THE INTEGRATED GRID
CAPACITY AND ENERGY IN THE INTEGRATED GRID
Table of Contents

Executive Summary .................................................... 3
Section 1: Capacity and Energy—Why Is It Important? ............. 7
Section 2: Emerging Supply and Demand-Side Technologies and Impact on Capacity and Energy ........................................ 8
Section 3: Cost to Supply Capacity ........................................ 14
Section 4: Emerging Wholesale Markets and Retail Rate Structures ........................................ 20
Section 5: Research and Development Needed to Enable Capacity from DER .... 28
Section 6: Conclusion .................................................... 32
References ................................................................. 33
Executive Summary

This paper addresses the role of capacity and energy in the Integrated Grid by providing insights from EPRI’s research in the following areas:

- How individual resources may contribute differently to the system’s capacity to deliver energy;
- How changing supply and load characteristics make it necessary to distinctly address both energy and capacity on wholesale and retail levels;
- The cost of capacity, based on an assessment of cost structures of several U.S. utilities;
- Emerging trends in wholesale markets and retail rate structures to value capacity and energy as distinct elements of those markets/structures; and
- Key research to enable DER to provide both capacity and energy.

The term “capacity” in the context of this paper is used in a simplified way to represent the supply and delivery capacity required to serve consumer demand. To meet that demand reliably, however, the system’s resources must be able to do more than just steadily generate power. Together they must be flexible enough to react to sudden changes in load or generation, such as when a large generator is suddenly forced offline. This required flexibility from generators may be characterized by a required capacity response.

The electric power system is beginning to change—rapidly in some areas—with the rise of generation from distributed energy resources (DER), such as small natural gas–fueled generators, combined heat and power plants, electricity storage, and solar photovoltaics (PV). Additionally, smart appliances, home energy management systems, and electric vehicles are providing new methods for consumers or third parties to manage energy use. To help realize fully the value of distributed resources while maintaining established standards of quality and reliability, EPRI introduced the concept of an Integrated Grid [1] and a framework to assess the benefits and costs of integrating DER [2].

EPRI’s Integrated Grid concept paper outlined the need for global collaboration in four key areas:

1. Interconnection rules and communications technologies and standards;
2. Assessment and deployment of advanced distribution and reliability technologies;
3. Strategies for integrating DER with grid planning and operations; and
4. Enabling policy and regulation.
A series of “ancillary services” such as frequency regulation, voltage support, load following/ramping, inertia, and several categories of operating reserves. Capacity, in this context, represents more than a power-output rating, having multiple attributes that are required by a power system for reliable operation. Furthermore, this capacity must be planned to support the expected maximum demand over a long-term horizon.

Informing all stakeholders on the importance of capacity and energy in an integrated grid will be an important step as various jurisdictions consider policy and regulation that reflects the influence of renewable resources, distributed generation, energy storage, and new, more efficient loads. This report does not intend to state or imply preference regarding any particular market rules or retail rate structures insofar as those rules/structures encourage sufficient capacity with the flexibility required to continually satisfy demand.

With respect to enabling policy and regulation, the Integrated Grid paper outlined the following four key requirements:

- **Capacity-related costs** are considered as a distinct element of the total cost of grid-supplied electricity to ensure long-term system reliability.
- **Power market rules** operate to ensure long-term adequacy of both energy and capacity.
- **Policy and regulatory frameworks** are developed so that costs incurred to transform to an integrated grid are allocated and recovered efficiently and equitably.
- **New market frameworks** use economics and engineering to equip investors and other stakeholders in assessing potential contributions of distributed resources to system capacity and energy.

---

**Key Terms Defined**

**Energy**: Measured in kilowatt-hours (kWh), is required to operate consumers’ lighting, equipment, appliances and other devices, often called loads. Energy is measured over a period of time.

**Demand**: Is a measure of power, of how much energy a consumer (or group of consumers) requires at a specific point in time. Though demand can be measured at any point in time, the peak demand level is often an important characteristic of a load, a premises, or an entire power system.

**Capacity**: Is the maximum capability to supply and deliver a given level of energy demand at any point in time. **Supply capacity** describes networks of generators designed to serve load as it varies from minimum to maximum values over time. **Delivery capacity** is determined by the design and operation of the power transmission and distribution systems [1].

---

1 Operating reserves may be “spinning” reserve—unused capacity of operating (“spinning”) generators—or non-spinning generators that can provide capacity within a short period of time (for example, 15 or 30 minutes).

2 Key terms energy, demand, and capacity refer specifically to electricity in the context of this paper.
Key Points – The Importance of Capacity and Energy in an Integrated Grid

Several key insights from this paper highlight the importance of capacity and energy in an Integrated Grid:

• New trends in interior climate control, such as heat pumps, are often more energy efficient devices but do not necessarily reduce peak demand.

• Completely displacing a consumer’s energy requirements with self-generation, as with a zero-net-energy residence, often does not alleviate the need for capacity from the utility.

• In some regions of the U.S. and other countries, peak system load is increasing at a faster rate than overall energy consumption.

• Variable, supply-side resources (such as wind and solar) can supply energy at low, or zero variable cost, but alone are generally not sources of firmly available capacity.

• Emerging energy storage and energy management systems (capable of reducing peak demand, for example) provide flexible capacity for short periods of time, but do not currently serve as long-term energy resources.

• Fixed capacity costs for supply and delivery could represent for individual utilities between 35% and 75% of their average residential electricity bill [3].

• In general, generating units designed to be quickly and sparingly dispatched receive a large percentage of their income from transactions outside of the energy market. For instance, in the ISO New England Market, it is estimated that 80% of revenue for a combustion turbine generating unit comes from capacity/ancillary service payments [4].
Key Points – The Importance of Capacity and Energy in an Integrated Grid (continued)

• Recently, an increasing portion of total central generation revenue is coming from capacity-related payments in some areas, especially those with significant variable generation. For large combined-cycle power plants in Spain, a significant percentage of total revenue comes from capacity payments because the plants are dispatched sparingly (less than 1000 hours per year) [5].

