in Scandinavian countries already use data center waste heat to warm nearby buildings; a SkyLight facility could provide similar benefit to surrounding development, converting computational byproduct into community amenity. The buildings' existing infrastructure can be shared: the electrical service that powers computation can provide backup power to neighboring buildings, the network connectivity can extend to local institutions, the facility can include community-accessible spaces that transform a formerly vacant tower into a neighborhood asset.

The economic model that makes these provisions possible reflects both photonic computing's efficiency and the tower conversion strategy's capital efficiency. Traditional data centers require billions in new construction; tower conversion repurposes existing assets at a fraction of the cost. Traditional data centers pay premium prices for rural land with adequate power and water; tower conversion uses urban buildings whose values have collapsed precisely because traditional uses no longer work. The savings compound: lower construction costs, lower land costs, lower operating costs from reduced power consumption. The margin created by these savings can fund community benefit that traditional data center economics cannot accommodate.

The honest acknowledgment that has characterized this book applies with particular force to the projections in this chapter. Nothing described here has been demonstrated at scale. The ninety percent power reduction has been measured in device testing, not facility operation.

The tower conversion model has been analyzed, not implemented. The stakeholder alignment has been theorized, not validated through actual transactions. The gap between device performance and facility performance is significant. A photonic processor that achieves ninety percent power reduction in laboratory testing must maintain that efficiency when deployed in production systems, integrated with electronic components for input and output, operating continuously under varying loads for months and years. System integration introduces losses and overhead that device measurements do not capture. Engineering estimates suggest that system-level efficiency of seventy to eighty percent power reduction—rather than the ninety percent at device level—is realistic for initial deployments. This reduced figure still represents a transformational improvement, still enables the architectural and operational changes described in this chapter. But it is a more conservative projection than the device specifications might suggest.

The tower conversion model involves assumptions about building suitability that will vary case by case. Not every stranded office tower will be appropriate for computational use. Some buildings have electrical service inadequate for the load. Some have structural limitations that prevent equipment installation. Some have ownership structures or existing commitments that complicate conversion.

The model describes a category of opportunity, not a universal solution applicable to every distressed commercial property. The stakeholder alignment depends on demonstrating value before it can generate support. Real estate owners will not advocate for photonic computing based on theoretical analyses; they will advocate based on successful conversions that restore value to their portfolios. Institutional investors will not commit capital based on projections; they will commit based on operating facilities that generate returns. The political support that aligned stakeholders could provide remains potential until the first conversions demonstrate that the model works.

True Photonic and SkyLight understand this requirement and are acting accordingly. Prototype site development is underway in Georgia and New Orleans, working toward the first operational demonstrations of the tower conversion model. The specific properties remain in negotiation, but the commitment to building proof points—not merely publishing analyses—reflects recognition that the model must be demonstrated before it can be scaled. The approach is deliberately vertically integrated. True Photonic is building its own photonic computing stack for initial deployments rather than depending

on external partners for critical components. SkyLight is developing its own facilities rather than licensing the model to third-party operators. This vertical integration sacrifices speed for control: by owning the technology, the manufacturing relationships, and the initial facilities, the companies can ensure that first deployments succeed on their own terms rather than failing due to execution problems outside their control. The first conversions must work. Everything that follows depends on them working. **The strategy reflects that reality.**

The regulatory environment for tower conversion is untested. Zoning codes i n most jurisdictions do not contemplate data centers in office buildings. Building codes may require modifications to address the specific characteristics of computational equipment in spaces designed for human occupancy. Permitting processes may introduce delays as officials encounter unfamiliar applications. These regulatory uncertainties add timeline risk to the tower conversion strategy. These uncertainties are not reasons for pessimism; they are reasons for caution and for staged implementation that validates assumptions before committing to scale. The technical foundation for Clean Compute is sound.

The tower conversion thesis addresses a real problem with a logical solution. The stakeholder alignment creates genuine shared interest. But the translation from concept to deployment involves countless details that cannot be fully anticipated, and some of those details will prove more challenging than current projections assume.

• • •

The infrastructure crisis described in Part I of this book is not a problem that will solve itself. The demand for computation continues to grow at rates that strain even optimistic projections. The constraints on power, water, and community acceptance continue to tighten as existing resources are exhausted and new development faces organized opposition. The gap between what current technology requires and what sustainable development permits continues to widen. Something must change, and the question is whether photonic computing can be that change. Clean Compute represents one vision of what that change could look like.

Computation that operates in converted urban towers rather than purpose-built rural installations. Facilities that generate wealth for real estate owners rather than opposition from neighboring communities.

Infrastructure that supports the continued growth of artificial intelligence while solving rather than creating problems for the cities where most people live.

The vision is ambitious, but it is not fantasy. It follows directly from the technical specifications validated in Part III of this book. If photonic processors achieve the efficiency that physics predicts and device testing has demonstrated, then the tower conversion model becomes economically and physically viable.
The transformation is not speculative; it is consequential. The question is not whether these characteristics are possible but whether the technology can be deployed at the scale required to address both the infrastructure crisis and the commercial real estate crisis simultaneously.

SkyLight Holdings exists to answer that question.

Its conversions will either demonstrate that Clean Compute in urban towers is practical—that photonic computing can restore value to stranded assets while delivering computational infrastructure—or reveal obstacles that require further engineering and alternative approaches. Either outcome advances understanding. Either outcome contributes to addressing the crises that motivated this book. The following chapters complete the examination of photonic computing's potential. Chapter 13 looks beyond infrastructure to the applications and industries that photonic efficiency could enable—the new capabilities that become possible when computational power is no longer constrained by thermal limits. Chapter 14 examines the path forward: the timeline for deployment, the broader network infrastructure that converted towers could support, and the reasons for believing that the next few years will determine whether photonic computing achieves its promise or joins the long list of technologies that were always almost ready.

The infrastructure is the foundation. Without facilities to deploy photonic processors, the technology remains academic—interesting to researchers but irrelevant to the crises it purports to address. Clean Compute is the bridge between laboratory achievement and practical impact—the means by which theoretical efficiency becomes measured benefit. The towers are waiting. The technology is advancing. **What happens next depends on whether the bridge can be built.**

CHAPTER 13

The Next Stack

The previous chapters examined photonic computing as infrastructure—processors, data centers, the physical plant of computation.

This chapter examines what becomes possible when that infrastructure exists.

The applications are not hypothetical extensions of current technology; they are capabilities that electronic computing cannot achieve regardless of how much power it consumes or how many chips are deployed. They represent new categories of what computation can do.

The Poovey Stack is not merely a faster processor. It is a platform—a foundation on which applications can be built that exploit photonic computing's specific characteristics: femtosecond switching speeds, negligible heat generation, extreme power efficiency, and the ability to operate in environments where electronic processors cannot survive. Each characteristic enables applications that were previously impossible, impractical, or prohibitively expensive. The platform creates a surface for innovation that extends far beyond data center efficiency. True Photonic's corporate structure reflects this platform logic. The parent company develops and licenses the core technology. SkyLight Holdings, as described in Chapter 12, deploys infrastructure. True Photonic Ventures creates and funds companies that build applications on the photonic platform—vertical-specific enterprises that take the technology into markets where domain expertise matters as much as technological capability. The ventures become tenants in SkyLight facilities, creating an ecosystem where infrastructure and applications reinforce each other.

• • •

I did not set out to reinvent money.

I did not set out to rebuild the internet.

I did not set out to transform national defense or revolutionize medicine.

I set out to solve a problem in computation, and the solutions kept pointing toward transformations I had never anticipated.

The pattern became clear as I learned what the Poovey Stack could actually do. Every time Del or Mark explained another characteristic of photonic switching—the power efficiency, the thermal profile, the speed—my mind raced to applications. Not incremental improvements to existing products, but wholesale replacements of systems the world assumes are fixed.

If it switches in femtoseconds, it can settle financial transactions faster than light crosses a continent.

If it generates no heat, it can run continuously without the cooling infrastructure that makes data centers environmental pariahs.

If it operates at one percent of the power draw, entire industries built on energy-intensive computation become obsolete overnight.

If it fits inside the body without warming tissue, it can monitor and treat conditions that electronic devices cannot touch.

The verticals emerged not from market research but from understanding the physics and asking a single question: **What does this make possible that was impossible before?**

I started filing patents. Not just for the core technology—Del and Gary were handling much of that—but for every application I could be made to understand.

Blockchain infrastructure.

Optical networking.

Solar energy conversion.

Cardiac monitoring.

Robotic actuation.

Defense systems.

Each vertical where photonic computing could replace electronic systems, I wanted True Photonic to own the intellectual property before the world understood what was coming.

This is my Photonic BlackRock strategy. Own the IP for as much of the coming market as possible before the market knows to take us seriously. The window will not stay open forever. Once we open the first Clean Compute Center, once the efficiency numbers become public, once the industry realizes that photonic computing actually works—*there will be a mad scramble.* Every major technology company, every chip manufacturer, every vertically integrated player will race to develop or acquire photonic capability. The patents we file now, for applications the market has not yet imagined, become the toll roads of that future.

Ten provisional patents filed between November and December 2025 - more being filed every week. The Poovey Stack itself. A Photonic Muscle system. The graphene substrate platform. TCP/IP-L networking (internet-in -light). Photonic-graphene solar cells. Relaxed Lithography manufacturing. Storage-in-Light memory. CardioLT wearable monitoring. PulseLT implantable devices. Substrate architecture. Each patent represents a vertical where we intend to be first—not just first to market, but first to **own the intellectual framework** that defines the market.

Our internal tagline captures the scope of the ambition: **The Next Generation of Everything**.

It is not hyperbole.

Every device that currently runs on electrons could, in principle, run on photons—and run better. Every system that generates heat could run cool. **Every** network that merely transports data could also compute during transit. **Every** financial transaction that takes minutes could settle in milliseconds. **Every** medical device that lasts years could last decades.

The platform is horizontal; the applications are limitless.

This chapter surveys four domain *examples* where the next stack enables transformational applications: a new architecture for trust and value exchange, a new model for the internet itself, sovereign infrastructure for national security, and—at the most intimate scale—devices that operate inside the human body. The examination is necessarily forward-looking; these applications are in development, not deployment. But the physics that enables them is validated, and the path from current capabilities to these applications is engineering rather than discovery.

Bitcoin mined coal.
Ethereum mined electricity.
We mint with photons.

That is not another slogan. It is a description of what becomes possible when the fundamental unit of computation changes from an **electronic transistor** to a **light-actuated light-switch**. The blockchain revolution promised to transform finance, governance, and trust itself—but it ran into a wall built of physics.

The energy required to secure distributed ledgers grew until Bitcoin alone consumed more electricity than Thailand. The heat generated by mining operations required cooling infrastructure that drained aquifers and strained power grids. The carbon footprint expanded until cryptocurrency mining produced more emissions than many industrialized nations. The very technology that promised to democratize finance became an environmental catastrophe that concentrated wealth among those who could afford industrial-scale power consumption.

The contradiction was not a bug in implementation. *It was a feature of the underlying physics.* Electronic transistors generate heat when they switch.

More computation means more switching means more heat means more cooling means more power.

Proof-of-work security requires computation that is deliberately wasteful—that is the point, the mechanism by which the network achieves consensus without central authority. But wasteful electronic computation is environmentally ruinous at scale. The mathematics of decentralized trust collided with the thermodynamics of electronic switching, and thermodynamics won.

What we call the **Light-Switch Ledger™**—a blockchain architecture built entirely on photonic infrastructure—eliminates this contradiction entirely.

The numbers are stark, and they deserve examination in detail.

A single Bitcoin transaction today consumes **approximately 500 kilowatt-hours of electricity.** That is enough energy to power an average American home for seventeen days. It generates roughly **500 grams of carbon dioxide**—equivalent to driving a gasoline car for a thousand miles, or taking a commercial flight from

New York to Chicago. Every time someone buys a coffee with Bitcoin, the planet absorbs the environmental impact of a cross-country road trip.

A single transaction on photonic infrastructure consumes less than 0.001 kilowatt-hours and generates less than 0.1 grams of carbon dioxide—**less than a single human breath.**

The reduction is not incremental. It is not a percentage improvement that compounds over time. It is three orders of magnitude, achieved immediately, as a direct consequence of the underlying physics. Photons do not generate heat when they switch. They do not require cooling. They do not fight thermodynamics; they exploit it.

Scale the comparison to global operations. Blockchain networks today—Bitcoin, Ethereum, and the thousands of smaller chains—consume approximately 200 terawatt-hours of electricity annually. That is comparable to the total electricity demand of Thailand, or Argentina, or the Netherlands. It represents roughly one percent of global electricity consumption, dedicated entirely to the maintenance of distributed ledgers. The associated carbon emissions exceed 100 million metric tons per year—comparable to the annual emissions of Belgium or the Czech Republic.

On photonic infrastructure, that same computational work shrinks to less than one terawatt-hour annually. The carbon footprint of an entire mid-sized nation becomes the carbon footprint of a city block. The cooling water requirements—billions of gallons annually for current mining operations—drop to zero. The massive industrial facilities that house mining operations, with their roaring fans and endless rows of heat-generating equipment, become unnecessary. A photonic mining operation fits in an office building and runs silently. But efficiency, however dramatic, is only the beginning. Speed transforms what blockchain can do.

