CONSIDERATIONS IN ESTABLISHING PERFORMANCE STANDARDS FOR LOADS ACTING AS RESOURCES IN ERCOT'S MARKET

Jay Zarnikau
Frontier Associates LLC[1]
January 26, 2001

ERCOT has made progress toward designing future ancillary services markets that will permit demand-side resources to contribute to system reliability and price stability. However, the proposed Protocols require further refinement to define appropriate performance standards for loads that wish to provide demand-side resources into ERCOT's ancillary services markets. It has been suggested that large industrial energy consumers should "bid conservatively" when offering to provide ancillary services. This paper identifies some of the factors that should be considered in designing standards to better define this concept and address some related ambiguities in the proposed Protocols. It is hoped that this document might serve as the basis for designing standards that balance the needs of the ISO with the needs of industrial energy consumers that wish to participate as loads acting as resources.

Traditional Role of Interruptible Service in ERCOT

The ERCOT utilities presently rely on around 3,000 MW of interruptible load to preserve system reliability during periods of generation supply shortages or transmission constraints. Some loads may be interrupted for economic reasons, as well.

Over one-third (i.e., 1,200 to 1,300 MW) of ERCOT's interruptible load is presently instantaneous interruptible load. Industrial firms that take instantaneous interruptible service have installed under-frequency relays, which interrupt service to

[1] Frontier Associates LLC has participated in and monitored the development of ERCOT's Protocols on behalf of a number of clients with an interest in demand-side resources (e.g., curtailment programs, interruptible service, small-scale renewable energy resources, and energy efficiency programs). This analysis was requested by Nucor Steel.

instantaneous interruptible energy consumers when system frequency dips below 59.7 Hz.

At least four types of Texas industries are particularly good candidates for instantaneous interruptible programs: steel mills, chlorine manufacturing plants, oil and gas production facilities, and air separations plants. These types of facilities can be interrupted with no advance notice with tolerable economic and safety consequences.

As a responsive reserve resource, loads hold some advantages over generation resources. Loads with under-frequency relays can respond to frequency dips faster than power plants. Fuel savings and emissions reductions can be achieved when loads acting as resources are used to displace generation spinning reserves. Finally, participation by loads is necessary to enhance the competitiveness of these markets.

Traditionally, instantaneous interruptible loads using under-frequency relay equipment have been used by the utilities in Texas to provide an offset to spinning or responsive reserve requirements. ERCOT's total responsive reserve requirement has been a minimum of 2,300 MW, and ERCOT has traditionally limited the portion of responsive reserves that can be "counted" from interruptible customers to 25% of the system total, or 575 MW. ERCOT has stated that this limit is presently being re-evaluated.

Protocol Requirements

Under the proposed ERCOT Protocols, large industrial interruptible loads will be permitted to participate in tomorrow's markets for balancing services, responsive reserves (if under-frequency relays are present), non-responsive reserves, and replacement capacity. Yet, industrial customers[2] face some uncertainties in meeting some of the performance standards for providing ancillary services in the markets that

[2] Virtually all loads fluctuate, which generates part of the need for ancillary services (generation fluctuation also creates part of the need for ancillary services). Some fluctuate to a greater degree than others. For example, residential and commercial loads fluctuate primarily due to heating and cooling requirements; industrial loads fluctuate based how the particular processes that they utilize actually use electricity as well as their production and operation schedules.

ERCOT is establishing, at least as a "load acting as a resource." This problem could be particularly acute for fluctuating loads that wish to provide a responsive reserve resource.

The system proposed by ERCOT requires that the QSE's resources (including the load level of a customer that wishes to provide an ancillary service) be accurately forecast one day in advance for each of the ninety-six 15-minute intervals of the following day. This enables the QSE to accurately plan the quantity of ancillary services available to be bid into an ancillary services market or self-arranged to meet the QSE's needs. Unfortunately, it may be impossible to forecast the load levels of most loads (and thus the exact amount of ancillary services they can provide) with the accuracy expected by ERCOT.

During the technical conferences on January 5, 2001 and January 19, 2001, it was suggested that fluctuating loads could participate if they provided "conservative estimates" of their possible demand reduction ancillary service to their QSE and the market. A customer who expected its load level to be somewhere between 50 and 100 MW during a given 15-minute period during the following day and could drop its load to zero in response to an instruction from the ISO would presumably conservatively bid 50 MW of demand reduction under this suggested approach. Indeed, if the load bid 100 MW of an ancillary service, but was actually at a load level of 50 MW during an interval, the QSE would be below its ancillary service requirement by 50 MW (unless another resource was available to the QSE to make up the difference). A problem with this conservative strategy is that it could result in unreasonably low compensation to loads for the value that they provide. This could discourage loads from participating in ancillary services markets.

