EnergyHub

Building trustworthy virtual power plants: The VPP Maturity Model

A framework for building VPPs that perform like — and outperform — conventional resources

© 2026 EnergyHub, Inc. All rights reserved.

CONTENTS

© 2026 EnergyHub, Inc. All rights reserved.

EXECUTIVE SUMMARY

After a decade of flat demand, U.S. electricity load is **climbing** while transmission and distribution (T&D) **costs rise** and reliability concerns **increase**. Traditional demand response (DR) programs have helped, but their delayed telemetry, seasonal constraints, and inconsistent output mean they can't fully meet grid needs.

Virtual Power Plants (VPPs) are the fastest, most cost-effective way to add flexible capacity to the grid.

VPPs are **40–60% cheaper** than traditional infrastructure with $15–35B in annual utility savings potential by 2032. They aggregate and optimize distributed energy resources (DERs) into grid-responsive assets that can be deployed faster, at lower cost, and with more adaptability than conventional plants. This white paper introduces two tools:

1. **The Huels Test** — A pass/fail "imitation game" assessing whether a VPP is operationally indistinguishable from a conventional power plant.
2. **The VPP Maturity Model** — Five measurable, progressive levels spanning operational maturity, autonomy, and grid impact.

This framework is designed to help channel future investment by technology providers, device manufacturers, and utilities into the capabilities that will scale VPP value: tighter data loops, increased automation, and more precise control.

As VPPs progress up the maturity scale, they will become **cheaper to build**, **faster to deploy**, and increasingly able to **outperform traditional power plants** when and where it counts.

© 2026 EnergyHub, Inc. All rights reserved.

INTRODUCTION: IT'S TIME TO ASK MORE OF OUR VPPS

Load growth is back, wholesale and retail electricity rates are rising, and distribution infrastructure costs continue to climb. Regulators have warned that without additional capacity, blackout risks will increase. In short: we need *reliable* new capacity that can be measured, scheduled, and deployed when and where it's needed most.

Policymakers and researchers agree: VPPs are essential to solving these challenges. The U.S. Department of Energy estimates that the U.S. needs to add 160 GW of flexible capacity to meet forecasted demand, while recent studies show 100 GW of new load could be added to the grid if those new loads were willing to reduce demand just 44 hours per year .

In practice, skepticism among grid operators about VPPs remains high. Years of delayed telemetry and limited control have left many utility planners unwilling to treat VPP capacity on par with conventional generation. **Concerns typically stem from recurring gaps that undermine confidence:**

Performance reports with long delays

For many VPPs, operators must wait days or weeks for measurement and verification data. Without near-real-time performance feedback, it's impossible to make reliable operational or procurement decisions.

Fixed, short-duration events

A large share of VPPs still deliver only one load shed profile, often capped at two hours, without the flexibility to pre-position load or manage recovery. This restricts their usefulness in multi-hour, multi-stage dispatch scenarios common in today's markets.

Uncertain year-round availability

Seasonal participation rules, customer fatigue, and device-class constraints mean many portfolios vanish outside of summer afternoons. For planners tasked with meeting winter peaks or shoulder-season reliability needs, this unpredictability is a deal-breaker.

© 2026 EnergyHub, Inc. All rights reserved.

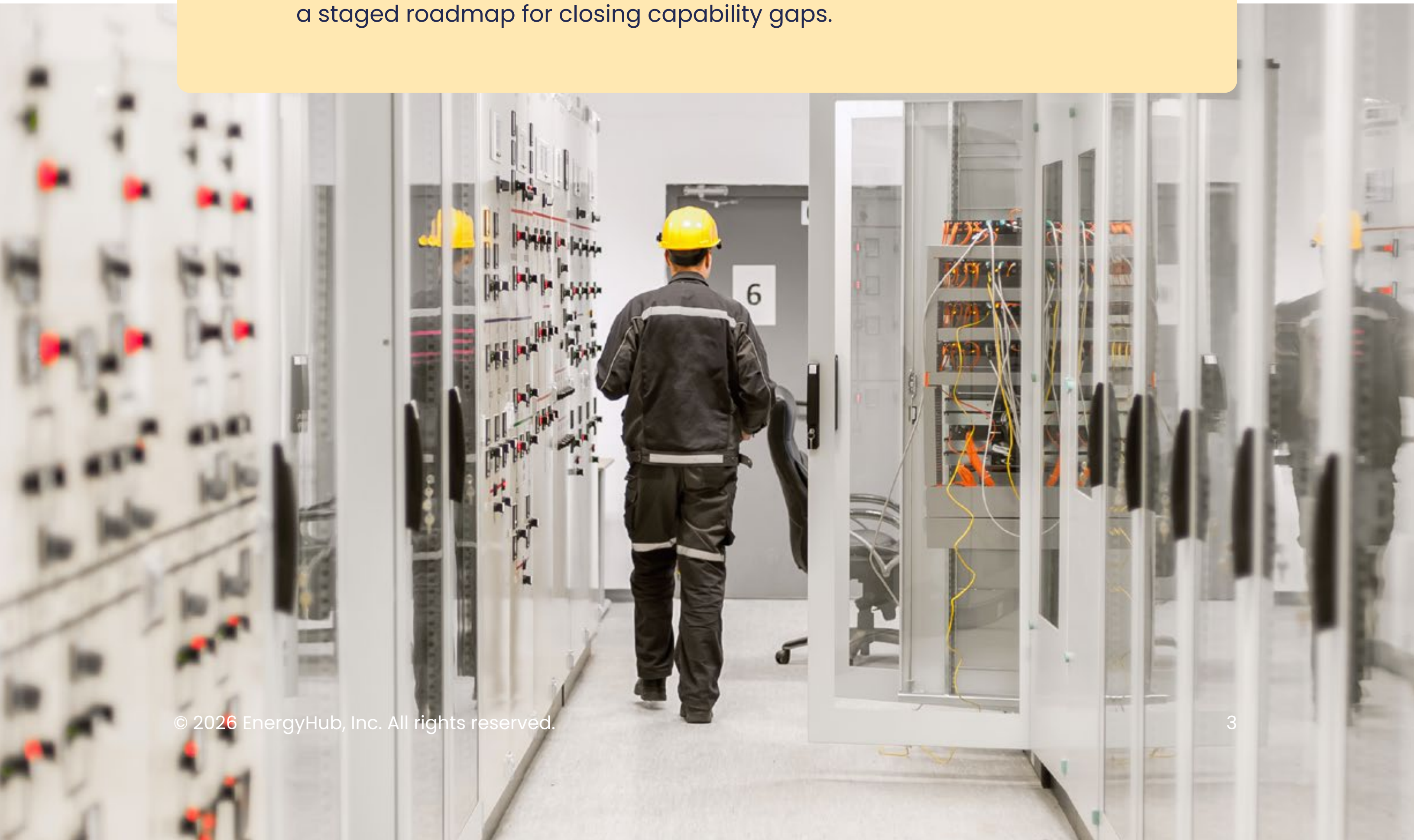
These issues persist for several reasons: OEM telemetry standards vary widely, integration with SCADA/EMS systems is inconsistent, and most portfolios were originally designed for legacy DR tariff structures rather than dynamic market participation. As a result, even technically capable DER fleets may be undervalued because they can't consistently meet operational expectations.

Markets like **CAISO** and **ERCOT** are now tightening VPP telemetry and verification standards, requiring faster data, higher accuracy, and proof of sustained availability. This raises the bar and provides a clearer definition of "good enough" for capacity accreditation.