Current blockchain architecture reflects the limitations of electronic computation. Bitcoin's ten-minute block interval exists because the network needs time to propagate transactions and reach consensus across thousands of nodes running electronic processors with nanosecond-scale switching speeds. The network must wait for information to travel around the world, for nodes to validate transactions, for miners to compete in the computational lottery that determines who

adds the next block. Ten minutes is not a design choice; it is a physical constraint imposed by the speed of electronic switching and the latency of global communication.

Ethereum improved on this but still measures confirmation times in seconds—and even then, true finality requires multiple block confirmations that extend the effective settlement time to minutes. Layer-two solutions and roll-ups have pushed throughput higher, but they do so by moving computation off-chain, sacrificing some of the decentralization that blockchain was designed to provide. **The fundamental constraint remains: electronic consensus is slow because electronic computation is slow.**

These latencies make blockchain unsuitable for applications requiring real-time settlement. Stock trades execute in microseconds; blockchain cannot compete. Payment authorizations must complete before a customer leaves the counter; blockchain cannot deliver. **Micropayments could enable entirely new economic models—paying fractions of a cent for individual articles, songs, or API calls—***but the transaction costs and confirmation times of current blockchain make such payments economically irrational.* The technology that promised to revolutionize finance remains too slow for most financial applications.

Photonic switching operates at 150 femtoseconds—five thousand times faster than the electronic transistors in current optical networks, and millions of times faster than the processors securing today's blockchains. A femtosecond is to a second as a second is to 32 million years. **In the time it takes an electronic transistor to complete one switching cycle, a photonic switch can complete thousands.**

This speed translates directly into consensus capability. Block intervals can shrink from ten minutes to one second, then to 100 milliseconds, then to single-digit milliseconds. The roadmap extends to block times measured in microseconds—thousands of blocks per second, each cryptographically secured, each representing irreversible consensus. Throughput can scale from the millions of transactions per second achieved by current roll-up solutions to billions of transactions per second on native photonic infrastructure. The limitation shifts from computation to communication; settlement becomes faster than the physical propagation of light across continental distances.

Micropayments settle faster than light can cross a continent. A customer in New York can pay a merchant in Los Angeles, and the transaction achieves cryptographic finality before a light beam traveling the same distance would arrive. DeFi applications, IoT economies, and gaming platforms can run entirely on-chain with latencies imperceptible to human users. The blockchain becomes fast enough to serve as infrastructure rather than novelty.

Security improves alongside speed. Attackers attempting to reorganize the chain—to reverse transactions by building an alternative history—face an impossible race. By the time an attacker could marshal resources to attempt a rewrite, legitimate consensus has already advanced by thousands of blocks. The attack surface shrinks not through algorithmic complexity but through raw physical speed. The honest network simply outpaces any possible adversary.

The architecture we are developing implements these capabilities through several integrated innovations, each grounded in physics rather than mathematics alone.

Our cryptographic primitive—LightHash-FS™—provides a 150-femtosecond hash function that is inherently resistant to quantum attacks. Current blockchain security relies on mathematical problems that classical computers find difficult but that quantum computers could solve efficiently. The threat of quantum computing has loomed over cryptocurrency for years, with the recognition that sufficiently powerful quantum machines could break the cryptographic foundations of existing chains. LightHash-FS™addresses this threat not through post-quantum mathematical algorithms but through the physics of photonic switching itself. The hash function exploits properties of light that quantum computers cannot accelerate—interference patterns, polarization states, and temporal correlations that have no mathematical shortcut. Security becomes a property of physics rather than a bet on computational difficulty.

Fiber-optic delay lines replace energy-intensive proof-of-work with physics-based verification. In current blockchain systems, difficulty is an arbitrary parameter adjusted to maintain consistent block times regardless of how much computational power the network accumulates. More miners mean more energy consumption to achieve the same result—a treadmill of waste that serves no purpose beyond making attack expensive. The Light-Switch Ledger replaces this waste with physical constraints that achieve the same security properties without

the environmental cost. Difficulty becomes a function of fiber cavity length—the physical distance that light must travel before consensus can form. This distance is measurable by anyone with a spectrometer. Verification requires no trust in reported hashrates or computational claims; it requires only measurement of physical properties that cannot be falsified.

Tamper evidence emerges from the physics itself, not from software attestations t hat sophisticated attackers can circumvent. Opening a sealed photonic module shifts its optical resonances in ways that cannot be concealed. The precise frequencies at which light resonates within the device depend on physical dimensions measured in nanometers; any physical intrusion alters these dimensions and changes the resonance signature. Compromised nodes do not merely fail verification—they self-destruct their cryptographic material before extraction becomes possible. The device knows it has been tampered with because the laws of physics change when tampering occurs.

This is not just cryptography. **It is cryptophysics.** Difficulty is physics, not smoke. Security is **measurement**, not *trust*. Verification is **observation**, not *computation*.

The token architecture reflects this new foundation and creates economic structures impossible on electronic infrastructure.

Photonic Compute Credits™—PCC—serve as the base utility token of the ecosystem. Unlike cryptocurrencies with arbitrary issuance schedules determined by software parameters, PCC is backed by physical infrastructure: one PCC represents one megawatt-hour of renewable-powered photonic compute capacity. Customers purchase PCC to access the platform; the token is consumed as jobs execute, with consumption rates varying by workload type—AI inference, blockchain validation, data processing, storage. Each transaction carries an audit capsule recording actual energy consumed, carbon equivalent, water used (zero), and latency achieved. The customer pays for outcomes; the ledger tracks physics. Compute becomes a metered utility with the transparency of a power bill and the efficiency that only photonics can deliver.

The Clean Compute™ stablecoin serves a different function: stable value storage rather than utility consumption. While PCC fluctuates with demand for photonic compute—rising when capacity is scarce, falling when abundant—the Clean Compute stablecoin maintains a fixed peg to one megawatt-hour of renewable

energy. Holders can redeem tokens for actual electricity or equivalent compute capacity at any time.

This peg is not maintained through algorithmic mechanisms that can fail under stress, as we have seen with previous stablecoin collapses. It is maintained through physical reserves of energy generation capacity and computing infrastructure. Enterprises use Clean Compute stablecoins for treasury management, long-term compute contracts, and cross-border settlement where price stability matters more than exposure to compute demand. PCC is what you spend; Clean Compute is what you save.

Staking in the Light-Switch Ledger™ operates on two tiers. Core validators commit physical infrastructure—fiber-ring segments that can be measured and verified through optical resonance. This is **proof-of-infrastructure**: validators must own and operate real photonic equipment to produce blocks and achieve consensus finality. Slashing for core validators is triggered by measurable drift in resonance integrity; the physics cannot lie even if the operator does. Community validators stake PCC tokens to participate in block attestation and network governance. They verify the work of core validators and vote on protocol upgrades, earning rewards proportional to their stake. The two tiers reinforce each other: infrastructure operators provide the physical foundation; token stakers provide the distributed oversight. The network is secured by physics at the core and community at the edge.

Even bandwidth becomes an asset class. Sub-nanosecond photonic time slots—windows of exclusive access to network capacity measured in billionths of a second—become tradeable instruments. High-frequency traders can purchase guaranteed latency for financial settlement. Global banks can reserve capacity for cross-border transactions. The network's capacity becomes a market, with prices reflecting actual demand for actual physical resources rather than speculative dynamics divorced from utility.

The implications extend beyond cryptocurrency speculation to the architecture of trust itself. A ledger that consumes no water, generates no heat, and operates at femtosecond speeds can be deployed in urban cores without the community opposition that has blocked electronic data centers. The cooling towers and industrial aesthetics that make data centers unwelcome neighbors simply do not exist in photonic facilities. A Light-Switch Ledger node can operate in a downtown office

building, drawing power from rooftop solar, invisible and inaudible to surrounding businesses and residents.

The technology can run in environments where electronic infrastructure cannot survive. Desert installations operate without the water that electronic cooling requires. Arctic facilities operate without the heating that prevents electronic components from failing in extreme cold—photonic devices are indifferent to temperature extremes that would destroy transistors. Orbital platforms become feasible; the vacuum of space eliminates convective cooling, making electronic data centers impossible, but photonic devices that generate no heat face no such constraint. The blockchain extends to environments that current technology cannot reach.

Money itself shifts from energy-burning proof to energy-backed trust. The value stored in the ledger corresponds to physical infrastructure that exists in the world—fiber-optic networks, photonic switching capacity, renewable energy generation. The abstraction of cryptocurrency reconnects to physical reality. Trust emerges not from the waste of energy but from the measurement of light.

Picture the skyline of Atlanta, New York, Los Angeles, New Orleans, or Washington at night. Not just glass and steel—but photonic towers glowing with light. Inside them, trust and compute run silently at femtosecond speed. No cooling towers sending plumes of steam into the urban sky. No smokestacks marking industrial processes. No water lines draining municipal supplies. Just photons racing through logic gates, securing value and intelligence at the heart of the city. This is not just infrastructure. It is sovereign-grade trust, embedded directly into the urban core.

The electronic age of hashing is closing. The photonic age of trust is rising—from the skylines of America to the networks of the world.

• • •

The internet was built for a world that no longer exists.
When the protocols that govern global communication were designed in the 1970s and 1980s, computation was scarce and expensive while bandwidth was the bottleneck. A mainframe computer cost millions of dollars. A megabyte of storage cost thousands. Network links measured capacity in kilobits per second—thousands of times slower than the connections we now consider

unacceptably sluggish. The engineers who designed TCP/IP, the fundamental protocol suite of the internet, made rational choices given these constraints.

Concentrate computation in central facilities where expensive machines can be shared. Use networks to move data to those facilities for processing. Design the network itself to be simple and stateless—a dumb pipe that forwards packets without understanding or operating on their contents.

This architecture made sense when the limiting factor was processing power. It makes no sense when computation is cheap and bandwidth is precious, when the constraint is not the cost of processing but the time and energy required to move data across distances. Yet the internet still operates on protocols designed for the opposite conditions, like a city whose roads were laid out for horse-drawn carriages and never redesigned for automobiles.

Consider what happens when you ask a question of an AI assistant. Your query travels from your device to a local router, through your internet service provider, across multiple autonomous systems operated by different companies, through undersea cables or continental fiber networks, to a data center that might be thousands of miles away. At the data center, your query enters a processing queue, waits for resources, undergoes computation on servers that generate enormous heat and consume enormous power, and produces a response. That response then travels the entire path in reverse, traversing the same networks, the same routers, the same cables, before finally reaching your screen.

Every step of this journey consumes time measured in milliseconds. Every step consumes energy. The data arrives at the distant data center unchanged from when it departed your device—the network has done nothing to it except move it. The network is a passive transport layer, adding latency and consuming energy without contributing to the actual work of answering your question.

Photonic computing inverts this architecture entirely.

TCP/IP-L—our redesign for photonic networking—implements computation within the network itself. Because the Poovey Stack processes information using light, and because optical networks already transmit information using light, the boundary between computation and communication dissolves. Network nodes become computational resources rather than passive relay points.

The routers that forward your traffic can simultaneously process it.

A signal traveling between two points need not remain unchanged during transit. Intermediate nodes can perform processing operations on the optical signal itself—filtering, pattern matching, preliminary inference, encryption, compression, transformation—without converting the signal from optical to electrical form and back again. The conversion between light and electricity is where electronic networks waste time and energy; optical-electrical-optical conversion at each network hop adds latency and consumes power. Photonic networks eliminate this conversion entirely. The signal remains light throughout its journey, and the nodes it passes through can operate on that light directly.

The data that arrives at a destination is not raw information requiring full processing but partially computed results requiring only completion. If your query to an AI assistant passes through ten network hops, each hop can perform a portion of the processing. By the time the signal reaches the data center, much of the work is already done. The data center completes the computation rather than starting from scratch. Response time shrinks not just because photonic switching is faster than electronic switching, but because the network itself has been doing useful work during transit.

The latency implications are profound and enable applications that current infrastructure cannot support.

When data travels through an electronic network, it moves passively. Routers examine packet headers to determine where to send the data next, but they do not examine or operate on the actual content. The data traverses each hop unchanged, and milliseconds accumulate with each hop. A packet crossing ten routers adds ten sets of processing delays, ten sets of queuing delays, ten opportunities for congestion to slow transit further. For a query traveling from New York to Los Angeles, round-trip latency typically exceeds 60 milliseconds even under optimal conditions—and real-world conditions are rarely optimal. For many applications, 60 milliseconds is imperceptible. A web page that loads in 60 milliseconds feels instantaneous to a human user. But for applications requiring genuine real-time response, those milliseconds are the difference between viable and impossible.

Consider remote surgery. A surgeon in Boston operating a robotic system on

a patient in rural Montana needs visual feedback that accurately represents the current state of the surgical field. If the video feed is 60 milliseconds old, the surgeon's hands are operating on anatomy that has moved since the image was captured. For delicate procedures—neurosurgery, ophthalmic surgery, microsurgical repair of tiny blood vessels—60 milliseconds of lag can mean the difference between success and catastrophic damage. Current telesurgery systems work only over short distances, with dedicated low-latency connections, limiting the technology to major medical centers that can afford specialized infrastructure. Patients in rural areas, in developing nations, in conflict zones cannot access surgical expertise that exists elsewhere in the world.