• Centralized capacity markets have existed in some areas for a decade or more, with the objective of long-term resource adequacy. While they have provided a mechanism for encouraging capacity additions when needed, they have exhibited some price volatility as market rules have evolved [6].

• Emerging capacity markets, such as the UK’s forward market, are providing longer-term price signals. While the market is designed for a delivery year that is four years ahead, it also provides contracts for new or refurbished units for an additional 15 years or 3 years following the delivery period, respectively. This is intended to provide long-term incentives for building of needed generation capacity to meet planning objectives.

• Some emerging residential electricity rates are attempting to address capacity-related costs through modifications to traditional rate elements.

• Significant R&D progress has been made in developing open standards, protocols, and tools for integrating distributed energy resources. Before such resources can be considered reliable for long-term capacity at scale, additional R&D is needed to assess the interoperability and cyber security of proposed solutions. A better understanding of consumer behavior is critically important for many of these resources, and is a key research imperative at EPRI.
SECTION 1: Capacity and Energy—Why Is It Important?

It is commonly understood that electric energy is generated at power plants and delivered to consumers over a network of power lines. Less familiar are those capabilities required from wires, transformers, generators, and other power-system equipment to constantly meet consumers’ dynamic demand for electric energy. Historically, utilities have engineered components and systems to meet the requirements of consumer demand, without highlighting their diverse and crucial capabilities.

These capabilities are often collectively referred to as capacity\(^3\). The system, as a whole, must have the capacity to generate and deliver the maximum energy required at every location within. Additionally, the system must have the capacity to react to small shifts in generation or load as well as large, unplanned disturbances such as the sudden outage of a large generator or transmission line. These capacity components are rarely unbundled or priced separately from energy that consumers purchase, but these capabilities are integral to the delivery quality and reliability that consumers expect. Also, many of the resources necessary to provide capacity are long-lived: expected to remain in service for 20–50 years.

Each resource may contribute differently to the system’s capacity to deliver energy. Depending on the system and the devices connected, resources may include central generators, distributed generation, energy storage, and controllable load. Individual resources’ contributions to system capacity can be characterized by availability and flexibility\(^4\).

Availability — describes the ability of a resource to be present over a period of time. In order to be considered available, individual resources must be either providing energy or capable of providing energy when called upon by the system operator. Resources that require extended maintenance each year and those with higher probabilities of unplanned outages or unavailable fuel supplies have less availability, and are expected to contribute less to meeting future load needs. Availability during peak demand is often most important as these periods require the most resources to supply energy and avoid involuntary load shedding.

Flexibility\(^5\) — describes the capability of resources to adjust their energy output as system conditions change. Because load variations and forced outages are not always known in advance, flexibility is needed to maintain system reliability. Flexibility can be measured in three main ways: the range of output adjustability, the speed that output can change, and the duration that output levels can be maintained. To address a variety of system conditions, flexibility is needed both for increasing and reducing energy output. Resources that lack the ability to adjust energy output, or to sustain adjustments over an extended period, contribute less to overall system flexibility\(^6\).

Because individual resources are unique, it is up to the system operator, market operator, and the resource owners to ensure that enough of each capability is present to ensure electricity is delivered reliably. Planners estimate future system needs based on forecasted system conditions and establish mechanisms to deploy a future portfolio that can meet those needs.

---

\(^3\) Qualifiers may refer to individual properties, like the “maximum output” capacity or “ramping” capacity of an individual generator.

\(^4\) For simplicity, the combination of flexibility and availability from resources is referred to generally as capacity throughout the rest of this report.

\(^5\) Flexibility can include many different specific attributes, and not only those related to active power or energy, but the whole collection of services that are needed to maintain a reliable, secure, and efficient electric power system.

\(^6\) Energy and capacity from individual resources must be deliverable and not significantly limited by transmission infrastructure. For resources connected to the distribution system, deliverability may be affected by the distribution infrastructure as well as transmission.
SECTION 2: Emerging Supply and Demand-Side Technologies and Impact on Capacity and Energy

Long-term capacity requirements are driven by anticipated consumer demand. However, even on a national scale, the nature of consumer demand is changing. In the past, energy use and peak demand each grew with some correlation to one another and often to econometric indices such as gross domestic product (GDP). In recent years total electricity use has grown slowly or even declined in certain regions of the United States and other countries. Researchers to a varying extent attribute this to changes in the economy, weather, or increased energy efficiency. However, peak demand has not followed the same trend. Figure 1 demonstrates this in EIA-reported data since 1992. Researchers have a less definitive understanding of this peak demand growth, and its relationship to energy consumption. Some, such as the U.S. Energy Information Agency (EIA), point to an increased use of interior climate control and a more service-oriented economy as likely causes [7, 8].

Regionally, a report from the New York Independent System Operator (NYISO) indicates a similar trend (shown in Figure 2) to the nationwide data. The report concludes that energy consumption in New York State has grown only at 0.23% annually for the past decade, while peak demand has grown at 1.99% annually. The report expects a similar trend over the next decade, projecting energy use to grow at 0.16% while peak demand is forecasted to grow at 0.83% [9].

In recent years total electricity use has grown slowly or even declined in certain regions of the United States and other countries... However, peak demand has not followed the same trend.

Figure 1: U.S. Energy consumption versus summer peak demand (Source: EIA Form 411)
Several technological trends could be influencing this observation at the system level. For example, while some energy efficiency measures also help reduce peak demand, many new loads, consumer behaviors (with home automation), and distributed generation sources reduce average energy consumption but leave peak demand from the grid nearly unchanged. In milder climates, for instance, heat pump technology is an efficient alternative to traditional resistance heating. Heat pumps extract heat from the outside air, rather than making resistance heat indoors. However, extracting heat from outdoor climates becomes much more difficult as temperatures drop. Not only is the work more difficult (making the unit less efficient), but the large temperature differential reduces the effectiveness of the compressor [10]. If the demand for heat is large enough, exceeding the compressor’s capability, an additional resistance element must meet the remaining heating load. At low outdoor temperatures (generally below freezing), this element could be required continuously for several hours [11].

