POWER TO THE PEOPLE

How will the integration of demand response initiatives affect the US power market? Ron McNamara, principal consultant at energy consultancy Utilicast, and former managing director at energy trader Saracen Energy Partners, discusses how to integrate end-users into the electricity market

Over the past decade, a new policy paradigm has emerged in the US power sector that challenges the *status quo* over how electricity is produced, consumed and regulated. The seeds of this challenge were planted 20 to 30 years ago when technological advances reduced the economies of scale enjoyed by large-scale electricity generation.

As the market and policy makers look to develop new routes to efficiency, six major areas of technological innovation have emerged:

- Demand response;
- Renewable energy;
- Storage;
- Microgrids;
- The smart grid; and
- Small-scale distributed generation.

There is little doubt that 20 to 30 years from now, we will look at these six technologies as having ushered in a new electricity era. If the proponents of each are correct, large-scale adoption of one or more of these technologies could have profound effects on the operation and hence the outcomes of bilateral and organised electricity markets in the US. Of the six, the one that represents the most fundamental threat to the existing paradigm, at least initially, is demand response, this article's focus.

Ron McNamara, Utilicast

Demand response

Background

In real-time there is no 'demand' for electricity – rather there is 'load'. The distinction between the two is important when it comes to increasing the amount of demand response (which is defined below). Load is simply physical consumption and has nothing to do with price, whereas demand indicates the amount of power that consumers are willing and able to purchase at different prices. So, in real-time, the so-called demand for electricity is assumed to be unresponsive to price, that is, it is completely inelastic. The job of the system operator is to use the transmission system to make sure that generation meets this (fixed) load at the lowest possible cost.

The consequence is that, other than in emergency situations, a reduction in consumption has not traditionally played a significant role in balancing supply and demand, and as a result, the electricity system is built, and operated, to meet load regardless of how expensive that becomes. So-called peaking units may operate for only a few hours per year when load spikes to very high levels. In an era of low economic growth and concern for the environment, government and industry leaders are asking whether it is more efficient for demand to be rationed than it is to support load regardless of the cost.

Defining demand response

Demand response refers to a system that allows the demand 'curve' to

> *There is little doubt that 20 to 30 years from now, we will look at these six technologies as having ushered in a new electricity era*

> *Things happen so fast in real-time on an electricity grid that prices cannot guide the actions of the participants quickly enough to maintain the stability of the system*

become more elastic – that is, making the demand for electricity more responsive to price. This includes shifting electricity usage from one period to another in response to price signals or simply reducing the amount demanded at any given price. In effect, the desire to increase and improve demand response is a direct attempt to allow demand to respond to real-time conditions – that is, to relax the notion that demand in real-time is inelastic.

A fast-growing aspect of electricity markets, demand response has become of increasing interest to utilities striving to include end-users in initiatives aimed at managing energy costs and boosting reliability in this energy-intensive era. Incentive-based demand response programmes and time-based retail rates combined could have supplied an estimated 6% of national peak demand in 2008, according to the US Federal Energy Regulatory Commission (FERC). In a June 2009 report, FERC estimated that as much as 20% of national peak demand could be met by demand response by 2019, roughly equivalent to the output of 2,000 average-sized peaking power plants.

The co-ordination problem

If demand response is introduced to that extent, the potential effects on the current state of the market will obviously be important. Whether or not demand response is truly disruptive to the status quo of the electricity markets depends in large part on the way in which it is integrated and compensated.

Understanding why this is so requires an understanding of some of the unique physical characteristics of electricity and how they have been dealt with in the past.