Consider autonomous vehicles. A self-driving car processing sensor data in a remote data center faces the same latency challenge. If the car detects an obstacle and sends the data to a distant server for processing, the response arrives 60 milliseconds later—during which time a car traveling at highway speed has moved nearly six feet. Six feet can be the difference between a safe stop and a collision. Current autonomous vehicle architectures address this by processing data locally, on the vehicle itself, but this limits the sophistication of the AI that can be deployed. A vehicle cannot carry the computing resources of a data center. The AI must be simplified to fit the available hardware, accepting reduced capability as the price of acceptable latency.

Consider augmented reality. A heads-up display that overlays digital information on the physical world must update that overlay in synchronization with head movement. If the overlay lags behind head motion, users experience disorientation, nausea, and visual artifacts that make the system unusable. Current AR systems struggle with this requirement, limiting applications to relatively static overlays that do not require real-time environmental understanding. The metaverse that technology companies have promised remains unrealized in part because the network infrastructure cannot support the latency requirements of immersive experience.

Consider financial trading. Markets move in microseconds. A trade executed 60 milliseconds late might as well not be executed at all; the opportunity has passed, the price has moved, the counterparty has disappeared. High-frequency traders spend billions of dollars on private fiber networks, microwave transmission systems, and colocation facilities that shave milliseconds from their transactions. The entire geography of financial infrastructure—the placement of data

centers, the routing of cables, the location of exchanges—is shaped by the race for lower latency. Those who cannot afford this infrastructure are systematically disadvantaged; they trade on stale information against counterparties who see the market as it exists now rather than as it existed moments ago.

A photonic mesh processes data during transit and transforms all of these applications.

A signal traveling from New Orleans to Atlanta need not arrive in Atlanta unchanged. Intermediate nodes—whether in converted office towers, in billboard-mounted repeaters along the highway, or in purpose-built network facilities—can perform operations on the optical signal without conversion delays. Filtering removes noise and irrelevant data before it consumes bandwidth. Pattern matching identifies relevant features that downstream processing needs to know about. Preliminary inference runs early stages of AI models, so that final stages require less computation. Encryption and decryption happen in transit rather than at endpoints. Compression reduces bandwidth requirements dynamically based on network conditions.

The signal that arrives in Atlanta is not raw data requiring full processing but partially processed information requiring only completion. The latency reduction compounds: the signal travels faster through optical media than electrical (light in fiber approaches the speed of light in vacuum, while electronic signals propagate at a fraction of that speed), and the processing time at the destination shrinks because computation occurred en route. Round-trip latency from New York to Los Angeles could shrink from 60 milliseconds to single-digit milliseconds—fast enough for remote surgery, fast enough for truly autonomous vehicles, fast enough for immersive augmented reality, fast enough for financial trading without geographic advantage.

That last point deserves emphasis. Today, financial firms pay billions to shave milliseconds by locating servers in New Jersey data centers adjacent to exchange matching engines. **Geography is destiny;** Manhattan and Chicago command financial infrastructure because latency demands proximity. A photonic mesh that delivers single-digit millisecond latency nationwide **erases that advantage.**

A smart governor in Mississippi or New Mexico or Montana could build a financial services hub that competes with Wall Street on equal footing—not

through tax incentives or regulatory arbitrage, but through physics. The infrastructure that currently concentrates wealth in coastal cities could distribute opportunity to communities that have never had access to high-speed financial markets. **Latency democracy.**

The infrastructure for this network already exists in embryonic form, waiting to be activated.

The United States has approximately 350,000 billboard structures. Many are positioned along highways and arterial roads that also carry significant data traffic—the same corridors where fiber-optic cables run, where wireless towers cluster, where the physical infrastructure of the internet concentrates. These billboard structures already have power connections; digital displays require electricity. Many already have fiber connectivity for updating those digital displays. They have physical presence at regular intervals along transportation corridors, providing coverage density that purpose-built infrastructure would require years and billions of dollars to replicate.

Our patent portfolio includes technology for photonic repeaters scaled to fit these existing structures—devices roughly the size of a cable box that regenerate optical signals across distances while performing computational operations on data in transit. A billboard-mounted repeater receives an optical signal, performs specified processing operations on the light itself, amplifies the signal to compensate for transmission losses, and retransmits it toward the next node. These devices attach to the back of the billboard structure, draw minimal power, and require no modification to the billboard's advertising function. The advertising continues unchanged; the photonic network rides along invisibly.

The economics are transformative. We are not buying advertising space; we are leasing attachment rights to unused structural real estate. We estimate attachment fees at less than one hundred dollars per month—pure incremental revenue for billboard owners who lose nothing and gain a small but steady income stream from the back surface of their structures. Compare this to the cost of building dedicated network infrastructure: a single cellular tower costs between one hundred fifty thousand and three hundred thousand dollars to construct, requires years of permitting and environmental review, and often faces community opposition. A photonic repeater on an existing billboard costs a fraction of that, requires minimal permitting because the structure already exists with

appropriate zoning, and generates no community concern because it changes nothing visible about the billboard.

Scale the comparison nationally. Three hundred fifty thousand billboards at one hundred dollars per month represents thirty-five million dollars annually for complete national mesh coverage—less than the cost of building two hundred conventional cell towers. The existing billboard infrastructure becomes the backbone of a photonic network without requiring a single new structure to be permitted, constructed, or defended against community opposition. Deployment that would take a decade through conventional infrastructure development can proceed in months.

Combined with SkyLight towers serving as regional computational hubs, this architecture creates what we call the **SkyLight National Photonic Mesh ™**—abbreviated **SNPM** in our planning documents. This is a network where computation and communication blur together, where data processing happens throughout the network rather than only at designated facilities, and where the physical infrastructure leverages existing assets rather than requiring greenfield construction.

The mesh addresses limitations that constrain every current cloud provider. Latency that makes real-time applications difficult becomes latency that enables them. Bandwidth constraints that throttle data movement relax as compression happens in transit. Energy consumption that scales linearly with data volume becomes energy consumption that scales sublinearly as processing distributes across the network. Concentration of critical infrastructure in vulnerable mega-facilities gives way to distribution across hundreds of smaller nodes, each individually less critical, collectively more resilient.

The network becomes the computer. The dumb pipe becomes an intelligent fabric. The internet transforms from a passive transport layer into an active computational layer that adds value at every hop.

This is not an upgrade to the existing internet. It is a replacement for what the internet can be—a recognition that the assumptions underlying current network architecture no longer hold, and that a fundamental redesign based on photonic computation can unlock capabilities that incremental improvement of electronic systems will never achieve.

Golden Dome needs timing that electronics cannot provide.

The phrase "**Golden Dome**" has entered public discussion as shorthand for the evolving architecture of integrated air and missile defense—the systems that must detect, track, discriminate, and intercept threats ranging from legacy ballistic missiles to hypersonic glide vehicles to autonomous drone swarms. The technical requirements of these systems have escalated beyond what electronic infrastructure can deliver, creating capability gaps that adversaries are actively working to exploit.

The core challenge is sensor fusion. Modern air and missile defense does not rely on any single sensor. Radar systems detect objects at great distances but struggle to characterize them precisely. Infrared sensors can identify the heat signatures of missiles and aircraft but have limited range and can be fooled by countermeasures. Space-based sensors provide broad coverage but face their own limitations of resolution and revisit time. Electronic intelligence intercepts communications and control signals but requires correlation with other data to be useful. Each sensor type provides partial information; effective defense requires combining these partial views into a coherent operational picture.

The quality of that combined picture depends critically on timing. Sensor data must be synchronized with precision measured in nanoseconds. A radar return from an object must be correlated with an infrared detection of the same object, and that correlation depends on knowing exactly when each sensor made its observation. If timing synchronization drifts by microseconds, correlation fails. Objects appear to duplicate; real targets merge with false ones; the operational picture degrades from clarity into confusion.

Electronic systems have pushed timing synchronization about as far as physics allows. GPS-disciplined clocks provide nanosecond-level accuracy. Atomic references improve on this further. Network protocols synchronize distributed systems with impressive precision. But the fundamental limits of electronic switching constrain how quickly systems can process and respond to timing information. A processor that switches in nanoseconds cannot fully exploit timing information resolved in picoseconds; the processing itself introduces uncertainty that blurs the precision of the underlying data.

Photonic systems measure time in femtoseconds—a thousand times finer

resolution than nanoseconds. This precision translates directly into sensor fusion capability. Track correlation becomes cleaner because timing windows become tighter. Spoofed sensor feeds—false data injected by adversaries to confuse the picture—fail time-and-spectrum checks that electronic systems cannot perform with sufficient precision. The ability to distinguish real threats from decoys, debris, and deliberate deception improves because the physics of discrimination operates at timescales electronic processing cannot match.

The command-and-control systems that coordinate defensive responses face parallel challenges. Modern C2 networks must be resilient against cyber attack, because adversaries understand that degrading command systems is more efficient than defeating weapons systems directly. Electronic C2 systems rely on cryptographic security that sophisticated nation-state attackers can potentially compromise. Software attestation can be fooled by supply chain attacks that insert malicious code before deployment. The integrity of the entire defensive system depends on the integrity of networks that face continuous, sophisticated assault.

Photonic infrastructure provides security properties that electronic systems cannot match. Hardware-bound identity ensures that devices are what they claim to be—the optical properties of a photonic module constitute an unforgeable fingerprint that cannot be cloned or spoofed. Continuous attestation verifies that systems remain uncompromised throughout operation; any physical tampering changes measurable optical properties and triggers automatic response. Optical data diodes provide one-way information flow that even the most sophisticated cyber attack cannot reverse; data can flow out of a secure enclave for logging and audit, but no data can flow in through the diode regardless of what software vulnerabilities might exist. These properties do not depend on algorithms that might be broken or implementations that might be flawed; they depend on **physics.**

The SkyLight National Photonic Mesh provides these capabilities as a ground layer for national defense infrastructure—faster to deploy than space-based assets, harder to spoof than electronic systems, easier to integrate into existing command structures than entirely new architectures. A forty-three-story tower in downtown Atlanta or a converted hotel in Washington can host SCIF-grade photonic floors close to decision-makers and fiber meet-me rooms, with sovereign custody of keys, compilers, and audit trails maintained entirely under U.S. control.

The Field-Programmable **Photonic** Array—FP**P**A—extends these capabilities to mission-specific applications. Traditional defense systems require purpose-built hardware for each function. Developing and deploying new hardware takes years, creating a lag between emerging threats and defensive capabilities. By the time a system designed to counter a specific threat reaches operational deployment, adversaries have often developed new threats that the system was not designed to address.

FPPA provides a reconfigurable photonic fabric that can be reprogrammed for different mission requirements without hardware changes. The same physical infrastructure that performs multi-sensor track fusion in the morning can perform adaptive signal filtering for electronic warfare in the afternoon and run post-quantum cryptographic protocols for secure communication in the evening. Mission teams can retarget computational kernels in hours rather than the years required to develop and deploy new electronic systems. Experimentation becomes safe; new approaches can be tested on reconfigurable hardware without commitment to fixed-function systems. When threats evolve, responses can evolve at software speed rather than hardware speed.

The applications are illustrative rather than exhaustive, and the details remain sensitive. Hypersonic vehicle discrimination benefits from track fusion with femtosecond timing windows and maneuver scoring that electronic processors cannot compute quickly enough to be operationally relevant. Electronic warfare benefits from adaptive filtering profiles that can be tuned in the field as adversary systems evolve. Signals intelligence benefits from spectral and temporal feature extraction that reduces the volume of data requiring human analysis. Secure communication benefits from line-rate post-quantum cryptography that hardens links without sacrificing throughput.

I am not a defense contractor, and True Photonic is not positioning itself to become one. But the capabilities that emerge from photonic computing have obvious implications for national security, and we would be naive to ignore them. The companies that dominate defense technology—Lockheed Martin, Raytheon, Northrop Grumman, General Dynamics—have the security clearances, the procurement relationships, and the program management expertise that a technology startup cannot replicate. Our role is to provide the underlying platform—**the physics**, the manufacturing relationships, the operational expertise—that makes new capabilities possible. The integration of those capabilities

into defense systems is work for established primes, in partnership with government programs that can navigate classification and security requirements.

The broader strategic point extends beyond any specific weapon system or defense program. A nation that controls femtosecond-class computational infrastructure—that can process information faster than any adversary, that can fuse sensor data with precision adversaries cannot match, that can secure communications through physics rather than just mathematics—occupies a position of technological advantage comparable to historical advantages in other domains. Air superiority transformed warfare in the twentieth century. Information superiority through photonic infrastructure could prove equally decisive in the twenty-first.

Unlike current chip manufacturing, photonic computing does not require the concentrated fabrication capability in Taiwan and South Korea that makes today's semiconductor supply chain a geopolitical vulnerability. The Relaxed Lithography approach described in Chapter 11 means that photonic devices can be manufactured using mature processes available at dozens of facilities in the United States, Europe, Japan, and other allied nations. The strategic advantage, if achieved, depends on technology and facilities that can be built, maintained, and expanded on sovereign territory—immune to supply chain disruptions, blockades, or conflicts that could paralyze electronic alternatives.

