AIONTRACKS

Linking data centers to stranded electrons™

AionLink: Meeting NOGRR282 ride-through with computational load modification

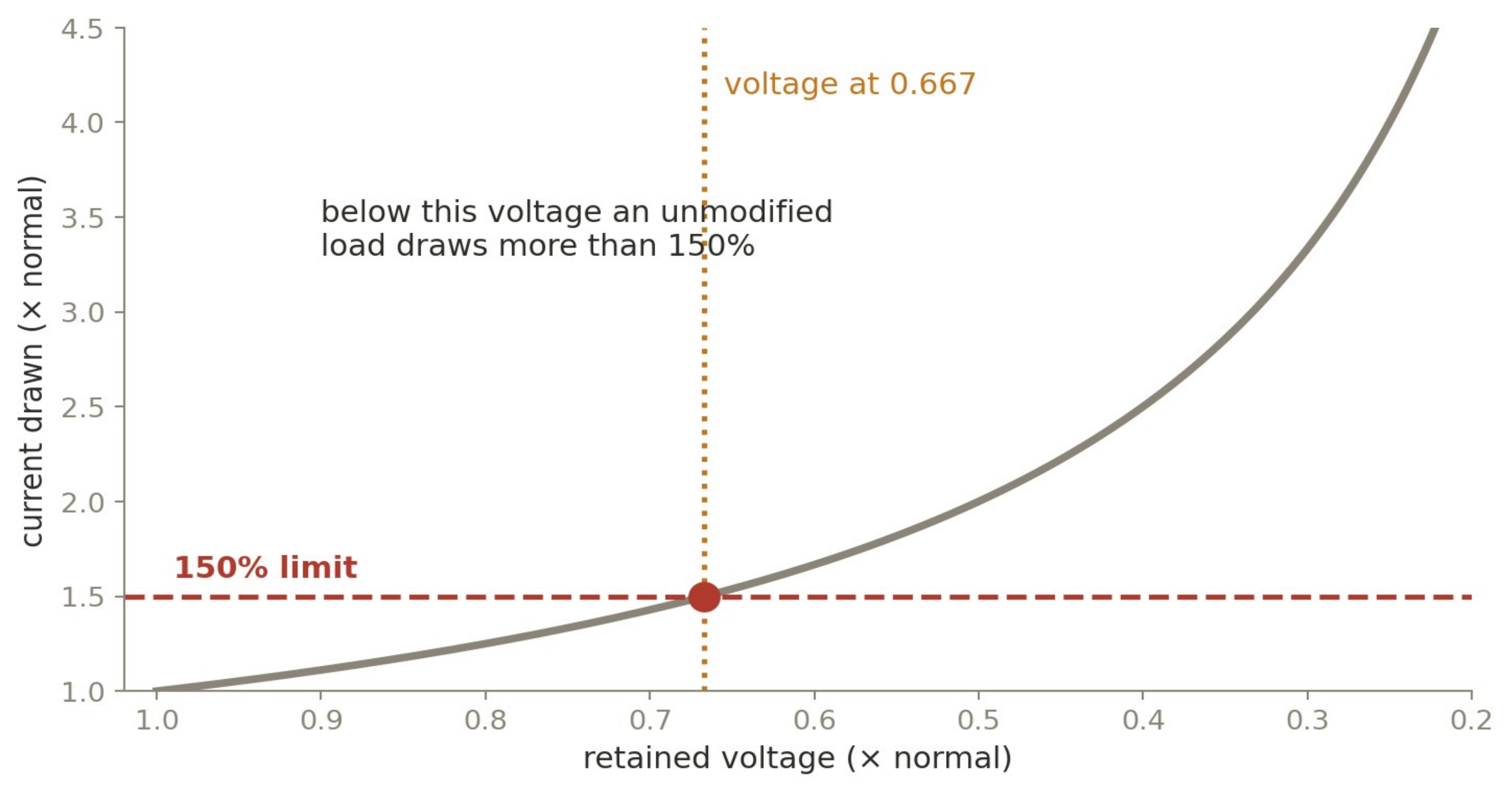
"You don't have to swing hard to hit a home run. If you've got the timing, it'll go."

— Yogi Berra

ERCOT Large Load Working Group · 21 August 2026 · Praneet Sharma, M.D., Founder & CEO

