Scarcity Pricing using ORDC for reserves and Pricing Run for Out-Of-Market Actions

David Maggio, Sai Moorty, Pamela Shaw

ERCOT Public

Agenda

1. History of the Operating Reserve Demand Curves (ORDC) at ERCOT
2. Determination of the ORDC
3. Calculation of Reserve Price Adders
4. Real-Time Pricing Methodology
5. Reliability Deployment Price Adders
6. Upcoming Changes for ORDC
7. Potential Future Changes: Real-Time Energy and Ancillary Service (AS) co-optimization using ORDC as basis for AS Demand Curves

History of Operating Reserve Demand Curve (ORDC) at ERCOT

- The Public Utility Commission of Texas (PUCT) opened Project 40000 to explore options to address Resource Adequacy and Scarcity Pricing Concerns at the end of 2011.

- In November, 2012, GDF Suez filed Dr. Hogan's paper describing the implementation of an Operating Reserve Demand Curve (ORDC) to potentially address Resource Adequacy concerns through appropriate scarcity pricing.
  - "Electricity Scarcity Pricing Through Operating Reserves: An ERCOT Window of Opportunity"
  - http://interchange.puc.state.tx.us/WebApp/Interchange/Documents/40000_355_741583.PDF

History of ORDC at ERCOT

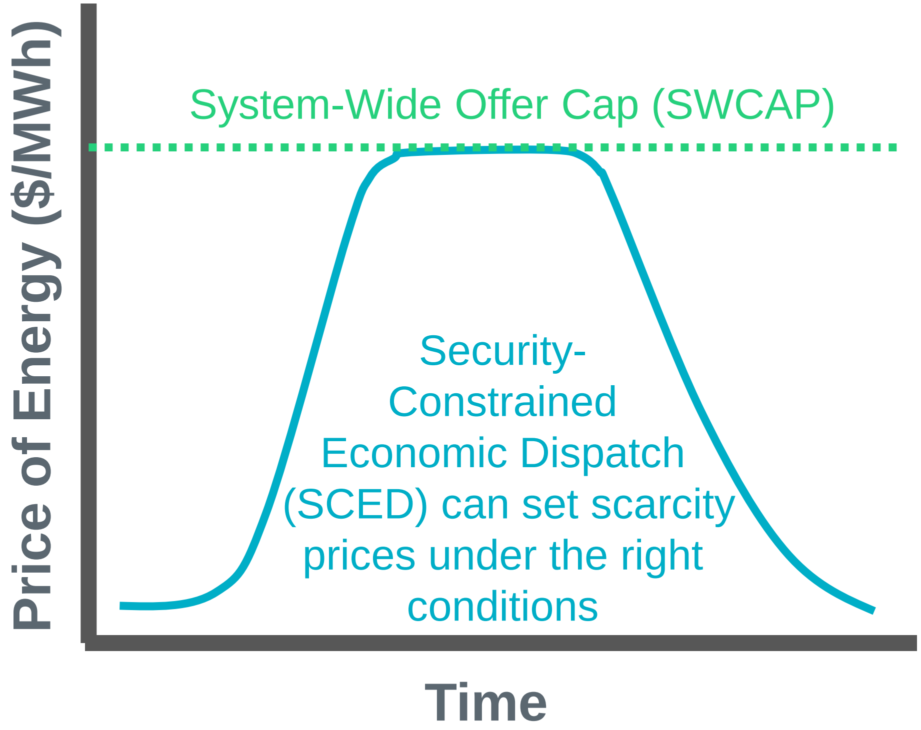
- Key points in Dr. Hogan's paper:
  - Improve scarcity pricing by reflecting the marginal value of available reserves in Real-Time energy and reserve prices.
  - The marginal value of available reserves is the product of the Value of Lost Load (VOLL) and the Loss Of Load Probability (LOLP) corresponding to this level of available reserves.
  - Suggested that ERCOT implement ORDC based scarcity pricing mechanism and evaluate its impact before considering other more radical changes.

History of ORDC at ERCOT

ERCOT is an energy-only market

Energy pricing must support an appropriate level of investment in Resources

- When operating reserves are low, the probability of having a scarcity event increases.
- If the scarcity event is severe, ERCOT may need to shed load.
- System prices should reflect the **marginal value** of available reserves, since reserves reduce the chance that ERCOT will have to shed load.



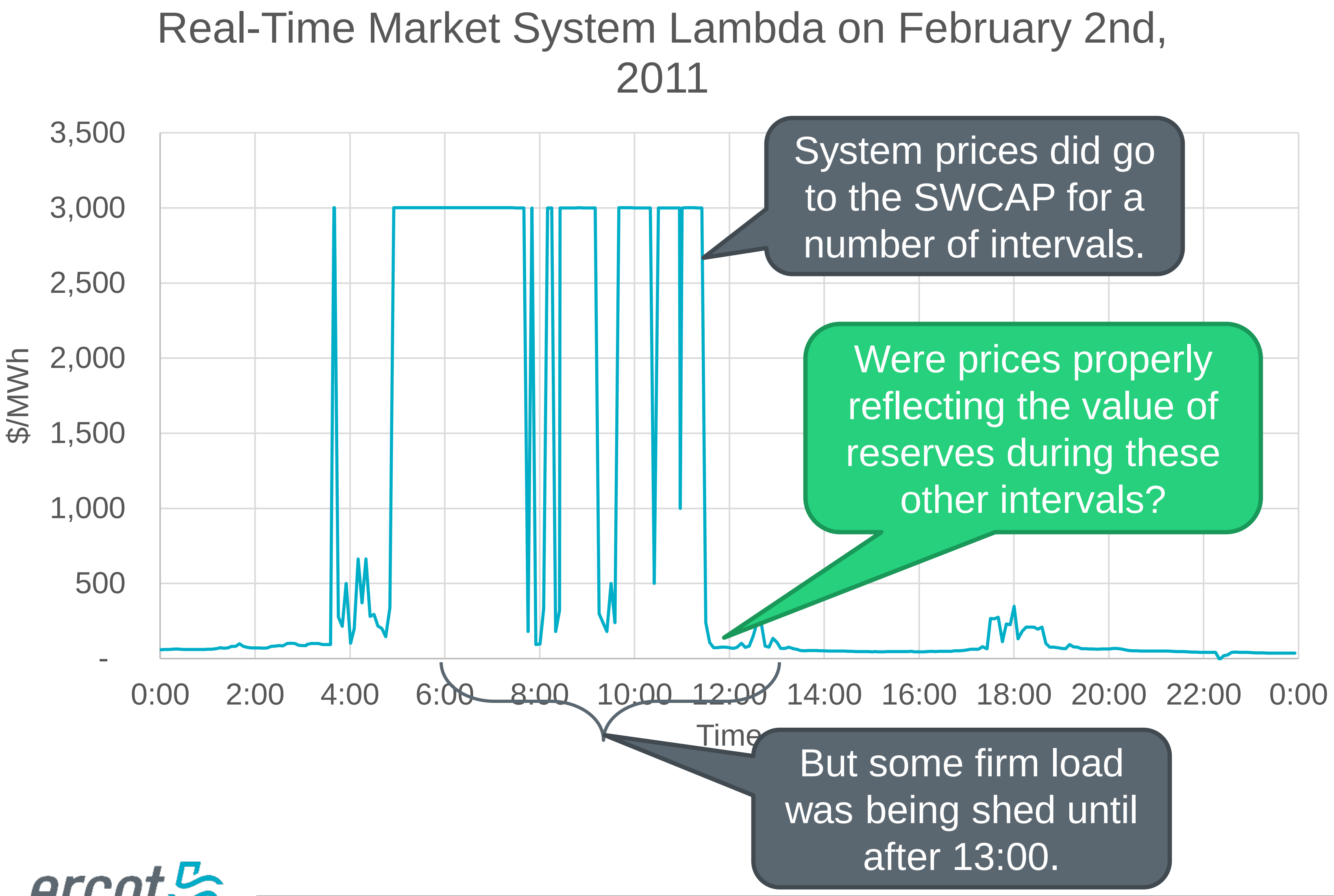
History of ORDC at ERCOT

- In Real-Time SCED, there is no energy and AS co-optimization
  - AS is procured in the Day-Ahead Market (DAM) and in the Adjustment Period, if adjustments are needed, via Supplemental AS Market (SASMs).
  - Resources with AS responsibility have their MW capacity reserved for AS and it is not available to Real-Time SCED unless released by ERCOT processes.