• • •

The applications described so far operate at the scale of nations and networks—systems that reshape global finance, transform the internet, and strengthen national defense. But the same physics that enables planetary-scale transformation also enables interventions at the most intimate scale imaginable: inside the human body.

The heart generates electrical signals that can be read from the skin's surface. This fact has been known for over a century; the electrocardiogram was invented in 1903 and has remained a cornerstone of cardiac diagnosis ever since. What has changed is what can be done with those signals once captured. Modern artificial intelligence can detect patterns in cardiac electrical activity that human cardiologists cannot perceive—subtle variations in waveform morphology invisible to the naked eye, timing relationships between beats that emerge only in statistical analysis, trends that develop over hours or days and predict events that have not yet occurred.

The limiting factor is no longer algorithmic capability; machine learning models can identify cardiac abnormalities with accuracy matching or exceeding trained cardiologists. The limiting factor is the hardware that runs these algorithms. Continuous AI analysis requires continuous computation. Continuous computation in electronic devices requires continuous power consumption and generates continuous heat.

Current wearable cardiac monitors face a fundamental constraint defined by the physics of electronic processing. A processor running sophisticated neural networks continuously will exhaust its battery in hours; the same lithium cells that power a smartphone for a day will power a neural network for a fraction of that time. The processor will also become warm—potentially uncomfortably warm when worn against skin, and certainly warm enough to reduce battery efficiency and component lifespan. Manufacturers therefore compromise.

One compromise runs simplified algorithms—rule-based detectors that look for gross abnormalities like extremely fast heart rates or completely irregular rhythms. These algorithms catch obvious problems while missing subtle patterns that more sophisticated analysis would detect. A patient developing atrial fibrillation might show warning signs in the fine structure of their heartbeat hours or days before the arrhythmia fully manifests; a simplified algorithm will miss those signs entirely.

Another compromise runs sophisticated algorithms intermittently rather than continuously. The device sleeps most of the time, waking periodically to analyze a few seconds of cardiac activity, then sleeping again to conserve power. This approach misses transient events that occur between analysis windows. Atrial fibrillation often appears in brief episodes—a minute here, thirty seconds there—and then resolves without symptoms the patient would notice. An intermittent monitor might happen to wake during an episode, or might happen to sleep through every one.

The clinical consequences of these compromises are measured in human lives. Atrial fibrillation is the most common serious heart rhythm disorder, affecting over thirty million people worldwide. Its primary danger is stroke; the irregular rhythm allows blood to pool and clot in the heart, and those clots can travel to the brain with devastating consequences. Approximately one in four strokes has no identified cause at initial evaluation; many of these are later attributed to atrial

fibrillation that was never diagnosed because it occurred in brief episodes that conventional monitoring missed.

Earlier detection enables prevention. Patients with known atrial fibrillation can take anticoagulant medications that dramatically reduce stroke risk. But patients whose atrial fibrillation remains undiagnosed cannot receive this protection. They walk around with a treatable condition, unaware, until a stroke announces the diagnosis in the most harmful possible way.

CardioLT—a True Photonic Ventures spinout developing what we call Cardio-in-Light technology—eliminates the compromise between computational sophistication and operational endurance.

A photonic processor generates negligible heat because photonic switching generates negligible heat. The same physics that eliminates cooling towers from data centers eliminates thermal concerns from wearable devices. The processor can run sophisticated neural networks continuously without becoming warm, without degrading battery life through heat-induced inefficiency, without creating discomfort for the wearer.

The same power efficiency that enables ninety percent reduction in data center energy consumption enables fourteen-day battery life in a wearable cardiac monitor measuring thirty-five by twenty by two millimeters. Comparable electronic devices running comparable algorithms last twenty-four to forty-eight hours before requiring recharge or replacement. The photonic device extends monitoring duration by an order of magnitude while simultaneously increasing algorithmic sophistication.

The patient wears a device smaller than a standard adhesive bandage. The device monitors every heartbeat, continuously, with AI sophistication previously available only in hospital intensive care units. Multiple sensing channels capture cardiac electrical activity from different vectors, providing three-dimensional information about electrical propagation through the heart. A photonic processor analyzes each beat in real time, comparing patterns against models trained on millions of annotated heartbeats, identifying not just arrhythmias that have occurred but patterns predictive of arrhythmias likely to occur.

The device communicates wirelessly with healthcare providers. When it detects

concerning patterns, it transmits alerts and diagnostic data without requiring patient action. The patient need not notice symptoms; the device notices patterns the patient cannot feel. The cardiologist receives information that would previously have required hospitalization to obtain—continuous, comprehensive cardiac monitoring delivered to their dashboard while the patient goes about daily life.

The regulatory pathway for CardioLT is 510(k) clearance—the FDA's process for devices that are substantially equivalent to existing approved products. Cardiac monitors exist and have been cleared through this pathway many times. CardioLT is a better cardiac monitor; it does what cardiac monitors do, only with longer battery life and more sophisticated analysis. The regulatory path is well-understood, typically requiring twelve to eighteen months from submission to clearance for devices with clear predicate equivalents.

CardioLT is pursuing seed funding to reach prototype completion and clinical validation, with 510(k) submission projected within twenty-four months. The market for ambulatory cardiac monitoring exceeds three billion dollars annually in the United States alone, with growth driven by aging populations and increasing awareness of atrial fibrillation's role in stroke.

• • •

Inside the body, the constraints become more severe and the opportunities more profound.

Implantable cardiac devices—pacemakers that regulate heart rhythm, defibrillators that shock hearts out of dangerous arrhythmias—have saved millions of lives since their introduction in the 1960s. They are also limited by constraints that have persisted for decades despite continuous engineering effort by the companies that dominate the market: Medtronic, Abbott, Boston Scientific. These companies have refined their devices through generations of incremental improvement, but fundamental limitations remain because the fundamental technology remains electronic.

Battery life determines how often patients must undergo surgery to replace depleted devices. Current pacemakers last seven to ten years; defibrillators, which maintain larger batteries for high-energy shocks, last five to seven years. A patient implanted at age sixty will face three or four replacement surgeries over

a normal lifespan—each carrying risks of infection, bleeding, lead dislodgement, and the complications that attend any procedure requiring anesthesia in aging patients. Device replacement is not a minor inconvenience. It is a significant medical procedure with measurable mortality risk. Patients dread it, physicians accept it as unavoidable, and the healthcare system absorbs billions of dollars in costs that serve only to restore functionality that was already present before the battery died.

Heat generation limits computational sophistication. An implanted device that becomes even slightly warm damages surrounding tissue—the heart muscle that the device is meant to protect. A temperature increase of a few degrees, sustained over time, causes cell death and scarring. Current devices therefore run the simplest algorithms that can achieve acceptable performance—rule-based discrimination between dangerous and benign rhythms that has not fundamentally changed in design philosophy for decades.

These algorithms examine features like heart rate and rhythm regularity, applying thresholds established through clinical experience. If the heart beats faster than a certain rate, it might be ventricular tachycardia, a dangerous rhythm requiring intervention. If the rhythm is completely irregular, it might be ventricular fibrillation, requiring immediate shock. The algorithms work well for clear-cut cases that obviously exceed thresholds. They struggle with ambiguous rhythms that fall near the boundaries—rhythms that require more sophisticated pattern recognition than simple threshold comparison.

The result is a persistent and deeply distressing problem: inappropriate shocks. Ten to fifteen percent of defibrillator discharges are triggered by rhythms that were not actually dangerous. The device mistakes a rapid but benign rhythm—perhaps caused by exercise, or by a supraventricular arrhythmia that does not threaten life—for a lethal ventricular arrhythmia, and delivers a shock designed to stop ventricular fibrillation.

These shocks are painful. Patients describe them as being kicked in the chest by a horse, or struck by lightning. They are also traumatic beyond the physical pain. The patient experiences sudden, unexpected violence from a device inside their own body. They lose trust in the device that is supposed to protect them. They live with the knowledge that at any moment, without warning, they might receive another shock for a rhythm that posed no actual danger.

Each inappropriate shock depletes the battery, accelerating the timeline to surgical replacement. Some patients develop anxiety disorders related to their devices, fearing the next random shock, avoiding activities that might trigger rapid heart rates and device response. Some patients request removal despite the life-threatening conditions the devices were implanted to treat. They would rather risk sudden cardiac death than continue living with a device they cannot trust. The metal leads that connect implanted devices to heart tissue create their own category of complications. These leads are thin wires that travel through veins into the heart, where they attach to cardiac muscle. They must flex with every heartbeat—tens of millions of flex cycles over a device's lifetime. Mechanical fatigue causes some leads to fracture. A fractured lead can deliver shocks erratically, or fail to deliver shocks when needed, or cause electrical artifacts that the device misinterprets as dangerous rhythms requiring intervention.

Leads can become infected. Bacteria colonize the lead surface, forming biofilms that antibiotics cannot penetrate. Lead infection is a serious, potentially fatal complication that sometimes requires complete removal of the lead and device—a dangerous procedure involving specialized tools to free leads that have grown into heart tissue over years of implantation. Sometimes percutaneous extraction fails and the patient requires open-heart surgery to remove an infected lead.

Leads create problems in the powerful magnetic fields of MRI scanners. Electrical current flows in response to changing magnetic fields; this is basic physics. A lead in an MRI scanner acts as an antenna, picking up energy from the scanner's magnetic pulses and converting it to electrical current. That current heats the lead, potentially damaging the heart tissue where the lead attaches. Most patients with cardiac devices face restrictions on MRI access despite the fact that three-quarters will need MRI imaging for unrelated conditions during their device's lifetime. They must accept inferior diagnostic alternatives or forgo imaging entirely.

PulseLT—developing what we call Pulse-in-Light technology for implantable cardiac therapeutics—addresses each of these constraints through photonic architecture.

A processor generating effectively zero heat can run sophisticated AI continuously without tissue damage. Deep learning models trained on millions of

cardiac rhythm examples can distinguish dangerous ventricular arrhythmias from benign rapid rhythms with accuracy far exceeding simple threshold comparisons. The AI learns patient-specific patterns, adapting to each individual's cardiac characteristics rather than applying population-derived thresholds that inevitably misclassify edge cases. Inappropriate shocks could be reduced by ninety percent or more—not through incremental algorithm improvement but through a fundamental change in what algorithms can be deployed inside the body.

Power consumption one-hundredth of electronic equivalents enables battery life exceeding twenty-five years—potentially the patient's lifetime. The calculation is straightforward: if a pacemaker's processor consumes one hundred microwatts rather than ten milliwatts, the battery that currently lasts seven years lasts much longer, limited then by battery chemistry and self-discharge rather than drain rate. Conservative engineering that accounts for self-discharge, defibrillation energy requirements, and safety margins still yields practical battery life of twenty-five to thirty years. A patient implanted at age sixty-five might never need replacement surgery. The device lasts until the patient does.

Optical fiber leads—thin, flexible glass fibers that carry light rather than electricity—eliminate the MRI heating problem entirely. Light does not interact with magnetic fields the way electrical current does. A patient with optical leads can undergo MRI scanning without restriction, without special protocols, without risk of tissue heating. The leads are more flexible than metal wires, potentially reducing the mechanical stress that causes lead fracture. Their biocompatibility may reduce infection risk. Their smaller diameter may simplify extraction if removal becomes necessary.

Optical sensing provides information beyond electrical signals. Fiber-optic sensors can measure blood oxygen saturation, tissue perfusion, and hemodynamic parameters that current electronic leads cannot detect. The device gains visibility into cardiac function that goes beyond rhythm monitoring to include actual pump performance and tissue health.

The regulatory pathway for PulseLT is more demanding than for CardioLT. Implantable cardiac devices are Class III medical devices requiring Premarket Approval—the FDA's most rigorous pathway, involving clinical trials demonstrating safety and efficacy in human patients. The timeline is measured in years,

not months. The investment required is substantial—tens of millions of dollars for the clinical trials alone. The pathway is navigable; pacemakers and defibrillators are approved regularly. But it requires sustained commitment and regulatory expertise.

PulseLT is pursuing seed funding to reach working prototype and large-animal proof-of-concept, positioning for subsequent funding rounds that would support the pivotal clinical trials required for PMA submission.

The market, if the technology succeeds, is substantial. Global sales of implantable cardiac devices exceed fifty billion dollars annually, dominated by established players who have not faced a fundamental platform challenge in decades. A device offering twenty-five-year battery life, near-zero inappropriate shocks, and unrestricted MRI access would command premium pricing and rapid adoption—not as an incremental improvement over existing products but as a categorical advance that renders current technology obsolete. The path is long. The investment required is significant. But the destination, if reached, transforms a major category of medical technology and eliminates suffering that current devices cause while trying to prevent it.