How AI compute behaves when voltage falls

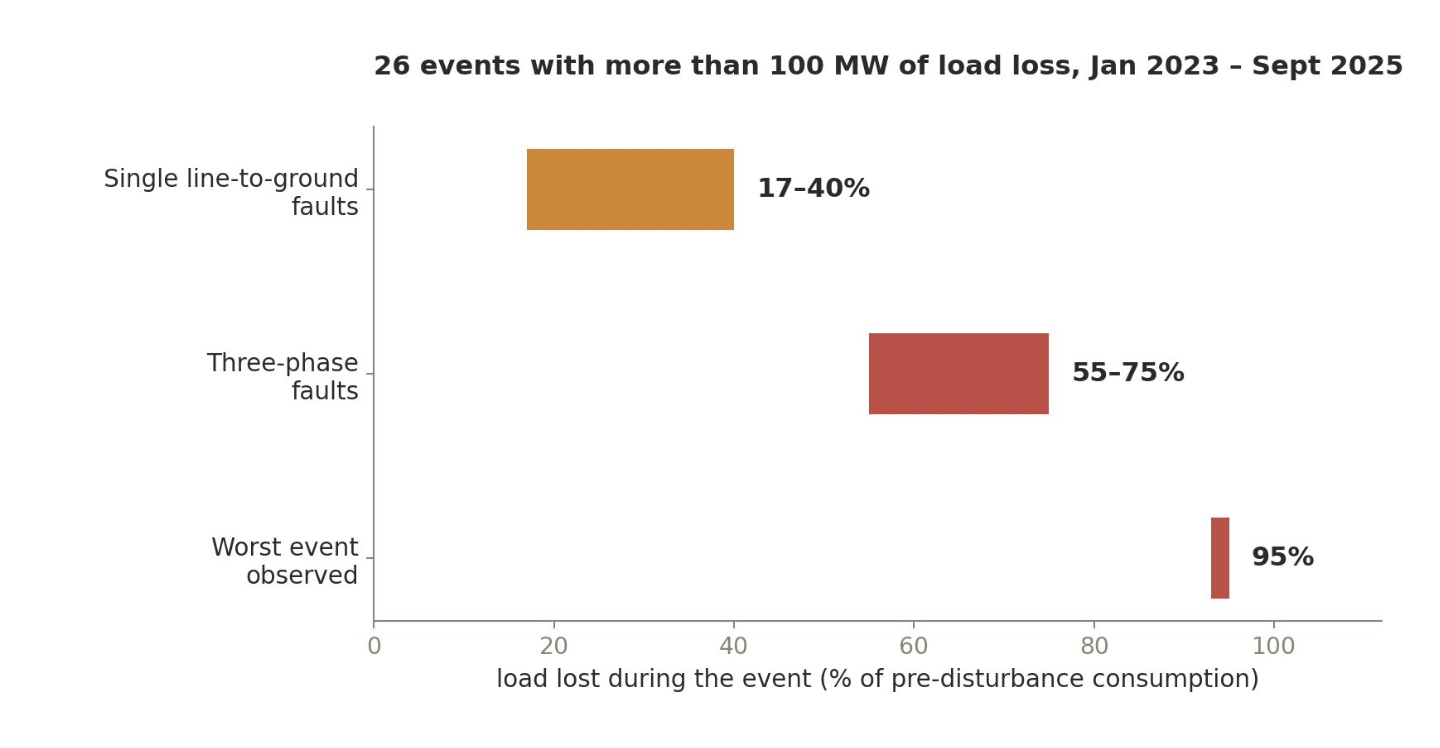
A GPU rack holds its power constant, so the current it draws rises as voltage drops.



A load that holds its power constant reaches 150% of normal current at 0.667 retained voltage. Below that, an unmodified load exceeds any 150% limit.

This is already happening in ERCOT

NERC reviewed 26 large-load ride-through events on this system in under three years.



What the review found

Constant-power supplies raise current as voltage falls — some facilities tripped their own breakers on overcurrent during a 50 ms dip

Current reached 150–160% of pre-fault at facilities studied

One event lost 19% of load from a dip to only 0.844 pu

Recovery ranged from five minutes to nearly two hours

NERC Event Analysis, Incident Review: Voltage-Sensitive Crypto Load Reductions, 9 January 2026. The facilities reviewed largely lack UPS; AI data centers typically have it, and that stored energy is what load-side control complements.

What NOGRR282 §2.15 asks of a large computational load

In force since 1 August 2026, as a condition of electrical service.

During a voltage sag

1. Stay connected and consuming for a set time — 2 seconds at 0.85, down to 150 milliseconds below 0.35
2. Keep current at or below 150% of the facility's normal current draw, measured at the point of interconnection
3. Return to at least 90% of normal within 2 seconds of the voltage recovering
4. Withstand at least six sags in 90 seconds

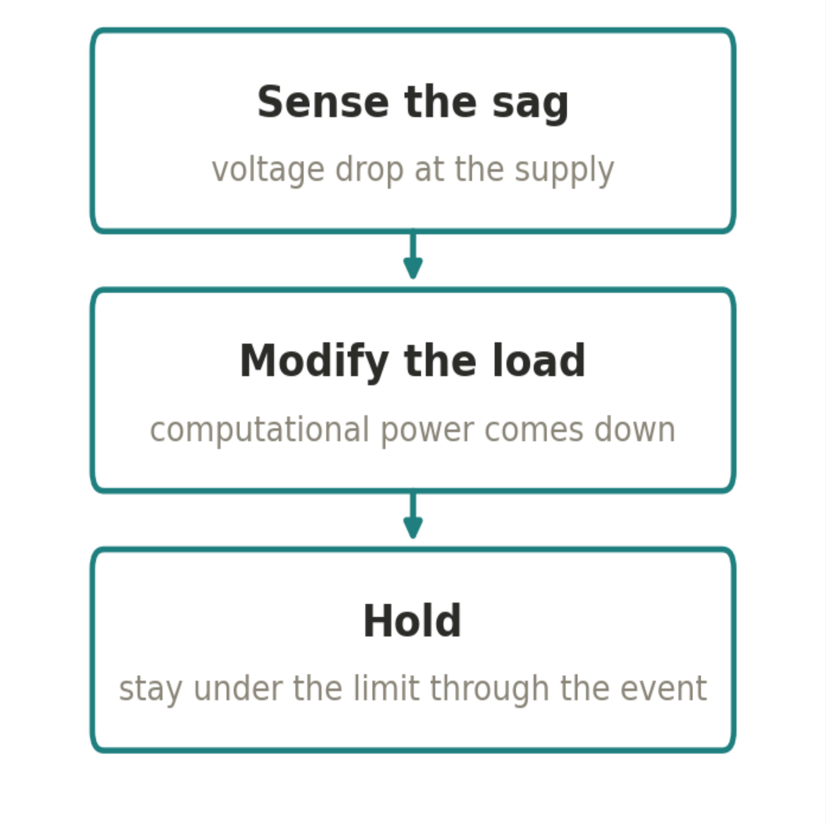
The rest of this presentation walks the evidence against this list.