Initially, some additional problems were identified with this strategy, as well. If the customer's bid of 50 MW was accepted, but the customer actually dropped 100 MW, then it could be guilty of an "uninstructed deviation." After some further discussion, it was concluded that a load that provided "too much" of an ancillary service would not be assessed an uninstructed deviation after all.[3] Indeed, over-supplying a demand-side

[3] This was discussed during the afternoon of the January 19, 2001 workshop.

resource would normally benefit the system during a period of generation shortage. Nonetheless, the proposed Protocols seemingly require some clarification in this area.

The problem of providing too much of a resource is particularly a problem in the responsive reserve market. If an under-frequency event occurs, the total amount of load on the circuit with the under-frequency relay will be curtailed. So if 50 MW is bid by a customer into the responsive reserve market or self-arranged by a QSE and an under-frequency event occurs, the total amount of the customer's load at the time (perhaps, 100 MW) will be provided. Because of the technology, it is impossible for the customer to limit the exact amount provided.

It has been suggested that if loads (particularly fluctuating loads) cannot be precisely forecast, then they have no value to the ancillary services market and should not be allowed to participate. Yet, loads have traditionally played an important role in the provision of responsive reserves in ERCOT. The ability of loads to drop off the system in a much faster manner than generation has preserved the reliability of the ERCOT system on many occasions in the past. It is particularly beneficial to drop fluctuating loads during times of impaired reliability, because their absence will make it easier for ERCOT's operators to return system frequency to desired levels. For many years, the ERCOT system has relied upon interruption of these loads to maintain system reliability – even though forecasts of their exact load levels have never been required by ERCOT in the past.

It has also been suggested that customers with load levels that cannot be accurately forecast should not be "loads acting as resources," but instead should consider participating in the market as BULs (balancing up loads). For some loads, this is an excellent suggestion. However, converting instantaneous interruptible loads to BULs would clearly be a wasted opportunity. Today's instantaneous interruptible loads have a successful history of rapidly responding to reliability problems. As BULs, they would actually be required to respond to reliability problems in a slower manner. Normally, this slower response would provide much less value to ERCOT. The under-frequency relays that permit their quick response would need to be disconnected if they were to become BULs, and ERCOT could lose much of the instantaneous interruptible load that the

system presently depends on. Further, the economics of the BUL market are quite uncertain. BULs only receive a payment from the market during times of system emergencies or when there is a shortage of other ancillary services. Resources providing responsive reserves receive a payment during all 15-minute intervals when the resource has been included in QSE's self-arranged plan or a bid has been accepted. As a result, such loads may not be able to economically justify participation in the BUL market, and the system would lose the benefit of these interruptible loads.

Suggested Principles

Because there remains uncertainty and confusion regarding how loads can participate as a resource provider in ancillary services markets, it would be desirable for ERCOT to establish some reasonable performance guidelines for loads acting as resources. Some suggested principles to guide the development of such standards include the following:

- Loads should have a full and fair opportunity to provide ancillary services.
- If uncertainties regarding the exact load level of a load acting as a resource affects its value to the ISO, then this could be reflected in the price paid for the resource.
- The change to the new market structure for ancillary services should not adversely impact system reliability. If changes in rules necessitate the widespread removal of under-frequency relays at industrial sites because these loads cannot comply with new market rules, then the likely impacts on system reliability of such changes must be first studied.
- There must be a *reasonable probability* that the load will provide at least as much of the ancillary service that it is scheduled to provide or that it has bid to provide.
- The standards should establish a "safe harbor;" i.e., the industrial customer should be able to have some assurance that if it complies with the standards, then it will not incur penalties for poor performance or uninstructed deviations, but rather will be fairly compensated for its bid.

