Nimaful Samuel, 2026, 14:2
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science,
Engineering and Technology
An Open Access Journal

Demand Response Participation of Data Centers: Technical Mechanisms, Market Integration, Governance Implications, and Research Directions

Nimaful Samuel, Hanyabui Augustine, Holison Faith Esther
Eastern Illinois University

Abstract- Demand response (DR)—the intentional modification of electricity consumption in response to grid conditions or price signals—has expanded from emergency curtailment into a portfolio of market-based services that can support reliability, integrate variable renewable energy, and reduce infrastructure costs (U.S. Department of Energy, 2006; International Energy Agency, 2023). At the same time, data centers have become a fast-growing and highly concentrated source of electricity demand, with U.S. data center electricity use rising to an estimated 176 TWh in 2023 (4.4% of U.S. electricity consumption) and projected to reach roughly 325–580 TWh by 2028 under scenario assumptions driven in part by accelerated AI server adoption (Shehabi et al., 2024). This co-evolution—rapid load growth and increasing need for flexibility—makes data center participation in DR both attractive and complex. This paper synthesizes peer-reviewed research, national laboratory field studies, standards documents, and selected utility/market operator materials to explain (a) what "data center DR" means in operational terms, (b) how data centers technically deliver flexibility, (c) how DR is communicated, controlled, and verified, (d) how participation is monetized across tariffs and capacity/ancillary markets, (e) what empirical case studies indicate about feasibility and outcomes, and (f) what regulatory, reliability, cybersecurity, and environmental considerations constrain or enable scaling. Key findings supported by the reviewed evidence include: (1) measurable, automation-friendly load flexibility exists in both cooling/infrastructure systems and IT workloads, but the magnitude and response time vary significantly by mechanism and facility type (Ghatikar et al., 2012; Wierman et al., 2014). (5) Recent deployments illustrate two emerging archetypes: "compute-aware DR" (shifting non-urgent tasks across time and sometimes geography) and "grid-interactive UPS/battery services" (fast frequency response and reserve-type services), with high-profile examples reported by large operators and aggregators in multiple regions (Google Cloud, 2023; Microsoft, 2022; Baringa Partners LLP, 2023).

Keywords- capturing the core of your paragraph include demand response (DR), data center electricity demand, load flexibility, grid reliability, renewable energy integration, cooling systems, IT workload shifting, automation (OpenADR), measurement and verification (M&V), capacity and ancillary services, compute-aware DR, UPS/battery storage, fast frequency response, market participation, energy tariffs, cybersecurity, and carbon accounting.

I. INTRODUCTION AND BACKGROUND

Definitions and scope

Demand response is widely defined as a change in electricity consumption by end-use customers—relative to an expected (baseline) level—in response to price changes or incentive-based signals (U.S. Department of Energy, 2006; Goldberg et al., 2013). In contemporary grid planning, DR is also framed as a form of "demand-side flexibility" that can participate in wholesale and ancillary markets to help balance systems with increasing shares of variable renewable generation (International Energy Agency, 2023). In U.S. organized wholesale markets, compensation and participation rules for certain DR services have been shaped by regulatory actions such as Federal Energy Regulatory Commission Order No. 745, which addresses compensation for demand response in ISO/RTO energy markets (Federal Energy Regulatory Commission, 2011).

A data center is a facility housing computing, storage, and networking equipment supported by power conversion/distribution, cooling, fire protection, physical security, and monitoring/control systems (Google, 2023). From a power-system perspective, data centers matter because they are (a) large, often relatively flat loads with high reliability requirements, (b) increasingly clustered geographically, and (c) partially flexible through both infrastructure controls and IT workload management (Wierman et al., 2014; Ghatikar et al., 2012). This paper addresses data centers across hyperscale, enterprise, high-performance computing (HPC), and colocation contexts, noting when evidence is region- or market-specific.

Why DR participation is rising in relevance for data centers

The accelerating growth of data center electricity consumption—especially linked to accelerated computing for AI—has heightened grid reliability and planning concerns. A recent U.S. national laboratory report estimated that U.S. data center electricity use increased to 176 TWh in 2023 (4.4% of total U.S. electricity consumption) and could reach roughly 325–580 TWh by 2028 depending on equipment shipment and operational scenarios (Shehabi et al., 2024). These projections sit alongside broader reliability concerns about peak demand growth and resource adequacy as load rises and generation portfolios evolve (North American Electric Reliability Corporation assessments summarized by Reuters, 2025).

From the data center operator perspective, DR can (1) reduce exposure to demand charges and coincident peak penalties, (2) earn incentive payments in capacity or ancillary service products, (3) reduce outage risk by contributing to local grid stability, and (4) support corporate decarbonization strategies by aligning consumption with cleaner supply conditions (Chen et al., 2014; Google Cloud, 2023). Yet participation is constrained by strict uptime expectations—often expressed as "five nines" reliability—alongside equipment thermal limits, battery lifetime tradeoffs, and cybersecurity risk tolerance (Microsoft, 2022; ASHRAE, 2016).