A framework for advancing VPPs

The industry needs a clear way to (1) decide whether a VPP is "good enough" to be planned and dispatched like a traditional peaker plant, and (2) benchmark how far a given portfolio is from that bar — and what to do next. To that end, we propose two tools:

1. **The Huels Test** — A test of VPP viability that answers the question: ***What does it look like to replace traditional generation with a VPP?***
2. **The VPP Maturity Model** — Similar to those used in other industries, offering a staged roadmap for closing capability gaps.



© 2026 EnergyHub, Inc. All rights reserved.

THE GOAL: PASSING THE HUELS TEST

Our goal is to build VPPs so capable, predictable, and responsive that grid operators call on them as confidently as they would a conventional peaker plant. **The Huels Test** — named after EnergyHub Data Scientist Matthias Huels — is inspired by Alan Turing's "[Turing Test]()," which asked whether a software program could be labeled as "artificially intelligent" based on whether or not an ordinary user could distinguish a machine's responses from those of a human.

Similarly, the Huels Test asks:

> *If a grid operator is presented with two resources — one being a traditional peaker plant and the other a VPP — **can they tell which is which in operation?***

If the operator cannot tell the difference, the VPP passes the Huels Test and achieves parity — the point at which it can be planned, dispatched, and credited like a traditional plant.

To reach parity, a VPP must match the **visibility**, **schedulability**, and **availability** of traditional grid resources over sustained periods and under real operations conditions, as described on the next page.

© 2026 EnergyHub, Inc. All rights reserved.

Three pillars of VPP parity:

1. **Visibility**
   The VPP should deliver accurate, high-frequency telemetry (e.g., five-minute or faster) with minimal delay, directly to the control room. Post-event, next-day estimates don't provide the certainty operators need for real-time dispatch and system balancing.

2. **Schedulability**
   The VPP must be able to follow complex schedules — including pre-positioning to build load before an event, flat reduction periods, and gradual recovery that avoids snapback surges. A portfolio that cannot dynamically meet grid needs cannot support the variety of dispatch profiles required in today's markets.

3. **Availability**
   Operators require multi-season, multi-hour flexibility windows — typically at least four to six hours long — across summer, winter, and shoulder seasons. A resource that is limited to a single season or short two-hour events cannot reliably substitute for conventional capacity.

Passing the Huels Test does more than demonstrate technical competency — it unlocks planner and operator confidence, market eligibility, and capacity crediting that approaches or matches conventional resources.

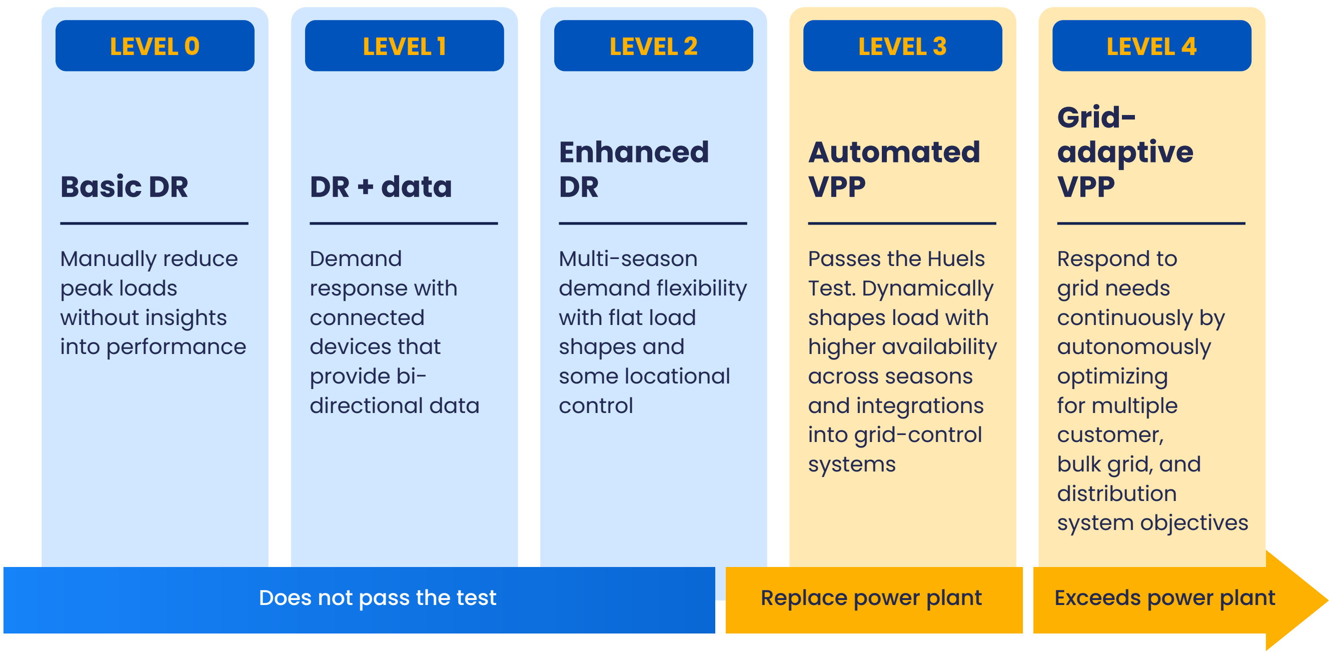
© 2026 EnergyHub, Inc. All rights reserved.

HOW TO MEASURE PROGRESS: THE VPP MATURITY MODEL

If the Huels Test tells us what success looks like, the VPP Maturity Model helps explain where we are now and what improvements need to be made to move forward. We won't be able to go from today's event-based demand response programs to exceeding power plant performance overnight. Instead, VPPs will progress through identifiable stages as data gets faster, automation gets smarter, and control becomes more precise.

The VPP Maturity Model captures those stages in five distinct levels, so planners, regulators, and program managers can evaluate their current reality, communicate gaps with precision, and fund the upgrades that matter most.

Figure 1: VPPs outperform a peaker plant as capabilities mature



© 2026 EnergyHub, Inc. All rights reserved.

What does VPP maturity mean for grid value?

At each level of the Maturity Model, the grid-services value that a VPP can deliver — per unit of capacity — goes up.

Consider a utility[1] with a combined generation and transmission capacity cost of about $100/kW-year and operating reserves purchased at $30/kW-year. A **Level 4 VPP** can deliver six-hour net load reduction events according to operator-defined schedules with near-perfect reliability, and dispatch almost instantly, delivering the full $130/kW-year in bulk grid capacity value.

Beyond that it can automatically mitigate overloads in distribution systems, adding roughly $50/kW-year in deferral value, and perform daily load shifting in response to market prices and customer rates, contributing perhaps another $30/kW-year in energy cost savings.