Figure 3 illustrates a situation in which a residential heat pump and resistance backup were monitored over a 12-hour period. The outdoor temperature [black line] begins at near-freezing before increasing gradually through the day. While the temperature is low, however, maintaining a near-constant indoor temperature requires the heat pump (blue line) to run continuously, but also requires the resistance backup (red line) for most of the morning. The total power consumed by the unit is the sum of the blue and red curves, for nearly a 14kW peak demand for a unit that requires only 4kW during operation of the heat pump alone.

A heat pump is more energy efficient than traditional alternatives; however, it has very similar demand to resistance heating during weather cold enough to require the backup element. EPRI is conducting research of next-generation heat pumps, using technology such as variable-speed compressors to reduce this peak demand [12].

Figure 2: New York State’s peak demand growing while energy consumption has flattened significantly (Source: NYISO)

7 Qualifiers may refer to individual properties, like the “maximum output” capacity or “ramping” capacity of an individual generator.
While consumers require less energy with more efficient appliances, they also may have the option of displacing some or all of their energy requirements with self-generation. However, though consumers may even completely displace their energy use over the course of a day, week, or year, often their capacity requirement remains unchanged. For example, Figure 4 shows a monitored residence that is designed to be zero-net-energy over the course of a calendar year. The net export from the premises is highlighted in green, while the net import is highlighted in blue. During the mid-day period, production from local photovoltaic generation far exceeds local demand, and the consumer requires the grid’s delivery capacity to export energy. Additionally, periods of peak demand often occur hours later than peak solar production, and the consumer requires both generation and delivery capacity from the grid in order to meet the demand of their end-use devices. Such consumers may be zero-net-energy, but still rely on the services provided by generation, transmission, and distribution infrastructure needed to serve their premises.

While consumers require less energy with more efficient appliances, they also may have the option of displacing some or all of their energy requirements with self-generation. However, though consumers may even completely displace their energy use over the course of a day, week, or year, often their capacity requirement often remains unchanged.
Not only is the demand side of the equation changing, but so is the supply. The capacity required to serve consumer loads has been provided almost exclusively from large, central-station thermal and hydroelectric generation. Depending on the fuel sources and operational characteristics of the plants themselves, each plant generally provides both energy and capacity in significant measure. Metrics such as availability and flexibility can be defined and calculated for units individually and comparatively as components of a fleet on a reasonably level playing field.

However, new resources that are coming on to the grid, even at the bulk system level, do not provide energy and capacity in the same way. Some, such as PV and wind power, have more consistent (rather than random) periods in which full capacity is not available (based on reduced wind speed or solar irradiance.) Metrics such as unforced capacity\(^8\), which reflect the contribution of different resources during periods of peak demand, show the significant differences that accompany these new resources. For example, Figure 5 shows the unforced capacity contribution of different resources on the PJM system. Because variable renewable resources don’t have full capacity at all times, they don’t contribute to fulfilling capacity requirements in the same way that a thermal generator does. Other resources, such as storage or demand response, are limited in how long they can be active or how often they can be used, which also reduces their contribution to capacity needs. As these variable or energy-limited resources become more common, system planners are presented with an increasingly complex

\[8\] Unforced capacity generally refers to a generator’s average available capacity, which is adjusted for periods of outage or forced derating.

Figure 4: Monitored zero-net-energy home over an average week shows that zero-net-energy does not equal zero capacity (Data Source: BiraEnergy)

Because variable renewable resources don’t have full capacity at all times, they don’t contribute to fulfilling capacity requirements in the same way that a thermal generator does.
task of blending the unique characteristics of these resources in planning for future capacity needs. Otherwise, the system may become overbuilt and expensive, or underbuilt and prone to costly service interruptions.

Variable renewable resources, both at the bulk power system level and on the distribution level, increase the system’s variability and uncertainty. Renewable resources are generally considered to be “must-take” resources because their variable costs are essentially zero. So under normal conditions these resources serve part of the electric system load, leaving a residual load shape to be served by other dispatchable capacity, either hydroelectric or fossil-fueled units. Since demand and generation must be balanced continuously, this new load shape may require greater flexibility from the remaining resources.

However, the existing mix of conventional generators has been developed for a less “peaky” load shape, and some generators may lack the flexibility needed as such conditions become more common. These plants may deteriorate more rapidly with the new requirements, and may need to be replaced sooner than planned. New capacity could potentially be of a more flexible type, which is usually less expensive per unit of capacity but are more expensive as generators of energy. Additionally, some new resources, such as energy storage and various forms of demand response, can often provide greater flexibility than existing generation because of their fast response rates. System operators will likely benefit from this flexibility when scheduling resources to meet the load’s requirements.

Figure 5: Unforced capacity (UCAP) values for PJM fleet in 2013 [* indicates default value taken for PJM RPM auction]
(Data Source: PJM)

9 Many bulk connected resources in restructured market areas actually offer bid costs and can be dispatched when prices drop very low or negative, due to transmission constraints or commitment constraints of other resources.

10 Some generating plants also become less efficient and may have higher emissions as their output becomes more variable.
The combined impact of adding capacity with less availability, as well as increased flexibility requirements, results in a power system with a total installed capacity that is increasing faster even than peak demand. This trend has been noticed in a number of countries around the world, particularly those with growing renewable generation, as shown in Figure 6.

Figure 6: Installed generating capacity is growing in many countries, even in excess of peak demand (Data Source: EIA)

The combined impact of adding capacity with less availability, as well as increased flexibility requirements, results in a power system with a total installed capacity that is increasing faster even than peak demand.