II. METHODOLOGY AND LITERATURE REVIEW

This research uses a structured, source-prioritized synthesis approach suitable for engineering-policy topics where evidence spans peer-reviewed studies, standards, and operational reports. Sources were collected and evaluated using four inclusion priorities:

First, peer-reviewed literature and widely cited technical reports that develop models of data center flexibility, workload shifting, and market participation (e.g., Wierman et al., 2014; Liu et al., 2013; Wang et al., 2016; Chen et al., 2020). Second, national laboratory field studies and engineering reports documenting measured load shed/shift outcomes and enabling technologies in operating facilities (e.g., Ghatikar et al., 2012; Ghatikar et al., 2010). Third, primary standards and official program documentation governing DR messaging, interoperability, security, and market participation (e.g., OpenADR Alliance documents; IEEE 2030.5 overview; ISO/RTO participation guides). Fourth, vendor/industry case studies and credible market analyses used to triangulate real-world deployments and quantify reported outcomes (e.g., Microsoft, 2022; Baringa Partners LLP, 2023; Google Cloud, 2023).

Because the user specified no geographic constraint, the synthesis is intentionally multi-jurisdictional. When conclusions rely on market rules that differ by region (e.g., DS3 in Ireland vs. CAISO demand response products vs. ISO-NE capacity participation), the relevant jurisdiction is stated explicitly (California ISO, n.d.; ISO New England, n.d.; EirGrid, 2014).

III. LITERATURE REVIEW SYNTHESIS

The literature converges on a shared conceptual model: data center DR is the coordinated use of (a) physical infrastructure flexibility (cooling, airflow, power conversion) and (b) IT/compute flexibility (server power states, virtualization, job scheduling, load migration) to shape net grid import while meeting service-level constraints (Wierman et al., 2014; Chen et al., 2020).

Early empirical work by Lawrence Berkeley National Laboratory and collaborators emphasized "Open Auto-DR" readiness in data centers—characterizing loads, controls, and enabling technologies—and identified both institutional and technical barriers (Ghatikar et al., 2010). Follow-on field testing demonstrated that sizable demand savings were achievable with minimal operational impact through strategies such as temperature setpoint adjustments, coordinated cooling reduction, and selected IT strategies including load migration (Ghatikar et al., 2012).

In parallel, computer systems and operations research communities built formal models linking electricity prices and grid signals to workload scheduling and geographic load balancing, showing that shifting delay-tolerant workload can reduce coincident peaks and costs, sometimes in combination with local generation or storage (Liu et al., 2013; Liu et al., 2014). More recent work extends this framing to proactive grid–data center coordination through pricing and two-way communications, highlighting that naïve "price chasing" by distributed data centers may create grid-side stability and congestion issues without coordinated mechanisms (Wang et al., 2016).

A third stream focuses on ancillary services and fast response: using UPS batteries and coordinated control to provide regulation or frequency response, while accounting for battery degradation costs and operational risk (Aksanli et al., 2014; Baringa Partners LLP, 2023). This stream is increasingly relevant as grids value sub-minute flexibility and synthetic inertia-like behavior under high renewable penetration conditions (EirGrid, 2014; Baringa Partners LLP, 2023).

IV. TECHNICAL MECHANISMS FOR DR PARTICIPATION

Data center DR is best understood as a portfolio of mechanisms spanning seconds-to-hours timescales. The mechanisms below are not mutually exclusive; high-performing programs commonly stack multiple levers under a single supervisory controller (Ghatikar et al., 2012; Wierman et al., 2014).

Load shedding

Load shedding reduces instantaneous power draw, typically for reliability events or peak periods. In data centers, operationally acceptable shedding often targets non-IT loads first (cooling fans, pumps, CRAC/CRAH units, chillers) or uses IT-level power capping rather than server shutdown to manage risk (Ghatikar et al., 2012; ASHRAE, 2016). Field tests documented strategies such as server and CRAC shutdown, temperature setpoint adjustments, and coordinated cooling relative to IT load reduction, with response and recovery characteristics measured across tests (Ghatikar et al., 2012).

A key engineering constraint is adherence to allowable inlet temperature/humidity envelopes for IT equipment. Industry thermal guidelines—frequently referenced via ASHRAE TC 9.9 materials—support energy savings by allowing warmer operating temperatures under controlled conditions, but DR-induced setpoint excursions must stay within allowable ranges to avoid increased failure risk (ASHRAE, 2016).

Load shifting

Load shifting changes when energy is consumed rather than only how much. Two distinct shifting modes exist:

Temporal shifting within a facility defers flexible jobs (batch analytics, video processing, training runs, non-urgent indexing) to off-peak hours or periods of lower system stress (Google Cloud, 2023; Wierman et al., 2014). Geographic shifting distributes workload across multiple sites, routing delay-tolerant tasks to data centers in regions with lower prices or lower carbon intensity, subject to latency, capacity, and data sovereignty constraints (Wang et al., 2016; Liu et al., 2014).

In practice, temporal shifting is typically easier to deploy broadly because it does not require cross-region data movement and can be integrated into existing schedulers. Google reported using its "carbon-intelligent computing" task-shifting capability (originally built for carbon-aware operations) to temporarily reduce power consumption at specific data centers during forecast grid events, rescheduling non-urgent work after the event, and in some cases rerouting to other grids (Google Cloud, 2023; Google, 2023).

V. THERMAL STORAGE AND THERMAL INERTIA

Thermal flexibility can be delivered through explicit thermal storage (e.g., chilled water tanks, ice storage) or by using the building's thermal capacitance and allowable temperature bands to "pre-cool" or "ride through" short DR events (Ghatikar et al., 2012; ASHRAE, 2016). In many data centers, short events can be served by temporarily relaxing cooling setpoints or staging mechanical cooling while maintaining inlet limits, which reduces compressor power and/or fan power (Ghatikar et al., 2012).

