Real-time Ancillary Service Co-optimization and Multi-Interval SCED

Comments Submitted by Luminant Energy Company LLC

Luminant Energy Company LLC (Luminant) offers these comments in response to ERCOT’s concept paper outlining Real-time Ancillary Service Co-optimization (RTASCOOP) and Multi-Interval SCED (MIS). Luminant believes that, on a conceptual level, RTASCOOP has merit and should be carefully considered. RTASCOOP may improve liquidity for intra-day ancillary procurement and may lead to more precise scarcity pricing for both Energy and Ancillary Services if properly designed and implemented.

Luminant is not optimistic that the benefits of MIS will outweigh the costs. In theory, robust load participation could improve price formation if loads are able to accurately represent their value of lost load on an individual basis. However the, efficacy of MIS is dependent upon ERCOT’s ability to accurately forecast system conditions in advance. There are significant challenges to predicting future system conditions at a fidelity that is sufficient for system dispatch, which we believe is critical for MIS to be successful. Moreover, MIS could harm the existing energy-only marketplace by creating costly make-whole payments to keep deployed load resources whole for an anticipated system condition that fails to materialize.

Luminant has long been supportive of market changes that would allow loads to participate in SCED. In 2013, Luminant submitted detailed comments in PUC Project No. 41061 outlining an approach that Luminant believed (and continues to believe) would be cost-effective, with minimal changes to ERCOT’s systems. Luminant also submitted detailed comments to NPRR 555 that, if they had been fully considered, could have offered a low-cost solution to incorporate additional demand response in SCED by providing the same treatment and benefits to load resources as are provided to quick start generation resources in the current market, without creating any additional uplift to load. Luminant is willing to re-visit these or other proposals whose costs are more in line with the expected benefits from integrating more loads in SCED.

Both RTASCOOP and MIS as proposed by ERCOT will require significant systems and market changes, and are expensive (Between $25,000,000.00 and $42,500.000.00 according to ERCOT’s initial estimate provided to the PUC in Project No. 40000). Luminant believes the following high-level concerns and suggestions will need significant attention and consideration before the market proceeds down the path of implementing either or both of these market design changes:

***Real-time Ancillary Service Co-optimization***

* Implementation of an ORDC in the DAM
* Improved Utilization of the Transmission System
* Balance of the Day AS Co-optimization (on no less frequently than an hourly basis)

***Multi-Interval SCED***

* Net Load Forecasting / Ramp-capable Resource Capacity Forecasting Accuracy
* Equitable Risk Profiles for all Resources
* Make Whole Payments and Uplift to Load
* Price Correction for Block Load Deployment

Thereafter, Luminant offers preliminary responses to the questions posed by ERCOT in the concept paper. Luminant believes that the questions posed by ERCOT must be sufficiently vetted by the stakeholder community. These salient details may make-or-break the initiative, and it is imperative that stakeholders are given adequate time to consider both key policy issues and detailed implementation concerns.

***A. Real-Time Ancillary Service Co-optimization***

***1. Implementation of an ORDC in the Day-Ahead Market***

Luminant believes that implementation of an ORDC in the DAM would enable more efficient pricing of Ancillary Services and should be designed such that the DAM is allowed to solve regardless of AS offer sufficiency. While ancillary service insufficiency at the close of the DAM is an unlikely outcome, the ORDC mechanism in the DAM would ensure proper pricing from the co-optimization between AS & Energy. Energy prices would reflect the opportunity cost of ancillary services as co-optimization seeks to maximize producer revenue and fill the ancillary service short requirements. This day-ahead process would mimic what is proposed in ERCOT’s concept paper in real-time co-optimization.

Ancillary Service insufficiency, however, may still occur if there are insufficient qualified offers, and must be permissible in the DAM function. The likelihood of an AS offer insufficiency is low since allowing a demand curve to drive up prices as scarcity of a product ensues will help prioritize fulfillment of that product while preserving the existing willing buyer / willing seller for energy component of the day-ahead financial market. ERCOT may announce its need and execute a SASM closer to real-time, or if that proves to still be insufficient, may issue a reliability unit commitment for ancillaries in the RT market based on the severity of the insufficiency. This concept also addresses one of ERCOT’s concerns (#4 below) posed in the white paper.