11 Total installed capacity reflects the sum of the peak output rating of each generator, regardless of timing, availability, or flexibility.
SECTION 3: Cost to Supply Capacity

In the Integrated Grid concept paper, national average costs for generation, transmission, and distribution components were summarized as costs related to serve the consumer with energy (kWh) and costs for the capacity that delivers energy and grid-related services. Based on the U.S. Department of Energy’s Annual Energy Outlook 2012, an average U.S. residential consumer consumes 982 kWh per month, paying an average bill of $110 per month, with the average cost of $70 for generation of electricity. That leaves $30 for the distribution system and $10 for the transmission system, as shown in Figure 7.

Calculating the total cost of capacity followed the analysis summarized in Figure 8. These values were based on the assumption that most costs associated with T&D are related to capacity (except for a small fraction representing system losses—estimated to be $3 per month per consumer from recent studies in California) [13]. Working with recent data from PJM [14] regarding the cost of energy, capacity, and ancillary services, it was estimated that 80% of the cost of generation is energy-related, leaving the rest for capacity and grid services. As illustrated, the combination of transmission, distribution, and the portion of generation that provides grid support averages $51 per month while energy costs average $59 per month.

Figure 7: National level data shows the average consumer consuming 982 kWh/month with an average bill of $110/month (Data Source: EIA)

Figure 8: National level data shows that the average bill can be broken down into roughly $59 for energy and $51 for capacity (Data Source: EIA, SCE, PJM)
Understanding Capacity Cost from Regulatory Filings

Additional analysis is required to better understand the nature of capacity and energy costs, as well as the range across different utilities. Helpful to this task is that fixed and variable costs for a FERC jurisdictional investor-owned utility can be readily estimated from FERC Form 1 filings. These filings provide detailed financial and operating statements, including the balance sheet, income statement, and cash-flow statement, along with other information such as energy sources and disposition. Fixed costs (those that don’t vary directly with sales or output) include return of and on capital (depreciation and return on equity), interest, taxes, and most operating and maintenance (O&M) expenses. Variable costs include fuel, variable O&M expenses, and expenses for power purchased from other entities.

To develop a range of estimates, ten investor-owned utilities of various sizes, organizational structures, and geographic territories were selected from across the United States. The reported annual costs of generation and delivery (transmission and distribution) are shown in Figure 9. For the selected utilities, the generation component ranged from 30% to 70% of the average residential electricity bill.

---

12 The most recent year with complete financial and detailed sales data at the time of the analysis was 2012.

13 Fixed costs do change as utilities invest in their systems, and costs associated with past investments decline as they are depreciated. Whether total fixed cost for a utility declines or falls depends on the difference between new investments and the total depreciation. The path of future investments is not considered fixed until plans are committed. Until then they can possibly be affected by DER, conservation, or other changes to demand and energy growth expectations.

14 Some estimation is required even for total fixed and variable cost; the split of variable and fixed O&M is not known, for instance, nor is the utility’s true cost of capital.
For integrated utilities, generation comprises fixed costs, such as equipment or fixed O&M, as well as variable costs, such as fuel or purchased power. Delivery costs are mostly fixed infrastructure- or customer-related costs, with a small, variable component due to losses. For utilities in structured markets, generation cost may instead be partially or even entirely purchased power. Power purchased in structured markets is treated as variable cost, but it contains a blend of spot purchases and contracts, which may incorporate some fixed components. Details of independent generators’ cost structures are not public and are generally not relevant to power purchasers.

Further breakdown of costs for the same ten utilities is included in Figure 10. Two of them are distribution-only utilities that own no generation. Their entire generation cost is power purchased from generators or other utilities. The remaining utilities are either integrated utilities or market utilities that have not fully divested their generation. Distribution-only utilities (whether they have any transmission or not) may have contractual commitments to generators, however, they are not otherwise committed long-term to any generation source.

Figure 10: Cost components for ten selected investor-owned utilities

---

15 The utilities that own both generation and delivery assets, but are not vertically integrated, are represented by the term “hybrid” in Figures 9-11.
Sorting all of the components into fixed and variable categories, a total estimate for fixed (or capacity) costs and variable costs, based on the utilities’ total costs is shown in Figure 11. As shown, the total fixed portion of the surveyed utilities’ costs ranged from roughly 35% to 75% of the average residential electricity bill.

**Wholesale Market Considerations**

With respect to dynamics on the wholesale side, the changing nature of energy resources is significantly altering market activity in different regions. In market-based systems, generators are free to sell power to retailers or marketers under bilateral contracts, and/or they can sell power into a structured spot market in which the operator essentially performs a continuous series of short-term auctions. The prices that result from these auctions vary throughout each day and depend strongly on the variable cost of fuel. For a generator, the minimum price expected to be offered would cover its short-term variable costs, such as fuel, but provide no return on investment. The price paid to the generator when it is dispatched, however, is the locational marginal price. The locational marginal price typically is determined by the bid of the highest-cost generator that would be required to serve an additional unit of load at each location. For any running units where their bid is less than the locational marginal price, this implies some

---

16 Variable costs might also include emissions allowances, reagents, and other costs that vary with output, net of any per-unit tax credits.

17 Transmission constraints complicate this picture. For instance, a low-cost generator in one area may be temporarily undeliverable to load in other areas, and high-cost generators may have to be used instead. Locational marginal price indicates the lowest cost to generate and deliver another increment of output to each location, respecting all transmission constraints.
contribution toward their fixed costs, and generating plants rely on these contributions to remain in business.

Prospective generators typically locate new units where they may expect to be inframarginal, having a minimum offer that is lower than the average price paid by the auction at that location. An efficient baseload unit can expect returns from energy sales on a regular basis, but peaking units depend largely on revenues obtained during peak hours. Given their relatively high fixed costs for little energy generated, they need to generate into very high prices. Peak loads and prices are highly weather dependent, and in most markets peak prices are capped. If capped values are too low, peaking generators may face revenues insufficient to remain in business based on energy payments alone.