Thermal strategies are especially relevant because non-IT energy can be substantial, and because cooling response can be automated through the building management system (BMS) with relatively low cyber-physical blast radius compared to IT shutdown (Ghatikar et al., 2012; Wierman et al., 2014).

UPS, battery storage, and "grid-interactive" energy systems

UPS systems provide power conditioning and ride-through during grid disturbances; the same batteries can potentially provide fast grid services when not needed for outages. An Enel X case study notes that data center UPS systems are built to respond quickly (sub-second for backup transfer) and thus can contribute to grid balancing through demand-side response/frequency programs (Enel X, n.d.).

Two distinct operating patterns matter:
Backup-only UPS, where batteries are kept near full state-of-charge with minimal cycling, prioritizing availability.

Grid-interactive UPS / battery dispatch, where controlled cycling provides services (frequency response, reserves) while preserving minimum reserve margins and respecting degradation constraints (Microsoft, 2022; Baringa Partners LLP, 2023).

Battery degradation economics can dominate feasibility for high-cycle products. A peer-reviewed analysis found that minimum regulation prices required to compensate for increased battery cycle costs can exceed observed prices in some markets, pushing operators toward targeted high-value events or strategies such as providing regulation during recharge intervals (Aksanli et al., 2014).

IT workload management, virtualization, and power capping

IT-side DR includes dynamic voltage and frequency scaling (DVFS), server consolidation, turning idle servers to low-power states, and power capping that reduces CPU/GPU power draw while attempting to preserve quality of service (Chen et al., 2014; Wierman et al., 2014). Virtualization and workload orchestration enable more granular control by migrating virtual machines/containers away from constrained clusters and prioritizing critical services over deferrable tasks (Wang et al., 2016; Google Cloud, 2023).

Academic work also emphasizes that workload shifting and local generation can jointly reduce coincident peaks and energy expenditure, but the optimal strategy depends on uncertainty in peak warnings and renewable availability (Liu et al., 2013).

On-site generation and microgrids

Many data centers have on-site generators for backup. These assets can support DR either by reducing grid import (behind-the-meter generation during events) or, where market rules and interconnection agreements allow, exporting power. The literature and field reports treat this as a technically powerful but institutionally sensitive pathway due to emissions, permitting, and program rules that distinguish emergency operation from economic dispatch (Liu et al., 2013; Wierman et al., 2014). In some jurisdictions, “green” alternatives to diesel (gas engines, fuel cells, battery systems) are being explored to deliver resilience and grid services with lower emissions (Microsoft, 2022).

Renewable integration

Renewable integration for data centers intersects with DR in two main ways: (1) operational alignment of demand with renewable availability via load shifting, and (2) hybridization with on-site renewables plus storage to shape net import/export. Studies highlight that DR is increasingly valuable as renewable penetration rises and supply becomes more variable, but market design must ensure incentives reflect grid value and avoid perverse outcomes (Wierman et al., 2014; International Energy Agency, 2023).

Comparative technology table

Table 1
Technical mechanisms for data center DR participation: characteristics, maturity, and risks

Mechanism	Typical response time	Typical sustained duration	Primary controlled assets	Key constraints	Maturity evidence
Cooling setpoint adjustment / staged cooling	Minutes	Hours	BMS, CRAC/CRAH, chillers, fans	Thermal envelopes, hot-spot risk, recovery management	Field-tested in operating data centers (Ghatikar et al., 2012).[4]
IT power capping / DVFS / server provisioning	Seconds–minutes	Minutes–hours	Servers, cluster schedulers	SLA/QoS risk, performance regressions	Modeled and evaluated in DR strategies research (Chen et al., 2014; Wierman et al., 2014).[42]
Temporal workload shifting	Minutes–hours (planning lead time)	Hours–days	Job schedulers, queues	Deadline constraints, customer expectations	Deployed in production DR pilots (Google Cloud, 2023).[48]
Geographic workload shifting	Minutes–hours	Hours–days	Multi-region orchestration	Latency, data residency, network limits	Formalized in proactive DR models (Wang et al., 2016).[28]

Mechanism	Typical response time	Typical sustained duration	Primary controlled assets	Key constraints	Maturity evidence
UPS/battery dispatch for frequency services	Sub-second–seconds	Minutes (often limited)	UPS/battery, power electronics	Battery SOC reserve, degradation, protection coordination	Demonstrated in grid-interactive UPS deployments and modeling (Microsoft, 2022; Baringa Partners LLP, 2023).[40]
On-site generation (backup genset dispatch)	Minutes	Hours	Diesel/gas gensets, switchgear	Emissions permits, noise, program eligibility	Considered in coincident peak/DR optimization (Liu et al., 2013).[44]
Thermal storage (explicit)	Minutes	Hours	Chilled water/ice systems	Capex/space, control integration	Mentioned as DR-enabling storage analogue in DR frameworks (Ghatikar et al., 2012).[4]

VI. CONTROL ARCHITECTURES AND COMMUNICATIONS

Control hierarchy: from grid signal to actuator

Data center DR systems commonly implement a layered architecture:

Grid/market layer: ISO/RTO or utility defines event triggers and product requirements (e.g., capacity commitment, reliability dispatch, fast frequency response).

Intermediation layer: aggregators coordinate portfolios of customers, often acting as the interface to market settlement and compliance (Microsoft, 2022).