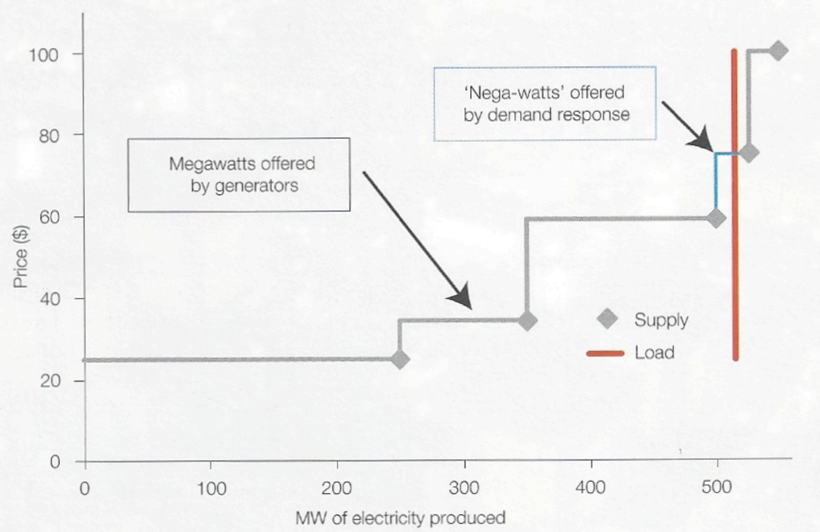
First, unless electricity is consumed at the point where it is produced, it will be exported to the transmission system and will travel to where it will be ultimately consumed

©iStockphoto.com/David Elfstrom

Energy end-users – Demand response

F1. Aggregate electricity supply and demand curves

Source: Ron McNamara PhD



– that is, electricity producers and consumers are interconnected through a transmission network. Second, electricity will follow the path of least resistance between two points. Third, electricity has to be produced and consumed simultaneously. These three basic characteristics imply that the actions of producers and consumers must be co-ordinated at virtually every instant in time within specific tolerances. Failure to do so will most likely cause the electricity system to 'trip' or stop working.

This 'co-ordination problem' is not the product of specific contracts, regulatory requirements, political boundaries or accounting principles. Nor is it something new. As long as generators and consumers have and continue to be connected to a transmission system, the co-ordination problem will remain an issue. Moreover, the extremely short time frame involved in co-ordinating the system, coupled with the fact that production and consumption occur almost simultaneously, means there are limits on how effective the price mechanism (the market) can be in co-ordinating supply and demand. In other words, things happen so fast in real-time on an electricity grid – demand changes, generation and transmission facilities become unavailable – that prices cannot guide the actions of the participants quickly enough to maintain the stability of the system.

For this reason, all electricity markets are a hybrid. Prior to 'real-time', the forward markets rely on prices and the market mechanism to co-ordinate the decisions of buyers and sellers. Once inside the time frame when electricity is actually being physically produced and consumed – usually five or 15 minutes – the system operator has the authority to take the actions necessary to reliably match supply and demand.

Integrating the demand side

As such, there is no question that enhanced demand response will lead to more efficient outcomes for the power market. The question is not, therefore, whether we should try to increase the level of demand-side participation, but rather how best to integrate the demand side of the market. To date, the primary and indeed the only real way that has been tried is to allow or encourage the demand side to 'offer' into the market as a negative generator. Whereas a generator offer is comprised of dollars per megawatt (MW), a demand response offer consists of dollars per 'nega-watt' (that is, megawatt not consumed) – on the basis that a megawatt produced to meet load is identical to a megawatt not consumed. Thus, as depicted in figure 1, the real-time electricity supply curve is developed from offers to produce energy (from generators) and offers to not take energy (from consumers).

The supply curve in figure 1 is made up of four sets of generator offers:
(1) 250MW at $25/MW,
(2) 100MW at $35/MW,
(3) 150MW at $60/MW and
(4) 25MW at $100/MW, and
one demand response offer:
• 25MW at $75/MW.