ERCOT Nodal Operating Guide §2.15, established by NOGRR282, approved by the PUCT 9 July 2026.

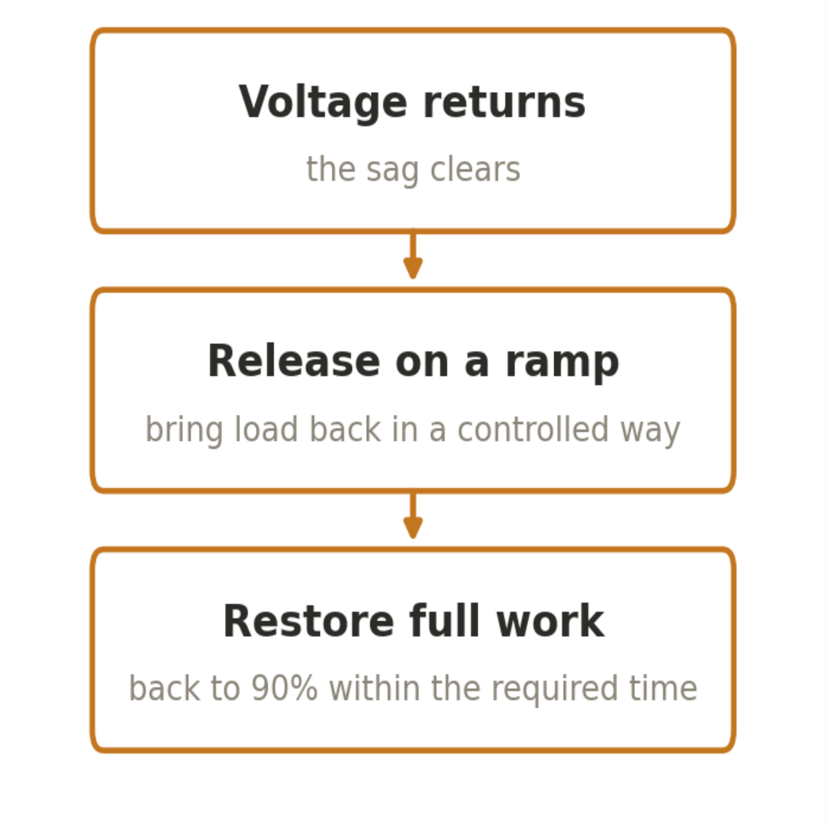
What AionLink does, and what we tested

Software that lowers computational power when the voltage drops, and brings it back when the voltage returns. Tested on production workloads through seventy voltage sags.

Responding to the sag



Recovering after it clears



Two MI300X GPUs running a BERT training job and a Qwen2.5-7B inference-serving job together — standard reference implementations. Seven depths, ten repetitions each, control off vs on.

How the electrical results were produced

The measured load data was applied to an independent model of a 9 MW facility.

AIONTRACKS supplied the measured load response. An independent engineer produced the electrical results using the open-source Data Center Model Library for PSCAD, modelling a 9 megawatt facility on a 69 kilovolt bus.