- With this design, higher prices during times of scarcity had to be driven by higher priced Energy Offer Curves.

- This dependency on higher priced Energy Offer Curves can fail to recognize the value that diminishing real-time reserves provide in avoiding potential load-shedding events on the system.

Real-Life Example from 2011



History of ORDC at ERCOT

- Dr. Hogan’s proposal involved the Real-Time co-optimization of energy and AS utilizing a LOLP based operating reserve demand curve

- An alternative approach (ORDC B+) was developed by ERCOT Staff and Dr. Hogan that could be implemented quickly and at a fraction of the cost but would capture system-wide scarcity pricing benefits with a full Real-Time energy and AS co-optimization

History of ORDC at ERCOT

- PUC directed ERCOT to file a Nodal Protocol Revision Request (NPRR) that would implement (ORDC B+) Real-Time Reserve Price Adder to LMP based on and ORDC with the following parameters:
  - VOLL= 9000 $/MWh,
  - Minimum Contingency Reserve = 2000 MW
  - Remove Price floors on submitted Energy Offers for the portion of capacity reserved for AS
  - Use the continuous form of the cumulative distribution function in calculating the LOLP

- ERCOT filed NPRR 568 on September 19th, 2013 to the Protocol Revisions Subcommittee (PRS) and the changes went live in production on June 1, 2014.

How Do You Determine the marginal value of reserves?

Sufficient real-time reserves help avoid load-shedding events

There is value in avoiding load-shedding events

Value of real-time reserves = *Value* of avoiding load-shed

Determination of ORDC

- The key in determining the ORDC is the calculation of the LOLP

- LOLP for a given level of available Real-Time reserves is the probability of an event/disturbance that can **exceed** the level of these available Real-Time reserves.
  - If such an event/disturbance occurs, it will lead to firm Load shed.

- LOLP depends on the following factors:
  - Errors in the forecasted demand
  - Forced outages of Resources
  - Errors in the forecast of Intermittent Renewable Resource (IRR) output

Determination of ORDC - LOLP

Option 1:

- Determine the probability distribution of aggregate events/disturbances by the convolution of the individual probability distributions of:
  - Forecasted demand error
  - Forecasted IRR output error
  - Forced outage Resource capacity after a reference time
- Calculate the LOLP as a function of Real-Time reserves using the probability distribution of the aggregate events/disturbances

Determination of ORDC - LOLP

Option 2: (ERCOT approach)

- ERCOT has adopted the approach of gathering the probability distribution of errors on the available reserves (forecasted reserves minus Real-Time reserves).
  - Inherently captures the three factors that were previously mentioned
- Analyzing historic "events," where an event is defined as the difference between the hour-ahead forecasted reserves and the reserves that were available during the Operating Hour
- The Hour-Ahead Reliability Unit Commitment (HRUC) study is the reference from which the probability distribution of errors in forecasted reserves is calculated.

Determination of ORDC - LOLP

Option 2: (ERCOT approach)

- Reserve Error = Hour-Ahead forecasted reserves
  – Hourly Average of RT-SCED reserves

- Exclude non-Wind Intermittent Renewable Resources, Nuclear Resources, and other Resources with a status or operating condition indicating they cannot provide reserves at that time
- Calculated for each Operating Hour from Nodal Go-Live (December 1st, 2010)

Determination of ORDC - LOLP

Option 2: (ERCOT approach) - continued

- Statistics (mean-M, standard deviation-SD) are calculated on:
  - Seasons (Winter, Spring, Summer, Fall)
  - 6 time-of-day blocks, each consisting of 4 hours
  - Errors are assumed to be normally distributed
- From this, the LOLP for a given level of available reserves (R) for a particular season and time block is

LOLP(M,SD,R) = 1 - CDF(M,SD,R)

Where CDF is the Cumulative Distribution Function of the normal distribution with mean M and standard deviation SD

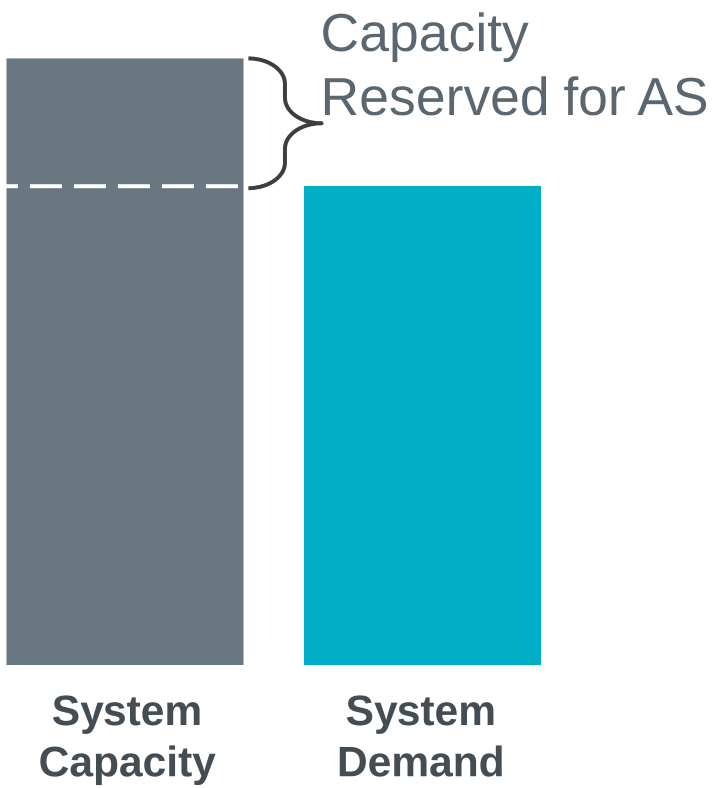
$$CDF(M,SD,R) = \frac{1}{2} \times \left[1 + erf\left[\frac{R-M}{\sqrt{2} \times SD}\right]\right]$$

$$\mathrm{erf}(z) = \frac{2}{\sqrt{\pi}} \times \int_0^z e^{-t^2} dt$$

Ancillary Services

- An AS is a reliability service procured by ERCOT and is capacity that is reserved on Resources to allow ERCOT to respond quickly to changing system conditions.

- The primary types of AS are:
  – Regulation Service (Reg-Up and Reg-Down) capacity that can be ramped within 5 minutes
  – Responsive Reserve Service (RRS), Frequency responsive, and capacity that can be ramped within 10 minutes
  – Non-Spinning Reserve Service (Non-spin), capacity that can be ramped within 30 minutes



Ancillary Services vs. Operating Reserves

- These AS are cleared in the DAM or a SASM and have specific rules on how they are utilized.
  - Again, ERCOT does not have energy and AS co-optimization in Real-Time.