• • •

The **Photonic Shape-Shifting Muscle patent** extends the platform beyond sensing and processing to actuation: **making things move.**

The technology uses light-activated piezoelectric elements to achieve mechanical response at femtosecond timescales—a million times faster than the fastest electronic motor controllers. This speed might seem academic; no mechanical system needs to respond in femtoseconds. Mechanical components have mass and inertia that limit how fast they can move regardless of how quickly they are commanded. A robotic arm weighing kilograms cannot accelerate instantaneously no matter how fast its controller operates.

But the speed of the controller translates into control bandwidth—how rapidly a system can sense, decide, and correct. A robotic arm controlled by electronic processors can update its position perhaps a thousand times per second, limited by the electronic control loop's latency. Between updates, the arm moves open-loop, without feedback, accumulating errors that the next update must correct. A photonically controlled arm can update a million times per second.

Errors have no time to accumulate. Corrections happen before deviations become significant.

The difference is not visible in smooth, planned motion where the arm follows a predetermined path. It becomes apparent in response to disturbances, in fine manipulation, in tasks requiring precision at the edge of what mechanics allow.

Consider a surgical robot positioning a scalpel. The surgeon's hand trembles slightly—an unavoidable consequence of human physiology. Hand tremor typically oscillates at frequencies between eight and twelve hertz, far slower than electronic control loops, easily compensable in principle. But compensation is only as good as the control bandwidth allows. Electronic control systems sampling at a thousand hertz can filter some of this tremor, but high-frequency components of the tremor—the jitter and vibration that create a ragged rather than smooth cut—persist between samples. Photonic control, operating a thousand times faster, can cancel tremor more completely, enabling steadier cuts, smaller incisions, access to anatomy too delicate for current robotic surgery.

The difference could make operable conditions that are currently inoperable. Tumors millimeters from critical blood vessels or nerve bundles currently require human surgeons at the peak of their manual skill—and even those surgeons sometimes cannot operate safely. Photonic surgical robots could operate with precision exceeding human capability, not by replacing human judgment but by providing tools that translate human intent into mechanical action with fidelity that human hands cannot achieve.

Consider a prosthetic hand grasping an egg. The challenge is not strength but control—applying enough force to hold the egg securely, not so much as to crush it. Current electronic prosthetics achieve this through careful calibration and slow, deliberate motion. Users learn to move carefully because the hand cannot respond quickly enough to dynamic changes. They approach objects slowly, grasp gently, move without sudden accelerations that might cause grip force to spike and damage fragile items.

Photonic prosthetics could achieve stable grasping dynamically. As the egg's surface deforms under pressure, the prosthetic's sensors detect the deformation, the photonic controller computes the appropriate response, and the actuators

adjust grip force—all within microseconds, continuously, thousands of times while the user's hand moves naturally. The prosthetic responds to incipient slip before the egg begins to fall, adjusting grip force preemptively rather than reactively. The prosthesis would feel more natural to its user because it would behave more like a natural hand—reactive, adaptive, fast enough that the user need not think about control.

Consider an exoskeleton augmenting a worker's strength for heavy lifting or extended physical labor. Current exoskeletons add force but also add lag—a perceptible delay between the user's intention and the machine's response that makes the device feel clumsy, unnatural, exhausting to use for extended periods. Users report fighting their exoskeletons rather than being helped by them. The machine does not move when expected; it moves a moment later, out of phase with the user's motion, requiring constant compensation that itself consumes effort.

The delay exists because electronic controllers cannot keep pace with the human nervous system's response time—approximately fifty milliseconds from sensation to motor response for trained reflexes. The exoskeleton's controller sees the user begin to move, computes the appropriate assistance, and commands the actuators—but by the time the command executes, the user has already moved on to the next phase of motion. The assistance arrives late, out of synchronization with need. Photonic controllers, operating in nanoseconds rather than milliseconds, could respond faster than human nerves transmit signals. The exoskeleton could predict and assist motion before the user's own muscles fully engage. The device would feel transparent—augmenting strength without introducing perceptible lag, making the exoskeleton an extension of the body rather than a machine the body must accommodate.

The insight that unifies these applications is that photonic muscles require photonic brains. An actuator capable of femtosecond response, controlled by a processor with microsecond response, wastes most of its capability—like connecting a racing car to steering with multi-second delays. The speed that matters is the speed of the complete system, from sensor to processor to actuator and back. The Poovey Stack provides the control system that allows photonic actuators to achieve their theoretical performance. The combination creates robotic and prosthetic systems operating in regimes that electronic systems cannot reach regardless of how much engineering effort is applied to electronic control.

Industrial applications may prove as significant as medical ones. Manufacturing robots that position components with sub-micron accuracy could enable new fabrication processes—semiconductor manufacturing, precision optics, nanotechnology applications where current robots lack adequate precision. Inspection systems that physically manipulate parts at high speed could increase throughput by orders of magnitude in quality control applications. Material handling systems that respond faster than objects can fall could reduce damage and increase automation in domains currently requiring human judgment and reaction time—warehouse operations, agricultural harvesting, food processing where delicate items require careful handling.

• • •

True Photonic Ventures exists to pursue these verticals. Not all of them—no single company could pursue them all—but enough to demonstrate that photonic computing is not merely more efficient infrastructure. It is a new category of capability, and we intend to own as much of that category as the patent system allows.

The structure is designed to create and support vertical-specific companies. The parent company provides technology licenses, technical support, and access to the photonic platform. The ventures arm provides capital, business development expertise, and the infrastructure for building companies in specific markets. The resulting portfolio companies combine photonic capability with domain expertise, pursuing opportunities that neither a pure technology company nor a pure domain company could capture alone.

The capital formation strategy includes pursuit of Small Business Investment Company designation from the U.S. Small Business Administration. SBIC status would allow True Photonic Ventures to leverage private capital with government-guaranteed debt at favorable rates, amplifying investment capacity significantly. For every dollar of private capital raised, an SBIC can access up to two dollars of SBA-guaranteed leverage. The program has supported the formation of companies including Apple, Intel, Costco, and FedEx in their early stages; it remains a significant source of growth capital for emerging technology ventures with domestic focus.

The physical infrastructure strategy creates additional alignment between portfolio companies and the broader True Photonic ecosystem. Portfolio companies

become tenants in SkyLight facilities, locating their computational workloads in Clean Compute Centers built around photonic infrastructure.

CardioLT's AI training runs on SkyLight servers. The Light-Switch Ledger validates on SkyLight infrastructure. PulseLT's device development uses SkyLight computing resources. The arrangement reduces their computing costs—they access photonic efficiency at infrastructure pricing—while increasing SkyLight facility utilization. The ecosystem reinforces itself: more applications create demand for more infrastructure; more infrastructure enables more applications; the portfolio companies and the infrastructure company grow together. CardioLT, PulseLT, the Light-Switch Ledger, and the companies that will pursue Photonic Muscle applications represent the first wave of ventures built on the photonic platform.

They are not the last.

The characteristics that make photonic computing attractive for cardiac monitoring—low power, no heat, extended operation—make it attractive for any wearable or implantable device: neural interfaces, drug delivery systems, diagnostic sensors, therapeutic stimulators. The characteristics that make it attractive for blockchain—speed, efficiency, physics-based security—make it attractive for any application requiring distributed trust: supply chain verification, identity management, voting systems, intellectual property registration. The characteristics that make it attractive for robotic control—speed, precision, bandwidth—make it attractive for any application where electronic control is the limiting factor: autonomous vehicles, drone swarms, precision agriculture, space systems. The platform creates a surface for innovation that extends beyond any single company's ability to pursue.

The tenant relationship between portfolio companies and SkyLight facilities serves purposes beyond cost reduction. Companies developing on the photonic platform benefit from proximity to the infrastructure team—the engineers who understand the Poovey Stack's characteristics, who can optimize workloads for photonic execution, who can troubleshoot integration challenges that arise when novel applications meet novel hardware. The physical collocation in converted towers creates an innovation campus without the traditional campus's isolation from urban markets and talent pools. A medical device company in a SkyLight tower in Atlanta has access to Emory's medical research complex and Georgia

Tech's engineering talent while running its computational workloads on photonic infrastructure in the same building.

• • •

The applications in this chapter span from the planetary to the personal. A new architecture for money and trust that eliminates the environmental catastrophe of current blockchain. A new model for the internet that transforms passive transport into active computation. Sovereign infrastructure for national defense that provides capabilities electronic systems cannot match. Medical devices that operate inside the body for decades rather than years. Robotic systems that respond faster than human nerves can transmit signals.

They are unified by the platform that enables them: **the Poovey Stack's combination of femtosecond speed, negligible heat generation, extreme power efficiency, and operation in environments hostile to electronics.**

Each application exploits these characteristics in different ways, but all depend on the same underlying physics. The same switch that enables a sustainable blockchain enables a twenty-five-year pacemaker. The same processing capability that hardens missile defense enables surgery more precise than human hands can perform.

The honest acknowledgment that has characterized this book applies to these applications as to the infrastructure chapters. The Light-Switch Ledger is in development, not deployment. TCP/IP-L is a patent filing, not a shipping product. The defense applications require partnerships and clearances we do not yet possess. CardioLT and PulseLT are pursuing regulatory pathways, not holding FDA clearances. Photonic Muscle is a patent portfolio, not a product line. Each application requires years of engineering, regulatory navigation, market development, and—for the defense applications—policy evolution before it can deliver on its potential.

But the potential is grounded in physics that has been validated, not speculation about undiscovered phenomena. The femtosecond switching speeds that enable the Light-Switch Ledger have been measured at Technion and in True Photonic's laboratories. The power efficiency that enables twenty-five-year implant battery life has been demonstrated at the device level. The control bandwidth that enables surgical precision has been characterized in working

systems. The applications follow from the technology; the question is execution, not feasibility.

True Photonic Ventures exists to pursue that execution across multiple verticals simultaneously. The portfolio structure allows parallel development in markets with different timelines, different regulatory requirements, different success factors. Some ventures will succeed; others may not. The blockchain application might reach market before the medical applications clear regulatory approval; the medical applications might find commercial traction while the defense applications await policy evolution. But the platform that underlies them all—the Poovey Stack, the SkyLight infrastructure, the photonic computing capability—provides a foundation for continued innovation regardless of which specific applications reach market first.

The next generation of everything. *Not hyperbole.* **Aspiration.**

The final chapter examines the path from current status to this envisioned future: the timeline, the milestones, the obstacles that remain, and the reasons for believing that **2026 may indeed become the year of light.**

The year 2026 may or may not become the year of light. What happens next depends on execution

CHAPTER 14

The Light Ahead

The previous chapters have described a technology that exists, at least in validated prototype form, and that has continued to advance since initial testing. The fundamental switching physics were validated at Technion University in Israel, where pump-probe spectroscopy measured switching speeds in the 150-200 femtosecond range—three to four times faster than anything previously published in the scientific literature, and five thousand times faster than the electronic transistors that power contemporary computing. Since that validation, the technology has continued to evolve. The original switch composition has been reformulated to eliminate lead, addressing both environmental concerns and regulatory requirements for commercial deployment.

The graphene-boron nitride-sapphire substrate developed with Astera Energy represents a further advancement, offering thermal conductivity exceeding 2,000 watts per meter-kelvin and enabling the dense three-dimensional architectures the Poovey Stack requires. The Relaxed Lithography approach offers a manufacturing pathway that bypasses the most constrained chokepoints in the global semiconductor supply chain.

The question that remains is not whether the underlying physics works in a laboratory setting, but whether the complete system can be brought to market, at what pace, and against what obstacles.

This chapter attempts to answer those questions honestly. The path from validated device physics to commercial deployment is never simple, and photonic computing faces challenges both technical and commercial. But there are also reasons to believe that the convergence of several factors—technology readiness, infrastructure crisis, market demand, and geopolitical pressure—may compress the timeline that new computing paradigms historically require. The proposition that 2026 could become the year of light is not a marketing slogan but an assessment based on the specific circumstances the industry now faces.
The stakes of getting this assessment right extend beyond any single company's commercial prospects. The computing industry has reached a point where the

next several years will determine whether artificial intelligence continues on its current trajectory of capability expansion or encounters constraints that force retrenchment. The hyperscalers— Microsoft, Google, Amazon, Meta—are each projecting AI infrastructure spending measured in hundreds of billions of dollars through the end of the decade. These projections assume that infrastructure can be built, that communities will permit it, that power grids can supply it, and that the fundamental physics of electronic computing can be extended far enough to justify the investment.

Each of these assumptions is now questionable in ways that were not evident even two years ago.

• • •

The path from current status to commercial deployment differs fundamentally from traditional semiconductor development in one critical respect: it proceeds on multiple tracks simultaneously rather than sequentially. This parallel structure is not accidental but reflects a deliberate assessment by experienced professionals who have evaluated the technology, the market, and the execution requirements, and who have concluded that concurrent development across the core platform, infrastructure deployment, and vertical applications is not merely possible but necessary to capture the opportunity the current moment presents.

Traditional deep technology development follows a waterfall model: prove the science, then develop the process, then build the factory, then find customers, then expand to adjacent markets. Each phase waits for the previous phase to complete. This model made sense when capital was scarce, when manufacturing required purpose-built facilities, and when market demand was uncertain.

None of those conditions apply here.