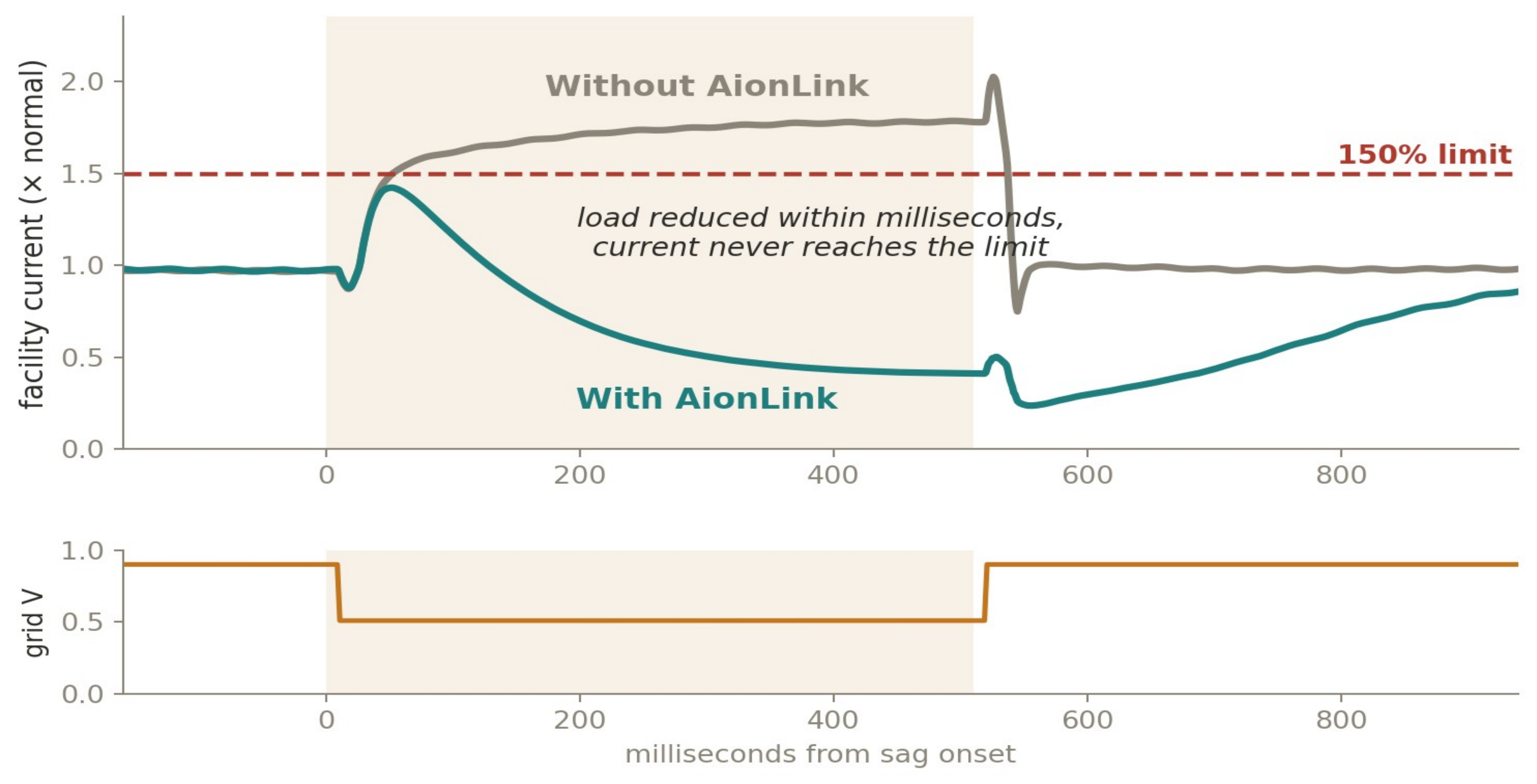
measured GPU load → independent 9 MW facility model → grid-side voltage, current, and internal supply behaviour

The author thanks Brett Ross of Pacific Northwest National Laboratory (PNNL) for providing advice on electromagnetic transient modeling and power electronics.

The model reports what the electrical quantities do. Its protection elements were disabled by design, so it does not report whether a supply trips.

The result at a moderate sag

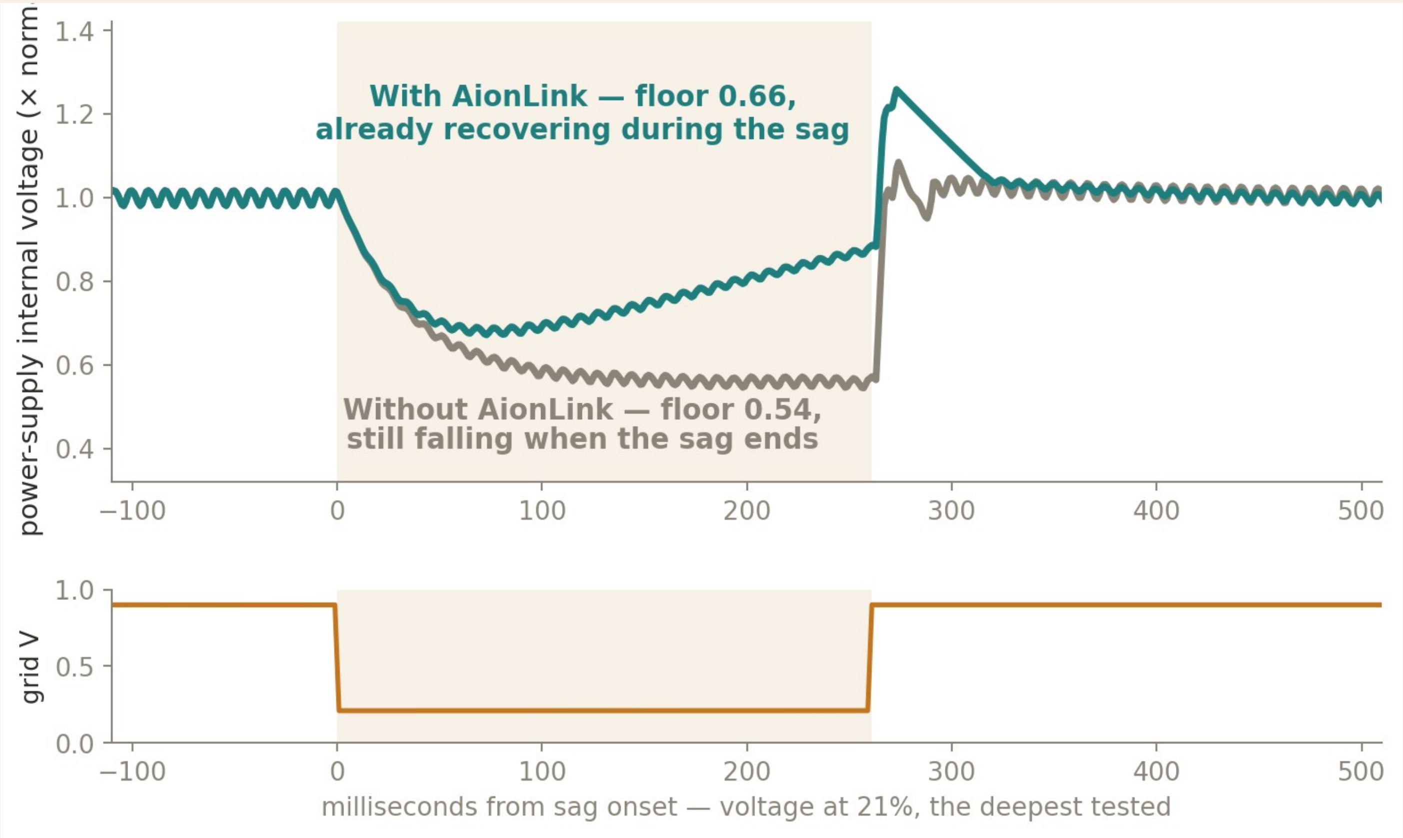
Voltage at 51% for half a second — the depth-class where most real events live.