- Operating reserves refers more generally to the extra capacity that is available in Real-Time and includes:
  - AS that has not been deployed
  - Headroom on Resources that are not providing AS
  - Extra headroom on Resources that are providing AS

Determination of ORDC - Multiple Reserves

- For Real-Time, available reserves are classified as either fast or slow.
  - Fast reserves are able to respond to events occurring any time during the given Operating Hour, including just after the event.
  - Slow reserves are only able to respond to events occurring in the last 30 minutes of the given Operating Hour.
- The time period of evaluation of the marginal value of reserves is the Operating Hour.
- Assumes that the event/disturbance are uniformly random across the Operating Hour.

Determination of ORDC - Multiple Reserves

- Two distributions are used
  - Whole Hour distribution (S+F) corresponding to the entire Operating Hour where both fast and slow reserves can respond to events/disturbances
  - First ½ Hour distribution (F) corresponding to the first 30 minutes where only fast reserves can respond to events/disturbances
- The Whole Hour distribution (S+F) is based on the probability distribution of the hourly errors in reserves i.e. mean-M_{S+F} and standard deviation SD_{S+F} described earlier
- The first ½ Hour distribution (F) is derived from the probability distribution parameters of (S+F) and its probability distribution is defined with
  - Mean $M_F=0.5*M_{S+F}$,
  - Standard deviation $SD_F=0.707SD_{S+F}$

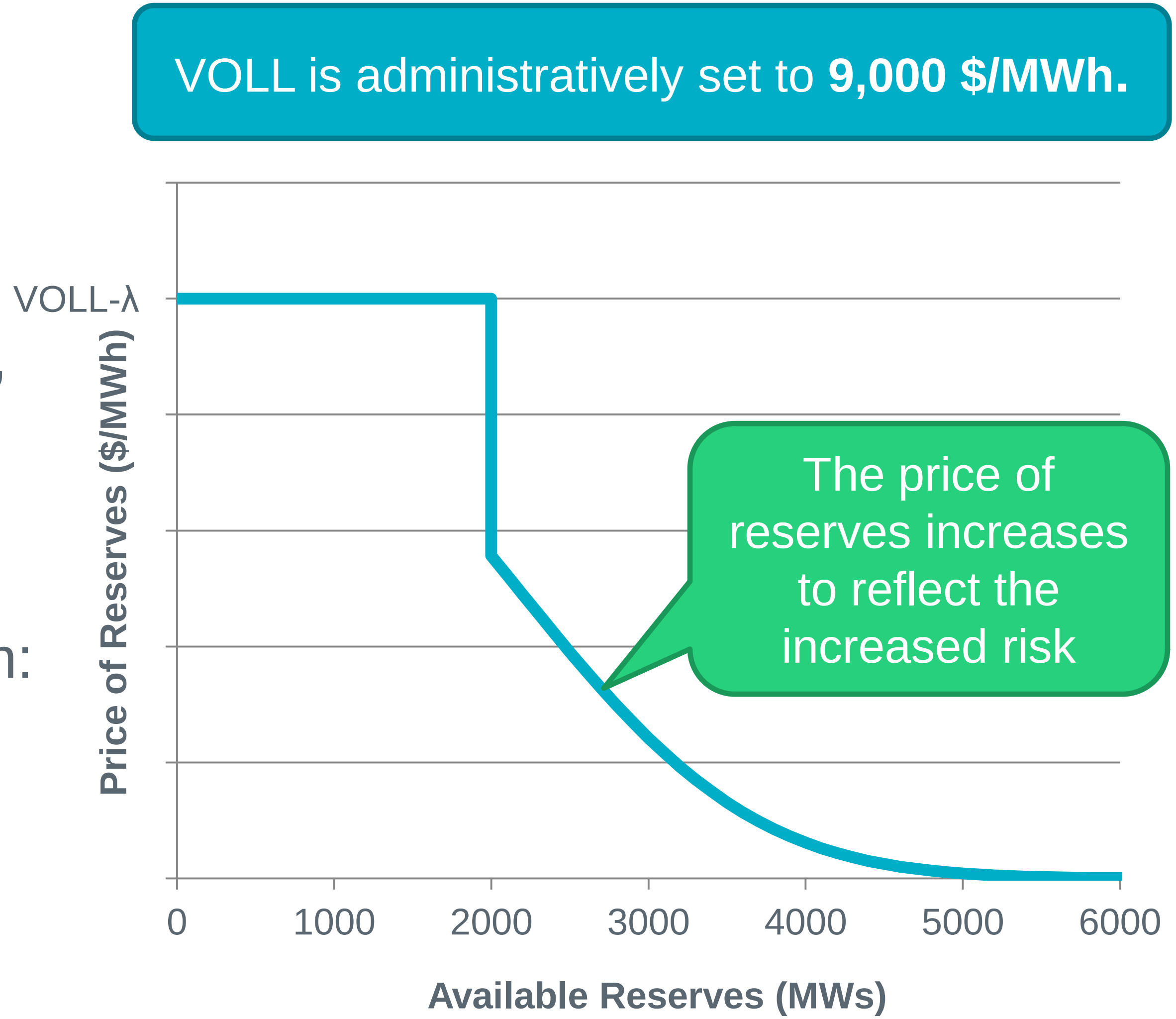
Determination of ORDC - Minimum Contingency Level X

- ERCOT will start taking out-of-market actions well before available reserves drop to zero.
- To factor in the price impacts of these out-of-market actions when reserves are steadily being depleted, a Minimum Contingency Level of 2000 MW was decided upon by the PUCT.
- The LOLP for reserve levels below the minimum contingency level (X) will be set to one. In addition, since ERCOT is at a higher risk of shedding firm load when reserves fall near or below the minimum contingency reserve level, the LOLP curve is shifted to the right by the minimum contingency level (X) amount.
  - LOLP = 1 when available RT reserves R <= X
  - When R > X, LOLP = 1 - CDF(M,SD,(R-X)), i.e. right shift distribution by X

The Operating Reserve Demand Curve

- ERCOT implemented an Operating Reserve Demand Curve (ORDC) on June 1, 2014.

- Reserve prices should be based on:
  - VOLL
  - LOLP



Calculating the Reserve Price Adders

RTORPA

- Represents the value of the Real-Time On-Line Reserve Capacity (RTOLCAP) being provided by Qualified Scheduling Entities (QSEs)
  - Capacity able to respond immediately following an event

RTOFFPA

- Represents the value of the Real-Time Off-Line Reserve Capacity (RTOFFCAP) being provided by QSEs
  - Capacity not currently available but could be in 30 minutes

Calculating the Reserve Price Adders

- With reserves broken up into two categories (RTOLCAP and RTOFFCAP), two separate prices have to be calculated.
- First is the calculation RTOFFPA:
  – Based on the probability of reserves falling below a minimum contingency level over an hour given the amount of reserves available in the latter 30 minutes of the hour.
    - Amount of reserves is the sum of RTOLCAP and RTOFFCAP
  – The probability value is multiplied by .5 x (VOLL-System Lambda), reflecting that the reserves are only available for half of the hour.
  – The minimum contingency value is administratively set at 2,000 MW.