Also included on figure 1 is the expected load of 515MW for the period. The market price in this example will be $75/MW and the demand response offer will be the marginal 'generator'. The final physical solution will consist of 500MW of actual generation and 15MW of foregone consumption. Total payments from consumers to generators and demand response providers would be:

$38,625 ($75 × 515)

If there had been no demand response, and load had stayed at 515MW, then the $100/MW generator would have been used and

> *Once inside the time frame when electricity is actually being physically produced and consumed, the system operator has the authority to take the actions necessary to reliably match supply and demand*

the total payments would have been:

$$\$51,500 (\$100/MW \times 515MW).^1$$

From this example it appears that the inclusion of demand response resulted in a relatively simple and more cost-efficient solution.

Factors to put in the mix

However, there are a few important issues surrounding this solution that are relevant to how effective demand response can be in mitigating high prices. How each of these issues are dealt with in terms of the market design and operation will determine the effectiveness of demand response.

Ownership

In the example, we assumed the provider of demand response actually owned the megawatts that they decided not to use – the nega-watts – when the price reached $75. But as previously discussed, electricity is a 'real-time' commodity, so physical production and consumption must occur simultaneously. Hence, there is no way that a demand response provider can 'own' the physical megawatts they intend to consume prior to when they are actually consumed. What they can own, through contracts or regulation, is the option to take the megawatts, if they are available, at a specific price.

While mechanisms have been developed to overcome this problem, they are by definition subjective. The most typical is to establish a baseline level of consumption based on historical usage and then treat demand response as a reduction from the baseline level.

Load forecast

The example also assumed the system operator would continue to use a load forecast of 505MW, when in fact there were up to 25MW of demand response offered into the market at $75. In the case where demand response has been offered into the market, should the load forecast be reduced accordingly? If so, the demand response will never be used!

In the example above, if the system operator initially forecast load at 505MW and then saw that there was 5MW of demand response available, should they have reduced their load forecast accordingly? Moreover, what load forecast amount should they use to determine how much spare capacity they need to carry in order to cover an emergency?

> *There is no way that a demand response provider can 'own' the physical megawatts they intend to consume – what they can own is the option to take the megawatts, if they are available, at a specific price*

Identification of market power

When demand is treated as inelastic, a generator that has market power has the potential to raise prices far beyond what it cost to produce the electricity. All electricity markets have sophisticated detection mechanisms in place designed to identify and stop monopoly practices by generators. But in the presence of demand response, how is market power defined? If demand is capable of responding to price signals, that should serve to mitigate monopolistic pricing by generators. But is the threat of demand response all that is necessary or does there have to be active demand reductions?

Definition of reliable operation

Treating demand as inelastic is equivalent to saying that any reduction in the amount consumed is involuntary. Defining reliable system operation in that environment is relatively simple – it becomes much harder when significant levels of demand are willing and able to respond to prices and implement voluntary shutdown.

Portfolio offers

In most cases, demand response by a single consumer is unlikely to meaningfully affect either price or quantity. What has been occurring, and will continue to do so, is that demand response providers serve as aggregators of small electricity consumers and create a more substantial portfolio of nega-watts. As the level of demand response increases, electricity market design and operation will have to reflect the need for aggregators to offer and operate in a portfolio environment.

Moving forward

In essence, demand response is the next logical progression in the electricity industry reform process. Technological improvements in the latter part of the last century mean large-scale distant generation plants no longer enjoy a significant cost advantage over local small-scale gas turbines. As a result, there is no longer any economic justification for linking transmission and generation within a monopoly structure. A competitive generation sector requires non-discriminatory access to the transmission system and specifically to the coordination service in order to allow competition to take place.

One of the positive consequences of this unbundling of the coordination service from the transmission service is the creation of transparent wholesale electricity prices. Those prices have now driven advances on the demand side (for example, smart meters and remote sensing) that will allow for greater participation of end users in managing their energy costs. As pointed out above, there are obstacles to overcome, but true demand-side participation will fundamentally change the provision of electricity and forever alter the way in which the industry operates. ■

1. There is currently debate in the industry as to exactly what price a nega-watt should receive, e.g. the locational marginal price (LMP) or the LMP less the retail rate paid. This example abstracts from that debate and assumes a nega-watt is paid the locational marginal price.