The uncontrolled trace flattens near 2.0 because each power supply limits its own current there. Left alone the load would demand 2.8 to 4.2 at deeper sags — the supply's limiter caps what reaches the grid, but well above the 150% the rule allows.

The deepest sag, inside the power supply

Where the supply’s own limiter binds both cases, the internal voltage is what separates them.



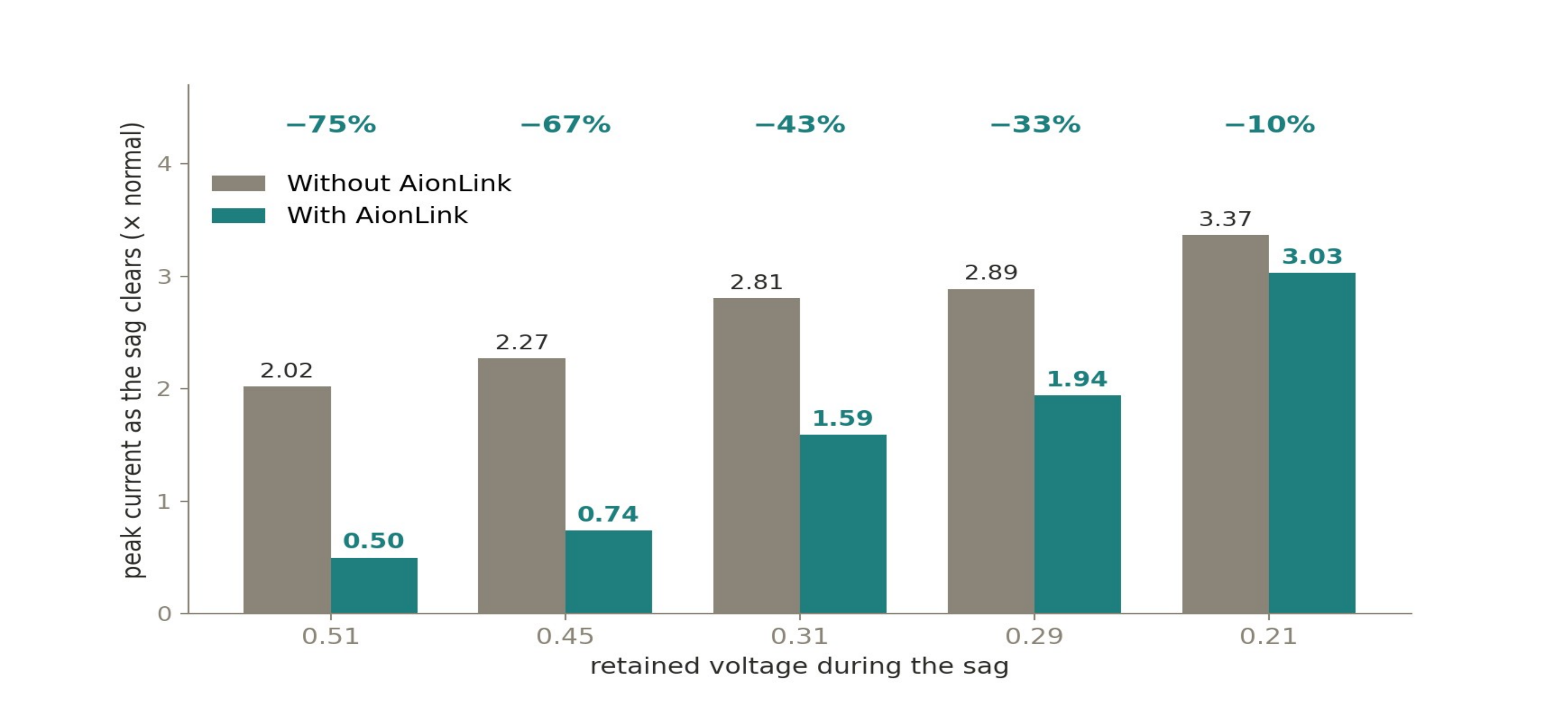
What this shows

- Internal voltage floor lifted from 0.54 to 0.66
- The controlled supply turns upward and recovers while the fault is still on the system
- The uncontrolled supply is still falling when the sag ends — it has no equilibrium, only a clock
- Whether either trips depends on the supply’s protection threshold and delay, which this model does not contain

The controlled supply reaches its floor and begins recovering about 50 ms later, during the sag. This is the quantity that determines whether a power supply stays up.