Calculating the Reserve Price Adders

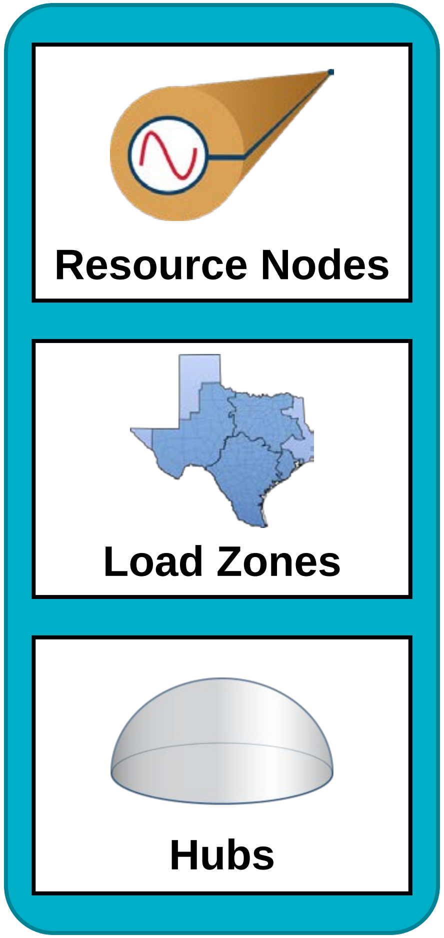
- Once RTOFFPA is known, then calculate the RTORPA:
  - Based on the probability of reserves falling below a minimum contingency level over 30 minutes given the total amount of reserves available in the first 30 minutes of the hour.
    - Amount of reserves is only RTOLCAP
  - The probability value is the multiplied by .5 x (VOLL-System Lambda), reflecting that 30 minute period
  - This product is then added to RTOFFPA to determine the final RTORPA

Idea is that RTOLCAP contributes to preventing load shedding in both halves of the hour.

Real-Time Pricing Methodology

- These price adders are determined during each SCED interval, however the impact still needs to be incorporated into the energy price.
- Locational Marginal Prices (LMPs) from SCED are then combined with the price adders to form real-time Settlement Point Prices (SPPs).
- QSEs will also be paid or charged for AS imbalances in real-time.

While LMPs are location-specific, the value of reserves is ERCOT-wide



Further Details

- Methodology for Implementing ORDC to Calculate Real-Time Reserve Price Adder
  - http://www.ercot.com/content/wcm/key_documents_lists/89286/Methodology_for_Implementing_ORDC_to_Calculate_Real-Time_Reserve_Price_Adder.zip

But Wait, There’s One More Adder

- Reserves need to be properly valued during times of scarcity.

- It’s also important to account for price distorting actions that may occur during energy scarce conditions, specifically those taken by ERCOT to maintain reliability.

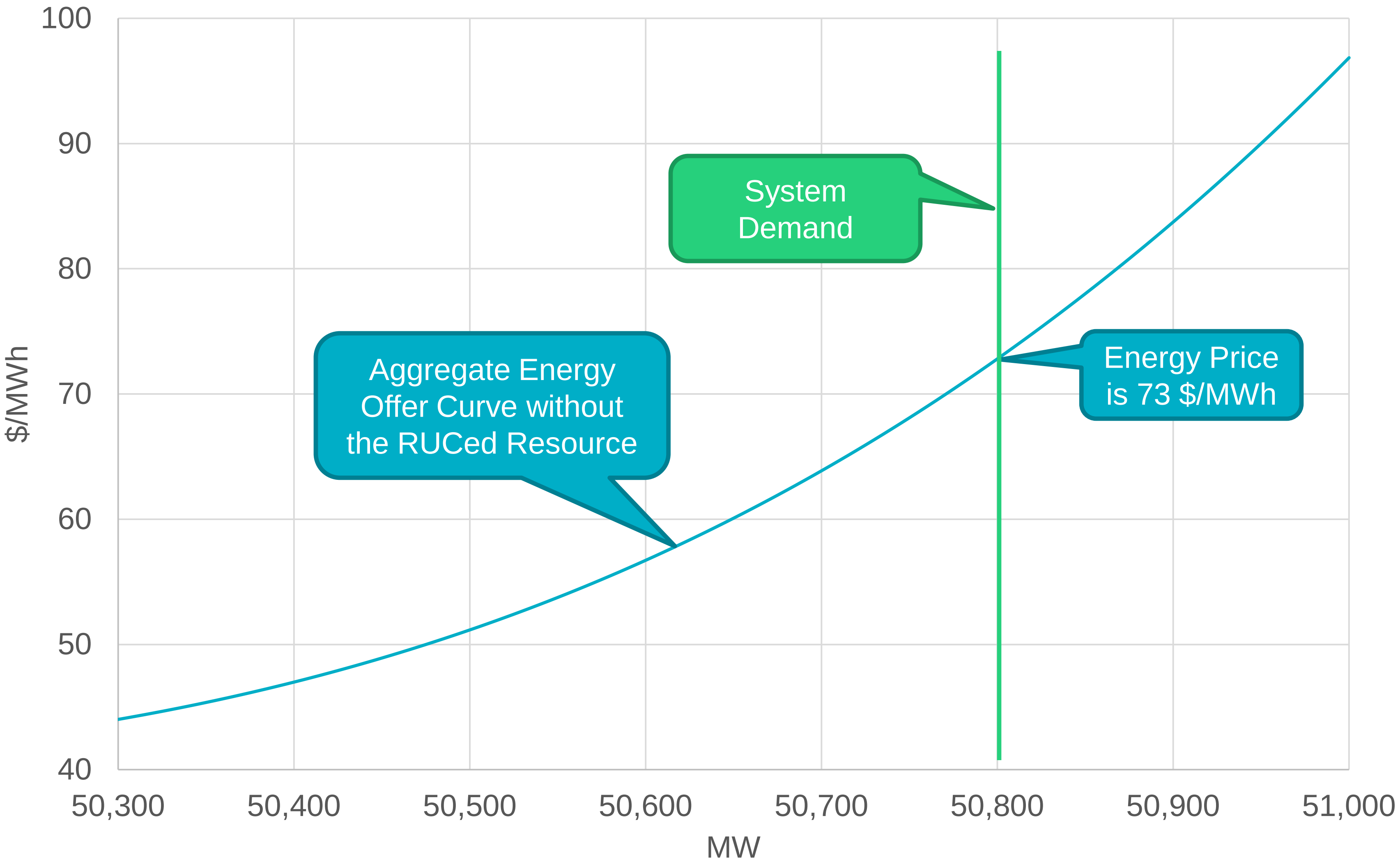
  – These “reliability deployments” could be suppressing prices at the exact moment that more Resource production is needed

What are the Reliability Deployments?