The Relaxed Lithography approach means manufacturing capacity already exists at mature foundries worldwide. The infrastructure crisis means demand dramatically exceeds supply. The validated switching physics means the fundamental science questions have been answered. What remains is engineering (*something we do well*), manufacturing scale-up, and market development—activities that can and should proceed in parallel.

The organizational structure embodies this parallel approach.

True Photonic Inc. develops the core technology platform: the Poovey Stack architecture, the switch physics, the logic family implementation, and the manufacturing processes. SkyLight Holdings develops the infrastructure to deploy that technology: identifying and converting stranded tower assets, securing power and connectivity, preparing facilities to receive photonic processing systems. True Photonic Ventures develops vertical applications that demonstrate specific use cases: CardioLT for cardiac monitoring, PulseLT for pacemaker technology, Photonic Muscle for robotics and prosthetics. Each entity has its own timeline, its own milestones, and its own path to revenue—but all three advance concurrently, with progress in each reinforcing the others. The logic is straightforward: when the machines are ready, the data centers must be ready to receive them. When the platform matures, the vertical applications must be prepared to utilize it. Waiting for sequential completion wastes the very time that market conditions make precious.

This structure requires assembling expertise across multiple domains simultaneously. The core platform team centers on the inventors and engineers who developed the technology: Dr. Gary Poovey, the plasma physicist whose twenty-five years working with leading semiconductor manufacturers culminated in the invention of the light-actuated light-switch that makes photonic logic possible; Del Wolverton, whose decades of experience in fiber optic switching and whose founding of Space Microwave established his credentials in commercializing high-speed optical systems. These are not academics theorizing about what might be possible but practitioners who have built and shipped optical technology before, who understand the gap between laboratory demonstration and production deployment, and who have made the assessment that this technology can cross that gap.

The depth of this experience deserves elaboration because it shapes the confidence underlying the parallel development strategy. Del Wolverton built Space Microwave from concept to acquisition, developing fiber optic switching technology that shipped in commercial telecommunications systems. He has navigated the full arc from laboratory demonstration through manufacturing scale-up through customer deployment through eventual exit—the complete journey that most deep technology efforts never complete. His assessment that the Poovey Stack can make this journey is not the optimism of someone who has never tried; it is the judgment of someone who has succeeded before and recognizes the characteristics that enabled that success.

Dr. Gary Poovey brings a different but complementary perspective: the plasma physicist and materials scientist whose quarter-century working with leading semiconductor manufacturers gave him intimate understanding of why previous photonic computing efforts failed at the physical level. His invention of the Poovey Light-Actuated Light-Switch—the breakthrough that enables light to control light at femtosecond timescales—emerged not from academic theory but from deep experience with the practical limitations that defeated earlier approaches. The intellectual property developed with True Photonic and Del Wolverton and Dr. Poovey represents new compositions and gate architectures that address those limitations directly—not extensions of prior work but fundamental innovations protected by the patent portfolio described throughout this book. As Chief Science Officer of True Photonic, Del leads the research and development that translates the fundamental switching mechanism into manufacturable devices. The graphene-boron nitride-sapphire substrate is not an incremental improvement on prior approaches; it is a fundamental re-architecting based on understanding why those approaches could not succeed.

Robert Switzer brings yet another dimension: the investment banker and technologist who serves as Chief Executive of Astera Energy. His dual perspective—understanding both what technology can achieve and what capital markets require—shapes the substrate development partnership in ways that pure technologists might miss. The graphene deposition processes that Astera Energy has developed are not laboratory curiosities but manufacturing capabilities designed from inception for commercial viability. Switzer's involvement signals that the substrate technology has passed the scrutiny of someone who evaluates ventures for a living, who understands the gap between demonstration and deployment, and who has concluded that this gap can be bridged.

The vertical applications require different expertise—domain specialists who understand the specific requirements of cardiac medicine, of pacemaker technology, of robotic actuation. CardioLT is not merely a photonic device shaped like a cardiac monitor; it must meet the clinical requirements that cardiologists expect, integrate with hospital information systems, satisfy FDA regulatory frameworks, and fit within reimbursement structures that determine commercial viability. Building this capability means recruiting people who have navigated these requirements before, who understand why previous cardiac monitoring approaches succeeded or failed, and who can apply that knowledge to a fundamentally new technological platform. The same logic applies to PulseLT in the

pacemaker domain and to Photonic Muscle in robotics. Each vertical assembles people who have made things work in that specific domain and who have assessed that photonic technology can improve upon what exists.

The infrastructure development through SkyLight Holdings proceeds on its own track, identifying and preparing facilities that can receive photonic processing systems as they become available. The company is actively engaging with multiple cities, each offering different advantages and different paths to deployment. Georgia has emerged as a major data center hub, with significant hyperscaler presence in the Atlanta metropolitan area creating both demand for computing infrastructure and a workforce familiar with data center operations. The state has demonstrated regulatory support for technology infrastructure through tax incentive programs and streamlined permitting processes.

New Orleans holds particular significance in this planning. The city offers substantial stranded commercial real estate from which potential conversion targets can be selected, a power grid with available capacity, and economic development incentives that align municipal interests with infrastructure investment. Local advisors with deep knowledge of New Orleans—including former Congressman William Jefferson, whose understanding of the city's political structures and development processes has been invaluable—have shaped the approach to community engagement and municipal partnerships.

New Orleans is also my city of birth, adding personal commitment to the strategic rationale. This is not mere sentiment; it creates accountability that abstract business relationships do not. Some of my family is still there. The people who will work in a SkyLight facility, who will benefit or suffer from its success or failure, include people I have known my entire life. I cannot treat New Orleans as an abstraction on a spreadsheet. If photonic computing can deliver what the physics promises, I want that delivery to begin in a place where I will personally witness the consequences—good or bad.

The aspiration is to make New Orleans the first city to deploy operational photonic computing infrastructure—to demonstrate that a technology promising community benefit actually delivers that benefit in a specific place with specific people. The city that recovery has rebuilt repeatedly could become the city that demonstrates a new model for computational infrastructure: one that creates jobs, preserves buildings, and operates without the resource burdens

that have made data centers unwelcome elsewhere.

The parallel engagement with other cities ensures that deployment proceeds wherever conditions first align. Georgia offers different advantages: proximity to hyperscaler operations that create immediate demand, a regulatory environment that has welcomed technology infrastructure, educational institutions that produce relevant talent. Other cities under consideration bring their own combinations of stranded real estate, available power, municipal interest, and strategic location. The goal is not to pick a single winner but to develop multiple viable paths and advance whichever reaches readiness first. New Orleans may be first; Georgia may be first; another city not yet in active discussion may emerge as the optimal initial deployment. The parallel structure accommodates this uncertainty while maintaining momentum across all possibilities.

The current status, as of late 2025, reflects this parallel development already in motion. The core technology includes validated device physics with independently confirmed switching speeds in the 150-200 femtosecond range, a complete Boolean logic family implemented in photonic architecture, and patent protection covering fourteen patent filings with over 600 claims spanning the core switching technology, manufacturing processes, substrate development, memory architecture, network infrastructure, and multiple vertical applications.

The intellectual property position builds on published prior art including the NIBLE BUS—an all-optical architecture for managing data transfer through photonic switches, steering components, and buffer memories—and a purely optical NAND logic gate that establishes the foundational building block from which all Boolean operations can be constructed. These publications, now part of the public record, demonstrate that the photonic logic concepts underlying True Photonic's technology have been reduced to practice and documented with sufficient specificity to support the patent claims built upon them.

The infrastructure track has identified prototype deployment sites and begun preliminary assessments. The vertical applications have reached the stage of detailed technical specifications and initial team assembly. None of these tracks is complete, but all are advancing concurrently.

The eighteen months following initial funding close see all tracks accelerate

simultaneously. The core platform demonstrates complete logic chains executing meaningful algorithms, then transitions to early production using established fabrication processes at the 90-180 nanometer node. The infrastructure track converts the first stranded tower assets, preparing physical facilities to receive systems as they emerge from production. The vertical applications advance through their own development milestones—prototype completion, regulatory pre-submissions, pilot program preparation. The convergence point, targeted for months twelve through eighteen, sees initial commercial deployment: photonic processors operating in converted facilities, running workloads that demonstrate real-world capability, generating revenue that validates the commercial model.

The coordination across tracks is not merely conceptual but operational. Consider a concrete example: CardioLT's regulatory submission requires clinical validation data, which requires prototype devices, which requires the Poovey Stack platform to reach sufficient maturity for medical-grade fabrication. Simultaneously, the clinical validation will generate computational workloads—AI analysis of cardiac signals—that can run on SkyLight infrastructure, demonstrating the platform's capability for medical applications while generating data that supports the regulatory submission. The CardioLT timeline does not wait for SkyLight deployment; the SkyLight deployment does not wait for CardioLT approval. Each advances on its own schedule while creating inputs the others require. The coordination is designed, not accidental—a recognition that interdependencies can accelerate rather than constrain if managed correctly.

This timeline assumes that device physics continue to perform as validated, that manufacturing yields meet projections based on mature process node experience, and that no insurmountable integration challenges emerge as system complexity increases. Any of these assumptions could prove wrong. The history of promising semiconductor technologies includes many that validated in laboratory settings but failed to achieve commercial scale. But the people making these assessments are not optimists hoping for the best; they are experienced professionals who have seen technologies succeed and fail, who understand what distinguishes the two categories, and who have concluded that this technology has the characteristics that enable success. Their collective judgment, informed by decades of relevant experience, is that this can be done—and that the parallel structure maximizes the probability of doing it within the window that market conditions have opened.

The network infrastructure vision extends beyond individual data centers to a distributed architecture that fundamentally reconfigures how computational resources connect to users and to each other. This vision builds on the stranded tower conversion strategy described in Chapter 12 but extends it to encompass wide-area photonic networks that link converted facilities into a coherent computational fabric. The economics of this network architecture depend on a characteristic of photonic processing that electronic systems cannot match: the ability to perform computation within the network itself rather than merely transporting data to centralized facilities for processing. Contemporary networks—whether fiber optic backbones or wireless last-mile connections—function as dumb pipes. They move bits from one location to another but cannot operate on those bits in transit. This requires data to travel to processing facilities, undergo computation, and then travel back to the point of use, with each leg of the journey consuming time and energy.

Photonic processing enables a different model. Because the technology computes using light, network nodes themselves can become computational resources. A signal traveling between two points need not remain passive during transit; intermediate nodes can perform processing operations without converting the signal from optical to electrical form and back again. This capability transforms the network from a transport layer into a distributed processing layer.

The practical implementation of this vision involves three infrastructure elements: converted towers that serve as major computational nodes, billboard-mounted repeaters that extend coverage and perform edge processing, and spectrum arrangements that provide the wireless connectivity linking these fixed assets. Each element addresses specific requirements of distributed photonic networking while leveraging existing infrastructure assets.

Converted towers function as regional hubs in this architecture. The office buildings described in Chapter 12—stranded assets whose original tenants have departed, now available at favorable terms—provide the physical space, power connections, and urban locations that hub facilities require. A forty-three-story tower in downtown Atlanta or a former hotel tower in New Orleans can accommodate substantially more computational capacity than traditional data center square footage because photonic systems generate dramatically less heat and require dramatically less cooling infrastructure. Billboard repeaters represent a novel element in this infrastructure equation. The United States

has approximately 350,000 billboard structures, many of them positioned along highways and arterial roads that also carry significant data traffic. These structures already have power connections and, in many cases, fiber connectivity for digital displays. The patent portfolio includes technology for photonic repeaters scaled to fit these existing structures—devices that can regenerate optical signals across distances while performing computational operations on data in transit. A network of billboard-mounted repeaters could provide both coverage density and distributed processing capability without requiring construction of purpose-built infrastructure.

The economics of billboard deployment illustrate the leverage that existing infrastructure provides. We are not buying advertising space; we are leasing attachment rights to unused structural real estate—small photonic repeaters, roughly the size of a cable box, mounted on the back of existing structures. We estimate attachment fees at less than one hundred dollars per month (use could be mandated by legislation - national security)—pure incremental revenue for billboard owners who lose nothing and gain a small but steady income stream from the back surface of their structures. Compare this to the cost of building dedicated network infrastructure: a single cellular tower costs between one hundred fifty thousand and three hundred thousand dollars to construct, requires years of permitting and environmental review, and often faces community opposition. A photonic repeater on an existing billboard costs a fraction of that, requires minimal permitting because the structure already exists with appropriate zoning, and generates no community concern because it changes nothing visible about the billboard. Scale the comparison nationally: three hundred fifty thousand billboards at one hundred dollars per month represents thirty-five million dollars annually for complete national mesh coverage—less than the cost of building two hundred conventional cell towers.

The spectrum component requires partnerships with entities holding relevant wireless licenses. Free-space optical communication—using light beams rather than radio waves to transmit data through the air—operates in spectrum bands that the Federal Communications Commission has left largely unregulated, creating opportunity for deployment without the lengthy and expensive license acquisition process that constrains traditional wireless buildouts. However, free-space optical has limitations, particularly sensitivity to weather conditions, that suggest hybrid approaches combining optical and licensed radio spectrum for reliability. The strategic path forward likely involves partnerships with existing

spectrum holders who can benefit from photonic network capabilities without requiring them to develop the technology independently.