The surge when the sag clears

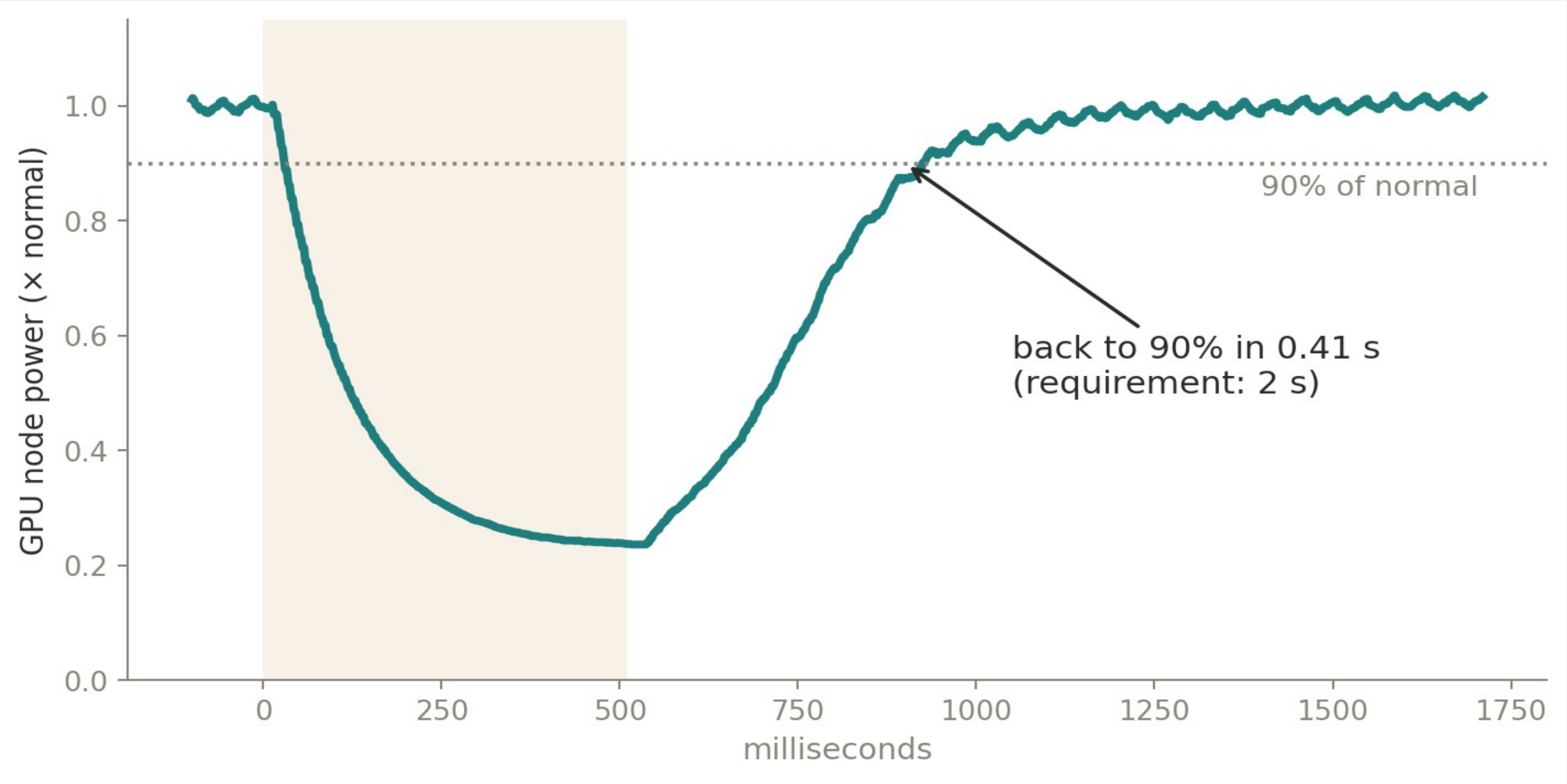
This is the mechanism that has actually been tripping facilities in ERCOT.



NERC’s January 2026 review records facilities tripping their own distribution breakers on overcurrent, during the fault and on the inrush when it clears. Reducing the load reduces this surge at every depth.

Returning after the sag

Consumption comes back well inside the requirement, on a bounded ramp.



90% of normal consumption restored in 0.41 s median, 0.49 s worst case, across all seventy sags — and the load never disconnected. ERCOT's January 2026 study found better system performance when large loads reconnect within 2 seconds of voltage recovery.