In Europe, a high uptake of renewable generation and increased international supply of relatively cheap coal has reduced wholesale market prices significantly [5]. That in turn has reduced the operating hours (and energy-based revenue) of central generating resources, particularly natural gas turbines and combined-cycle (CCGT) units. This is most obvious in Spain, where the average operating hours for CCGT plants has declined steadily to less than 1000 hours per year (Figure 12). Plants necessary to provide for system reliability will rely increasingly on out-of-market capacity or ancillary service payments for revenue.

This trend is not unique to Europe. In a number of U.S. markets, central generators (particularly those relied upon during peak hours) often depend on revenue from

![Figure 12: Combined-cycle gas plants in Spain have been decreasingly dispatched, making them more reliant on capacity and ancillary service payments (Source: ACER/CEER).](image-url)
capacity and ancillary service payments. Figure 13 shows anticipated revenue in the New England ISO market for existing power plants, and indicates that upwards of 80% of a peaking units' operating revenue is likely to come from non-energy related sales [4].

In Europe, a high uptake of renewable generation and increased international supply of relatively cheap coal has reduced wholesale market prices significantly. That in turn has reduced the operating hours (and energy-based revenue) of central generating resources, particularly natural gas turbines and combined-cycle (CCGT) units.

Figure 13: For certain technology types, most of their revenue comes from direct capacity or ancillary service payments (Source: Entergy Research)
As power system resource portfolios continue to include more low- or zero-variable cost resources, changes may be expected in the way in which both new and existing resources earn revenue. A few studies have evaluated the higher penetrations of variable generation and the continued impact on reducing system energy prices. The resulting price impact will likely make many existing generators uneconomic, and will impact the economics of the renewable resources themselves. In some cases the generators that become uneconomic are no longer required to meet the system’s energy and electric reliability needs, and the market forces appropriately drive those resources to retirement. However, in other cases these generation plants may continue to be necessary for reliability. In such cases it is necessary to understand how a market’s structure may result in these plants recovering insufficient revenue to remain in business. Below are some examples of existing markets and how they are changing to address these and other issues:

**Centralized Capacity Markets**

In theory, a structured spot market can provide sufficient opportunity for prospective generators to construct and operate the capacity necessary for reliability. However, some of the highest-cost generators (typically peaking units) in such a market can be expected to generate for relatively few hours each year, only in the very highest-priced hours. If generator bids are capped too low during these hours, then these marginal generators may not produce enough revenue to remain in business. This effect is known in the industry as the “missing money” problem [15–17], referring to the revenue needed to support reliability that cannot otherwise be obtained within the spot market structure.

Some electric systems with structured spot markets have attempted to address this issue by adding centralized capacity markets to their designs. The intent is to ensure that capacity is procured sufficient to meet designated resource adequacy targets, with all cleared resources paid the same clearing price. One such market is the PJM Reliability Pricing Model (RPM), which has operated since 2007 [18]. The RPM and similar capacity markets have changed over years to meet locational needs and include resources with different characteristics. Though the goal is to provide price signals that increase the certainty in payments needed to drive investment in capacity, there has been significant price volatility from year to year, as shown in Figure 14. This volatility is somewhat inherent due to changing fuel prices and imperfect knowledge of future reliability needs. However, additional volatility has resulted

---

18 The studies typically assume all resources are bidding true variable costs into the market, often ignore the effect of prices during shortage periods, and ignore any bilateral sales or centralized capacity market payments.
Market rules, which can have a significant impact on clearing prices, and which have changed significantly over the market's history [6]. These variations include the process for setting the demand curve, net cost of new entry calculations, penalties for non-performance, and rules for demand response and variable generation. Recently, the United Kingdom followed the example of eastern U.S. markets by introducing a four-year ahead forward capacity market. However, the additional capacity procured during a forward auction is not immediately available. Therefore the UK system operator recently retained 2.56GW of generation and interruptible load via out-of-market contracts to ensure sufficient reserves during the 2015 winter months [19]. Examples: New York, PJM, New England, United Kingdom.

The intent is to ensure that capacity is procured sufficient to meet designated resource adequacy targets, with all cleared resources paid the same clearing price.
Forward Period, Commitment Period and New Entrant Contract Length of Centralized Capacity Markets

Capacity markets exist in a number of U.S. electricity markets and now are gaining interest in European electricity markets as well (some examples are given in Figure 15). Three important components of the capacity market are:

- The forward period: How long ahead of the capacity period the auction is held;
- The delivery period: The length of the period for which the capacity commitment is for; and
- Contract length: How long cleared resources are guaranteed the price.

Longer commitment periods provide greater lead-time for building capital-intensive plants, but periods that are too long can cause uncertainty in the capacity need. Longer delivery periods can create greater price certainty but may create distortions for periods of time that have different needs (i.e., winter period vs. summer period). Finally, in the UK capacity market, new entrants can receive a 15-year contract, characterized by greater certainty over a good portion of the unit’s life. Consumers, however, may incur additional risk given the long-term nature of these contracts [20].

![Figure 15: Various forward period, delivery period and contract lengths](image-url)

© 2015 Electric Power Research Institute (EPRI), Inc.
**Administratively Set Capacity Payments**

In some jurisdictions, fixed investment costs are partly covered by administrative payments in addition to revenues earned from the energy market. The payments and eligibility can vary, but the typical approach is that such administrative payments equal the capital costs of peaking units that cannot be recovered in energy markets. In Spain this comprises two payments: an investment incentive and an availability (firmness) payment. The availability payment is paid exclusively to combined cycle plants at 10,000 Euros per MW per year for the first 20 years of its life. The firmness payment is paid to combined cycle plants and coal plants at some percentage of the plant’s capacity (for example, 80–90%) with the rate based on expected energy market revenues (for example, 5,000 Euro per MW). **Examples:** Spain, Ireland, Chile.

