Chasing the Wrong Holy Grail

How the industry's obsession with speed became a one-dimensional answer to a multi-dimensional problem — and the unnecessary crisis it is creating

By Rob Cain — Head of Strategy & Finance, PhotonWeave

Every field has its Holy Grail — the single prize everyone agrees is worth chasing. In the race to build the physical infrastructure of artificial intelligence, the industry has crowned one Grail above all others: *speed*. Faster electrical lanes, faster serializers, faster lasers, more gigabits forced down every channel. Tens of billions of dollars are pouring into the pursuit, and the direction is treated as settled.

I want to argue that it is the wrong Grail. Speed is a one-dimensional answer to a problem that is profoundly multi-dimensional — and that single mismatch, a narrow solution aimed at a wide problem, is the hidden source of the sprawling and largely unnecessary crisis now forming around AI infrastructure: the exploding demand for power, water, and land; the pollution and public-health costs; and the social backlash that follows all of it.

That word — *unnecessary* — is the one I most want to defend. We are paying an inherently avoidable price to achieve intelligence at scale. A great deal of it is the price of an architectural choice. Change the choice and most of the crisis recedes. I come at this not only as an engineer's problem but as a strategist's: when you step back far enough to see the whole system, the pattern is unmistakable. We have optimized a single variable so relentlessly that the costs have nowhere to go but outward — into the grid, the watershed, the air, and the neighborhood.

A one-dimensional answer

Consider how the industry actually buys bandwidth. It does it the only way it has known for thirty years: push more bits per second through a channel. When copper signaling hit its wall — when you go beyond roughly 200 gigabits per second per lane, an electrical link can barely cross a server rack — the response was not to rethink the frame but to double down on it: move the optics closer to the compute, drive the light with lasers, and keep the high-speed serializer-deserializer escalator climbing.

Every one of those moves optimizes the same lever: speed. And every one pays for it in the same currency: power and heat, concentrated onto an ever-smaller patch of silicon. That is what “one-dimensional” means in practice. A single-lever strategy has only one thing to pull, and pulling it harder makes everything downstream worse. The laser compounds the problem, because a laser is not silicon — it cannot be bonded directly to the compute die, it runs hot, it is expensive to build at volume, and independent analyses of co-packaged optics now identify the absence of a native silicon laser as the primary obstacle in the entire field. The industry's own Grail is fighting it.

The bill that never makes it onto the pitch deck

Follow the single-variable strategy downstream and you arrive at a cascade of costs that are usually discussed as if they were separate, unrelated crises. They are not. They are the shadow cast by one upstream decision.

Power. Gartner projects global data-center electricity consumption will rise about 26% in 2026, to roughly 565 terawatt-hours, with AI-optimized servers on track to draw more power than all conventional data-center hardware combined by 2027; the IEA expects data-center demand to more than double by 2030. Racks that drew 5–15 kilowatts a few years ago now exceed 100. Power availability, not chip performance, has become the binding constraint on whether AI can grow at all.

Land and real estate. A single hyperscale facility can span hundreds of acres, and each one drags behind it transmission corridors, substations, and — as the Lincoln Institute of Land Policy has documented — the eminent-domain fights needed to route all of it. In the first months of 2026 alone, more than 75 projects representing over $130 billion in investment were blocked or contested.

Water. Cooling all that concentrated heat consumes staggering volumes of water — globally, data centers are projected to draw between 4.2 and 6.6 billion cubic meters annually by 2027, and in the United States they already rank among the top ten industrial water users. In drought-exposed regions, that pits a data center directly against agriculture and households.

Pollution and public health. Because the grid cannot keep up, facilities increasingly lean on on-site gas turbines and diesel generators, which emit fine particulate matter and nitrogen oxides linked to asthma, heart disease, and premature death. The American Lung Association's 2026 “State of the Air” report named data centers a growing air-quality concern. In Memphis, more than thirty gas turbines at a single AI facility drew a Clean Air Act challenge from residents and the NAACP. One analysis pegged the annual health damage from a single Northern Virginia facility at $53–$99 million, and a broader study estimated that data centers imposed roughly $25 billion in health and environmental costs on the U.S. economy in a single year — with billions tied specifically to AI workloads.

Social cost. The people who live nearby absorb the rest: electricity bills reported up more than 250% in some neighborhoods near data centers, constant low-frequency noise, and the sense that the burden and the benefit have been split between different people. Local opposition is now rising across the country, and it is not irrational. It is a community doing the accounting the industry declined to do.