The criteria, so far

✓ **Ride-through duration (Table A)**

met at all seven depths — the load stayed connected and consuming throughout

✓ **Current ≤ 150% (§2.15(3)(d))**

met to 0.45 retained voltage

✓ **Return to 90% within 2 s**

0.49 s worst case — met at every depth

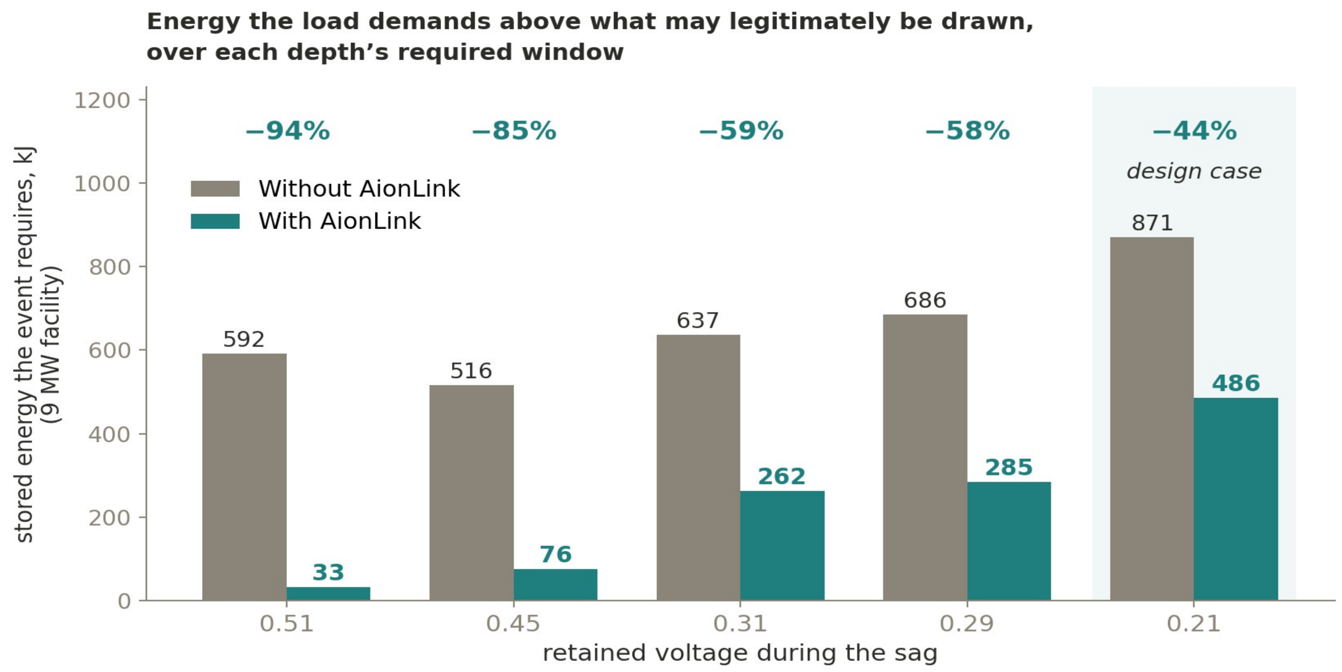
✓ **Withstand 6 sags in 90 s**

7 sags in 64 s — met

Deep-sag behaviour — below 0.45 retained voltage — on the next slide.

How much stored energy the requirement needs

And how much less of it when the load takes part.