**Enhanced Scarcity Pricing**

In energy-only markets, contributions toward a generator’s fixed costs are obtained entirely from the energy market without dedicated capacity payments. The magnitude of these contributions depends entirely on how a plant’s energy costs compare with the market price when they are running. The lowest-energy-cost generators such as nuclear plants may receive some fixed-cost contribution in virtually every hour they operate because the market price is based on bids from higher-cost units. For the most expensive units, the price must rise higher than its cost for enough time for it to generate sufficient revenue. Rather than explicitly controlled reserve margins, development of sufficient capacity is encouraged through scarcity pricing mechanisms. Scarcity pricing occurs when capacity is scarce, although it is typically induced when the system is short on reserves rather than risking a load shedding event. While many markets have various forms of scarcity pricing, those in energy-only markets must have higher scarcity price caps and other unique rules to account for the limited side-payments. For example, the Australian National Electricity Market may perform numerous simulations to determine what price cap may lead to sufficient profit of marginal peaking resources to meet a pre-determined reliability standard of less than 0.002% unserved energy. The price cap recently was as high as $12,500 per MWh [21]. In the United States, the Electric Reliability Council of Texas (ERCOT) currently holds a price cap of $9,000 per MWh. ERCOT recently implemented a market mechanism referred to as the Operating Reserve Demand Curve. Its design includes a "price adder"\(^{19}\) equal to the operational probability of lost load multiplied by the estimated value of lost load (VOLL), that is applied to the energy market clearing price [22]. Therefore, the price is increased not only when the system is truly scarce of reserve, but also when the probability of reserves becoming scarce is not zero. This “adder” is paid to online resources as well as offline resources who can be available within 30 minutes. By comparison, in energy markets with capacity payments, these scarcity prices are typically around $1,000 per MWh and are applied only during reserve shortages [23]. **Examples:** ERCOT, Alberta Electric System Operator, Australian National Electricity Market.

---

\(^{19}\) A “price adder” is, in this instance, an administrative increase in the wholesale price of electricity, resulting in additional revenue for power providers and additional cost to power purchasers.
When considering supply and demand issues in the energy sector, natural gas is commonly compared and contrasted with the electricity system. The network of transmission and distribution pipelines delivers natural gas produced by thousands of merchant wellheads. Short-term and long-term wholesale prices are determined at market centers across the U.S. and Canada, and infrastructure investments for upstream gas exploration and production are based on market prices for natural gas.

However, with gas markets, capacity (availability and flexibility) for hour-to-hour variation in demand is supplied by pipelines and compressor stations that compress the gas to a high pressure (through a process known as line-pack). This line is then unpacked (reducing pressure) to meet hourly and daily demand variations. This is somewhat analogous to central generation plants changing their output based on a signal from control centers to ensure that generation at any instant is equal to electricity load. During periods of high natural gas demand such as extremely cold weather, large underground gas storage can be called upon for the needed flexibility. This is roughly analogous to reserve margins in the electricity sector, which are expected to meet maximum demand over a longer term.

The natural gas infrastructure that provides capacity includes pipelines, compressor stations, and storage that is separate and distinct from the infrastructure that supports upstream exploration. Generally, market prices drive investment for upstream activities, while the elements that provide the capacity are mostly procured through long-term bilateral contracts.

In electricity, power generators provide energy as well as the availability and flexibility required for reliable operation. Given generation’s multi-dimensional role, it is more difficult to design a capacity mechanism that provides an adequate price signal for building generation infrastructure, based on long-term planning.
Retail Rates

As mentioned previously, retail rate design is equally important during this transformation. Often subject to regulatory oversight, rates normally are designed to accomplish three objectives:

- Enable cost recovery of prudent utility investments and purchases;
- Provide incentives or price signals for certain customer behaviors that may reduce cost of service or environmental impact; and
- Avoid burdening low-income customers with high costs [24].

Aside from redistributive issues, rates generally are considered equitable if customers pay an appropriate share of total utility costs, in proportion to the costs they cause, along with a percentage of common costs. Rates encourage efficient consumption decisions if they reward or penalize these decisions in proportion to their resulting costs or savings. In either case, such cost causation is an important principle for both allocating costs and designing rates that guide consumptive decisions20.

A customer’s load profile can generally be characterized by two quantities: total energy consumption (in kWh) and the contribution to peak demand21 (in kW). The resulting costs they create are a function of both quantities. For most customers, the total energy consumption is easily measured, and the contribution to peak demand is assumed to be proportional to that measurement. Utilities often then segment customers by class and rate schedule, with each rate schedule carrying with it an expectation of load shape based on limited factors. Especially for residential customers, there are generally only a small number of load shapes used estimating contributions to peak demand22, with the expectation that customer load profiles under the same rate class are largely similar. Specifically because residential customers are generally only billed based on total energy consumption, there is an implied expectation that total energy use corresponds directly with a certain contribution to peak demand.

While these two variables may have been linked more strongly in the past, new technologies and consumer behaviors are changing that relationship. A customer with electric vehicles, PV, or storage operating behind the meter...

...because residential customers are generally only billed based on total energy consumption, there is an implied expectation that total energy use corresponds directly with a certain contribution to peak demand... new technologies and consumer behaviors are changing that relationship.

20 Though cost causation is an important principle, methods of assigning costs are inexact and often subject to significant debate.
21 The term “peak demand” is used generically here. In the power system, certain costs may be related to premises, feeder, substation, or system peak demands.
22 Often there is only one residential rate, although some utilities have all-electric rates or rates that differ according to whether the premises uses electricity for heating.
will likely have a net load profile much different from a residence with only electric HVAC and other common appliances. Though total energy consumption has been the measurement to determine the bills of residential customers for more than 100 years, metering technologies have advanced considerably. Many modern-day meters are capable of recording and transmitting consumption over 15-minute intervals for any customer, and these capabilities may be needed to capture the important changes occurring within customer premises.