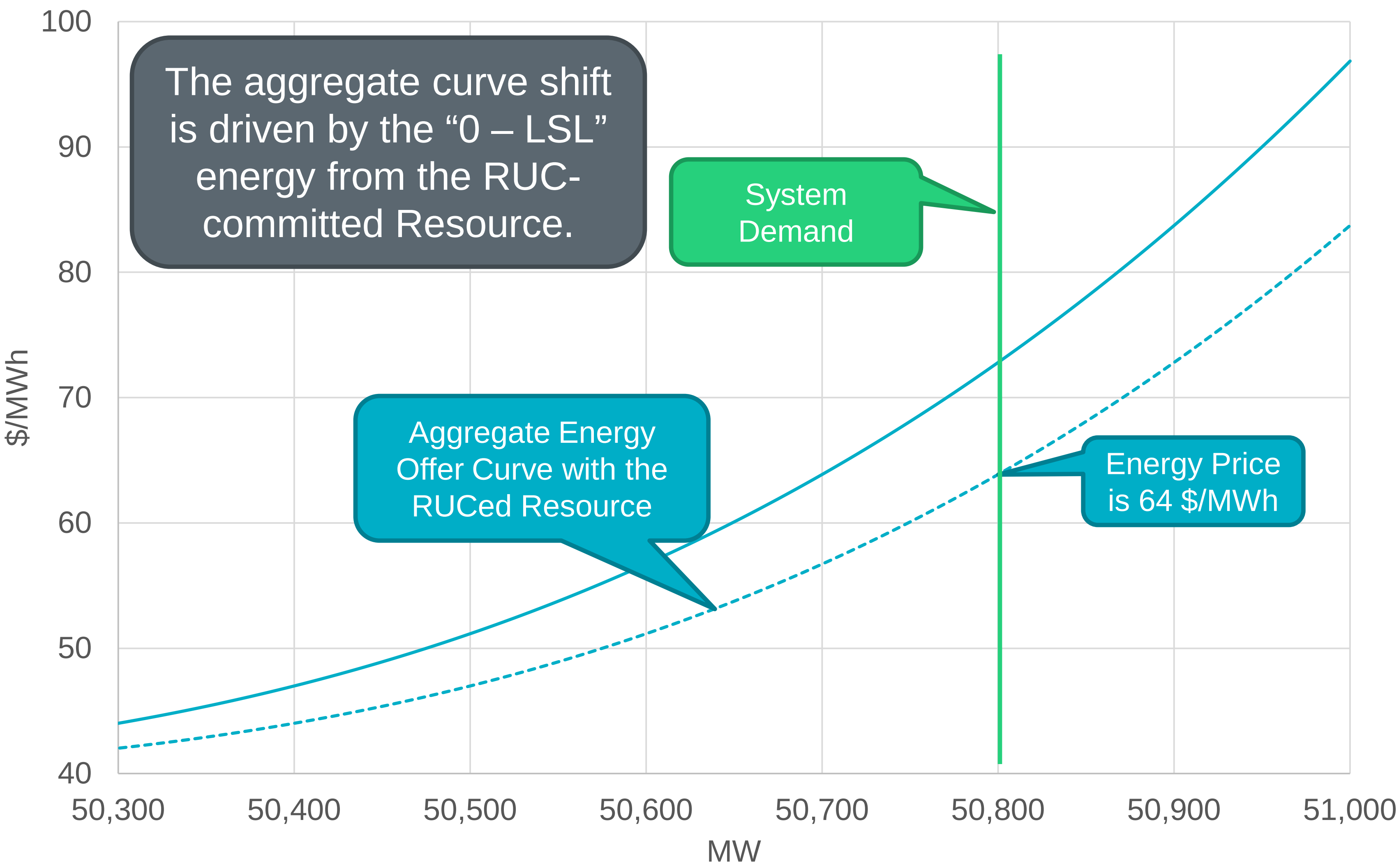
- The commitment of Resources through Reliability Unit Commitment (RUC)
  - This would also include Reliability Must Run (RMR) Resources or Resources contracted for capacity.
  - Relaxes the low limits to negate the impact of the "price-taking" capacity in the second (Pricing Run) of RT SCED
- Deployment of Load Resources other than Controllable Load Resources (Increase demand by deployed amount in second Pricing run of RT SCED)
- Deployment of Emergency Response Service (ERS) (Increase demand by deployed amount in second Pricing run of RT SCED)

Additional types of reliability deployments will be added as part of Nodal Protocol Revision Request (NPRR) 768