Site supervisory control: an energy management system (EMS) or data center infrastructure management (DCIM) platform translates DR requests into actionable setpoints and constraints.

Local control loops: BMS, UPS controllers, and IT orchestrators execute commands and maintain safety limits (Ghatikar et al., 2012; Wierman et al., 2014).

Interoperability standards and protocols

Open standards reduce integration cost and enable scalable automation. OpenADR Alliance describes

OpenADR as a standardized way to convey DR signals from system operators to end-use customers, supporting automated participation (OpenADR Alliance, 2012). OpenADR documentation describes a server–client model with Virtual Top Nodes (VTNs) and Virtual End Nodes (VENs), and emphasizes secure transport (including TLS) with certificate-based authentication (OpenADR Alliance, 2022; OpenADR Alliance, 2015).

Utility implementations explicitly reference these concepts. Southern California Edison describes sending DR event or price notifications through a Demand Response Automation Server (DRAS) that uses OpenADR to communicate with a customer's OpenADR device or EMS, enabling automatic load reduction during program events (Southern California Edison, n.d.).

At the distribution/DER interface, IEEE 2030.5 defines an application-layer protocol over TCP/IP supporting utility management functions including demand response and time-of-day pricing, alongside DER management (IEEE, 2018). In practice, OpenADR has been widely associated with DR event signaling, while IEEE 2030.5 is often discussed in the context of DER and customer energy environments; hybrid architectures may arise where data centers combine load flexibility, batteries, and other behind-the-meter resources.

Security, resilience, and operational guardrails

OpenADR profile specifications describe security mechanisms including TLS 1.2, certificate requirements, and optional XML signatures for higher-assurance non-repudiation use cases (OpenADR Alliance, 2015). Cybersecurity guidance for smart grid functions—including demand response—has also been developed by National Institute of Standards and Technology in smart grid cybersecurity guidelines (National Institute of Standards and Technology, 2014).

From a reliability standpoint, the architecture must also include fail-safe behavior: if communications fail or signals are malformed, the system should default to "safe load" modes that preserve uptime, battery reserve, and thermal margins. Industry narratives around grid-interactive UPS emphasize that customer reliability remains paramount and that rigorous testing and compliance verification are required before providing grid services (Microsoft, 2022; Data Centre Review, 2024).

Economics, Incentives, and Valuation

Participation pathways: tariffs vs. wholesale markets
Data center DR economics are driven by where value is paid:

Retail tariffs and demand charges: Many large customers face demand charges based on peak kW and, in some cases, coincident peak structures, making peak shaving financially valuable even without explicit DR programs (Chen et al., 2014).
Incentive-based DR programs: Utilities and aggregators pay customers for verified load reductions during dispatch events or for enrolled capacity availability (U.S. Department of Energy, 2006; ISO New England, n.d.).

Wholesale participation: In organized markets, demand-side resources may provide energy reductions and certain ancillary services under defined products and rules. California ISO describes demand response participation frameworks and registration systems for Proxy Demand Response and Reliability Demand Response resources (California ISO, n.d.). ISO New England documents how demand response assets map to demand response resources and can participate in capacity and energy/reserve markets under specific structures (ISO New England, n.d.).

For the U.S., regulatory actions such as Federal Energy Regulatory Commission Order No. 841 (storage participation) and Order No. 2222 (DER aggregations) shape broader participation pathways for storage and aggregated distributed resources, which can affect how behind-the-meter batteries and flexible loads at data centers enter wholesale markets (Federal Energy Regulatory Commission, 2018; Federal Energy Regulatory Commission, 2020).

Value stacking and product fit

DR value “stacks” when a facility can simultaneously avoid retail charges, earn incentive payments, and provide high-value grid services. But stacking is constrained by operational coupling: a battery used for frequency response may not be fully available for outage ride-through, and aggressive cycling can shorten battery life (Aksanli et al., 2014; Microsoft, 2022).

Peer-reviewed work evaluating participation in regulation services and frequency control suggests significant potential electricity cost reductions in modeled scenarios while meeting QoS/SLA constraints, but outcomes depend on utilization scenarios, market prices, and control policies (Chen et al., 2014). Conversely, battery wear costs can make continuous regulation uneconomic at prevailing prices in some markets, encouraging participation in higher-priced emergency events or limited windows (Aksanli et al., 2014).

Measurement and verification: baselines as the “currency” of DR

Compensation depends on credible quantification of load reduction relative to a baseline—the estimated load that would have occurred absent the DR event. The U.S. National Action Plan on Demand Response M&V report emphasizes that DR products define how reduction is valued and measured, and that baseline methodologies are central across programs (Goldberg et al., 2013). The same report highlights diverse uses of DR measurement—from settlement to planning—and the importance of consistent protocols to reduce disputes and improve comparability (Goldberg et al., 2013).

For data centers, baseline complexity increases when workloads are dynamically shifted across time or geography; the “counterfactual” baseline must account for what the facility would have consumed given workload and environmental conditions. This makes telemetry (IT load, cooling power, outside air conditions) and transparent M&V design especially critical (Wierman et al., 2014; Goldberg et al., 2013).

Cost-benefit analysis framework for data center DR investments

A publishable valuation framework typically includes:
Capital costs: controls integration, metering/telemetry, EMS/DCIM upgrades, battery system upgrades, thermal storage capex.
Operational costs: engineering labor, maintenance, battery degradation, increased failure risk, potential performance penalties.