Ambiguities Presently in the Protocols

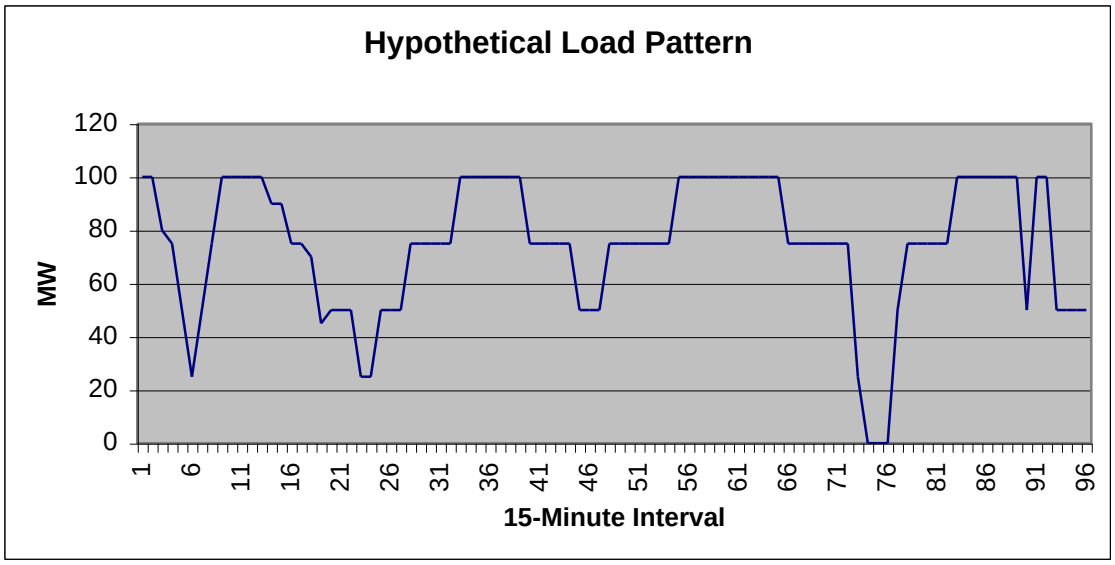
The proposed Protocols presently contain a number of ambiguities that need to be addressed so that these principles can be pursued. In particular:

- The penalties for unsatisfactory performance or uninstructed deviations are not well defined. It appears as though such penalties will be determined solely by ERCOT and will be "extremely high," based on comments made at the January 5th workshop. The absence of better information about how these penalties will be calculated and applied makes it difficult for a load to compare the benefits of providing an ancillary service against the potential penalty that might result if the load level is unexpectedly lower than anticipated.
- There remains some confusion regarding how the capacity provided by a load will be calculated. If the load level fluctuates within a 15-minute interval, will the average load level within the interval determine the amount of the capacity resource provided? Or could the load acting as a resource be out of compliance at any time the load level momentarily dips below the bid or self-arranged amount of capacity? The discussion of this issue at the January 5th workshop suggests that this is not yet clear.
- At the January 19th workshop, ERCOT and some key stakeholders indicated that a load acting as a resource would not be penalized if it inadvertently provided too much of a resource. Yet, particularly for smaller QSEs where a load might constitute a significant share of its resource portfolio, there remains a risk that the QSE could violate the Section 6.10.5.4 performance standards if a load over-provided a capacity resource in an ancillary services market.[4]

Proposed Standards for Bidding Fluctuating Loads

To demonstrate how performance standards might be applied, some data for a hypothetical fluctuating load (a rather extreme case) are presented below. It is assumed that the customer does not know its exact load pattern for the following day, but the load pattern for a typical day is provided on the graph. Some alternative qualification and performance standards are discussed below. These approaches could be utilized to fashion a reasonable safe harbor for bidding by loads.

[4] This section of the proposed Protocols states that satisfactory performance requires that: "Not less than 95% nor more than 120% of the RRS requested, subject to the declared capabilities of the QSE, is provided within ten minutes of ERCOT's deployment Dispatch Instruction and maintained until recalled or expiration of the QSE's service obligation . . . ."



LOAD CHARACTERISTICS

Average Load (MW)	75
Minimum Load (MW)	0
Maximum Load (MW)	100
Contract Load (MW)	100
Load Factor	75%
Probability Load is at Minimum Level	3%
70% of Average Level (MW)	52.5
Probability Load is at least 70% of Average Level	74%
60% of Average Level (MW)	45
Probability Load is at least 60% of Average Level	93%
50% of Average Level (MW)	37.5
Probability Load is at least 50% of Average Level	93%
70% of Maximum Load Level (MW)	70
Probability Load is at least 70% of Maximum Level	74%
60% of Maximum Level (MW)	60
Probability Load is at least 60% of Maximum Level	74%
50% of Maximum Level (MW)	50
Probability Load is at least 50% of Maximum Level	92%

Bidding the Absolute Minimum

If the customer were required to bid[5] the minimum load level it experiences on a typical day, it would not be able to participate in the market. In this example, the customer's minimum load level is 0 MW. However, this minimum level is experienced only 3% of the time.