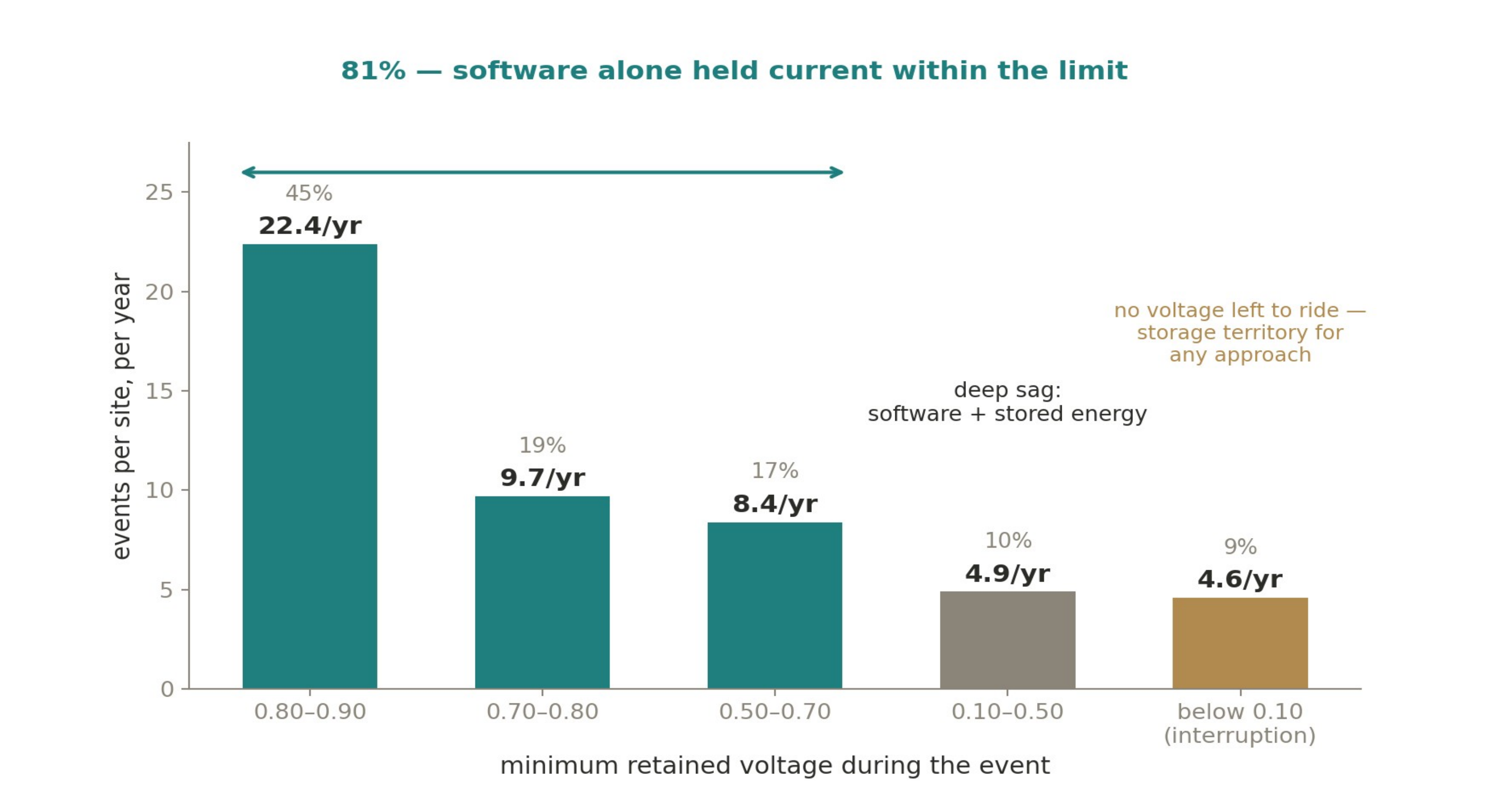
What this means

- 44% less at the deepest sag — the case a facility sizes storage for
- 385 kJ avoided on a 9 MW facility at that depth
- 85 to 94% less across the moderate depths, which are most real events
- Stored energy does not lower the current drawn at the meter — transferring to battery lowers it by dropping off, which is what the rule forbids

Integral of demand above the compliant draw (1.5 × V) over each depth's required ride-through window, scaled to a 9 MW facility. Storage is sized for the deepest event a facility must ride through, so the design case governs a capital figure.

How often these events happen

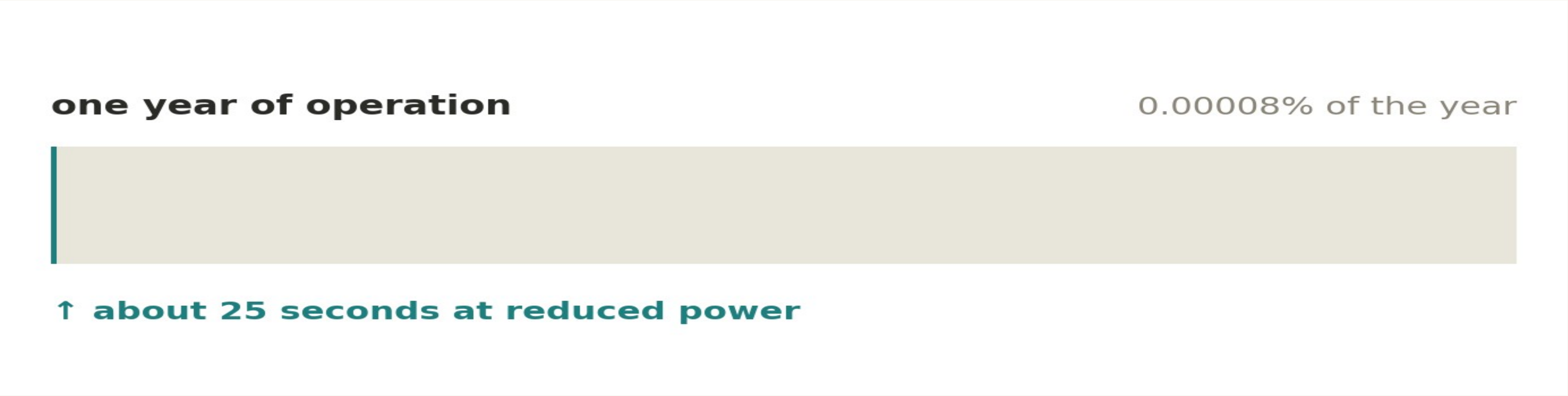
Four out of five sags fall in the range software alone covered. The unmet region is a minority of a minority.



EPRI DPQ survey, US distribution average per site, one-minute aggregation, 1993–95 — the only US survey with a published magnitude histogram; site-to-site variance is large and a transmission-served load generally sees fewer events. NERC's January 2026 review confirms shallow sags still cause real load loss today: one ERCOT event lost 19% of a facility at 0.844 pu.

What this costs in compute

Roughly fifty events a year, each lasting well under a second including the ramp back.



Through eight sags in a sealed test run, 21,907 inference requests were served with a median response time of 47.9 milliseconds. Service never stopped. Work is deferred by a fraction of a second, not lost.

About 50 sag events per year per site, each producing reduced power for the fault duration plus a recovery ramp of roughly 0.4 seconds.

For the data center operator

Cost in compute	About 25 seconds of reduced power per year — under 0.0001% of operating time
Capital avoided	Less stored energy to install, and existing supplies carry more of the requirement
Workloads preserved	21,907 inference requests served through eight sags, median response 47.9 ms, service never stopped
Broad hardware support	Validated on NVIDIA A100, H100, GH200 and AMD MI300X
Fast to deploy	Software on existing GPUs — no electrical work, no procurement lead time
Works with your UPS	An oversized UPS has the same current problem; reducing the load is what brings it within limit

Thank you

“If you ask me anything I don’t know, I’m not going to answer.”
— Yogi Berra

Praneet Sharma, M.D. · praneet@aiontracks.com · AIONTRACKS, Inc., Phoenix