Benefits: avoided demand charges, DR incentive payments, capacity payments, ancillary services revenue, avoided curtailment costs, reputational value from sustainability claims.

Risk-adjusted reliability cost: expected value of downtime costs times incremental probability of downtime due to DR actions (Wierman et al., 2014; Microsoft, 2022).

Given the “tail risk” of outage costs, many operators adopt conservative control policies: limiting DR to non-critical loads or preserving fixed ride-through margins for UPS systems, which can reduce theoretical revenue but improve risk-adjusted net value (Microsoft, 2022; Ghatikar et al., 2012).

Case Studies and Pilot Programs

This section summarizes documented DR participation examples across hyperscalers, colocation providers, and research/field pilots. Reported outcomes are not directly comparable because programs differ by product definition, baseline rules, and grid context; the comparative table therefore emphasizes mechanism, scale indicators, and measured/claimed outcomes with citations.

Field-tested DR strategies in operating data centers
A landmark field study evaluated DR strategies and enabling technologies at four California data centers, including cooling and IT equipment strategies and load migration approaches (Ghatikar et al., 2012). The authors report that with minimal or no impact to operations, demand savings of ~25% at the data center level or ~10–12% at the whole-building level were achievable through tested strategies (Ghatikar et al., 2012). This result is important because it

anchors DR feasibility in measured outcomes rather than only simulations.

Compute-aware demand response in hyperscale operations

Google reports deploying and piloting a DR approach that builds on its carbon-intelligent computing platform: upon receiving notice of a forecast grid event, its planning system generates hour-by-hour instructions limiting non-urgent compute tasks at specified data centers, rescheduling after the event and sometimes rerouting tasks to other grids (Google Cloud, 2023). Google reports actions across multiple regions: daily peak-period reductions across several European countries during winter 2022–23 energy scarcity, daily summer peak reductions in Taiwan under a program run by Taiwan Power Company in summers 2022 and 2023, and reductions during extreme weather-related stress events in parts of the U.S. (Google Cloud, 2023; Google, 2023).

Separately, Reuters reported that Google entered agreements with two U.S. utilities—Indiana Michigan Power and Tennessee Power Authority—to curtail AI data center load during peak demand periods, describing these as Google's first formal DR agreements involving temporary curtailment of machine learning workloads (Reuters, 2025).

Grid-interactive UPS and frequency services in Ireland

In Ireland, EirGrid developed DS3 (Delivering a Secure, Sustainable Electricity System) to securely operate the power system with increasing levels of variable, non-synchronous renewables (EirGrid, 2014). DS3 includes frequency and system services products designed to maintain system stability around 50 Hz under renewable variability (EirGrid, 2014).

A prominent colocation example is Digital Realty's participation—via Enel X—in DS3-related dynamic frequency balancing using UPS battery systems. A published interview describes millisecond-by-millisecond frequency tracking with hardware that meters UPS and sends power requests, alongside grid compliance testing and a stated estimate that adding 1 MW of capability could save approximately 4,000 tonnes of CO_2 per year, with 6 MW initially integrated (Data Centre Review, 2024). A separate Enel X case study similarly frames UPS as an instantly responsive asset that can draw from batteries when frequency drops (Enel X, n.d.).

Complementing these operator narratives, Baringa Partners LLP modeled the system-level decarbonization and consumer cost impacts of grid-interactive UPS (G-UPS) systems providing DS3 services instead of fossil fuel-fired plants. The report estimates that using G-UPS to provide DS3 services could save around 1.5 million tonnes of CO_2 in the Republic of Ireland power sector in 2025 and 0.5 million tonnes in Northern Ireland, with associated consumer cost savings (Baringa Partners LLP, 2023).

Grid-interactive UPS deployments by technology firms

Microsoft describes deploying grid-interactive UPS technology in a Dublin data center, emphasizing that UPS batteries provide continuous protection for servers and can be controlled to provide services back to the grid with secure communication and testing (Microsoft, 2022). Microsoft cites analysis suggesting that in Ireland and Northern Ireland, grid-interactive UPS replacing fossil-provided grid services could avoid about two million metric tons of CO_2 in 2025, referencing Baringa's commissioned analysis (Microsoft, 2022; Baringa Partners LLP, 2023).

Comparative case table

Table 2
Selected data center DR case studies and pilots (multi-jurisdictional)

Provider type	Program / market context	Mechanism(s)	Indicative scale	Reported outcomes
Research + enterprise/HPC sites	Utility collaboration and	Cooling/IT strategies; load	4 data centers tested	~25% data center-level or ~10–12%

Provider type	Program / market context	Mechanism(s)	Indicative scale	Reported outcomes
(California)	field tests	migration		whole-building demand savings with minimal/no operational impact (Ghatikar et al., 2012).[4]
Hyperscale operator	Utility/grid partner DR pilots (multiple regions)	Non-urgent compute task shifting; rescheduling; some geographic reroute	Multi-region pilots; peak-hour reductions reported in Europe and Taiwan	Peak-period reductions during winter 2022–23 (Europe), summers 2022–23 (Taiwan), extreme events (U.S.) (Google Cloud, 2023; Google, 2023).[37]
Hyperscale operator	U.S. utility agreements (reported)	Curtailment of ML workloads during peak	Not disclosed	First formal DR utility agreements reported for AI workload curtailment (Reuters, 2025).[75]
Colocation operator + aggregator	DS3 frequency response (Ireland)	Grid-interactive UPS; dynamic frequency balancing	6 MW initially integrated (reported)	Millisecond frequency tracking; stated ~4,000 tCO_2/year per MW potential; 6 MW initial integration (Data Centre Review, 2024).[80]
Technology firm	Grid services market in Ireland (via aggregator)	Grid-interactive UPS; battery dispatch	Not disclosed	UPS controllability with secure comms; cited ~2 million