$130/kW-year
Full bulk capacity value

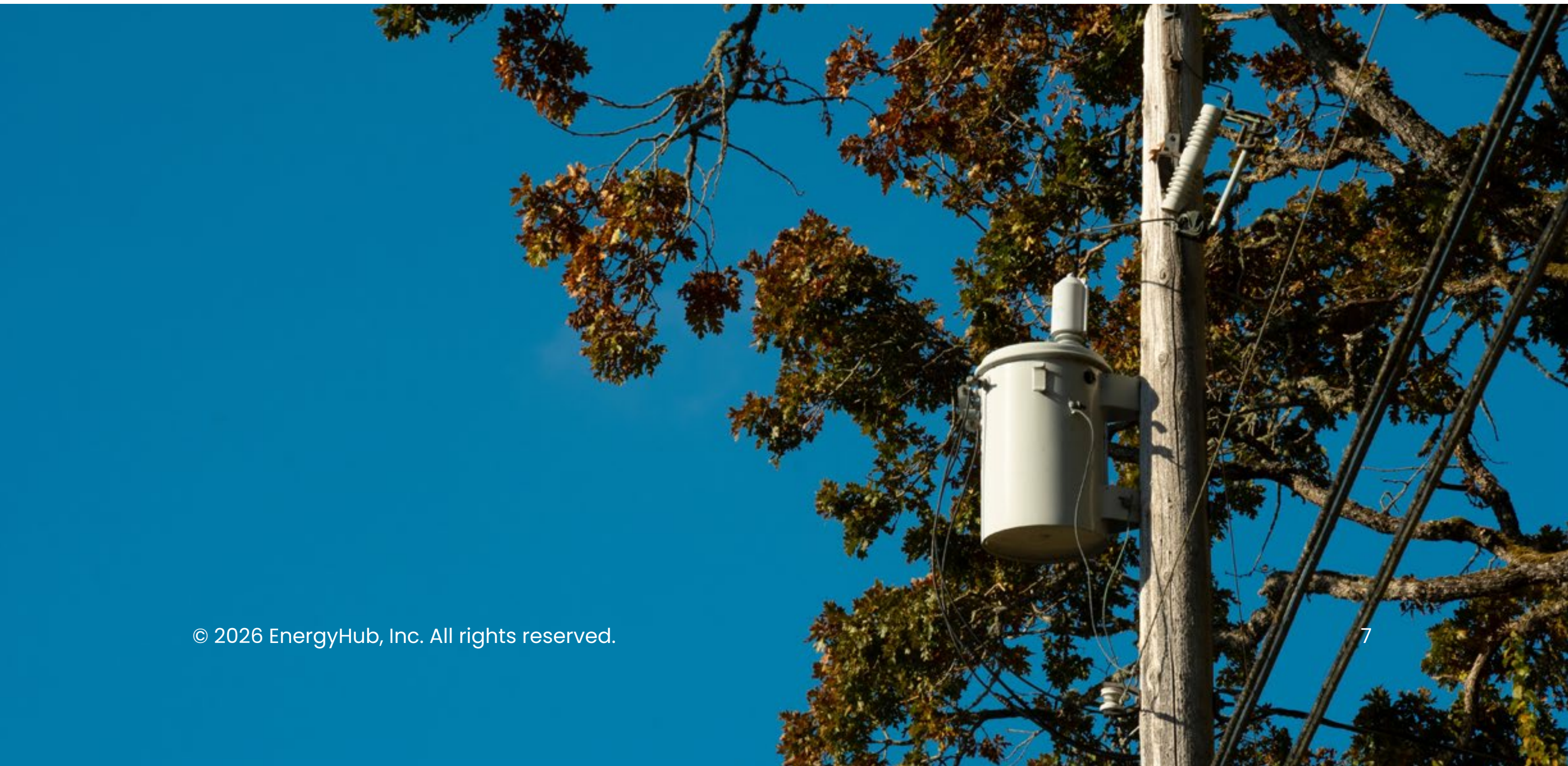
+

$80/kW-year
Additional distribution and energy savings

=

In total, a Level 4 VPP delivers about
$210/kW-year

[1] While the precise values here are illustrative, they align with values published in prior studies, such as **California's Virtual Power Potential**, **New York's Value of Distributed Energy Resources** proceeding, and **Xcel Colorado's AVPP tariff**.

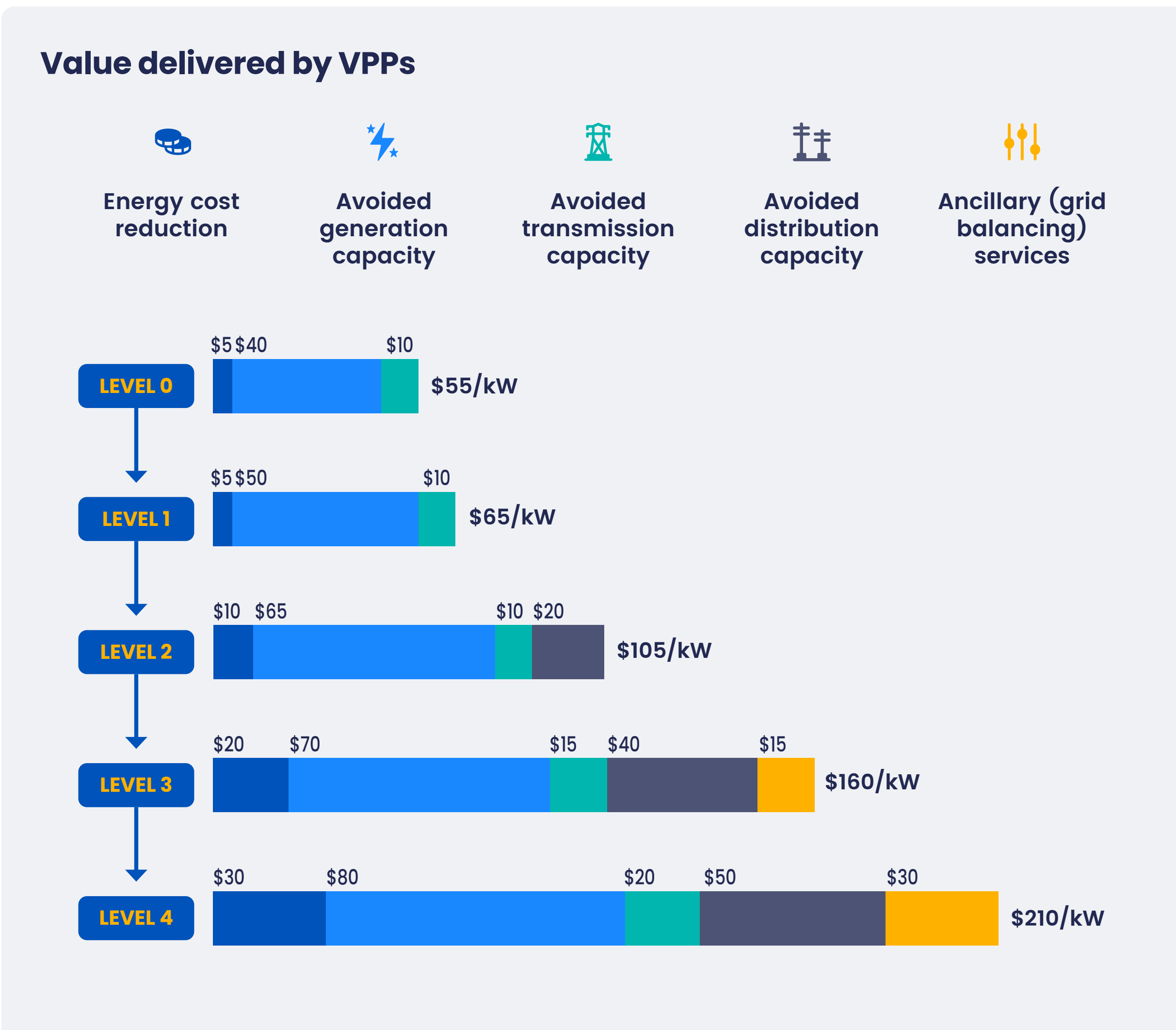


© 2026 EnergyHub, Inc. All rights reserved.

In contrast, a **Level 0 VPP** is unlikely to earn grid operators' full trust. DR events for C&I assets require up to six hours of advance notice and are typically limited to two or three hours in duration. Legacy direct-load control switches for air conditioners or water heaters can be triggered via FM radio signal, but the fraction still operational is often unknown, forcing utilities to keep significant thermal capacity on standby in case the program under-delivers.

Such constraints drastically reduce planning value. Grid planners often derate a Level 0 VPP's transmission and generation deferral benefits to roughly half the full capacity cost, and estimate only minimal energy cost reductions — since the 10 DR events per year shift load away from high-cost hours but have limited reach and flexibility.