***2. Improved utilization of the transmission system***

Real-time co-optimization allows for more efficient use of the transmission system by selecting resources providing ancillaries for energy as energy prices rise at the plant bus. This is an improvement from the status quo, as resources with ancillary service obligations currently have their capacity fulfilling ancillary service awards unavailable to SCED.

**3. *Balance of the Day AS Co-optimization at least hourly***

Luminant believes that real-time co-optimization of Energy and Ancillary Services should include an hourly re-optimization of AS Reserves for the balance of the day. This would encourage existing bilateral transactions to occur in structured, more efficient markets.

**Luminant comments to ERCOT posed questions for Ancillary Service Co-Optimization**:

1. What is the proper timeline for updating Ancillary Service Offers and Energy Offer Curves within Real-time AS Co-optimziation?

*For the 5 minute ORDC Co-optimization, offers should be allowed to be updated in real-time. Possibly, the offer price quantity pairs could be locked down ahead of the hour but quantities available to be selected can change in real-time via telemetry (similar to the way HSL and Energy Offer Curves are handled today). For the balance of the day hourly offers, a deadline could be enforced ahead of when the Optimization is scheduled.*

1. How should ERCOT “share” ramp rates between energy and ancillaries? (Similar to ramp sharing between SCED and LFC?)

*Luminant believes that the optimization engine should consider 100% of the ramp-rate available at the current unit output. The amount of regulation assigned to a unit from a co-optimization run should limit the amount of ramp available for energy deployment in SCED. All ramp capacity is available for the co-optimization but the awarded regulation ramp is effectively reserved based upon the constraints of the resource and its current operating point. Other AS reserves (RRS, Non-spin) would be deployed by SCED in future SCED runs, therefore sharing of ramp rates between current SCED energy needs and future AS needs is permissible.*

1. ERCOT states that a resource’s entire capacity (between LSL and HSL are available for allocation between energy and AS.) Is it possible for a QSE to express their willingness to offer AS as something less than HSL-to-LSL?

*The ability to offer all or part of a resource to provide ancillaries should be governed only by the QSE representing that the resource, or some portion of its output, is limited and cannot meet the protocol requirements for an Ancillary Service.*

1. How will ERCOT handle AS insufficiency in the DAM? Can AS clear insufficient and be rectified in real-time?

*As stated above in these comments, Luminant believes that AS insufficiency should be allowed in the DAM, assuming that an ORDC is used for appropriate pricing in the DAM. ERCOT has the ability to run a SASM nearer real-time, and if the AS Insufficiency is not resolved in that market, ERCOT has the ability to RUC for capacity.*

1. Is it practical or feasible to expect UFRs to arm and disarm for intervals in which they were not selected to provide RRS?

*Luminant is not aware of any technical limitations preventing resources controlled by an under-frequency relay from arming and dis-arming on a five minute basis via QSE initiated SCADA instructions. However, if QSE representing a Load Resource deployed by a UFR wishes to avoid the selection uncertainty associated with real-time ancillary service co-optimization, then that QSE could offer as a price taker with a $0 bid. (A practice widely embraced today in day-ahead RRS markets.)*

1. Coordination of the PBPC, ORDC, and VOLL.

*Luminant believes that, as suggested by ERCOT, the ORDC, PBPC and VOLL must be coordinated to maintain the desired outcome as currently designed and ensure a preference to fill energy demands over maintaining a reserve in times of scarcity. This may require action from the Public Utility Commission.*

1. Disaggregation of the ORDC:

*Luminant agrees with ERCOT’s application of the disaggregated ORDC to differing ancillary products.*

1. For Non-Spin, should the statistical distribution of Online Reserves be used or Online+Offline Reserves or some combination? Note that the price on the Non-Spin demand curve will be lower if the statistical distribution of the Online+Offline Reserves is used.