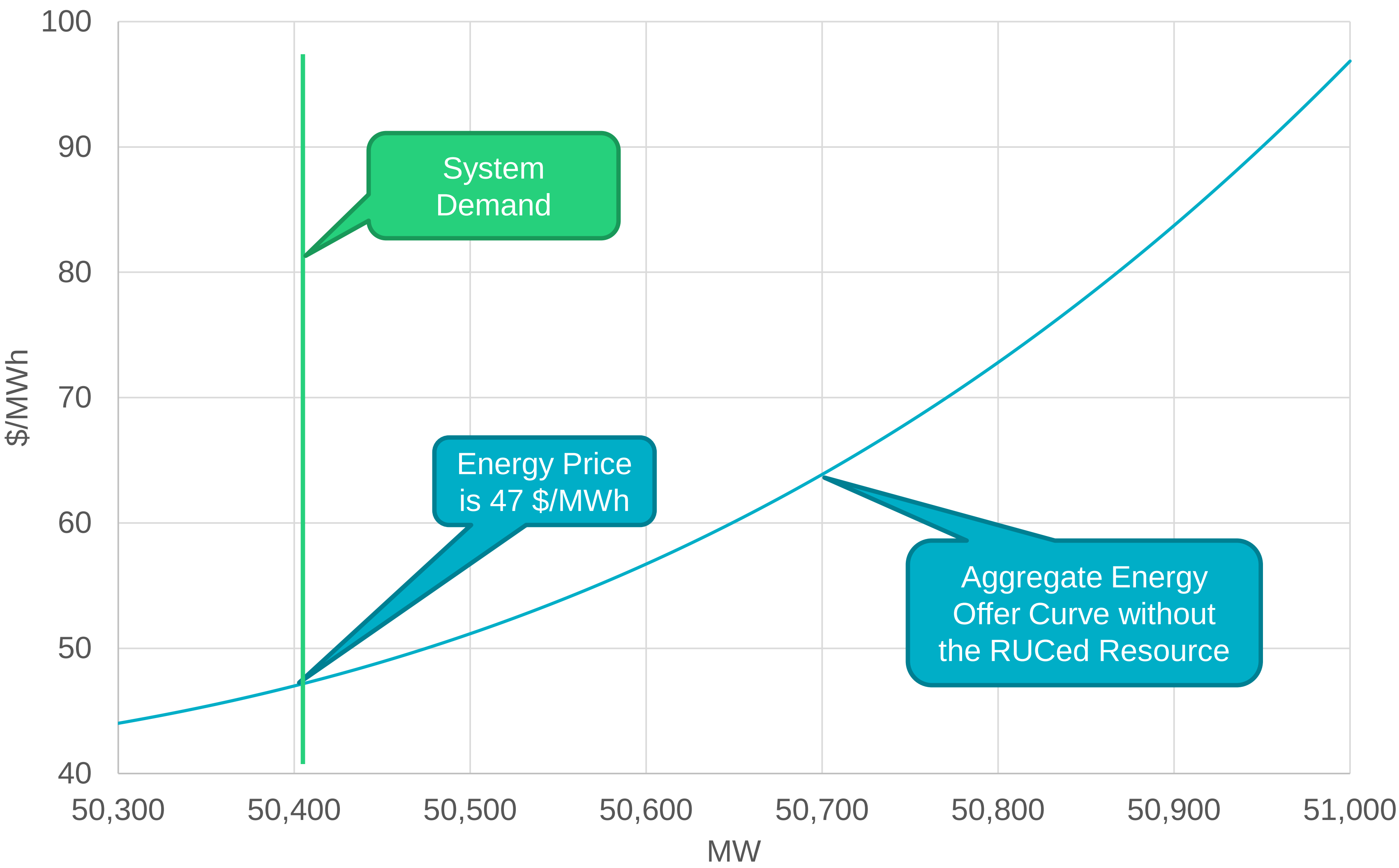
One Example of Price Suppression - RUCs



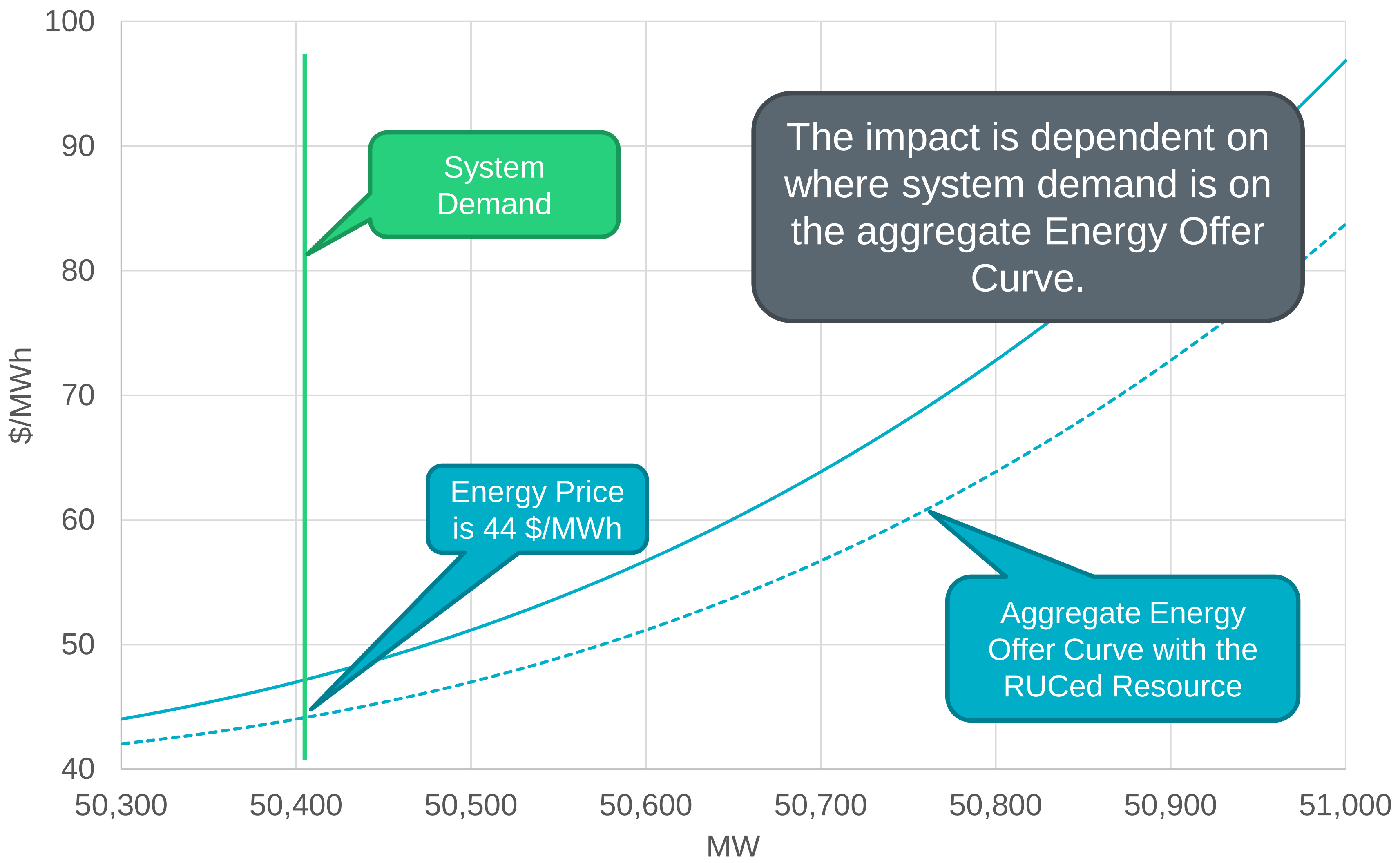
One Example of Price Suppression - RUCs



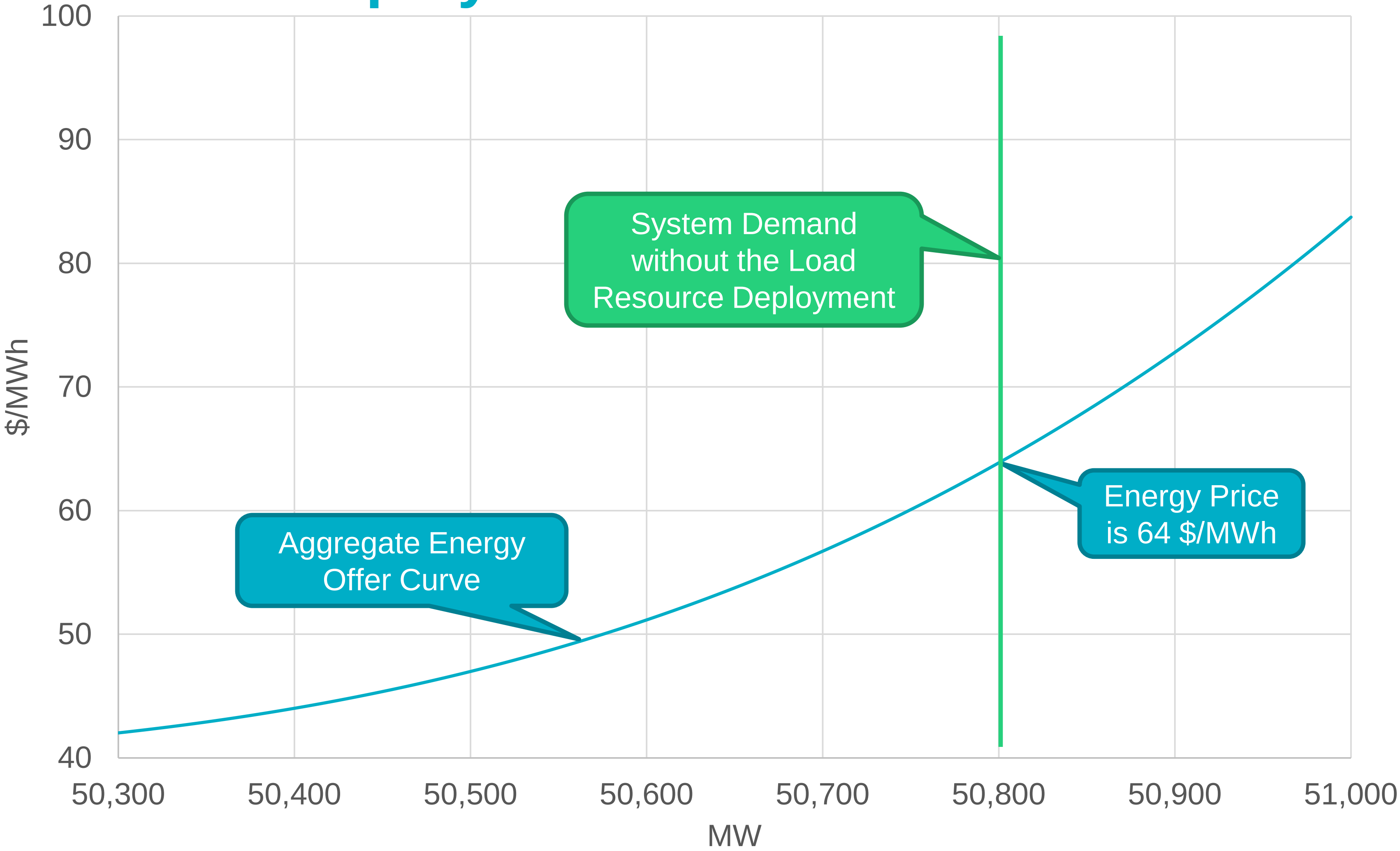
One Example of Price Suppression - RUCs



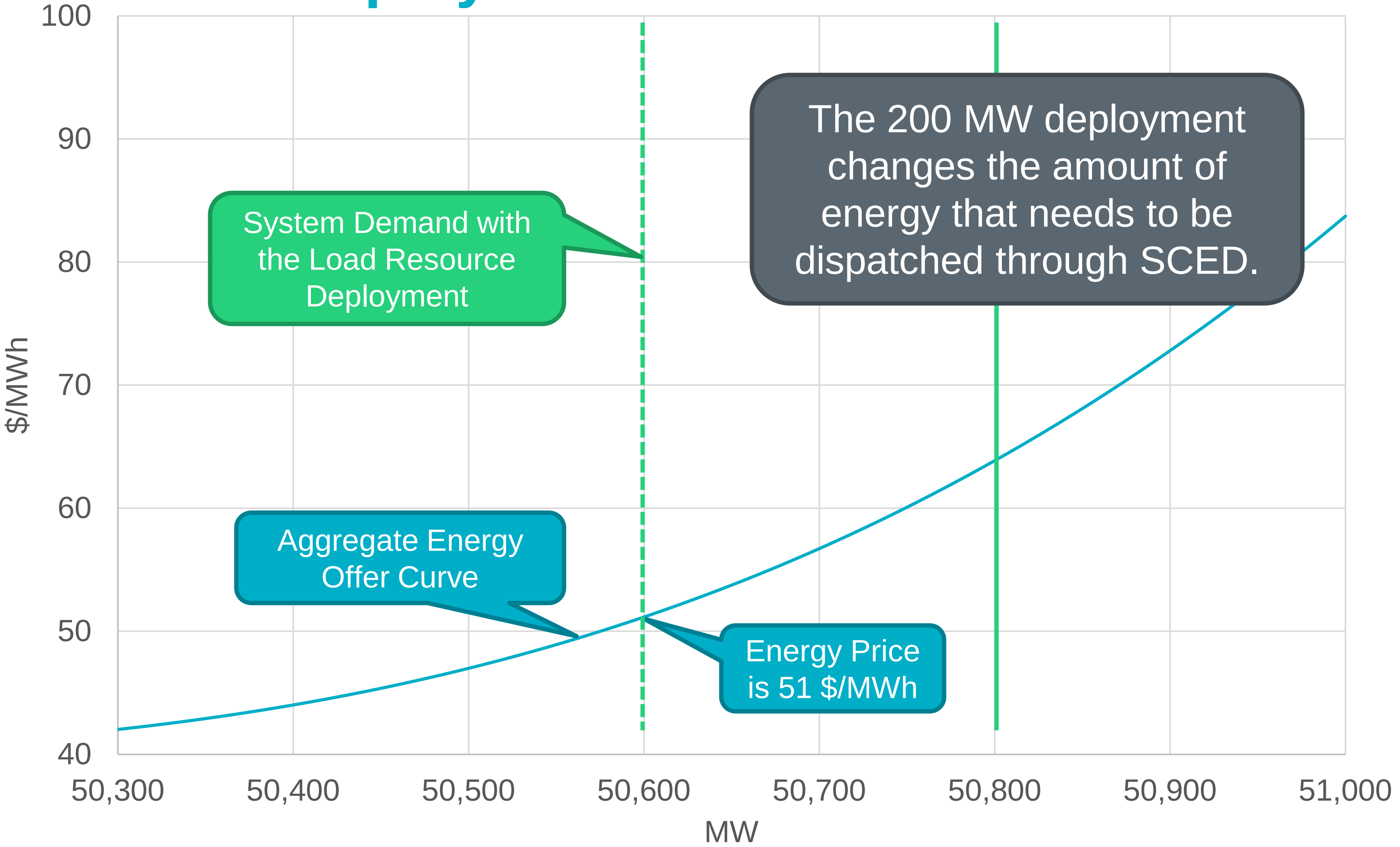
One Example of Price Suppression - RUCs



Another Example of Price Suppression – Load Resource Deployment



Another Example of Price Suppression – Load Resource Deployment



The Reliability Deployment Price Adder

- The RTORDPA was implemented on June 1, 2015 to counteract the price suppressing effect of reliability deployments.