Every one of these is being managed as its own emergency, by its own specialists, with its own line item. But they share a single root. The machine is too hot and too power-hungry, and it is too hot and too power-hungry because we decided to buy bandwidth with speed. You cannot cool your way out of a problem you designed in.

A multi-dimensional problem demands a multi-dimensional answer

Here is the reframe. Nature does not give us only one way to carry information. It gives us several: not just how *fast* a signal switches (time), but how *many* signals we run in parallel (density and space), in what

phase, and at what *wavelength*. Speed exploits exactly one of these and leans on it until it breaks. A multi-dimensional architecture spreads the work across all of them, so that no single dimension — and no single square millimeter of silicon — has to carry the whole load. Distribute the burden and it stops concentrating into heat. And heat, remember, is the root of the entire cascade above.

What that looks like in practice

This is the architecture we are building at PhotonWeave, and the point I most want to land is that it does not require inventing new physics. It requires refusing to solve for speed.

Start with density instead of speed. A microLED, unlike a laser, is CMOS-friendly — it bonds directly to the driver silicon, it is grown by the thousands per wafer so it is cheap, and it runs cool. The reflexive objection is that LEDs are slow. They are; and it does not matter. A laser is about a millimeter across; a microLED at the pitch we target is about ten microns. That is on the order of ten thousand times more emitters in the same area — so each one can run ten thousand times slower and deliver the same aggregate bandwidth. Where a laser link is pushed toward hundreds of gigahertz, our elements operate in the low hundreds of megahertz — and still reach terabit-per-square-millimeter-class density. We buy bandwidth with volume, not velocity, and slow parallel light is cool, cheap light that mature, inexpensive CMOS can drive.

Pack emitters that densely and their light collides — the obvious objection, and the problem we exist to solve. Our approach, PhaseWeave, is deterministic scheduling: adjacent emitters never fire in the same window. Rather than fighting an analog channel to recover order, we define the order first and communicate within it. The physics problem becomes a scheduling problem, and software is very good at scheduling.

The most overlooked piece is the receiver — and it is where most of the field is quietly stuck. Conventional detection leans on a large analog front end that cannot shrink to microLED density and often cannot even sit on the same die, so even teams with an excellent transmitter can rarely show a complete, integrated signal chain on silicon. They have half a link. Our receiver is engineered precisely around that gap: it matches the density of the emitter, recovers the signal digitally rather than through a bulky analog stage, and builds on device technology already manufactured in high volume on mature-node CMOS. The specifics are the part of the stack we hold closest — the core of our advantage, and a trade secret we are glad to walk through with serious partners and investors under the appropriate cover. What I will say in the open is this: the hard half of this problem, the half the industry keeps deferring, is the half we have solved.

That last fact is what makes the crisis I described genuinely unnecessary. The alternative is not a decade-out moonshot demanding exotic materials and bleeding-edge fabs. It is a different assembly of components the industry already knows how to manufacture — which is also why it reaches proof at a fraction of the capital the speed-first path requires. The architecture is the advantage.

The crisis is a choice

The backlash against AI infrastructure — the lawsuits, the moratoria, the utility bills, the health studies — is easy to read as proof that AI's physical footprint is simply the cost of progress. I read it as proof that we are chasing the wrong Grail. We picked one dimension, speed, optimized it until the burden spilled out into everything around it, and then treated each spillover as a separate misfortune to be managed.

Pursue the right Grail instead — bandwidth synthesized across many dimensions, cool and dense and deterministic — and you attack the root rather than the symptoms. The power curve bends. The water draw falls. The land footprint shrinks. The pollution and the health bill and the neighborhood opposition ease, because the thing generating all of them was addressed at the source. A field that has invested enormously in a single approach will be slow to concede that the frame itself is the problem. But the most valuable move available right now is not to run the existing race faster. It is to question whether it is the right race.

The future of AI does not have to be a story about building more power plants to feed a hotter and hotter wire, at the expense of the communities unlucky enough to live beside it. It can be a story about weaving light — slowly, densely, and deliberately — into something the physics was ready to give us all along.

PhotonWeave is building the multi-dimensional communication architecture that turns best-in-class, already-manufacturable devices into real-world optical links — deterministic scheduling, calibration, integration, and a proprietary receiver architecture protected by a growing patent portfolio. If you are thinking about the physical layer of AI — as a device maker, a system builder, or an investor weighing where this goes next — I would welcome the conversation.