*Existing ORDC Methodology includes online and offline reserves. Luminant believes that when calculating the LOLP that is associated with Non-Spin, ERCOT should analyze historic events utilizing both online and offline reserves. When calculating the LOLP that is associated with RRS , only historic online reserves should be used. RRS can only be provided by 10-minute resources and demand curves for this product should only consider online resources.*

1. How should the demand curve for Reg-Down be determined? Is it the same as the Reg-Up demand curve?

*The demand curve for regulation should be statistically calculated based on the need for regulation rather than historical view of all reserves. This is true for both RegUp and RegDown and therefore the market needs to develop a methodology for calculating the risk for running out of those reserves.*

1. Do we need to model AS demand curves based on ORDC in the DAM?

*As previously mentioned in question 4 and in Luminant’s discussion of RTASCOOP, modeling AS demand curves improves AS pricing and improves convergence between markets.*

1. If we model AS demand curves in the DAM, as the DAM is voluntary, are there any issues? Resources do not have to participate since it is not mandatory to submit Resource specific Offers for energy and AS.

*AS insufficiencies in the DAM are a possibility today even without a demand curve. Adding AS demand curves allows DAM to solve but doesn’t increase the likelihood of insufficient offers.*

1. If AS demand curves are **not** modeled in the DAM, then AS procurement is effectively given higher priority than energy — that is, in DAM, all energy demand has a price and hence can be curtailed, whereas *the AS penalty is extremely high.* This ensures that the AS Plan MW Requirements are procured. However, in RT, AS demand curves are modeled based on ORDC. The amount of Non-Spin procured depends on the intersection of the Non-Spin demand curve and the offer stack for Non-Spin. Hence, the procured amounts of Non-Spin can exceed or be less than the AS Plan MW Requirement for Non-Spin. If there is more Non-Spin procurement than the AS Plan, then loads are charged this amount on a Load-Ratio-Share. This is the same as the current process of AS Imbalance Settlements

*As stated previously, Luminant suggests that AS demand curves are to be modeled in the DAM. The demand curve for non-spin may yield higher quantities procured during times of high-volume, low-priced quantity. This represents the willingness of load to accept a higher-than historic level of reliability, represented by the incremental reserves, if the incremental cost is low. Hedging a load ratio share of non-spin may change, as the quantity procured may increase beyond targets at certain times. However, Luminant is confident that products to hedge non-spin shorts will be developed should the appetite for this risk aversion develop.*

***B. Multi-Interval SCED***

***1. Net Load Forecasting / Ramp-Capable Resource Capacity Forecasting Accuracy***

As noted by ERCOT in its concept paper, the “accuracy of the short term load forecast is critical.” Equally important is the short term wind forecast. When summed together, the two forecasts provide a net-load forecast that is currently suitable for a 5-minute dispatch. ERCOT’s ability to forecast the amount of ramp-capable capacity is also critical. Capacity available to ramp is determined by units committing and de-committing, as well as the current online fleet ability to ramp based upon each individual unit’s expected output. Similar to forecasting net-load, forecasting ramp capability for 30 minutes in the future is problematic, as it is based upon uncertain expectations of resource availability and performance. If the combined estimate of ramp-capable capacity and net-load forecast is erroneous, then resources will likely be inappropriately dispatched. Recognizing that no estimate of net load and resource ramp capability will be perfectly accurate, Luminant suggests that the accuracy of the combined estimate must be sufficiently accurate to avoid inappropriately dispatching a block resource. For example, if the smallest load eligible to participate in Multi-Interval SCED were 50MW, then the net estimate of ramp capability and net load must be prone to less than 50MW of error. If smaller resources participate, (e.g., ~5MWs,) then the accuracy of the forecast must be better than 5MWs. Otherwise, make-whole payments may be necessary to compensate the dispatched resource based on erroneous predictions. Although ERCOT seems to have confidence in its ability to improve the net-load forecast, doing so will most certainly be an arduous task. This must be complemented with an initiative to predict ramp capability within an acceptable tolerance, as illustrated above, such that forward commitment of resources is appropriate within a 30-minute window.