Figure 2: Illustration of the increasing value of VPP capacity resulting from increasing VPP maturity. While the exact numbers will vary by location, these values align with prior studies.



© 2026 EnergyHub, Inc. All rights reserved.

VPP Maturity Model levels explained

LEVEL 0

Basic demand response

Level 0 represents the earliest stage of demand response — before connected devices enabled two-way communication. Resources here include phone-call DR, behavioral DR, and one-way load control switches.

A Level 0 VPP can reduce coincident peaks, but there is little to no feedback about actual performance, so planners can't confidently substitute these megawatts for conventional transmission or generation capacity. Behavioral DR typically requires six to 24 hours of notice, responses vary event to event, and without Advanced Metering Infrastructure (AMI), grid operators struggle to validate performance. One-way switch programs add some broadcast automation, but still offer no device-level feedback, leaving operators without visibility into either event success or device health.

Characteristics of a Level 0 VPP

	Dimension	Metric
OPERATIONAL MATURITY	**Passes the Huels Test?**	No
OPERATIONAL MATURITY	**Availability**	<10 short duration events per year
OPERATIONAL MATURITY	**Schedule tracking**	No ability to customize
OPERATIONAL MATURITY	**Dispatch latency**	>6 hours
OPERATIONAL MATURITY	**Telemetry**	None
OPERATIONAL MATURITY	**Data accuracy**	N/A
OPERATIONAL MATURITY	**Forecast accuracy**	N/A
OPERATIONAL MATURITY	**Integrations**	None
AUTONOMY		Manual event scheduling
GRID IMPACT		Only all-call events. No control over distribution impacts

© 2026 EnergyHub, Inc. All rights reserved.

LEVEL 1

Connected demand response

In 2012, EnergyHub launched the first demand response program that used Internet-connected thermostats with bidirectional communications, breaking away from switches and phone calls. These Bring Your Own Thermostat® (BYOT) programs enabled utilities to call DR events and quickly see how devices responded, even without AMI interval data. Because the system provides direct feedback, these resources can rightfully be classified as VPPs, not merely DR programs.

A Level 0 VPP can reduce coincident peaks, but there is little-to-no feedback about performance in real time. Limits on how many events a utility may call are common — often no more than 20 two-hour events per season — and the vast majority of Level 1 VPPs are active only in the summer months. Duration limits and seasonal constraints significantly reduce flexibility and overall availability. While this stage is a clear improvement over Level 0, generation and transmission capacity value is still routinely discounted compared to conventional resources, such as a natural gas peaker plant with fully proven, high-quality telemetry.

Characteristics of a Level 1 VPP

	Dimension	Metric
OPERATIONAL MATURITY	**Passes the Huels Test?**	No
	Availability	<20 short duration events per year
	Schedule tracking	Can schedule only start and stop times
	Dispatch latency	>1 hour
	Telemetry	Hourly data, with >1-hour latency
	Data accuracy	Uncertain
	Forecast accuracy	Forecast available but unknown accuracy
	Integrations	None
AUTONOMY		Manual event scheduling
GRID IMPACT		Only all-call events. No control over distribution impacts

© 2026 EnergyHub, Inc. All rights reserved.

LEVEL 2

Enhanced demand response

At Level 2, VPPs are close to passing the Huels Test for straightforward use cases. This stage introduces the ability to execute shaped schedules that go beyond simple start/stop events. Operators can, for example, specify a two-hour period of load building (pre-positioning), followed by a flat load reduction for four hours, and then a gradual two-hour recovery. Availability typically expands to multiple seasons (e.g., summer and winter), and both telemetry and forecasting improve. Dispatch can also account for customer requirements, such as time-of-use (TOU) rate periods.

Level 2 also marks the transition to partial autonomy. Some grid operators may choose to integrate a VPP into their centralized economic dispatch unit commitment systems. Others may prefer a VPP that self-dispatches based on operational signals — for instance, monitoring day-ahead energy market prices or system load levels and triggering a DR event when load or prices exceed a set threshold.

Characteristics of a Level 2 VPP

	Dimension	Metric
OPERATIONAL MATURITY	**Passes the Huels Test?**	20+ longer duration events in each of multiple seasons per year
OPERATIONAL MATURITY	**Availability**	Can deliver a sustained, flat net load reduction over ~4 hour period
OPERATIONAL MATURITY	**Schedule tracking**	Load shaped events with managed load build and recovery periods.
OPERATIONAL MATURITY	**Dispatch latency**	<15 minutes
OPERATIONAL MATURITY	**Telemetry**	15-minute interval data with <15 minute delay
OPERATIONAL MATURITY	**Data accuracy**	<10% error
OPERATIONAL MATURITY	**Forecast accuracy**	<15% error
OPERATIONAL MATURITY	**Integrations**	Possible integration with EMS
AUTONOMY	**Peak management**	Trigger event from a system load signal
AUTONOMY	**Market prices**	Trigger event from ISO/RTO prices
AUTONOMY	**Customer data**	Avoid increasing load during TOU on-peak periods
AUTONOMY	**Multiple grid services**	Manually select between services
GRID IMPACT		Some ability to reduce load using targeted locational groups of devices, but groups are static

© 2026 EnergyHub, Inc. All rights reserved.

Today, most advanced VPP resources operate at or near Level 2, making this stage a critical benchmark before the industry can widely achieve higher levels of automation, duration, and year-round reliability.

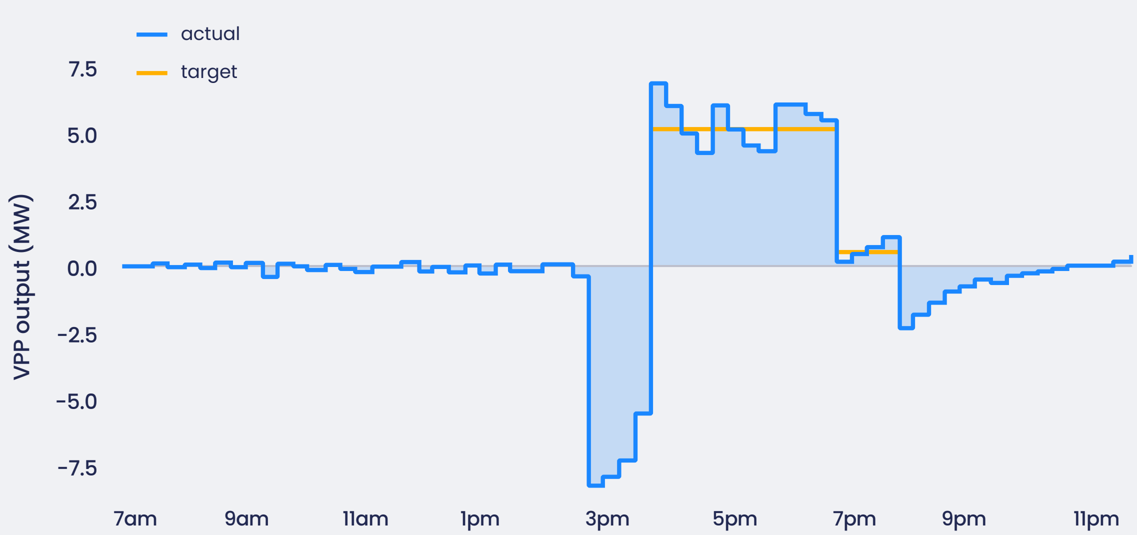
CASE STUDY

Intelligent load shaping and year-round reliability

One Mid-Atlantic utility has demonstrated what Level 2 performance looks like in practice. It tested whether its VPP could combine multiple short dispatch windows into a single, longer event that maintained a predictable, sustained load shape.