Also important to consider is the emphasis that rates place on each of the two variables. Consider a certain rate schedule as existing somewhere along a spectrum as illustrated in Figure 16, with one end of the spectrum entirely based on contribution to peak demand, and the other entirely based on total energy consumption. Today, because of the way most residential customers are billed, their rates are to a great extent or even entirely based on total energy consumption.

A total consumption-based rate naturally puts more emphasis on energy efficiency and conservation. However, it can also result in peak demand-related costs being avoided through certain kinds of energy efficiency measures and behind-the-meter generation [25]. Likewise, a rate that’s biased too heavily toward peak demand may result in variable costs being avoided by load management or energy storage. Creating a more balanced emphasis between total energy consumption and contributions to peak demand provides incentives for both energy efficiency and demand management, as well as the technologies that address each [26]. How these two elements are balanced depends on the jurisdiction, service territory, and existing asset profile. For instance, based on the ten utilities surveyed, the amount of peak demand-related costs could range from 30% to 75% of their average residential electricity bill.

Below are a few examples of utilities that are modifying rate structures to account for changes in technology and customer behavior:

Example: Time-of-Use Rate Pilot (Sacramento Municipal Utility District)

In 2011 and 2012 the Sacramento Municipal Utility District (SMUD) conducted a pilot study of how different treatments of load control and customer information affected residential energy consumption and peak demand reduction. While the study included the fielding of both customer- and utility-controlled thermostats and real-time feedback on energy usage, participants were also offered a time-of-use (TOU) rate with critical peak pricing (75¢/kWh) that could be called for up to 12 events per year. The rate otherwise had three different prices, ranging from 7.21¢/kWh during off-
peak periods to 27¢/kWh during weekdays from 4–7PM. The study found that participants who controlled their own usage decreased 32% of their load during normal weekday peak periods and 58% of their load during peak events—93% more reduction than those on a standard tiered rate with direct load control [27].

Example: Increased Fixed Charges (Multiple Utilities)

Customer bills typically include some “customer” charge, that is fixed monthly, to cover metering and billing costs. Recently, a number of utilities across the country have applied to extend their fixed charges or minimum bills to cover system-related costs as well. These requests are typically coupled with a corresponding reduction in the energy-based rate, reducing the sensitivity of the customer’s bill to changes in total energy consumption. However, these rate structures do not include direct incentives for customers to modify the magnitude or timing of their peak demand. For the residential rates that have been approved with these changes, fixed charges have been in a range of $16–$19 per month [28].

Example: Three-Part Rate (Salt River Project)

The board of directors at the Salt River Project (SRP) headquartered in Phoenix, Arizona, approved a new customer generation rate in February 2015 [29]. The new rate has three components: a fixed charge to cover metering and billing costs of $18–$20 per month, a reduced energy charge that roughly reflects the marginal cost of fuel and purchased power from SRP (anywhere from 3–6¢/kWh), and a demand charge that is intended to collect fixed costs of generation, transmission, and distribution capacity. The demand charge is tiered and based on the customer’s peak demand that is coincident with the system peak hours during that time of year. For a distributed generation customer with a peak demand of 8.5kW, the demand charge would range from approximately $41 in the winter, to $126 per month at the peak of the summer [30]. According to SRP, this new rate is intended to adapt to non-solar distributed generation technologies, as well as spur new technologies such as customersited storage, energy efficiency, and demand management [31].

Example: Demand Subscription (Spain)

In Spain, customers subscribe to a certain peak demand level with their retail provider and distribution company. The customer’s monthly “subscription cost” is proportional to the requested peak demand allotment. The customer demand isn’t allowed to exceed this allotment, regardless of the monthly energy use. The subscription level is enforced by use of an Interruptor de Control de Potencia, or ICP switch, installed in a residence’s main breaker panel [32, 33]. The ICP measures the peak demand and compares it to the amount contracted from the supplier. If the contracted amount is exceeded, service is interrupted until the customer can manually reset it. Spanish utilities Iberdrola and Unión Fenosa Distribución will incorporate the ICP into their smart meter deployments, set to be completed by 2018. This will enable the utility to monitor peak demand and to update remotely the contracted power amount [34].
SECTION 5: Research and Development Needed to Enable Capacity from DER

With greater deployment of distributed resources, especially variable generation, the need for flexible capacity could continue to grow and in the future could represent a much higher portion of overall system costs. Numerous pilots and studies [35–37] have investigated enabling distributed resources\(^{23}\) to provide both flexible capacity and ancillary services.

Market rules are evolving to value resources that provide the system with greater flexibility [38–40]. When properly integrated, distributed resources could also supply some of the capacity and ancillary service needed for load serving entities. Utility models have emerged where the distribution system includes an enabling communication platform for integrating distributed resources and becomes a market operator for distributed resources on the distribution side [41].

Reforming the Energy Vision in New York

Authorities in New York are embarking on a thorough evaluation and transformation of the utility concept as part of their “Reforming the Energy Vision” (REV) initiative. The stated goal of the effort is to “increase in the efficiency, responsiveness, and resilience of the system, with reductions in costs and carbon emissions, and increases in customer value.” In support of the initiative, it is being proposed that utilities will act as Distribution System Platform (DSP) providers, and will be tasked with ensuring reliability and enabling active third party and customer engagement.

The responsibilities of the DSP fall into three general categories:

- Integrated System Planning (including grid modernization)
- Distribution Operations
- Market Operations

System planning and distribution operations are intended to support traditional utility mandates of providing reliable, affordable electric service. However, the mandate to utilize distributed resources for system reliability will “increase the complexity and importance of distribution planning and operation.” However, the DSP in support of distribution and market operations “will also provide or sell a set of products and services to customers and service providers.”

According to the New York Public Service Commission’s report, detailed work remains to create uniform technology standards and market rules. In response, two working groups have been created, one on Market Design and the other on Platform Technology, to identify the necessary functional and business architecture for the DSP [41].