Inaccuracy between the net-load forecast and the ramp-capable capacity forecast could: (1) dispatch resources out-of-merit, (2) suppress real-time prices (absent a 3rd and 4th SCED run price correction), and (3) create uplift to load. Luminant suggests that the accuracy of the difference between the net-load forecast and the ramp-capable capacity forecast must be demonstrated before MIS with make-whole payments to dispatched blocky resources is considered. ERCOT and the stakeholder community must determine the following:

1. A threshold for the amount of uplift that is acceptable to load. Said another way, determine when the benefit of MIS exceeds the amount of uplift. This may be dependent upon the achievable forecasting accuracy.
2. A method for estimating the amount of uplift given the accuracy of the net-load forecasts and the ramp-capable capacity forecasts. ERCOT and stakeholders must anticipate, with reasonable certainty, the amount of uplift given the accuracy of these forecasts.

There are alternative design considerations for Multi-Interval SCED that could be adopted to eliminate much of the risk of uplift to the market from ERCOT’s commitment decisions. For the proposed Multi-Interval SCED solution, Luminant considered the following:

***2. Equitable risk between Load and traditional Resources***

The current draft paper protects Load Resources for commitments made for anticipated prices that fail to materialize. This creates an inequity between traditional resources, who bear the risk of price uncertainty when making commitment decisions and price their offers to reflect that risk.

Luminant believes that ERCOT should continue with its current Energy-Only Market design where qualified scheduling entities offer resources while bearing the risk of not recovering the cost of that commitment. This risk creates an incentive for QSEs to imbed the cost of those risks in their offers/bids. This would eliminate costly make-whole payments and allow traditional resources and loads to face equitable price risk. Luminant believes this topic demands an in-depth ERCOT and Stakeholder discussion.

***3. Make-whole payments and Uplift to Load***

The ERCOT draft paper suggests resources may be committed due to the difference between the net-load forecast and the ramp-capable capacity forecast. If the committed capacity is not needed from that resource and it does not contribute to price formation in Real-time by SCED, then it may be eligible for a make-whole payment. This make-whole payment covers the costs associated with the deployment of the dispatched resource and guarantees the revenue based on the resource offer. The make whole payment is only for **that** resource deployed by MIS. ERCOT’s paper specifies X intervals as eligible for make-whole payments, before the deployed resource will be settled exclusively on the real-time LMP. The paper details X as something between 3 and 6 five minute intervals. Luminant believes this decision is somewhat arbitrary, as resources eligible for a make-whole payment must consider this risk when formulating their offer. For example, if a resource were not eligible for a make-whole payment, such as is the case for all generation resources in the market making real-time commitment decisions, that resource must carefully consider its opportunity costs and the risks of unnecessary commitments from ERCOT associated with anticipated deployments and submit offers accordingly. Longer duration forward commitments, with a make-whole guarantee, reduce risk to the individual resource, which will distort that individual resource’s offer relative to other offers in the market and increase uplift to load. Luminant suggests that market participants intending to offer blocky resources consider their exposure and offer accordingly, rather than receive a make-whole payment. Market participants will then be responsible for accurately determining their willingness to perform based upon price uncertainty and contribute to efficient price formation. Including these costs in the LMP rather than a make-whole ensures proper price signals and promotes proper incentives.

***4. Price Correction for Block-Load Deployments***

Luminant believes that ERCOT’s proposed price corrections for block load deployments is critical (also known as a NPRR626 style price correction, or 3rd and 4th SCED runs) is critical. A load resource deployment, unlike the deployment of a generation resource, appears to SCED as a reduction in load. Correspondingly, the marginal load resource deployed is no longer on the margin, as it appears as a reduction in load in the Generation to be Dispatched (GTBD) calculation. NPRR 626 corrects for this condition, and should be applied to deployed blocky load resources. Alternatively, deployed load could be added back to GTDB to achieve the same ends.