By linking back-to-back two-hour dispatches for targeted groups of devices, the resource delivered a three-hour event with a flat net load reduction — closely matching the operator-specified schedule and offering early evidence of passing the Huels Test for simpler load profiles.

Figure 3: Average VPP output compared to target



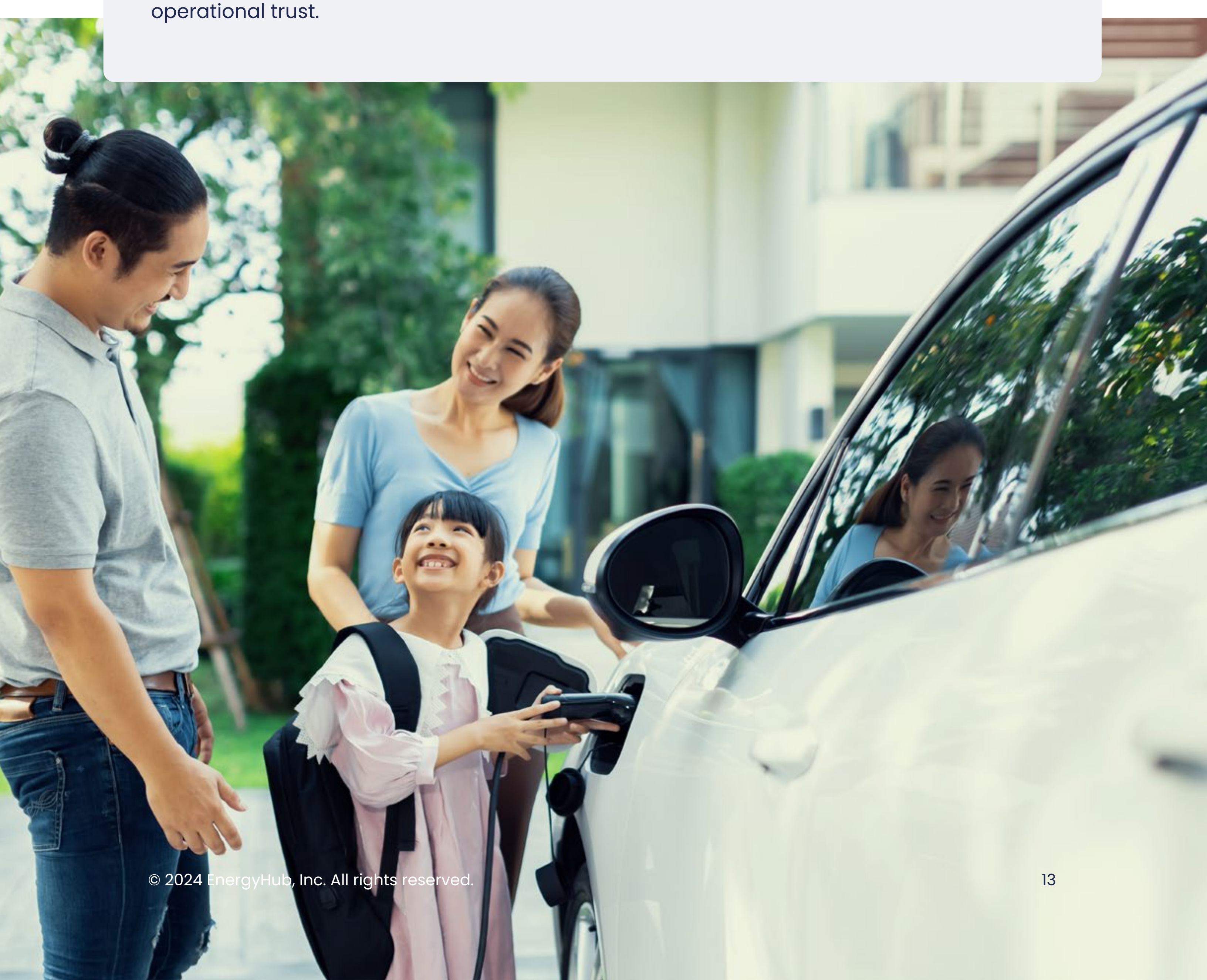
© 2026 EnergyHub, Inc. All rights reserved.

CASE STUDY

The utility has also validated **multi-season delivery** by running both summer and winter events, including a pilot where events were triggered daily in an effort to shift energy usage to less carbon-intense periods.

Regarding telemetry, the utility is making progress toward the 15-minute streaming expectations associated with Level 2. Development work is underway to establish the identifiers and data pathways needed to map device-level telemetry to specific dispatch intervals. With these capabilities in place, the portfolio will be positioned to deliver the visibility of a Level 2 VPP.

Together, these results demonstrate how a Level 2 VPP can reliably track shaped schedules, perform across seasons, and provide enhanced telemetry — key steps toward earning full operational trust.



© 2024 EnergyHub, Inc. All rights reserved.

LEVEL 3

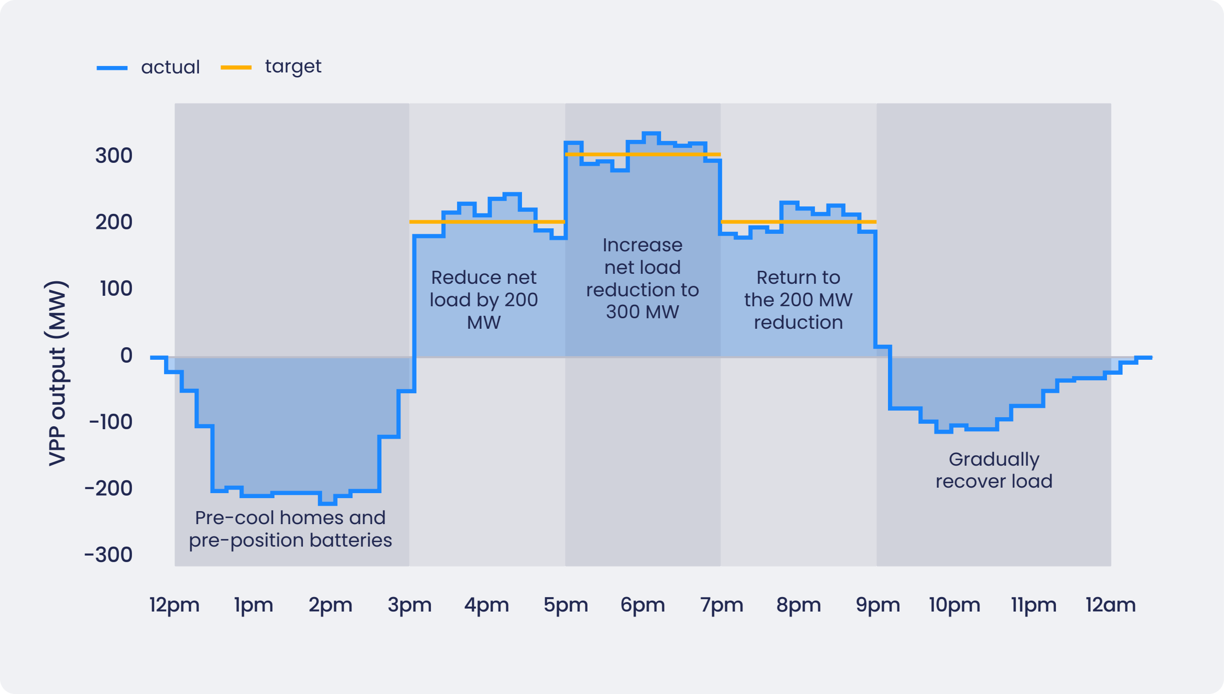
Automated VPP

While the ability to call 40, four-hour events per year is valuable, a Level 2 VPP still only uses DERs for grid services less than two percent of the time. As variable wind and solar penetration rises, and as new load from data centers and electrification adds stress to the grid, operators will need resources that provide more than peak reduction — they will need year-round grid balancing.

