



Incorporating Data Center Flexibility into NV Energy's 2026 Integrated Resource Plan

SUMMARY

OCTOBER 2026

Summary

In 2026 Telos Energy and GridLab conducted Phase 2 analysis of the *Bringing Data Center Flexibility into Resource Adequacy Planning: A case study of NV Energy* study. The updated study split NV Energy into its two operating utilities (Nevada Power 'NVP' and Sierra Pacific Power Company 'SPPC') to test whether the location of flexible data center load matters for capacity accreditation, and then carries the accredited capacity into a capacity expansion model to measure the resource-mix and cost impacts. Key results include:

- **Data center flexibility has capacity value:** Flexible data center load offers significant capacity value for NV Energy, with an ELCC of 67% – 96%, depending on deployment size and location. The second GW of participation in a single region had a marginal ELCC of 43.5%
- **Location matters:** 1 GW of flexibility in SPPC, where reliability risk is concentrated, earned a 96% ELCC versus 69% for the same flexibility in NVP, because the Greenlink/ON-Line interface is at its limit during the hours that matter.
- **Data center flexibility could defer gas additions:** Accrediting ~1.3 GW of flexibility (2 GW participating, about half the forecasted 2032 data center load) deferred the near-term CT build by one to four years and avoided 773 MW of combined cycle gas over the horizon.
- **Firm capacity from data center flexibility creates significant savings:** The resulting present-value savings are roughly \$650 million within 2028–2046 and \$1.3 billion including end effects, before any payment to participating data centers.



TELOS ENERGY

Overview

In Phase One of the GridLab Data Center Flexibility NVE Case Study, we examined how data centers could provide accredited firm ‘capacity’ (i.e., ELCC) through load flexibility by. A resource adequacy model was developed to calculate the ELCCs of flexibility, a capacity expansion model was developed to calculate the change in cost and resource mix if flexible data centers were accredited at those ELCCs (resulting in fewer traditional generators needing to be built), and a production cost model was developed to characterize the operational impacts of the revised portfolios. We used NV Energy’s (NVE) system for this case study. The original Phase 1 study represented NVE as a single zone with no transmission constraints.

The goals of the Phase Two update were fourfold:

1. To update the data center forecast from the estimate used in Phase 1, to the data center forecast used in the 2026 Integrated Resource Plan.
2. To determine whether the locational aspects of data center flexibility impacts capacity accreditation, specifically between Sierra Power in the north and Nevada Power in the south.
3. To quantify how accrediting that flexibility changes NVE’s resource mix and system costs, using updated 2026 IRP assumptions.
4. To characterize how different forms of data center participation (e.g. # calls per year, max call duration, % availability requirements etc.) impact accreditation (this will be discussed in a future publication).

Phase 2 builds on the Phase 1 model, updates it to reflect the latest 2026 IRP information, and enhances the resource adequacy/ELCC model by splitting NVE into its two subsidiary utilities, Nevada Power (NVP) and Sierra Pacific Power Company (SPPC). The Greenlink/ON-Line transmission system connects them and is represented as an interface between the regions, limited to 3,800 MW of flow. By representing NVE as two separate regions, we can answer our first question: whether location matters.

Four scenarios were run through the resource adequacy model to help us answer these questions. 2032 was selected as our study year for the ELCC analysis. Across all scenarios, data center flexibility is modeled as an energy-limited resource with a maximum annual curtailment cap of 0.5%. Mechanically, these flexibility resources were modeled as batteries that couldn’t recharge, with their output in any given hour limited by the level of data center load assumed to be enrolled in the program. Once the energy that represents 0.5% of its annual demand is depleted, the resource stops responding.

- Scenarios A and B were run to assess the locational impacts on accreditation (adding 1 GW in NVP vs. 1 GW in SPPC)
- Scenarios B and C were run to assess saturation effects (going from 1 GW to 2 GW in SPPC)

- Scenario D was run as a more distributed case. We assumed 2 GW of participation (roughly 50% of NVE's ~4 GW of forecasted data center load participates in 2032). This participation was split two-thirds in SPPC and one-third in NVP, in line with NVE's expected forecast of data centers in each region. The accredited capacity calculated from this case was added into the capacity expansion model — this allowed us to see how the resource mix would change as a result of this accreditation.