Requiring the load to bid its minimum possible value would seem unreasonable. Such a standard would be somewhat analogous to requiring a generation resource to have a zero forced outage rate in order to participate in the market.

Bidding an Expected Value or the Scheduled Amount

The customer's QSE would probably schedule 75 MW to meet the generation needs of the hypothetical customer in each of the ninety-six 15-minute intervals (the average or expected load value) if the customer's distribution of load levels is randomly distributed within the day. If the expected value or amount of generation scheduled by the QSE were used as the customer's bid into an ancillary services market, then 75 MW would be counted as the customer's resource. During an interruption, this is the expected value of the (unobservable) load level that would have occurred absent the interruption.

It is important to keep in mind that the generation scheduled to meet the "expected load" would be fully available to the system once the load is removed, whatever the level of the load prior to interruption. In fact, if the load is already at a reduced level, this reduced level frees up scheduled generation (as balancing energy) and would serve to reduce the likelihood of interruption.

This approach would fairly compensate the load if payments were based either on the bid amount (the expected value) or actual load levels in each of the intervals, but could involve some risk that the ISO will fail to achieve the targeted level or reserves. However, if numerous loads acting as resources are selected to provide the reserves

[5] The reference to loads as bidders is designed to simplify this discussion. Actually, QSEs will bid on behalf of loads acting as resources or will use loads acting as resources to self-arrange ancillary services needs. The QSE can also bid or self-arrange a block of resources that includes one or more loads acting as resources. But here it is assumed that a single load is being independently offered to simplify the discussion.

during the same interval, then load diversity should diminish this risk. If expected values are properly calculated or if generation is properly scheduled, and if sufficient loads participate to achieve the required load diversity, then this approach should result in the ISO achieving its targeted resource level.

This approach offers simplicity and a reasonable safe harbor for loads that wish to provide ancillary services. Using the expected value (the amount of generation scheduled) would eliminate any questions regarding the amount provided, but would require QSEs to separately schedule generation to meet the needs of loads acting as resources.

Probabilistic Approaches

Although less appealing from a simplicity standpoint, an alternative might be to use some sort of probability analysis to determine in advance the maximum bid a load could make within the safe harbor approach. Based on an analysis of the load's historical pattern, the probability that it will exceed certain levels can be estimated. The ISO could require the load to submit a bid that has a predetermined minimum probability of being met or exceeded. In effect, the ISO would evaluate the load pattern and determine a "safe harbor" amount based on an acceptable probability that can be maximum bid for each period.

In the example above, if the ISO required a 74% probability that the load level was met or exceeded, then the load could bid as much as 70 MW. If the ISO required a 92% probability that the bid provided at least as much resource was bid, a 50 MW bid might be submitted.

This approach could provide the ISO with a minimum acceptable level of confidence that its target levels of ancillary services will be met or exceeded. However, loads acting as resources could end up receiving payments based on bid amounts that are much less than the resource amounts that are actually provided. If a high probability value is adopted and multiple bids from loads acting as resources are accepted, then there

will also be a very high probability that too much of the ancillary service will be procured.

Choosing an appropriate probability standard or confidence level could be difficult. Power plant performance standards might include a simple minimum forced outage rate. But for loads acting as resources, other factors must be considered, including distances from targeted levels and load diversity if multiple loads acting as resources are selected to provide an ancillary service.

Conclusions

Loads acting as resources face uncertain qualification and performance standards under the proposed Protocols, as they are presently written. Uncertainty regarding the standards that loads acting as resources will be held to and the possible penalties for violating such standards will likely discourage participation by loads in ancillary services markets. Consequently, this is an issue that requires further consideration prior to the opening of the markets. A safe harbor must be created whereby the amount of resources that a load is providing can be reasonably determined and set.

Some considerations in the design of qualification and performance standards are suggested here. The risk to the ISO of a load (or a power plant, for that matter) under-providing a resource for a given time period (when the load level is below its expected value) must be balanced against the customer's need for some certainty in performance standards, the benefits of over-providing during other periods, and the likely diversity among loads acting as resources, which can net out much of any under-provision by a single load.