Provider type	Program / market context	Mechanism(s)	Indicative scale	Reported outcomes
				tCO₂ avoidable in 2025 if broadly deployed (Microsoft, 2022; Baringa Partners LLP, 2023).[40]

VII. GOVERNANCE, SUSTAINABILITY, RECOMMENDATIONS, AND FUTURE DIRECTIONS

Regulatory and market design implications

DR participation rules differ substantially across jurisdictions. In the U.S., ISO/RTO participation products—such as those described by California ISO (Proxy Demand Response, Reliability Demand Response) and ISO New England (demand response assets mapped to market resources)—create different entry points for large loads and aggregations (California ISO, n.d.; ISO New England, n.d.). Federal policy has also shaped the compensation/integration landscape for demand response and storage in wholesale markets (Federal Energy Regulatory Commission, 2011; Federal Energy Regulatory Commission, 2018; Federal Energy Regulatory Commission, 2020).

In Ireland, DS3 reflects a system operator response to renewable integration constraints, shifting value toward fast-acting services that can maintain frequency under high non-synchronous generation conditions (EirGrid, 2014). This helps explain why grid-interactive UPS services can be economically and institutionally salient there, particularly when policy targets drive rapid renewable penetration (Data Centre Review, 2024; Baringa Partners LLP, 2023).

Recommendation (market design): expand and standardize flexibility products that match what large digital loads can safely offer—especially products that value fast response and short duration (UPS/batteries) and products that value scheduled reductions with lead time (compute shifting). Align telemetry/M&V requirements and allow participation through aggregators while preserving accountability (Goldberg et al., 2013; OpenADR Alliance, 2015).

Reliability system implications

Large load growth from data centers has been linked in public reliability discourse to heightened risk during extreme events when supply margins are tight (Reuters summarizing NERC, 2025). DR participation can mitigate these risks if it is predictable, fast enough for the system need, and operationally reliable. However, it can also introduce new risks if many large loads respond simultaneously to the same signal without coordination, potentially creating load oscillations or rebound peaks (Wang et al., 2016; Wierman et al., 2014).

Recommendation (reliability integration): require staged ramping, rebound management, and coordination protocols for large-load DR, especially when using automated or price-driven schemes across multiple sites (Wang et al., 2016).

Cybersecurity implications

DR expands the cyber-physical attack surface: external signals can influence physical and IT operations, and telemetry/remote control interfaces become targets. Security mechanisms in OpenADR include TLS-based secure channels and certificate-based authentication; profile specifications emphasize cipher suites and upgradability (OpenADR Alliance, 2015). Broader smart grid cybersecurity guidelines (e.g., NIST smart grid cybersecurity guidance) stress risk management for communications, identity, and interoperability interfaces relevant to DR implementations (National Institute of Standards and Technology, 2014).

Recommendation (cybersecurity): treat DR interfaces as critical OT/IT integration points: implement strong identity and certificate lifecycle management, segmentation between DR gateways and core IT workloads, and validation logic that enforces safe operating envelopes even if external commands are malicious or erroneous (OpenADR Alliance, 2015; Microsoft, 2022).

Environmental impacts and carbon accounting

DR's environmental impact is not automatically positive: the emissions effect depends on what generation is displaced and when. Research on marginal emissions emphasizes that marginal CO_2 emissions vary by region and hour, and are critical for evaluating interventions involving shifts in electricity demand (Holland et al., 2022). This matters for data center DR because shifting work or reducing load during specific hours may avoid high-emitting marginal generation, while shifting to other times or locations could increase emissions if the receiving grid is dirtier at that time.

For corporate reporting, Greenhouse Gas Protocol Scope 2 Guidance distinguishes location-based (grid-average) and market-based (contractual) methods for accounting emissions from purchased electricity (Greenhouse Gas Protocol, 2015). Data center operators increasingly discuss time- and location-sensitive approaches (e.g., hourly matching) and the use of carbon-aware scheduling to align operational decisions with real-world grid conditions (Google Cloud, 2023; Greenhouse Gas Protocol, 2015).

A national laboratory report estimated that in 2023, U.S. data center electricity use corresponded to a national-average indirect water consumption intensity of 4.52 L/kWh and an emissions intensity of 0.34 kg CO_2e/kWh for the electricity mix supplying data center locations, compared with U.S. overall averages of 4.35 L/kWh and 0.35 kg CO_2e/kWh (Shehabi et al., 2024).

Recommendation (carbon accounting for DR): where DR is claimed as a decarbonization action, quantify impacts using marginal (time- and location-sensitive) emissions factors for operational decision support, while separately maintaining Scope 2 reporting conformance. Transparently report whether DR actions involved on-site fossil generation (Holland et al., 2022; Greenhouse Gas Protocol, 2015).

Barriers, risks, and mitigation strategies

Major barriers repeatedly highlighted across research and field work include:

Operational risk aversion and reliability: DR programs not designed for data centers' risk tolerance and criticality can deter enrollment (Wierman et al., 2014).

Program mismatch and product design: limited event frequency or overly abrupt requirements fail to capture available flexibility (Wierman et al., 2014).