- This effect is estimated by executing the SCED process a second time (Pricing Run) with the inputs changed to simulate the scenario in which the reliability deployments hadn't taken place.
  - The impact is calculated by taking the difference between the System Lambda resulting this second execution of the SCED process and the System Lambda from the first execution of the SCED process (before the inputs were changed)

Back to the Real-Time Pricing Methodology

- Like the reserve price adders, the RTORDPA also needs to be incorporated into the SPPs

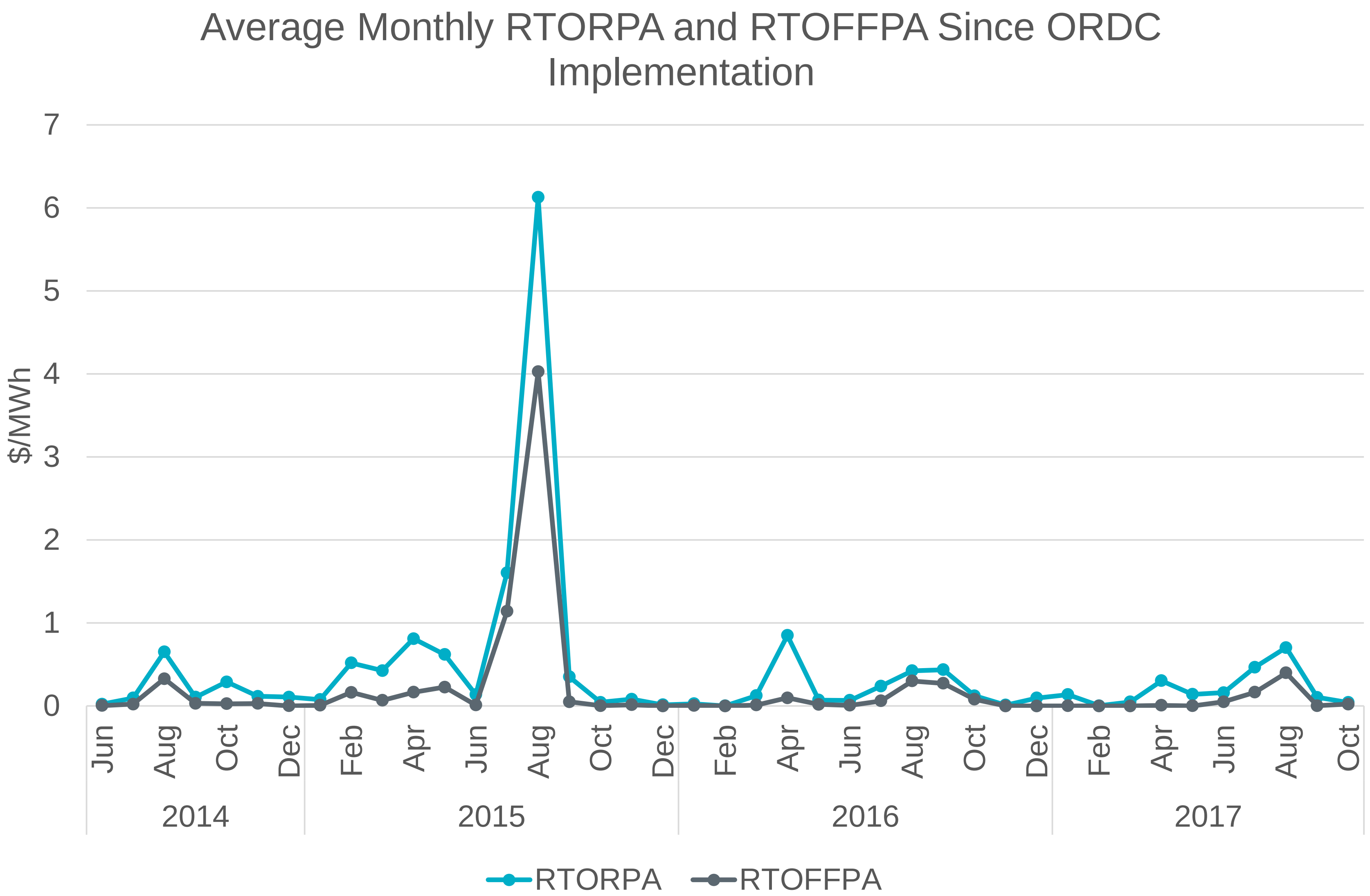
Settlement Point Prices
= Avg. LMPs +
Avg. RTORPA +
Avg. RTORDPA

- RTORDPA is also applied ERCOT-wide.

