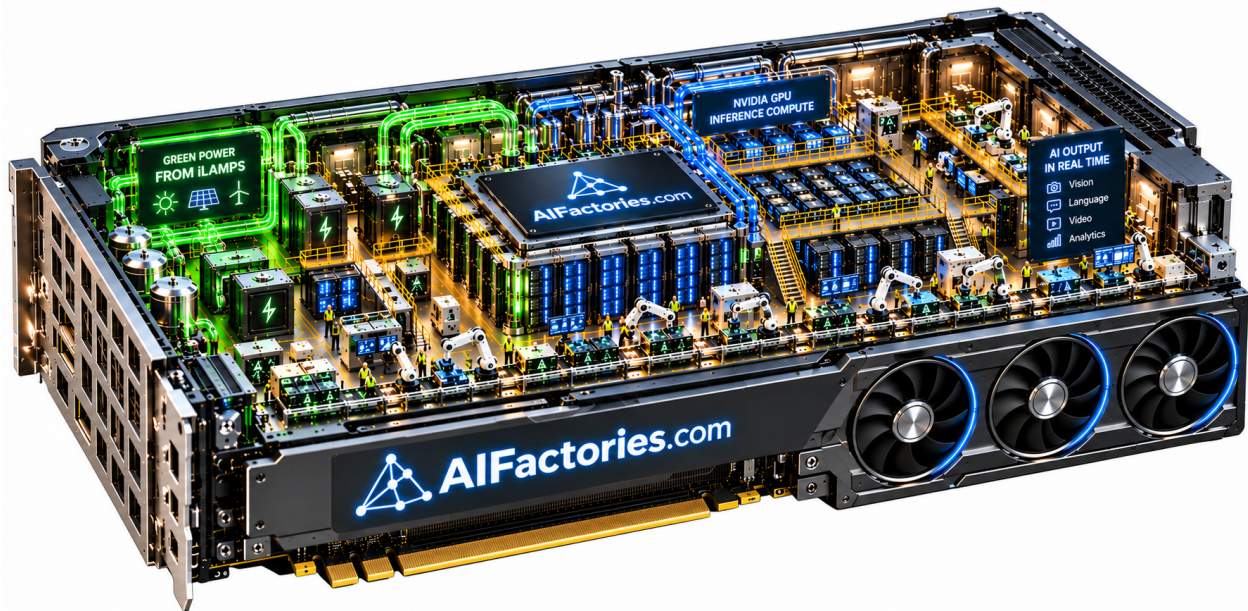




News Release: iSCE and Energy-Aware Job Processing

Conflow Power Group sets out the two-layer software architecture that makes AI inference work on an off-grid street pole

The iLamp Sparse Compute Engine fits the model to the pole through its key partnership with AI Factories .com

A new energy-aware job processing layer sorts, dispatches and returns the work across estates of poles that have no connection to one another.

London, UNITED KINGDOM, 21st October 2026 Conflow Power Group Limited (CPG) has set out the software architecture behind distributed AI inference on its iLamp platform: the iLamp Sparse Compute Engine (iSCE), and an energy-aware job processing layer now in build under the internal reference CPG-TS-ORCH-001.

Between them the two layers address the problem that has held distributed edge inference at pilot scale: how to sort, process and deliver inference work across thousands of autonomous, solar-powered nodes that are not connected to one another and do not have a fixed power budget from one hour to the next.

Why data centre orchestration does not transfer to a street pole

Conventional cluster and data centre orchestration rests on three assumptions: constant mains power, reliable inter-node networking, and physical security of the rack. None of the three holds on a public highway.

An iLamp generates approximately 600 watts of solar from a flexible array wrapped as a cylinder around the pole and carries its own storage in the base. Onboard services, lighting, cameras, sensors, WiFi and controllers, consume roughly a third of daily generation. What is left over is the compute headroom, and that figure moves with the hour, the season, the weather and the shading from whatever is growing next to the pole. A scheduler that treats each node as an always-on server will do one of two things, both bad: overcommit poles that cannot finish the work or hold the whole estate back to the capability of its worst node.

CPG has treated that constraint as the design problem rather than as an operational nuisance and has built two layers against it.

Edward Fitzpatrick CPG CEO comments: “Three million iLamps on Thor-class modules carry about three exaOPS of nameplate inference compute, of the order of two-thirds of a one-gigawatt Blackwell AI factory and roughly half of a Rubin-generation one, on 1.8 gigawatts of their own solar instead of one gigawatt of grid, with no cooling plant and no water. The factory’s revenue has to pay for its power before it pays for its debt; ours pays for its debt first, because the power is free at the point of use.”