---

\(^{23}\) For purposes of aggregation, distributed resources might include active energy storage (chemical and thermal), passive thermal energy storage (buildings/water heaters), distributed generation, electric vehicles, or demand response.
While the opportunity for using DER as a grid resource has been clearly demonstrated, the scale of deployment for significant contribution at bulk power level will require work in several areas outlined below. EPRI is working with the U.S. Department of Energy (DOE) and its national labs, as well as universities and manufacturers to address technology gaps through research, development, and demonstration (RD&D). The goal of these efforts is to enable the aggregation of DER as a safe, reliable, and affordable source of capacity and ancillary services. Below are a few examples of current research areas:

**An Open Language for DER**

For many DER, the interface between the resource and the grid is an inverter (an electronic device that converts dc to ac power). Particularly for PV and battery storage any advanced functionality required for grid support (including providing firm capacity or flexibility) is generally governed by the inverter’s design and programming. While most inverters have advanced functionality, including communications and control, it has proven a challenge to use these functions in a utility setting. Due in part to proprietary existing protocols, a significant gap has been noted for providing robust and more secure methods for communicating among and potentially aggregating DER simultaneously. EPRI worked with inverter manufacturers worldwide to define a communication specification for inverter-utility interfaces.

The specification includes:

- **Common Functions**: Inverters may provide several services that support an integrated grid. Efficient operation requires a standard definition for each of these services, satisfying both utility and owner/operator needs for the benefit of the public. For example, all inverters may be capable of managing VARs. In order for DER to be considered an effective VAR resource, however, there must be consistent methods by which VAR production is requested, implemented, and verified from diverse types, brands, and sizes of DER [42].

- **Object Models**: Independent of the architecture and the protocol, object models are required that govern the types and formats of the data to be exchanged. Some data is mandatory, while other information is considered optional or conditional on inverter types, vendors, or implementation decisions. These object models allow self-description and common understanding of the data being exchanged.

- **Protocols**: Specific communication protocols at the network/transport and application layers allow interoperability of inverters in many different environments. Existing protocols are identified, the applicable sections highlighted, and options selected to remove ambiguity.

EPRI’s work has led to different standards across the spectrum of DER as shown in Figure 17.

**Market rules are evolving to value resources that provide the system with greater flexibility.**
Verifying Interoperability of DER

If multiple DER claim to use the same language, the next step is to verify that these devices are indeed “interoperable,” having the ability to exchange and interpret received information accurately. EPRI, in collaboration with NREL and others, is developing a framework for testing the interoperability of distributed resources. Though focused on individual grid support functions, the framework considers goals at a higher level at which devices are expected to work in concert. In order to accomplish these goals, a consistent interpretation of utility requests must be established across multiple communication interfaces and architectures (for example, CEA-2045, SEP2, DNP3, and transactive energy).

This is part of a broader effort called “Integrated Network Testbed for Energy Grid Research and Technology Experimentation” (INTEGRATE) initiated by NREL and coordinated with different DOE departments related to Energy Efficiency and Renewable Energy (EERE) technologies.

Figure 17: A number of open protocols are being developed for accommodating DER in the power system infrastructure
Enabling Advanced Distribution Management and Communication

A key enabler for integrating distributed resources will be new functions that define how DER can be monitored and managed in aggregate, and coordinated with future Distribution Management Systems (DMS). These include modules such as DERMS as shown in the following figure. DERMS modules handle the direct management of individual DER plants, potentially thousands on a single feeder, and present the capabilities of those DER as a simpler set of services that can be utilized by DMS.

In the past, communication with various distribution system devices, including automation, regulating equipment, workforce management, and metering, has involved separate communications systems. Each system has its own commands and protocols, and occupies its own portion of the RF spectrum. New systems are deployed as-needed, increasing complexity, inefficiency, and duplication of functions, ultimately leading to higher costs. The Field Area Network (FAN) concept attempts to bring different distribution equipment technologies onto a singular wireless platform that is ubiquitous, more secure, and reliable, serving a broad range of smart grid applications, beyond just integrating DER. A flexible, multipurpose communication infrastructure at the grid edge could enable more rich and reliable communication with distributed generation and end-use devices at an added cost lower than that of deploying independent infrastructure. EPRI is working with a number of utilities, including Duke Energy and SRP, to develop FANs in their respective service territories [43].

While the opportunity for DER to contribute to future capacity and flexibility needs is real, significant research by all stakeholders is necessary to address key issues to scale this integration more reliably and securely.

Figure 18: Levels of enterprise integration being investigated by EPRI
SECTION 6: Conclusion

Sufficient capacity with the flexibility to provide a variety of services is required for reliable delivery of electric energy that is sufficient to satisfy demand at all points in the power system. Given the real-time nature of electricity, capacity must be both available and flexible. While such requirements have existed over the history of the power system, new developments in both distributed generation and customer loads create challenges and opportunities. However, the proportion of costs between capacity and energy are also changing, requiring a careful look at how costs are to be allocated and recovered at wholesale and retail levels. While possible to create reliable capacity with distributed resources, significant research and standardization through open protocols are necessary for cost-effective integration preserving system reliability.
References


The Electric Power Research Institute, Inc. (EPRI, www.epri.com) conducts research and development relating to the generation, delivery and use of electricity for the benefit of the public. An independent, nonprofit organization, EPRI brings together its scientists and engineers as well as experts from academia and industry to help address challenges in electricity, including reliability, efficiency, affordability, health, safety and the environment. EPRI also provides technology, policy and economic analyses to drive long-range research and development planning, and supports research in emerging technologies. EPRI’s members represent approximately 90 percent of the electricity generated and delivered in the United States, and international participation extends to more than 30 countries. EPRI’s principal offices and laboratories are located in Palo Alto, Calif.; Charlotte, N.C.; Knoxville, Tenn.; and Lenox, Mass.

Together… Shaping the Future of Electricity