Battery degradation and economics: continuous cycling can be uneconomic without sufficient prices or product structures (Aksanli et al., 2014).

Telemetry and baseline disputes: dynamic workloads complicate baselines, requiring more sophisticated M&V (Goldberg et al., 2013).

Cybersecurity and governance: expanded control interfaces require stronger security engineering (OpenADR Alliance, 2015; National Institute of Standards and Technology, 2014).

Mitigation strategies include conservative "flexibility budgets" (e.g., cap DR to a % of load), hard constraints on minimum UPS state-of-charge, staged ramping and rebound smoothing, simulation-backed thermal safety envelopes, and red-team testing of DR control interfaces (Ghatikar et al., 2012; Microsoft, 2022).

Policy and market recommendations

Policy and market reforms that would plausibly accelerate safe, scalable data center DR include:

Interconnection and planning integration: treat flexible load contributions as part of load interconnection and resource adequacy planning, particularly where data center growth is concentrated and rapid (Shehabi et al., 2024; Reuters, 2025).

Standardized automation and interoperability: encourage adoption of standardized signaling (OpenADR) and interoperable DER/load controls to reduce vendor lock-in and integration cost (OpenADR Alliance, 2012; Southern California Edison, n.d.).

Clean flexibility incentives: structure DR incentives to favor low-emission flexibility (batteries, load shifting, thermal methods) over dispatch of high-emitting backup generators, using clear eligibility rules and reporting requirements (Baringa Partners LLP, 2023; Holland et al., 2022).

Fast services and telemetry standards: adopt or expand fast frequency response and reserve products with telemetry and performance scoring that matches inverter-based, battery-backed flexibility (EirGrid, 2014; Data Centre Review, 2024).

Research gaps and future directions
High-priority research gaps emerging from the synthesis are:

Verified, generalizable flexibility curves: more public datasets from operating hyperscale/colocation environments quantifying response time, duration, rebound, and failure impacts by mechanism (Ghatikar et al., 2012; Shehabi et al., 2024).

Baseline design for compute-shifting DR: robust M&V methods that fairly measure "counterfactual" load when workloads can be moved across time and geography (Goldberg et al., 2013; Wang et al., 2016). Battery lifetime economics under real market conditions: empirical field evidence connecting cycling, degradation, warranty terms, and revenue in fast services (Aksanli et al., 2014; Baringa Partners LLP, 2023).

Cyber-physical risk modeling: formal threat models and resilience engineering for DR control planes that connect grid signals to large critical digital infrastructure loads (OpenADR Alliance, 2015; National Institute of Standards and Technology, 2014).

Carbon accounting harmonization: methods that reconcile corporate reporting frameworks with marginal, time-sensitive emissions evaluation to avoid overstating environmental benefits of DR (Greenhouse Gas Protocol, 2015; Holland et al., 2022).

German Energy Agency's comparative work on data center flexibility across jurisdictions underscores that regulatory incentives and flexibility market maturity are uneven, particularly outside well-developed DR markets, reinforcing the need for region-specific policy pathways (German Energy Agency, 2025).

REFERENCES

1. Aksanli, B., Rosing, T., & others. (2014). Providing regulation services and managing data center UPS energy for effective power capping (conference paper). (Available PDF). https://aksanli.sdsu.edu/papers/DATE14_Aksanli.pdf
2. ASHRAE. (2016). ASHRAE TC 9.9 data center power equipment thermal guidelines and related guidance (white paper). https://www.ashrae.org/file%20library/technical%20resources/bookstore/ashrae_tc0909_power_white_paper_22_june_2016_revised.pdf
3. Baringa Partners LLP. (2023). Irish decarbonisation and end consumer benefits of grid-interactive UPS systems (report). https://www.baringa.com/contentassets/8771b5d7c2ee4f1598d746b5ceafbcd3/baringa_irish_decarbonisation_g-ups_study_report_v4_0.pdf
4. California ISO. (n.d.). Demand response registration and participation overview (program pages and presentations). https://www.caiso.com/generation-transmission/load/demand-response
5. Chen, H., Hankendi, C., Caramanis, M. C., & Coskun, A. K. (2014). Reducing the data center

electricity costs through participation in smart grid demand response programs (conference paper). https://www.bu.edu/peaclab/files/2014/03/IGCC2014.pdf