The combined architecture—hub towers, billboard repeaters, and spectrum partnerships—creates what we call the SkyLight National Photonic Mesh™: a network where computation and communication blur together, where data processing happens throughout the network rather than only at designated facilities, and where the physical infrastructure leverages existing assets rather than requiring greenfield construction. This mesh architecture addresses several limitations of current cloud computing: latency that makes real-time applications difficult, bandwidth constraints that throttle data movement, energy consumption that scales linearly with usage, and concentration of critical infrastructure in vulnerable facilities.

The computation-in-transit capability deserves particular emphasis because it represents functionality that electronic networks cannot replicate regardless of investment.

When data travels through an electronic network, it moves passively—routers examine headers and forward packets, but no processing of payload content occurs during transit. The data arrives at a destination, undergoes computation, and travels back. Each leg of this journey consumes time measured in milliseconds that compound across hops. For applications requiring real-time response—autonomous vehicles, remote surgery, financial trading, augmented reality—these milliseconds determine whether the application is viable.

A photonic mesh processes data during transit. A signal traveling from New Orleans to Atlanta need not arrive in Atlanta unchanged; intermediate nodes can perform operations on the optical signal—filtering, pattern matching, preliminary inference—without converting to electronic form. The signal that arrives in Atlanta is not raw data requiring full processing but partially processed information requiring only completion. The latency reduction compounds: not only does the signal travel faster through optical media than electrical, but the processing time at the destination shrinks because some processing occurred en route. Applications that cannot tolerate round-trip latency to centralized facilities become viable when computation distributes across the network path.

The obstacles that remain are substantial, and intellectual honesty requires acknowledging them directly rather than minimizing their significance. Some obstacles are technical, some commercial, and some relate to the inherent difficulty of displacing entrenched technologies regardless of the new technology's merits. The technical obstacles begin with integration complexity. Individual photonic switches have been validated, and the physics of combining them into logic gates is well understood. But building complete processors—with billions of interconnected elements, clock distribution, power management, and interface circuitry—involves engineering challenges that become apparent only at scale. The transition from working devices to working systems has defeated promising technologies before. Photonic computing is not immune to this risk.

Manufacturing yield presents another technical challenge. The Relaxed Lithography approach reduces yield risk by using mature processes with well-characterized defect densities, but photonic devices impose requirements that differ from electronic transistors in ways that could introduce unexpected yield loss mechanisms. The graphene-boron nitride substrate, while offering remarkable thermal and optical properties, must be deposited with consistency across large wafer areas—a process that has been demonstrated but not yet optimized for volume production.

The commercial obstacles may prove more challenging than the technical ones. Established semiconductor companies—Intel, AMD, NVIDIA, and others—have massive resources, existing customer relationships, and strong motivation to defend their market positions. They could attempt to acquire photonic computing technology to control its development pace, develop competing approaches that leverage their existing manufacturing assets, or simply compete on price by accepting lower margins on electronic products to delay photonic adoption. The history of technology transitions includes many examples where incumbents successfully delayed disruption for years or decades.

NVIDIA in particular presents a formidable competitive position. The company's CUDA ecosystem has created software lock-in that makes switching architectures costly for customers even when alternative hardware offers theoretical advantages. NVIDIA GPUs are not merely products but platforms, with thousands of engineers trained on their specific programming models and millions of lines of code optimized for their specific capabilities. Any alternative computing approach must either integrate with this ecosystem

or offer advantages so compelling that customers willingly accept the switching costs. Photonic computing's performance advantages may eventually reach that threshold, but the path from laboratory validation to ecosystem displacement is neither short nor certain.

Customer adoption presents its own obstacles. Even if photonic processors deliver the performance and efficiency advantages the physics predicts, customers must modify their software, retrain their engineers, and accept the risk of deploying unproven technology in production environments. Enterprise technology decisions involve committees, procurement processes, and risk assessments that favor proven vendors and established approaches. Breaking into this market requires not just superior technology but reference customers willing to take early-adopter risk, case studies demonstrating real-world results, and often—unfairly or not—the endorsement of established players.

The financing environment creates additional complexity. Deep technology development requires patient capital willing to accept long development timelines and binary outcomes. The venture capital industry has shifted toward software investments with shorter cycles and more predictable returns. While there are investors who specialize in hard technology and understand semiconductor development timelines, they are fewer than in previous generations, and competition for their attention is intense. True Photonic's approach—parallel development of infrastructure through SkyLight Holdings and vertical applications through True Photonic Ventures—attempts to address this by creating revenue-generating businesses that can operate alongside the core technology development, but this structural complexity itself requires explanation to investors accustomed to simpler company structures.

Regulatory and certification challenges affect different parts of the business differently. Medical devices like CardioLT and PulseLT require FDA clearance processes that add years to development timelines. Data center facilities require permitting that varies by jurisdiction and has become increasingly contentious as community opposition grows. Network infrastructure deployment involves the FCC, state public utility commissions, and local zoning authorities. None of these obstacles is insurmountable, but together they constrain the pace at which even successfully developed technology can reach the market.

The FDA pathway for medical devices illustrates both the challenges and the

opportunities that regulatory frameworks create. A novel cardiac monitoring device like CardioLT would typically require either a 510(k) clearance, demonstrating substantial equivalence to an existing device, or a more rigorous Pre-Market Approval process if the technology is sufficiently novel to lack a predicate device. The photonic architecture underlying CardioLT is genuinely novel—no existing cardiac monitor operates on similar principles—but its outputs are familiar: electrical signals representing cardiac function that physicians already understand how to interpret. The regulatory strategy involves framing the technology as a better way to capture familiar signals rather than as a fundamentally different approach to cardiac monitoring, which may enable the faster 510(k) pathway.

• • •

Against these obstacles, why suggest that 2026 could become the year of light? The proposition rests not on technological optimism alone but on the specific convergence of circumstances that the computing industry now faces.

The infrastructure crisis described in Part I of this book has reached an inflection point. Power constraints have become the limiting factor for AI development at every major technology company. Microsoft's Stargate project requires nuclear power plants. Google data centers draw more electricity than some countries. Meta, Amazon, and OpenAI have all announced infrastructure investments measured in hundreds of billions of dollars, yet executives at each company acknowledge that they cannot build fast enough to meet projected demand. This is not a theoretical future problem; it is a constraint affecting product decisions today. Any technology that can deliver comparable computational throughput at dramatically lower power consumption commands immediate strategic interest regardless of its maturity level.

The scale of unmet demand creates market opportunity that historical semiconductor transitions never enjoyed. When Intel introduced the microprocessor, the personal computer market did not yet exist; the company had to create demand for its new capability. When NVIDIA pivoted to AI accelerators, the machine learning market was nascent and uncertain. Today's AI infrastructure market is neither nascent nor uncertain—it is constrained by supply rather than demand. Companies are paying premium prices for computing capacity, signing multi-year commitments for hardware that does not yet exist, and still facing allocation limits from vendors who cannot manufacture fast enough to meet orders.

Into this environment, a technology that offers ten times the throughput per watt does not need to create demand; it needs only to demonstrate that it works. The supply chain vulnerabilities have become impossible to ignore. TSMC's concentration in Taiwan represents a geopolitical risk that corporate boards and government agencies now discuss explicitly. The ASML chokepoint creates bottlenecks that no amount of spending can quickly resolve. The CHIPS Act has distributed billions of dollars to established semiconductor companies, but the facilities funded by these investments will not reach volume production until the end of the decade, and even then they will not eliminate the fundamental architectural limitations of electronic computing. An alternative approach that sidesteps these constraints—that can manufacture using widely available process nodes rather than contested leading-edge capacity—addresses strategic concerns that have escalated dramatically in recent years.

The community backlash against data centers has transformed the regulatory environment. Projects that would have sailed through permitting five years ago now face organized opposition, extended review processes, and outright rejection. More than one hundred billion dollars in planned data center investment has been blocked, delayed, or significantly modified due to community concerns about water consumption, energy usage, noise, and tax incentives. This opposition will not disappear; it reflects genuine conflicts between the industry's resource requirements and community interests. But a technology that can deliver data center capability without the resource footprint that generates opposition changes the calculus. Communities that reject electronic data centers might welcome photonic facilities that require no cooling towers, minimal water consumption, and a fraction of the grid capacity.

The validated technology base provides a foundation that did not exist even recently. Photonic computing has been pursued for decades, but previous attempts foundered on switching speeds that could not compete with improving electronic transistors, materials systems that could not be manufactured at scale, or architectures that could not implement complete computational functionality. The Technion validation—150-200 femtosecond switching, independently measured and documented—establishes a performance level that the roadmaps of electronic computing cannot match. The Poovey Stack architecture demonstrates that complete Boolean logic can be implemented in photonic devices. The Relaxed Lithography approach shows that manufacturing can proceed without the constrained resources that bottleneck the semiconductor industry.

These are not theoretical possibilities but demonstrated capabilities. The investment environment, despite its challenges, includes actors who understand the strategic significance of the moment. Sovereign wealth funds, pension systems, and strategic corporate investors recognize that the current trajectory of AI development is unsustainable—that something must change, whether through more efficient hardware, better algorithms, or reduced ambitions. Photonic computing offers the possibility that the ambitions need not be reduced, that the trajectory can continue if the underlying technology improves sufficiently. For investors positioning themselves for the next decade of technology development, this possibility commands attention.

The year 2026 represents a convergence point where these factors align. Technology readiness reaches the threshold for initial commercial deployment. Market demand exceeds what current technology can supply. Infrastructure constraints force strategic reconsideration. Alternative approaches gain credibility as traditional roadmaps falter. This does not guarantee success—the obstacles enumerated above remain real, and execution must be exceptional to overcome them. But it creates conditions more favorable for fundamental technological transition than the computing industry has seen in decades.

• • •

The vision that emerges from this analysis is neither triumphalist nor modest. It suggests that the computing industry stands at an inflection point comparable to the transitions from vacuum tubes to transistors, from transistors to integrated circuits, from mainframes to personal computers. Each of those transitions seemed unlikely to many observers even as it was occurring. Established players had reasons to believe their positions were secure. New approaches faced technical obstacles that skeptics considered insurmountable. The transitions happened anyway, driven by fundamental economics that eventually overwhelmed institutional resistance.

The economics driving this potential transition are straightforward even if the technology is complex. A photonic processor that consumes one-tenth the power of an electronic processor to perform the same computation—if such a device can be built and manufactured at reasonable cost—eventually displaces the electronic processor regardless of installed base, existing relationships, or customer inertia. A data center that requires no cooling towers and minimal grid capacity can operate in locations and under conditions where electronic data centers

cannot. A network that processes data in transit rather than merely transporting it enables applications that today's infrastructure cannot support.

The strategic framework that True Photonic has constructed attempts to capture this economic opportunity through multiple, reinforcing mechanisms. The core technology company develops and licenses the fundamental intellectual property. SkyLight Holdings deploys infrastructure that creates immediate demand for photonic processing. True Photonic Ventures develops vertical applications that demonstrate specific use cases and generate revenues while the core platform matures. The patent portfolio protects this position across the relevant technical domains. Each element of this structure supports the others, creating resilience against the specific risks that any single approach would face alone.

The cognition-as-export model described in Chapter 13 provides a framework for thinking about what success might mean at geopolitical scale. If the United States and its allies control the most advanced computational technology—not through regulatory restriction but through technological leadership—the outputs of that technology can reach global markets while the infrastructure remains sovereign. This is not isolation but rather the kind of competitive advantage that nations have historically sought to maintain in strategically significant industries. The parallel to energy policy, pharmaceutical manufacturing, and aerospace should be apparent.

Yet the appropriate posture for those advancing this technology is humility rather than certainty. The gap between laboratory validation and commercial success has swallowed many promising innovations. The obstacles are real, the timeline is aggressive, and the competition includes some of the world's most capable and well-resourced technology companies. Nothing in this book should be taken as a guarantee that photonic computing will succeed, that True Photonic specifically will succeed, or that the timeline suggested here will prove accurate.

What can be said with confidence is that the current trajectory of computing infrastructure is unsustainable. Power consumption cannot continue increasing at current rates. Water usage cannot continue depleting aquifers and straining municipal systems. Community opposition cannot be indefinitely overridden. Supply chain vulnerabilities cannot be wished away. Something will change, either because better technology enables a different approach or because demand for computation contracts to fit what current technology can supply.

The proposition of this book is that light offers a better path than retreat. The physics of photons enables computational capabilities that electrons cannot match. The materials science has advanced to the point where practical devic-

es can be fabricated. The manufacturing approach leverages available infrastructure rather than requiring constrained resources. The deployment model transforms stranded assets into productive facilities while resolving rather than creating community conflicts. The vertical applications address real medical needs with technology that could not exist in electronic form.

The title of this book—**Keep Computing**—contains both a description and an aspiration. It describes a technology that enables computational progress to continue beyond the limits of electronic approaches. It aspires to a future where the benefits of artificial intelligence, machine learning, and advanced computation remain available to those who seek them, without requiring environmental sacrifice, community disruption, or dependence on vulnerable supply chains. Whether that future arrives through photonic computing specifically or through some other approach not yet anticipated, the need it addresses is real and the effort to meet it is worthy.