Consider electric vehicles and heat pumps. If we use these devices merely for demand response and load shifting for TOU rates, the snapback after high-cost periods can create additional stress on the distribution network, compounding infrastructure needs. To avoid this and capture the full value potential, DERs must operate year-round with automated, multi-signal responses to grid conditions across generation, transmission, and distribution layers. VPPs at Levels 3 and 4 are designed to meet this need.

To be a fully capable substitute for a conventional power plant, a Level 3 VPP must support extended events — typically six hours or longer — with precise scheduling. For example, a utility preparing for a peak on a hot August afternoon might anticipate it between 5:00 PM and 7:00 PM but plan for it to arrive earlier or later. To manage the entire peak load afternoon period, they could orchestrate the following schedule:

Figure 4: Average VPP output compared to target



Utilities on the EnergyHub platform have **[already achieved this level of precise load shaping](#)**.

© 2026 EnergyHub, Inc. All rights reserved.

It is highly unlikely that single-resource VPPs (e.g., only thermostats or only batteries) could accomplish this. Level 3 VPPs will almost always be multi-DER portfolios — batteries, thermostats, C&R DR, and EVs — dispatched through a DERMS capable of combining these assets to meet the aggregate target schedule.

High-quality data is critical. Level 3 operation requires streaming telemetry from the DERMS into the control room at five-minute intervals with less than five-minute delay to interact with markets operating in five-minute settlements. While many batteries can meet this target, thermostats and EVs may require innovation to achieve equivalent visibility.

By this stage, many portfolios will operate year-round and automatically, co-optimize for customer TOU rates, wholesale costs, and operator-set constraints. With accurate data, flexible schedules, and continuous reliability, operators can use Level 3 VPPs in much the same way that they use conventional power plants: meeting the necessary information, scheduling, and dependability standards to pass the Huels Test.

Characteristics of a Level 3 VPP

	Dimension	Metric
OPERATIONAL MATURITY	**Passes the Huels Test?**	Yes.
	Availability	100+ 6-hr events. Can take actions in any season.
	Schedule tracking	Follow daily, hourly schedules
	Dispatch latency	<5 minutes
	Telemetry	5 minutely data with <5 minute delay
	Data accuracy	<5% error
	Forecast accuracy	<10% error
	Integrations	Yes, integrated with SCADA and/or EMS
AUTONOMY	**Peak management**	Peak forecasting + optimal event times
	Market prices	Optimize for day-ahead prices
	Customer data	Optimization for retail rates or customer comfort
	Multiple grid services	System recommends actions to user
GRID IMPACT	**Targeting**	Hierarchical DER groups periodically synced with grid models
	Locational forecasts	Available, with significant uncertainty
	Grid optimization	Optimize for static grid limits

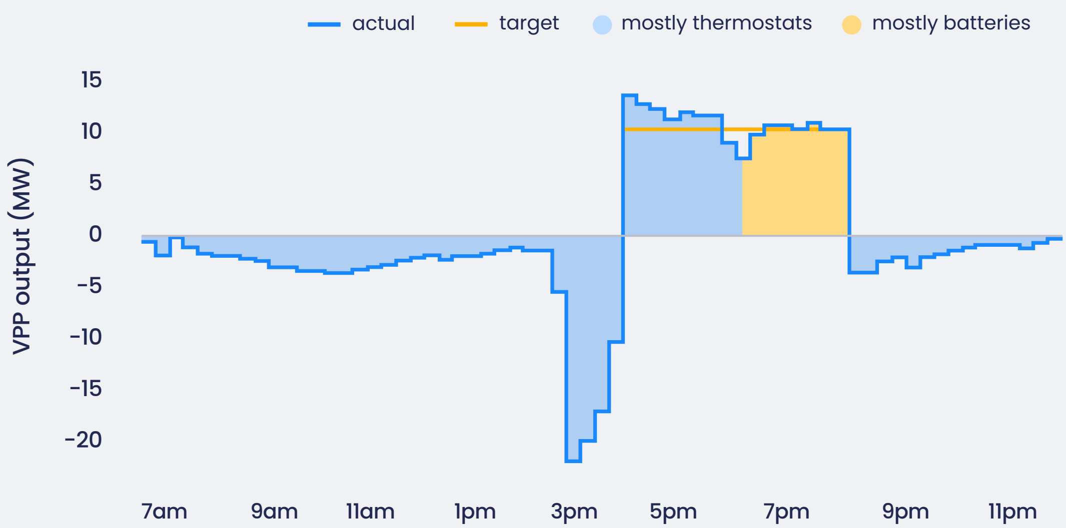
© 2026 EnergyHub, Inc. All rights reserved.

CASE STUDY

National Grid pilots cross-DER orchestration

National Grid illustrates how combining multiple DER types can extend event duration and refine load profiles — a foundational capability for Level 3 VPPs. In this pilot, National Grid dispatched thermostats and residential batteries in combination, producing a **sustained four-hour net load reduction** with a noticeably flatter shape than either could achieve alone. Although the event fell short of the six-hour benchmark expected at Level 3, the test demonstrated the operational benefits of multi-DER coordination and the potential for longer, more reliable events in the future.

Figure 5: Average VPP output compared to target (National Grid)



National Grid has also tested **locational dispatch** — dispatching targeted groups of devices to meet localized grid needs — as part of its non-wires alternative (NWA) program. In this work, grid operators dispatched separate groups of devices on Nantucket and on the mainland in parallel, demonstrating the ability to target resources geographically and produce **regulator-ready reporting**. This capability marks early progress toward the **hierarchical grouping** needed for full Level 3 VPP operation.

Together, these experiments show how utilities can begin combining heterogeneous DER profiles to produce longer events, targeted impacts, and higher-confidence reductions — all building blocks of Level 3 capabilities.

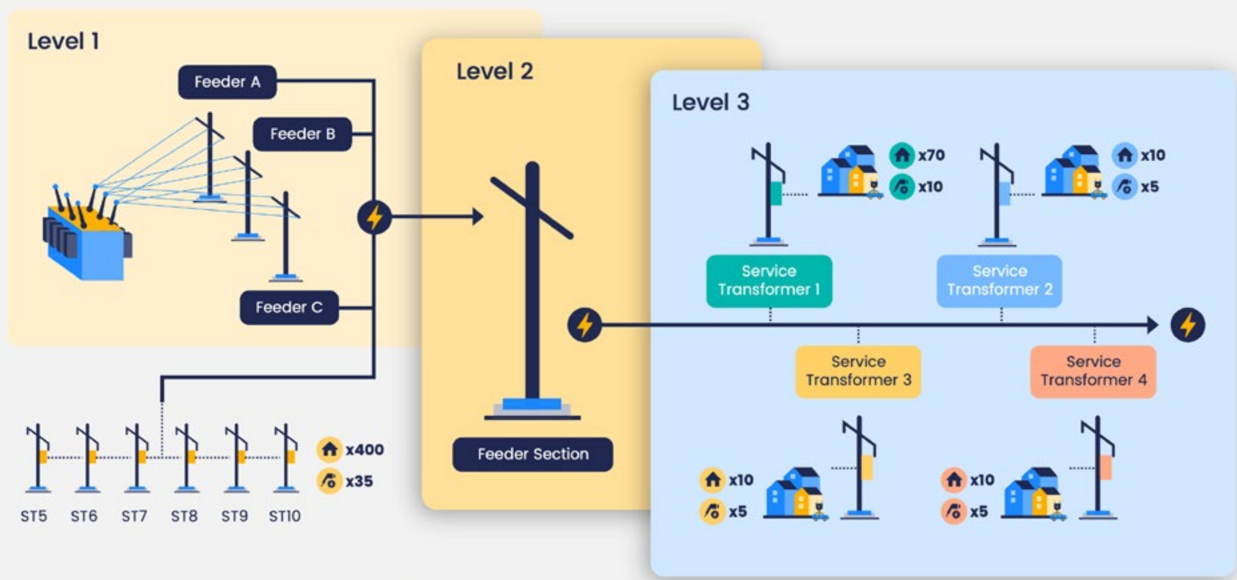
© 2026 EnergyHub, Inc. All rights reserved.