TABLE 1. Phase 2 resource adequacy scenarios and accreditation results

SCENARIO	CASE	TOTAL DC FLEX PARTICIPATION	ACCREDITED CAPACITY	ELCC (%)
A	1 GW in NVP	1 GW	690 MW	69.0%
B	1 GW in SPPC	1 GW	960 MW	96.0%
C	2 GW in SPPC	2 GW	1,395 MW	69.8%
D	1.33 GW in SPPC / 0.67 GW in NVP	2 GW	1,340 MW	67.0%

Methodology and Assumptions

Phase 2 follows the same four-step framework as Phase 1: (1) run a capacity expansion baseline model without flexibility, (2) in parallel, calculate demand-side ELCCs in a probabilistic resource adequacy model, (3) re-run the capacity expansion model with flexibility accredited at those ELCCs, and (4) compare the net present value of expansion plans with and without flexibility to characterize net benefits. Full details of the datasets and the resource adequacy methodology are documented in the Phase 1 report¹ and will be updated in the forthcoming full paper. Key assumptions, including those that differ from, or extend, Phase 1 are as follows:

- **Data center demand:** NVE's 2026 IRP forecast assumes faster load growth driven by data centers than what was assumed in Phase 1. NVE estimates ~4 GW of data center demand will be reached by 2032, whereas the Phase 1 study assumed it wouldn't hit that level until 2039. An updated 8760 data center load profile, provided by NVE in its 2026 IRP, was also used.
- **Topology:** The resource adequacy model represents NVP and SPPC as separate regions connected by the Greenlink/ON-Line interface, limited to 3,800 MW. The capacity expansion model represents NVE as a single zone.
- **Resource adequacy:** A sequential, Monte Carlo production cost simulation run across 18 weather years, each receiving 28 random outage draws for a total sample count of 504, calibrated to a 0.1 day/year LOLE in the 2032 study year; the ELCC of flexibility is the amount of firm load that can be added back to the system, after flexibility is enabled, before LOLE returns to 0.1 day/year LOLE.

¹ See https://gridlab.org/datacenter_flex/

- **Flexibility product modeling:** an energy-limited resource willing to curtail 0.5% of its annual energy (equivalent to a 99.5% uptime requirement), with its hourly response capped by the data center load in that hour. It bids as the ‘resource of last resort’ and is called upon only if the model would otherwise have to shed load.
- **Capacity expansion:** a 12.5% planning reserve margin over 2028–2046, with expansion candidates for gas CT, gas CC, solar, wind, battery storage, geothermal, and market purchases (“open positions”), and a 20 MMT/year CO₂ cap that binds in the early 2030s. Current policy targets a 45% reduction from 2005 levels by 2030 (~11 MMT/year); we held the cap at 20 MMT/year instead, consistent with NVE’s own IRP modeling of relaxing or removing the emission constraint, because the 2026 load forecast is so significantly higher than the previous trajectory. Flexibility is credited at its 2032 ELCC and held at that value through 2046, phased in with the data center load forecast before 2032. Costs are derived from NREL’s ATB, with the capex of gas resources increased by 40% due to supply chain shortages and delays.
- **Cost accounting:** production cost, fixed O&M, and annualized capital cost (overnight cost spread over each resource’s economic life) are discounted to 2026 at NVE’s 7.08% IRP discount rate. Value beyond 2046 is captured by carrying the 2046 annual cost difference forward as a perpetuity.

Key Findings

1. Data center flexibility has a high ELCC: Total ELCCs of 67%–96%, with the marginal ELCC of the second GW falling to 43.5% – with variation due to amount (saturation), location (North vs South), and availability (# of calls / MWh per year)

a. ELCCs can saturate: the first GW of participation in SPPC yielded a 96% marginal ELCC, while the second GW yielded a 43.5% marginal ELCC

The saturation effect is a product of how the flexibility resource was modeled and the shape of the resource adequacy risk it is responding to. Each GW of participation is willing to curtail 0.5% of its annual energy, and its response in any hour is capped by the data center load in that hour (roughly 80–90% of nameplate). The first GW is therefore an energy-limited resource responding to a system whose loss-of-load events, at the 1-day-in-10-year calibration point, are shallower and shorter — most events are a few hundred MW deep and under five hours long, consistent with the event histograms from Phase 1. A 1 GW resource with a limited energy budget can cover nearly all of those events in full, which is why its ELCC approaches its nameplate (96%).

At 2 GW of participation, the data center flexibility resource begins facing deeper and longer events. This is driven by the ELCC methodology, which adds blocks of firm load to calculate accreditation; these raise the load factor of the system and produce longer and deeper events that are harder for an energy-limited resource to fully resolve. Each incremental MW of response therefore contributes less to LOLE reduction. The behavior is analogous to the saturation observed for other energy-limited resources

such as storage; as more energy-limited capacity is added, the remaining unserved energy increasingly sits in events the resource cannot cover without exhausting its limited energy budget, and the marginal accreditation falls. The key point is that a smaller portfolio of flexibility resources can cover most reliability events, while a larger portfolio becomes harder to fully utilize because the system's reliability challenges become more demanding (deeper and longer duration) as additional firm load is added during the ELCC calculation.