The computing industry has faced existential challenges before. The vacuum tube era seemed permanent until transistors proved otherwise. The mainframe paradigm appeared unassailable until personal computers transformed the market. The x86 architecture dominated until mobile devices created new winners. Each transition disrupted established players, created new fortunes, and reconfigured competitive dynamics in ways that were easier to see in retrospect than they were to predict in advance. Those who correctly anticipated transitions—or who positioned themselves to benefit from multiple possible outcomes—prospered. Those who assumed the status quo would persist often did not.

The transition now approaching may be more consequential than its predecessors. Previous computing transitions changed which companies won and lost within an industry that continued to grow. The current inflection point raises the question of whether the industry can continue to grow at all under existing physical constraints. **If the answer is no—if electronic computing has genuinely reached its practical limits—then either a new approach emerges or computational progress slows.** The implications of that slowdown would extend far beyond the technology sector to every industry that depends on continued improvement in processing capability.

The light ahead may illuminate that path, or it may prove to be another promising technology that falls short of its potential. The honest answer is that no one can know with certainty until the attempt is made. But the attempt is worth making, and the moment to make it has arrived.

The year 2026 may or may not become the year of light. What happens next depends on execution, on resources, on factors beyond any single organization's control. But the physics is sound, the technology is validated, and the need is urgent.

That is enough to justify the work ahead.

AFTERWORD

Infrastructure That Shares

In the summer of 2021, residents of Mesa, Arizona discovered that their water rates would increase by twenty-five percent over the next five years. The reason was infrastructure investment—not for their homes or businesses, but for data centers. Google, Apple, and Meta were building facilities in the region, drawn by cheap land and cheaper power. The aquifers that served the community would now serve the cloud.

Mesa's residents had no stake in the facilities draining their water. They would not share in the profits when those data centers processed queries, trained models, and ran the computations that powered trillion-dollar companies. They would simply pay more for water while the economic value of that water flowed to shareholders in California.

This pattern is not new.

It is the pattern of every extractive economy in history.

• • •

In the coal regions of Appalachia, communities hosted an industry that powered the nation. They absorbed the pollution, the mine collapses, the black lung disease. The coal left on trains and did not come back. The profits accumulated elsewhere. When the industry declined, it left behind poisoned streams, collapsed local economies, and populations too sick and too poor to leave.

In the oil towns of Texas and Oklahoma, the pattern repeated. Roughnecks worked the rigs, their towns swelled with temporary prosperity, and the wealth flowed to Houston, to shareholders, to anywhere but the communities that hosted extraction. **When the wells ran dry, the towns became shells.**

The digital economy was supposed to be different.
Information does not deplete. Data centers do not scar the land like strip mines. The technology industry promised a cleaner, more distributed prosperity.

That promise has not been kept.

Today's data centers extract water from Arizona aquifers, electricity from Oregon dams, and land from rural communities across America. The computation they perform generates trillions of dollars in value. That value accrues to the companies that own the infrastructure—Amazon, Microsoft, Google—*and to the shareholders who own those companies.*

The communities that host this infrastructure receive jobs, temporary construction booms, and modest tax revenues negotiated from positions of weakness.

They do not receive ownership. They do not share in the upside when artificial intelligence models trained in their facilities generate billions in revenue. They bear the costs and watch the benefits flow elsewhere. This is the extractive model. It is the model that has governed every strategic resource in history, from coal to oil to compute. And it is a *choice*, not an *inevitability*.

• • •

The question of who benefits from strategic infrastructure is not incidental to the infrastructure's development. It shapes whether communities welcome or resist new facilities. It determines whether the political coalition supporting a technology is broad or narrow. It decides, ultimately, whether an industry is seen as a partner or a parasite.

When Intel announced plans for new fabrication facilities in Ohio, the company understood that community support was not automatic. Intel's pitch included not just jobs and tax revenue, but educational partnerships, supplier development programs, and commitments to local hiring. This was progress, but it was still the old model dressed in new language. The community would benefit from Intel's success, but it would not own any piece of that success. The ownership—the equity, the enduring stake—would remain with shareholders who had never set foot in Columbus.

There is a reason the model persists. Ownership requires capital, and communities do not have the capital to invest in semiconductor fabs or data centers. They cannot write billion-dollar checks. They can only offer what they have: water, power, land, and regulatory accommodation. They trade resources for jobs and hope the exchange is fair.

But what if the exchange could be structured differently?

What if access to infrastructure could be distributed and owned by the communities that host it?

• • •

The model we are building at True Photonic is designed around this possibility. Photonic Compute Credits™ represent access to compute capacity on our infrastructure. One PCC equals one megawatt-hour of renewable-powered photonic compute capacity—real infrastructure, measurable and verifiable.

Credits can be granted, purchased, held, transferred, and redeemed. This is not a cryptocurrency designed for speculation—it is a mechanism for distributing ownership of infrastructure access rights in a way that traditional equity structures cannot achieve.

A town of fifty thousand people cannot collectively purchase shares in a private company. The legal structures do not exist. The minimum investment thresholds exclude them. But a town of fifty thousand people could hold credits representing access rights to compute infrastructure. A local community foundation could acquire or be granted credits, hold them, and resell access to regional businesses at market rates, keeping the margin as ongoing revenue. A municipality could accept credits as partial payment for hosting facilities and use them to subsidize compute access for local economic development.

When a SkyLight Clean Compute Center comes online, a portion of the credits generated are allocated to community ownership. **This would not be charity**. It is not corporate social responsibility.

It is structural participation in the value the facility creates.

The community does not merely host the infrastructure—*it owns a piece of the access rights that infrastructure generates.*

Community-allocated credits can be held by local foundations, municipal authorities, non-profit organizations, or community development financial institutions. The specific structure depends on local conditions. What matters is that the credits represent an ongoing stake, not a one-time payment.

Community holders can then do what any owner can do: hold the credits and

accumulate more over time, sell them to generate immediate revenue, or use them to provide discounted compute access to local institutions. A community foundation might sell credits to regional businesses and use the proceeds to fund local programs. A municipality might use credits to subsidize compute costs for the local school district. A rural development organization might provide credits at cost to agricultural cooperatives implementing precision farming.

The community becomes a participant in the compute economy, not merely a host for its infrastructure.

I anticipate the objection: this sounds like corporate philanthropy with extra steps.

It is not. *The difference is structural.*

Corporate philanthropy is discretionary. A company can fund community programs one year and cut them the next. A change in leadership, a downturn in profits, a shift in priorities—any of these can end the giving. Communities know this, which is why they often discount corporate promises. They have been burned before.

Credit-based community ownership would be contractual. The allocation to community entities is encoded in the facility's operating structure from day one. It does not depend on continued goodwill. It does not require annual renewal. Once the credits are allocated, they belong to the community. The company cannot take them back. This difference matters.

Discretionary giving creates dependence. Structural ownership creates agency. A community that depends on corporate donations must maintain good relations with the donor. A community that owns access rights can do whatever it wants with those rights—sell them, hold them, donate them, or use them to attract businesses that need compute capacity. It has options that dependent communities do not have.

The goal is not to make communities grateful.

The goal is to make them powerful.

The first industrial revolution concentrated wealth in ways that shaped politics and society for two centuries. The owners of coal mines, steel mills, and railroads became a class with interests distinct from—*and often opposed to*—the communities that hosted their enterprises.

This concentration generated social instability, labor conflict, and political movements that reshaped the world.

The digital revolution is repeating this pattern. The technology industry has created more concentrated wealth than any industry in history. This concentration is not incidental to the technology. **It is structural.** Software scales infinitely, network effects compound, and winner-take-all dynamics ensure that gains flow to a small number of companies and their shareholders.

We have already seen the political consequences. Populist movements on both left and right draw energy from the perception that the technology economy benefits a narrow elite while leaving everyone else behind. This perception is not entirely wrong. The gains from digital technology have accrued disproportionately to a small number of regions, companies, and individuals. The costs—disrupted industries, displaced workers, strained communities—have been broadly distributed.

Photonic computing will create enormous value.

The question is whether that value will be captured by a new set of concentrated owners or distributed more broadly. The technology itself does not determine the answer. The ownership structures we build around the technology determine the answer.

We are choosing to build structures that distribute value broadly.

Not because we are better people than other technology founders. Because we believe it is the only model that is politically sustainable over the long term. A technology that takes from communities without giving back will eventually face resistance that no amount of lobbying can overcome. A technology that makes communities genuine partners has the possibility of enduring support.

This is self-interest, properly understood.

We want photonic computing to succeed not just for a quarter or a year but for decades. That requires a coalition broader than shareholders. It requires communities that see themselves as **beneficiaries**, not *hosts*.

• • •

I want to be clear about what we are not claiming.
We are not claiming that this model solves inequality. The structural forces that concentrate wealth are larger than any single company's policies. Credit-based community ownership is a mechanism, not a revolution. It redistributes access rights to a single company's infrastructure, not access to the technology economy as a whole.

We are not claiming that we have all the answers. The model I have described is still in development. The specific allocation percentages, governance structures, and community eligibility criteria will be refined through experience.

We will make mistakes. *We will learn from them.*

We are not claiming purity of motive. True Photonic is a business. We need to generate returns for our investors, or we will not exist long enough to implement any model at all. The community ownership structure is designed to be compatible with investor returns, not to replace them. We believe a broadly beneficial model is also a profitable model, but we do not pretend that profit is irrelevant to our decisions.

What we are claiming is simpler: **the choice exists.** The extractive model is not the only option. Technology companies can structure ownership to include the communities that host their infrastructure. The tools to do this exist today. What has been lacking is will.

We have the will.

• • •

In the opening chapter of this book, I described **water that did not come back**. A river diverted to cool servers, a community's allocation strained to serve a cloud it would never own. That image stayed with me as we structured True Photonic. Not because we can avoid using water—even photonic computing

might require a tiny bit of cooling—but because it captured something essential about the relationship between infrastructure and community.

The water belonged to the community.
The computation belonged to someone else.

The benefits **flowed away**, like the water itself.

I do not want True Photonic to be that kind of company.

In the Prologue to this book, I described how Gary Poovey came to join us. What I did not fully explain was the condition he placed on that partnership.

Gary prayed with his family before he committed. When he called back with his answer, it came with a request: *he wanted to work with people who would build an ethical company*. He had spent his career in an industry that optimized for shareholder returns above all else. He had watched good technology serve narrow interests. He was entrusting us with his life's work—the materials science breakthrough that made everything in this book possible—*and he wanted assurance that we would do some good in the world with it.*

I made him that promise.

I am not a particularly virtuous person. I am a businessman who has spent forty years chasing opportunities wherever they appeared.

But a promise is a promise.

And beyond the obligation, I have built enough businesses to understand that extraction is not sustainable. Eventually, the community pushes back. Eventually, the political winds shift. Eventually, someone asks the question that should have been asked from the beginning: **Who benefits from all this?**

When that question is asked of photonic computing, I want the answer to be: **everyone who helped make it possible.** The communities that host our facilities. The early believers who took risk when success was uncertain. The researchers who dedicate careers to problems that may not be solved in their lifetimes. The workers who build and operate infrastructure that makes computation possible.

And yes—the shareholders and founders who provide capital and *bear risk*.

Their returns are legitimate. But they should not be the only returns. The value photonic computing creates is too large, and the communities that make it possible are too important, for the old extractive model to be acceptable.

We are building something different.
We are building infrastructure that shares.

Gary Poovey trusted us with his breakthrough because we promised to do good with it. The community ownership model is how we keep that promise.

Not through charity, not through foundations that bear our name—*though we have that too*—but through structures that give communities real power over real assets. **Ownership**, not gratitude. **Stakes**, not handouts.

The technology is ready.

The ownership model is ready.

What remains is execution—and communities willing to be partners rather than hosts.

We are looking for those communities now.

We built the cloud to escape
the physical world.
The physical world is
sending the bill.

About the Author

Derek W. Bailey is CEO and Co-Founder of True Photonic, Inc., a photonic computing company developing femtosecond-scale optical switching technology validated at 150-200 femtoseconds—three to four times faster than any previously reported result.

Bailey has been an entrepreneur for more than four decades, building ventures across hospitality, real estate, insurance, staffing, and technology. His companies include OnPoint Staffing, Curbside Chef, International Broadband Technologies (video over copper wire), Popa Media (digital billboards with integrated VoIP), and WolvertonBailey, Inc., which developed the patented Counterpoise engine. He is also founder of Derek Automotive Technologies, developing parallel-hybrid self-charging electric vehicles.

His twenty-seven-year partnership with Del Wolverton has produced multiple patents across ventures. When Dr. Gary Poovey entrusted them with his breakthrough materials science—the organic semiconductor architecture that makes photonic computing practical—Bailey structured True Photonic around a promise: to build an ethical company that shares value with the communities hosting its infrastructure.

True Photonic now leads a multi-entity structure including the core platform, SkyLight Holdings (Clean Compute Centers), and True Photonic Ventures (vertical applications in medical devices and beyond). The patent portfolio spans fourteen filings with more than six hundred claims.

Keep Computing is his first book.
Bailey lives in Atlanta, Georgia.

Contact: dbailey@truephotonic.com