CASE STUDY

Multi-level distribution load optimization with EVs

Utilities on the EnergyHub platform are showing EVs can operate as automated, year-round grid assets. Through **multi-level distribution load optimization**, EVs on the platform are grouped hierarchically based on grid topology to respond to **simultaneous grid constraints** (e.g., system peak hours, feeder limits, and service transformer capacity) while meeting customer charging needs.

Figure 6: EVs are connected to multiple levels of the distribution grid



Across these deployments, results show:

- **95% off-peak charging,** among the highest rates achieved by any EV program in the country.
- **50%+ peak load reduction in load** on constrained assets during the top 1% of peak hours, easing stress on feeders, transformers, and substation equipment.
- **100% charging fulfillment** for drivers plugging in with adequate time, confirming that grid-aware charging can scale without creating customer pain points.

These outcomes prove EVs can deliver continuous, multi-signal optimization, a key indicator of VPP maturity — reducing load when and where it matters most, protecting distribution infrastructure, and maintaining customer satisfaction without manual intervention.

© 2026 EnergyHub, Inc. All rights reserved.

LEVEL 4

Grid-integrated VPP

At Level 4, VPPs operate as fully grid-aware, autonomous assets, able to self-dispatch 24/7/365 across the full value stack:

- **Generation capacity**
- **Transmission capacity**
- **Distribution capacity**
- **Ancillary services** frequency regulation, spinning reserves
- **Wholesale and retail energy costs**

These VPPs maintain complete situational awareness of transmission and distribution constraints, market conditions, customer rate structures, and renewable availability. This enables year-round, multi-service value delivery with reliability that exceeds that of a conventional plant.. Deep integration with utility and grid-operator systems is a core requirement at this stage. Typically, this involves:

- Edge DERMS — optimizing behind-the-meter resources
- Grid DERMS — combining AMI data with distribution system models to identify locations where DERs can solve for constraints

One way for these systems to work together is for the Grid DERMS to compute dynamic power limits (or "Dynamic Operating Envelopes") for groups of assets and send those limits to the Edge DERMS. The Edge DERMS can then use these grid limits as constraints in an optimization problem that co-optimizes for multiple conditions including:

- Distribution limits (via a dynamic operating envelope from the Grid DERMS)
- Wholesale energy prices from ISO/RTO markets
- System peak likelihood based on regional load forecasts
- Customer rates
- Availability of renewable generation

> When a VPP can meet each of these requirements — full-stack optimization, locational awareness, wholesale and retail market integration, and ultra-fast telemetry — it becomes a preferable substitute for conventional generation.

© 2026 EnergyHub, Inc. All rights reserved.

Customer data integration ensures dispatch decisions maintain alignment with retail rate designs and individual customer preferences. Ultra-low latency telemetry is another defining capability of Level 4. These systems can deliver two- to six-second telemetry and execute near-instant dispatch changes, enabling participation in second-by-second services such as frequency regulation. Pilot projects with single-OEM resources have proven the feasibility of this; scaling it across heterogenous, multi-OEM VPPs remains a key industry challenge.

Characteristics of a Level 4 VPP

	Dimension	Metric
OPERATIONAL MATURITY	**Passes the Huels Test?**	Yes
	Availability	364 days a year, 24-hours a day, with power/energy limits similar to a grid-scale battery
	Schedule tracking	Follows real-time grid signals
	Dispatch latency	<1 minute
	Telemetry	~2-6 seconds
	Data accuracy	<2% error
	Forecast accuracy	<5% error
	Integrations	Yes, integrated with SCADA, EMS, Grid DERMS
AUTONOMY	**Peak management**	Co-optimization for peak loads and other grid constraints using data feeds
	Market prices	Co-optimization for real-time prices and other grid constraints
	Customer data	Co-optimization for retail prices, customer comfort, and other constraints
	Multiple grid services	System can automatically select and perform appropriate grid services
GRID IMPACT	**Targeting**	Can dispatch any group of devices with any topology using real-time updates
	Locational forecasts	High quality with high degree of certainty
	Grid optimization	Optimize for dynamic grid limits

© 2026 EnergyHub, Inc. All rights reserved.

A ROADMAP TO LEVEL 4

While many OEMs and DERMS providers have demonstrated elements of the technology required for Levels 1–4 of VPP operations, much work remains to deliver this vision at scale.

1. Tighten the data loop

The biggest constraint most portfolios face is data quality. Telemetry that arrives late, at coarse intervals, or with wide error margins forces operators to discount the resource. Advancing maturity demands faster, more granular data with low uncertainty.

Key actions:

- Expanding OEM data-sharing agreements to improve granularity and speed
- Combining device data, AMI data, and machine learning models to enhance forecasts and telemetry
- Providing operators with real-time visibility via DERMS integrations with control-room software and operator-friendly visualizations

2. Automate for shape, not just shed

Most DR portfolios can shed load; fewer can sustain a shaped load profile for four or more hours. Achieving this requires device-level dispatch capability plus portfolio-level AI-driven optimizations.

Key actions:

- Deploying event-optimization engines capable of pre-positioning, multi-hour management, and gradual recovery
- Smoothing snapback for localized device groups to protect distribution assets and preserve net reductions
- Coordinating across multiple device classes (thermostats, batteries, EVs, C&I sites, etc.) so each contributes maximum value
- Respecting customer comfort, rate plans, and opt-out behavior in all schedules

© 2026 EnergyHub, Inc. All rights reserved.

3. Lower the customer barrier

A VPP is only as reliable as its active participants. Reducing friction at enrollment and during operation expands and stabilizes the portfolio, directly increasing its planning value.

Key actions:

- Streamlining enrollment with OEM integrations, partner marketing, and utility customer-data integration
- Supporting broad device and C&I site participation to widen program reach
- Designing incentives for long-term, reliable participation
- Using control algorithms that co-optimize for customer rates, comfort, and grid needs

4. Start local, then scale

Level 3 and particularly Level 4 capabilities hinge on locational awareness. Even the most well-orchestrated portfolio falls short if it cannot manage feeder or substation constraints. Integrating information about distribution systems is essential.

Key actions:

- Identifying stressed feeders with AMI data, SCADA trends, or planning models
- Quantifying deferral value to guide deployment decisions
- Beginning with simple locational dispatch (e.g., dispatching devices for constrained feeders via an Edge DERMS) and evolving toward sophisticated capabilities, like Grid DERMS integration
- Develop a long-term plan for integrating grid models and dynamic operating envelopes into the orchestration layer

© 2026 EnergyHub, Inc. All rights reserved.

5. Guide regulatory changes

Technical maturity and regulatory evolution are moving in the same direction. Preparing for Level 4 means shaping the rules for dependable DER participation. Early alignment avoids costly redesigns and secures higher capacity value as accreditation frameworks evolve.