Furthermore, because battery storage and load flexibility from data centers are both energy-limited resources, there are interactions between these resource classes when it comes to capacity accreditation. While not evaluated in this study, the saturation to data center flexibility accreditation will also accelerate saturation for battery storage ELCC. This interactive effect warrants further evaluation in future study work.

b. ELCCs are locational: Given constraints with Greenlink and the location of data center development, Northern data center flexibility is more valuable than Southern flexibility (96% vs 69% ELCCs), based on 1 GW of modeled flexibility in either region.

Comparing cases A to B, we see that the location of data center flexibility does have a meaningful impact on accreditation. Data center flexibility added in SPPC, where all RA events occurred in our model, was most valuable to the system with a 96% ELCC vs. 69% if added in NVP.

The ELCC difference reflects reliability risk location coupled with transmission limitations. All RA risk appeared in SPPC, where data center load growth is concentrated, and the local resource mix is thinnest relative to demand, so the region relies on imports across the interface during its tightest hours. As a result, there are times when data center flexibility in NVP is available, but the interface is at its limit. This prevents NVP resources from alleviating the unserved energy event.

In practice, data center flexibility on NVE's system would be spread across sites in both NVP and SPPC, which raises a dispatch question. Because each participating data center can only be called a limited number of times (or for a limited amount of energy) per year, those calls are a scarce resource. When load risk emerges in SPPC and the Greenlink interface still has headroom, an operator may want to call on NVP data centers first and let the interface carry their response north. That preserves the SPPC calls for the hours when the interface is at its limit and only local flexibility can relieve the shortfall.

c. ELCCs depend on availability (see Phase 1 study for more details) – as demonstrated in Phase 1, data centers willing to curtail 0.5% of their annual energy received a higher ELCC than those willing to curtail only 0.1% (87% vs. 67%), based on 1 GW of modeled flexibility

In Phase 1, the availability requirement (the uptime data centers require, expressed as the share of annual energy they are willing to curtail) was modeled as an annual energy budget, so each GW of participation at 99.9% uptime can respond for about 9 GWh per year versus 44 GWh at 99.5%. Table 2 summarizes the resulting ELCCs for the 2030 study year.

TABLE 2.

Phase 1 ELCC results by availability requirement (1 GW and 2 GW of participation)

DC UPTIME	ENERGY BUDGET PER GW (GWH/YR)	1 GW TOTAL ELCC	2 GW TOTAL ELCC	MARGINAL ELCC (2ND GW)
99.9%	8.8	67%	46%	25%
99.7%	26.3	86%	66%	46%
99.5%	43.8	87%	77%	67%

Source: Phase 1 study (GridLab/Telos Energy, September 2025).

At 1 GW, there was almost no difference between 99.5% and 99.7% uptime because either budget was sufficient to cover nearly all loss-of-load events; the binding constraint at those levels was the depth of response (1 GW, derated by the hourly load factor), not energy. At 99.9%, the budget was small enough that the resource exhausted itself during long or repeated events in the worst weather years, costing roughly 20 percentage points of ELCC. A tighter budget also saturates faster as participation grows, as the 2 GW results show. Phase 1 also found that the burden implied by these availability levels was modest — the 1 GW, 99.5% case required an average of 6.6 hours of fleet-wide response per year (about 3 hours for an individual 100 MW data center if events are rotated across the fleet), concentrated in summer evenings, with the large majority of events shorter than five hours. The Phase 2 cases all assumed the 99.5% availability level.

2. Data center flexibility can defer new gas additions: When utilized as a firm capacity resource in our capacity expansion model, data center flexibility deferred 880 MW of near-term CT gas additions by one to four years and reduced the total amount of combined cycle gas built over the horizon by 773 MW

The 2GW of data center flexibility participating in Scenario D resulted in 1,340 MW of accredited capacity. Adding this resource to the capacity expansion model reduced the firm capacity the model needed to build by 1,340 MW to meet its 12.5% reserve margin requirement. This resulted in a reduced CT gas build in the early years, and an overall reduction in combined cycle additions across the study horizon.