Let that sink in, a solar powered distributed AI network on the EDGE delivering power, light, services and compute literally everywhere there is a streetlight.

Layer one: iSCE fitting the model to the pole

The iLamp Sparse Compute Engine is CPG’s model-side stack. It has four components:

- Model Compression Foundry, which prepares production models for an energy-constrained runtime.
- Hybrid Sparsity Compiler, which combines structured and unstructured sparsity for the target silicon.
- Edge Runtime, implemented as a TensorRT plugin for the in-pole accelerator.

- Fleet-Level Inference Routing, which presents an estate of individual poles to a caller as one addressable resource.

Each iLamp carries an NVIDIA Jetson AGX Orin edge compute module rated at 275 TOPS. CPG's internal modelling indicates that iSCE delivers an effective compute multiplier of approximately 4.5 times against an uncompressed baseline on the same silicon. That figure is modelled and has not yet been independently measured. CPG is arranging third-party verification and will publish the result when it is available.

Layer two: energy-aware job processing

Compression alone does not tell a request where to go. That is the function of the job processing layer, and it is the part of the architecture CPG considers genuinely new.

The broad principle is straightforward to state and difficult to implement: the energy state of a node is treated as a first-class scheduling input, ranked alongside latency and queue depth rather than beneath them. Work is placed where there will be power to finish it, thermal margin to run it and a delivery path to return it. The estate as a whole therefore behaves as a single elastic resource whose capacity rises and falls predictably with daylight and season, rather than as a set of individually unreliable machines.

The specific allocation method is held as proprietary and CPG does not intend to publish it. The architecture is defined, and the layer is in build. CPG is currently closing the scope boundary between this layer and the tooling supplied by its data centre partners. Consistent with standing policy, CPG does not name its compute off-takers or partners or disclose their commercial terms.

Where CPG believes it is ahead

CPG's position rests on three claims that it invites the market to test.

- The constraint was designed in, not retrofitted. The iLamp platform has been in development for approximately six years, with generation, storage, thermal management and compute developed as one asset. Orchestration written for grid-connected racks has to have the energy constraint added afterwards, which is a much harder retrofit than it appears.
- Vertical integration gives the scheduler real data. Because CPG owns the generation and the storage in the same unit as the accelerator, the scheduler works from the actual state of charge and generation forecast of each node rather than inferring capacity from observed behavior.

- The economics improve with fleet size. An energy-aware scheduler is close to pointless across ten poles and decisive across tens of thousands, because the diversity of the estate is what smooths the supply curve. CPG's deployment pipeline stands at 2.7 million poles, against a global installed base of roughly 350 million street lights that currently sit in daylight doing nothing.

Comment

Edward Fitzpatrick, Founder, Chairman and Chief Executive of Conflow Power Group, said:

"The industry has spent a decade making models smaller and almost no time asking where, physically, the small model is supposed to run. We think the answer is the street light, because it is already everywhere, it already has a power source, and nobody is using it between dawn and dusk. Getting there means accepting that the pole is not a rack. Our two layers exist because of that one fact."

"We are publishing the shape of the architecture, not the recipe. The market can judge the shape now and judge the numbers when somebody other than us has measured them."

ENDS

Notes to editors

- Conflow Power Group Limited is a UK-based smart infrastructure company and the developer and global licensor of the iLamp platform: autonomous solar-powered smart street lighting with onboard edge compute. CPG holds the Frost and Sullivan Global Company of the Year 2026 award.
- iSCE, the iLamp Sparse Compute Engine, is documented internally as CPG-WP-iSCE-001 (white paper), CPG-ENG-iSCE-001 (engineering implementation plan) and CPG-PAT-iSCE-001 (patent disclosure).
- The job processing layer is documented internally as CPG-TS-ORCH-001. Status at the date of this release: architecture defined, in build.
- The 4.5 times effective compute multiplier is a modelled figure derived from CPG internal work. It is quoted as modelled and should be reported as such until independent measurement is published.
- CPG does not disclose the identity of its compute off-takers, their commercial terms or their service-level commitments, and does not disclose iLamp unit build costs.

- Interview and briefing requests, including technical briefings under NDA, should be directed to the contact below.

Media contact

Edward Fitzpatrick, CEO, Conflow Power Group Limited. +44 7773 727272. Edward@iLamp.com.
conflowpower.com

Copyright Conflow Power Group Limited.