Key actions include:

- Regulatory innovations to enable utilities to invest in increasingly mature VPP systems that deliver higher value grid services
- Monitoring emerging telemetry and verification standards that prioritize five-minute (and faster) data over M&V reports
- Enable flexibility in program design, such as allowing more frequent and longer events and multi-season availability
- Incorporating feeder/substation constraints into operations, anticipating formal recognition of distribution-level performance

6. Align key players

A VPP's maturity is ultimately constrained by its weakest link — often OEM telemetry quality or DERMS integration. Advancing maturity requires **coordinated improvement across device manufacturers, platforms, and utility systems**. This roadmap works best when all stakeholders share the same targets and measure progress through the same parity lens.

BUILDING A SHARED VISION

This white paper centers on a clear idea: grid operators and planners can fully rely on VPPs when they know exactly what to expect from them. **The Huels Test** defines that expectation; the **VPP Maturity Model** turns it into actionable milestones — stronger telemetry, greater availability, advanced automation, and distribution-aware control.

Progress along these milestones is not hypothetical — it is achievable with coordinated effort across utilities, OEMs, DERMS partners, and regulators. Success means building VPPs that operate with the same dependability, flexibility, and responsiveness as conventional grid resources, while delivering more value across the full stack.

If your organization is ready to benchmark its current capabilities, identify gaps, and plan a pathway to the next level of VPP maturity, we'd love to talk. Reach out to our team today to start the conversation.

EnergyHub manages more than **2.5M+ DERs** across all device classes, helping over **170 utility clients** improve the safety, reliability, and resilience of the grid.

EnergyHub

info@energyhub.com | www.energyhub.com

APPENDIX

System requirements for VPP maturity

	Availability	Scheduling	Telemetry	Accuracy	Grid Impact	Autonomy
LEVEL 0	Summer season	None	None	No guarantees	Bulk grid	Manual events
LEVEL 1	Summer season	Start and stop times	>1 hour	No guarantees	Bulk grid	Manual events
LEVEL 2	Multiple seasons	Flat load shed for 4-hours	15 minute	<10-15% error	Bulk grid & static locational groups	Some human oversight
LEVEL 3	Year-round	Follow daily, hourly schedules	5 minute	<5-10% error	Mapped to distribution grid	Minimal human oversight
LEVEL 4	24 hours, 365 days	Real-time response	~2-6 second	<2-5% error	Fully locational grid service	Fully autonomous across systems

© 2026 EnergyHub, Inc. All rights reserved.

VPP Maturity Model criteria

We organized the VPP Maturity Model around the way the control room operators experience these resources in real life. The model includes three evaluation categories, each with its own set of dimensions, and together they define what it truly means for a VPP to progress toward — and ultimately surpass — the Huels Test.

1. Operational maturity

“Can we trust this capacity?”

Operational maturity is the foundation of grid planner and operator confidence — the currency of grid management. Without it, even technically capable portfolios will be undervalued and underutilized. This category assesses whether a VPP delivers the reliability, precision, and consistency needed for it to be treated as a dependable grid resource.

Dimensions of operational maturity

OPERATIONAL MATURITY

Dimension	Description
Passes the Huels Test?	Can the VPP perform at a standard where a grid operator would consider it interchangeable with a conventional power plant?
Availability	Over what fraction of the year is the VPP available for grid services? Can it reliably support long, multi-hour dispatch events, or is it constrained to short durations?
Schedule tracking	Can the VPP follow a schedule or a grid signal in detail — including pre-positioning and staged recovery — or is it limited to fixed-shape DR events?
Dispatch latency	How quickly can the VPP respond to dispatch instructions and reach its new setpoint? Does it require day-ahead notice, or can it respond within seconds?
Telemetry	Does the VPP provide accurate, granular data for grid operators? At what interval (e.g., five-minute, hourly), and with what delay?
Data accuracy	How accurate is the data provided by the VPP, and how can operators verify it?
Forecast accuracy	Can the VPP offer precise, forward-looking estimates of capacity for operations planning?
Integrations	Can the VPP integrate seamlessly with core control room platforms, such as SCADA (supervisory control and data acquisition), CIS (customer information system), and EMS (energy management system)?

© 2026 EnergyHub, Inc. All rights reserved.

2. Autonomy & intelligence

"Will it do the right thing without a human intervening?"

This category evaluates whether a VPP can act intelligently in real time, adapting its operations to meet grid and customer needs without requiring manual scheduling for every action. Higher autonomy transitions a VPP from an occasional resource to an active, daily contributor to grid stability and economics.

Dimensions of autonomy

AUTONOMY

Dimension	Description
Peak management	Can the VPP automatically dispatch to reduce peak loads when system conditions warrant?
Market prices	Can the VPP dynamically respond to day-ahead and real-time wholesale electricity market prices?
Customer data	Can the VPP operate within customer-defined parameters such as comfort bands or rate plans?
Multiple grid services	Can the VPP autonomously choose or combine grid services to deliver optimal value at any given time?

3. Grid impact

"Does it deliver value where we need it?"

Grid impact is the ability to convert raw capacity into targeted, location-specific value. It's what turns "a lot of kilowatts somewhere" into avoided infrastructure costs, local reliability improvements, and true system efficiency.

Dimensions of grid impact

GRID IMPACT

Dimension	Description
Grouping	Can the VPP dispatch intelligent groups of DERs to alleviate local grid constraints?
Locational forecasts	Can the VPP forecast not only system-wide capacity, but also the localized potential to address distribution grid constraints?
Grid optimization	Can the VPP use transmission and distribution model data to identify and correct system constraints in real time?

© 2026 EnergyHub, Inc. All rights reserved.

EnergyHub

Full VPP maturity level criteria

		LEVEL 0	LEVEL 1	LEVEL 2	LEVEL 3	LEVEL 4
		Basic DR	**DR + Data**	**Enhanced DR**	**Automated VPP**	**Grid-adaptive VPP**
OPERATIONAL MATURITY	**Passes the Huels test?**	No	No	Almost	Yes	Yes
	Availability	~10 short events, one season	~20 short events, one season	~40 4-hr events, multi-season	100+ 6-hr events, year-round	24/7/365 operations
	Schedule tracking	Requested start/ stop times	Defined start/stop times	Events with flat load shed	Follow daily, hourly schedules	Follow real-time grid signals
	Dispatch latency	Hours	Hour+	<15 minutes	<5 minutes	<1 minute
	Telemetry	None	hourly with delay	15 minutely	5 minutely	~2-6 seconds
	Data accuracy	None	Uncertain	<10% error	<5% error	<2% error
	Forecast accuracy	None	Uncertain	<15% error	<10% error	<5% error
	Control room integrations	None	Minimal	Basic dispatch	EMS/SCADA for forecasts, dispatch, and telemetry	Level 3 + Grid DERMS/ ADMS
AUTONOMY	**Peak management**	Minimal	Manual events	Trigger from system load feed or system alert	Peak forecasting + optimal event times	Co-optimization for peak loads with other feeds
	Market prices	Manual events	None	Trigger from price feed	Optimize for day-ahead prices	Co-optimization for real-time prices with other feeds
	Customer data	None	None	Adjust dispatch based on typical retail rate	Optimization for retail rates or customer comfort	Co-optimization for retail prices, customer comfort, and other feeds
	Multiple grid services	None	None	User manually chooses grid service	System recommends settings to user	System can optimally select among grid services
GRID IMPACT	**Grouping**	None	None	Limited number of static groups, manually updated	Heirarchical groups periodically synced with grid models	Any group with any topology, with real-time updates
	Locational forecasts	None	None	Approximate	Available	High quality
	Grid optimization	None	None	Trigger locational events	Optimize for static grid limits	Optimize for dynamic grid limits