The accredited flexibility is phased in with the data center load forecast and reaches its full value by 2032. Its first effect is on the firm gas build. In the Base case, the model adds 880 MW of CTs in 2031 to meet the reserve margin; with flexibility, half of that is deferred by one year and the other half by four years, and both cases converge on the same 1,320 MW CT fleet by 2035. Flexibility does not eliminate the CTs entirely, but it delays them to future years in the study horizon. This results in near-term capital savings, which have a larger influence on net present value of portfolios compared to later resource builds. This is also critical given current supply chain disruptions and the elevated costs for new gas CTs. Deferring these resources a few years through low, or zero-cost load flexibility options can help buy time until supply chains can return to equilibrium.

Early renewable additions shift as well. In the Base case, the model leans on wind and solar partly for their firm capacity contribution. With flexibility providing ~1.3 GW towards

the reserve margin, early solar is trimmed by roughly 1.7 GW through 2034 and partially replaced with wind, and solar catches back up by the end of the decade as it is needed for energy rather than capacity.

In the middle of the horizon (2036–2041) the model turns to combined cycle gas to serve the growing energy need.² The flexibility case starts its combined cycle additions later and builds 773 MW less across the horizon — the largest permanent change in the resource mix.

After 2041 both cases shift to large-scale solar and storage additions as existing gas retires, and the two portfolios begin to converge. The flexibility case front-loads solar and defers roughly 2.5–3 GW of battery storage until the final year of the horizon, ending with essentially the same storage and solar totals as the Base case. By 2046 the only material difference between the two portfolios is the combined cycle capacity that was never built (Table 3, Figure 1).

This finding mirrors the Phase 1 finding that flexibility relieves the capacity constraint but not the energy constraint: once the system becomes energy-limited, the same quantity of energy resources is required regardless of accredited flexibility, and the value of flexibility lies in deferring builds and avoiding firm gas rather than shrinking the ultimate size of the portfolio.



² Note: the first year that combined cycle resources can be built in NV Energy's capacity expansion is 2036.

TABLE 3. Cumulative capacity additions by resource type (MW)

RESOURCE	2031	2035	2041	2046
Base Case				
Gas CT	880	1,320	1,320	1,320
Gas CC	–	–	3,251	3,251
Battery Storage	–	–	–	11,369
Solar	470	5,512	9,568	17,428
Wind	338	407	500	3,750
Market Purchases	500	500	500	500
Total	2,188	7,739	15,139	37,618
2 GW DC Flex Case				
Gas CT	–	1,320	1,320	1,320
Gas CC	–	–	2,478	2,478
Battery Storage	–	–	–	11,576
Solar	749	4,357	9,561	17,539
Wind	65	774	774	3,750
Market Purchases	500	226	500	500
Total	1,314	6,677	14,633	37,163
Change (DC Flex – Base)				
Gas CT	(880)	0	0	0
Gas CC	0	0	(773)	(773)
Battery Storage	0	0	0	207
Solar	279	(1,155)	(7)	111
Wind	(273)	367	274	0
Market Purchases	0	(274)	0	0
Total	(874)	(1,062)	(506)	(455)

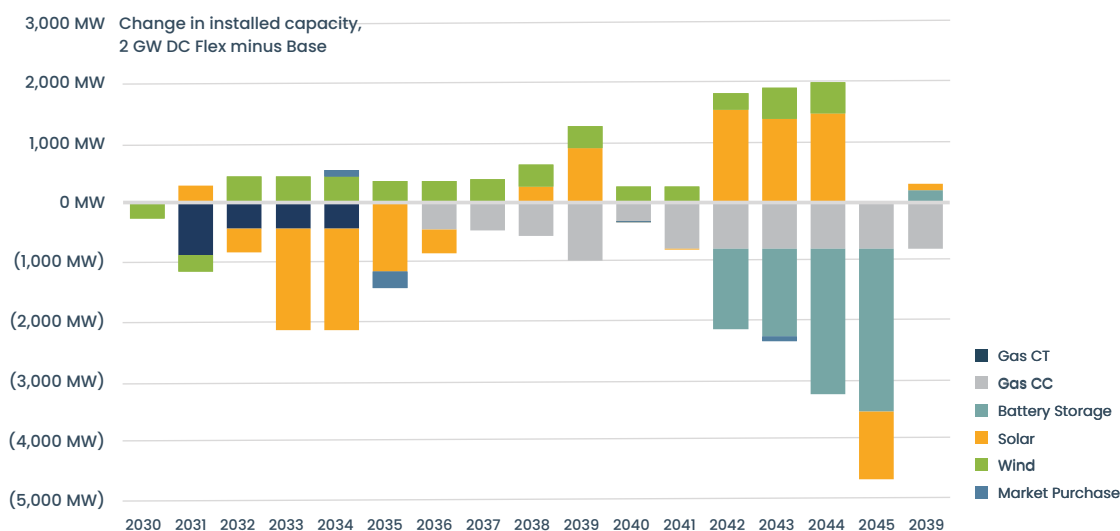


FIGURE 1.