6. Chen, M., Gao, C., Song, M., Chen, S., Li, D., & Liu, Q. (2020). Internet data centers participating in demand response: A comprehensive review. Renewable and Sustainable Energy Reviews, 117, 109466. https://www.sciencedirect.com/science/article/pii/S1364032119306744
7. Data Centre Review. (2024, June 26). In the spotlight… Digital Realty & Enel X's DS3 dynamic frequency project (industry interview). https://datacentrereview.com/2024/06/in-the-spotlight-digital-realty-enel-xs-ds3-dynamic-frequency-project/
8. EirGrid. (2014). Delivering a secure, sustainable electricity system (DS3) programme overview (publication). https://cms.eirgrid.ie/sites/default/files/publications/DS3-Programme-Overview-2014.pdf
9. Enel X. (n.d.). Grid-interactive UPS supports Ireland's renewable energy transition: Digital Realty case study (case study). https://www.enelx.com/content/dam/local-uk/pdfs/enel-x-digital-realty-case-study.pdf
10. Federal Energy Regulatory Commission. (2011). Order No. 745: Demand response compensation in organized wholesale energy markets (final rule). https://www.ferc.gov/sites/default/files/2020-06/Order-745.pdf
11. Federal Energy Regulatory Commission. (2018). Order No. 841: Electric storage participation in markets operated by RTOs/ISOs (final rule). https://ferc.gov/sites/default/files/2020-06/Order-841.pdf
12. Federal Energy Regulatory Commission. (2020). Order No. 2222: Participation of distributed energy resource aggregations in markets operated by RTOs/ISOs (final rule). https://www.ferc.gov/media/order-no-2222
13. Ghatikar, G., Ganti, V., Matson, N., & Piette, M. A. (2012). Demand response opportunities and enabling technologies for data centers: Findings from field studies (LBNL report). https://eta-publications.lbl.gov/sites/default/files/LBNL-5763E.pdf
14. Ghatikar, G., & others. (2010). Demand response and open automated demand response opportunities for data centers (LBNL report). https://eta-publications.lbl.gov/sites/default/files/demand_response_and_open_automated_demand_response_opportunities_for_data_centers_lbnl-3047e.pdf
15. Goldberg, M. L., & others. (2013). Measurement and verification for demand response (National Action Plan on Demand Response report). https://www.energy.gov/sites/default/files/NAPDR%20M%26V.pdf
16. Google. (2023). Data center impact report (Taiwan) (report PDF). https://www.google.com/about/datacenters/static/pdf/english-google-tw-impact-report-ext-final.pdf
17. Google Cloud. (2023, October 3). Supporting power grids with demand response at Google data centers (blog post). https://cloud.google.com/blog/products/infrastructure/using-demand-response-to-reduce-data-center-power-consumption
18. Greenhouse Gas Protocol. (2015). GHG Protocol scope 2 guidance (standard). https://ghgprotocol.org/sites/default/files/2023-03/Scope%202%20Guidance.pdf
19. Holland, S. P., Kotchen, M. J., Mansur, E. T., & Yates, A. J. (2022). Why marginal CO_2 emissions are not decreasing for U.S. electricity: Estimates and implications for policy. Proceedings of the National Academy of Sciences. https://pmc.ncbi.nlm.nih.gov/articles/PMC8872735/
20. IEEE. (2018). IEEE 2030.5-2018 (standard overview page). https://standards.ieee.org/standard/2030_5-2018.html
21. International Energy Agency. (2023). Demand response (topic page). https://www.iea.org/energy-system/energy-efficiency-and-demand/demand-response
22. ISO New England. (n.d.). About demand resources and participation structures (market

guidance). https://www.iso-ne.com/markets-operations/markets/demand-resources/about

23. Liu, Z., Wierman, A., Chen, Y., & others. (2013). Data center demand response: Avoiding the coincident peak via workload shifting and local generation. Performance Evaluation, 70(10), 770–791. (Available PDF). https://niangjunchen.github.io/papers/dcdr.pdf
24. Liu, Z., Liu, I., Low, S., & Wierman, A. (2014). Pricing data center demand response. ACM SIGMETRICS Performance Evaluation Review, 42(1), 111–123. (Available PDF). https://smart.caltech.edu/papers/DCDRpricing.pdf
25. Microsoft. (2022, July 7). Microsoft datacenter batteries to support growth of renewables on the power grid (feature story). https://news.microsoft.com/source/features/sustainability/ireland-wind-farm-datacenter-ups/
26. National Institute of Standards and Technology. (2014). Guidelines for smart grid cybersecurity (NISTIR 7628 series). https://csrc.nist.gov/publications/detail/nistir/7628/rev-1/final
27. OpenADR Alliance. (2012). The OpenADR primer (overview). https://www.openadr.org/assets/docs/openadr_primer.pdf
28. OpenADR Alliance. (2015). OpenADR 2.0b profile specification (v1.1) (specification). https://gridlinktechnologies.com/wp-content/uploads/2021/05/OpenADR_2_0b_Profile_Specification_v1_1-4.pdf
29. Reuters. (2025, August 4). Google agrees to curb power use for AI data centers to ease strain on U.S. grid when demand surges. Reuters.
30. Shehabi, A., Smith, S. J., Hubbard, A., Newkirk, A., Lei, N., Siddik, M. A. B., Holecek, B., Koomey, J., Masanet, E., & Sartor, D. (2024). 2024 United States data center energy usage report (report). https://legalectric.org/f/2025/09/lbnl-2024-united-states-data-center-energy-usage-report_1.pdf
31. Southern California Edison. (n.d.). Automated demand response with OpenADR for business (program page). https://www.sce.com/business/save-costs-energy/savings-strategies-for-businesses/what-is-demand-response/automated-demand-response-with-openadr
32. U.S. Department of Energy. (2006). Benefits of demand response in electricity markets and recommendations for achieving them (report to Congress). https://www.energy.gov/oe/articles/benefits-demand-response-electricity-markets-and-recommendations-achieving-them-report
33. Wang, H., Huang, J., Lin, X., & Mohsenian-Rad, H. (2016). Proactive demand response for data centers: A win–win solution. IEEE Transactions on Smart Grid, 7(3), 1584–1596. (Available PDF). https://ncel.cuhk.edu.cn/sites/default/files/07346506.pdf
34. Wierman, A., Liu, Z., Liu, I., & Mohsenian-Rad, H. (2014). Opportunities and challenges for data center demand response (survey paper). (Available PDF). https://smart.caltech.edu/papers/dcdrsurvey.pdf