Change in cumulative capacity additions, 2 GW DC Flex case minus Base case. Negative values indicate capacity deferred or avoided with flexibility.

3. Data center flexibility saves hundreds of millions in cost for the system: Assuming 50% of the data centers online in 2032 participate (~2 GW) and are willing to curtail 0.5% of their annual energy, NVE's system costs can be reduced by ~\$1.3 billion on a present-value basis (~\$650 million within the 2028–2046 horizon, plus end effects)

To estimate the cost impact, we compared the Base and 2 GW DC Flex capacity expansion solutions across three cost streams over 2028–2046: production cost (fuel, variable O&M, start-up, and emissions costs), fixed O&M, and capital cost. Capital cost is included on an annualized basis – each build's overnight cost is spread over its economic life. This allows us to only impose the cost of resources built late in the horizon to the years those resources operate in the study period. Each year's costs were discounted to 2026 at the 7.08% discount rate used in NVE's IRP. To capture value beyond 2046, the final-year annual cost difference between the two cases (the Flex case is roughly \$170 million per year cheaper in 2046, mostly from avoided combined cycle and deferred storage) was carried forward as a perpetuity. Table 4 summarizes the results.

TABLE 4. Present value of system costs, 2028–2046 (\$ millions, 2026\$)

COST COMPONENT	BASE	2 GW DC FLEX	CHANGE
Production Cost	16,780	17,043	262
Fixed O&M	1,700	1,561	(139)
Annualized Capital Cost	8,102	7,325	(777)
Total – Within Horizon	26,582	25,929	(653)
End Effects (perpetuity of 2046 difference)	–	–	(606)
Total – Including End Effects	–	–	(1,259)

Negative values indicate savings relative to the Base case.

Within the planning horizon, flexibility reduced the present value of annualized capital costs by \$777 million and fixed O&M by \$139 million, while raising production costs by \$262 million as the deferred CTs, lower early solar build, and smaller combined cycle fleet were replaced by more gas generation and market purchases. This nets to \$653 million of in-horizon savings. Adding the perpetuity for end effects brings total savings to approximately \$1.26 billion, or about \$1.50 per MWh of load served over the period. No cost was assigned to the flexibility itself, so this figure represents the gross system value available to compensate participating data centers.

Why This Matters

Data center load is now the dominant driver of NVE’s load growth, and the same is true for a growing number of utilities. **The conventional planning assumption – that this load is inflexible and must be backed by firm capacity at the planning reserve margin – is the most expensive one available.** This study shows that a modest, well-defined flexibility product (0.5% of annual energy, a handful of calls per year concentrated in summer evenings) earns a capacity credit approaching a new gas peaker, and that for NVE, 2 GW of participation would defer or avoid roughly \$0.65–1.3 billion of system cost in present-value terms. That value is large relative to the likely cost of the flexibility itself, which suggests there is room for a deal that benefits both ratepayers and data center customers.

The locational result adds a second lesson. Flexibility is worth about 40% more where reliability risk and transmission limits coincide, so programs and tariffs that treat all large-load flexibility the same will leave value on the table. Accreditation, incentives, and interconnection terms should recognize where on the system the flexibility sits, in the same way that supply-side resources are already valued differently by location.

These findings suggest that NV Energy and other investor-owned utilities should begin building this into their own planning and market design now, while large loads are still being

interconnected and before contracts lock in inflexible service:

1. calculate a demand-side ELCC for large-load flexibility in the IRP, using the same reliability tools already applied to supply-side resources;
2. carry that accreditation into the capacity expansion model so the avoided-capacity value is visible in the preferred plan;
3. translate the avoided cost into large-load rate design, interconnection agreements, and incentive levels that are set before the load comes online; and
4. design the flexibility product — depth, availability, and call limits — around the event characteristics the reliability model reveals, so the burden on participating customers is bounded and predictable. The methods demonstrated here are built from standard IRP tools and can be adopted without new modeling platforms or processes.

In addition to these findings, future work by the study team will evaluate how a portfolio of data centers, each with its own operating constraints, call limits, and duration requirements, can be effectively scheduled so that the aggregate response to the system meets resource adequacy needs. Additional work can also evaluate the interactive effects between data center flexibility and other energy-limited resources like battery energy storage, hydro, and conventional demand response.