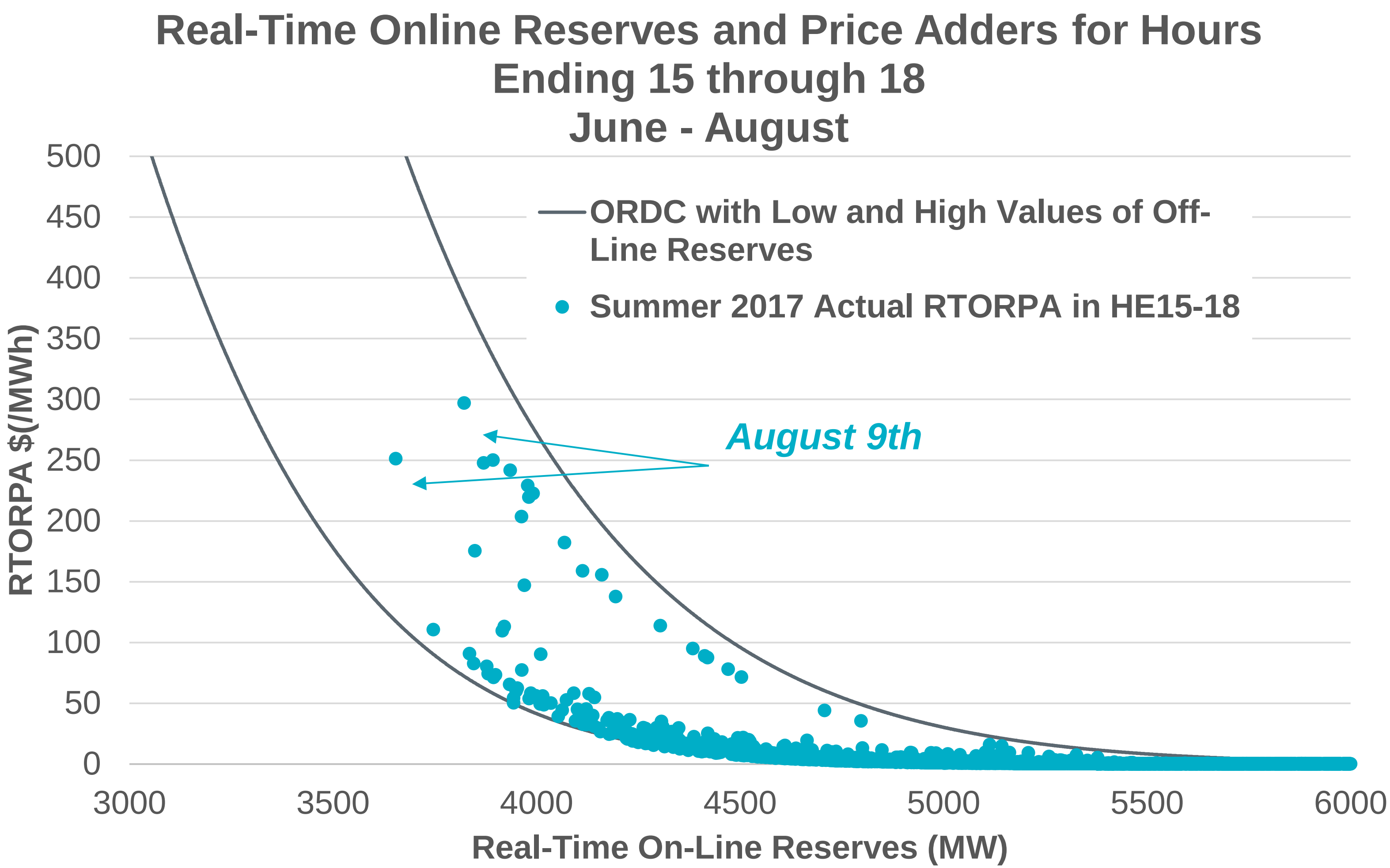
RTORDPA does not account for the potential local impacts of reliability deployments.

- RTORDPA is also paid or charged based on AS Imbalances in order to make the Resources indifferent to energy prices.

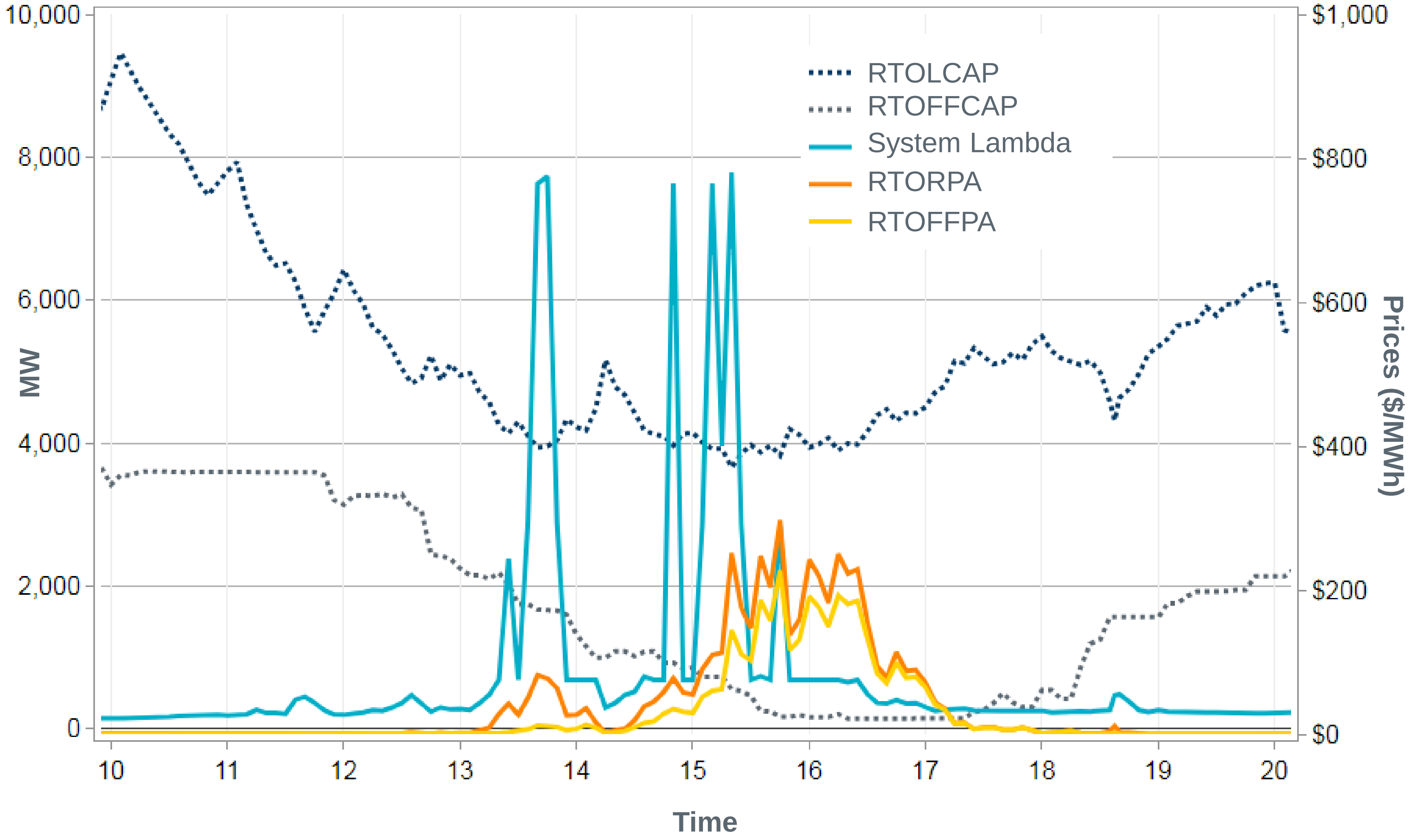
Historic Results – Reserve Price Adders



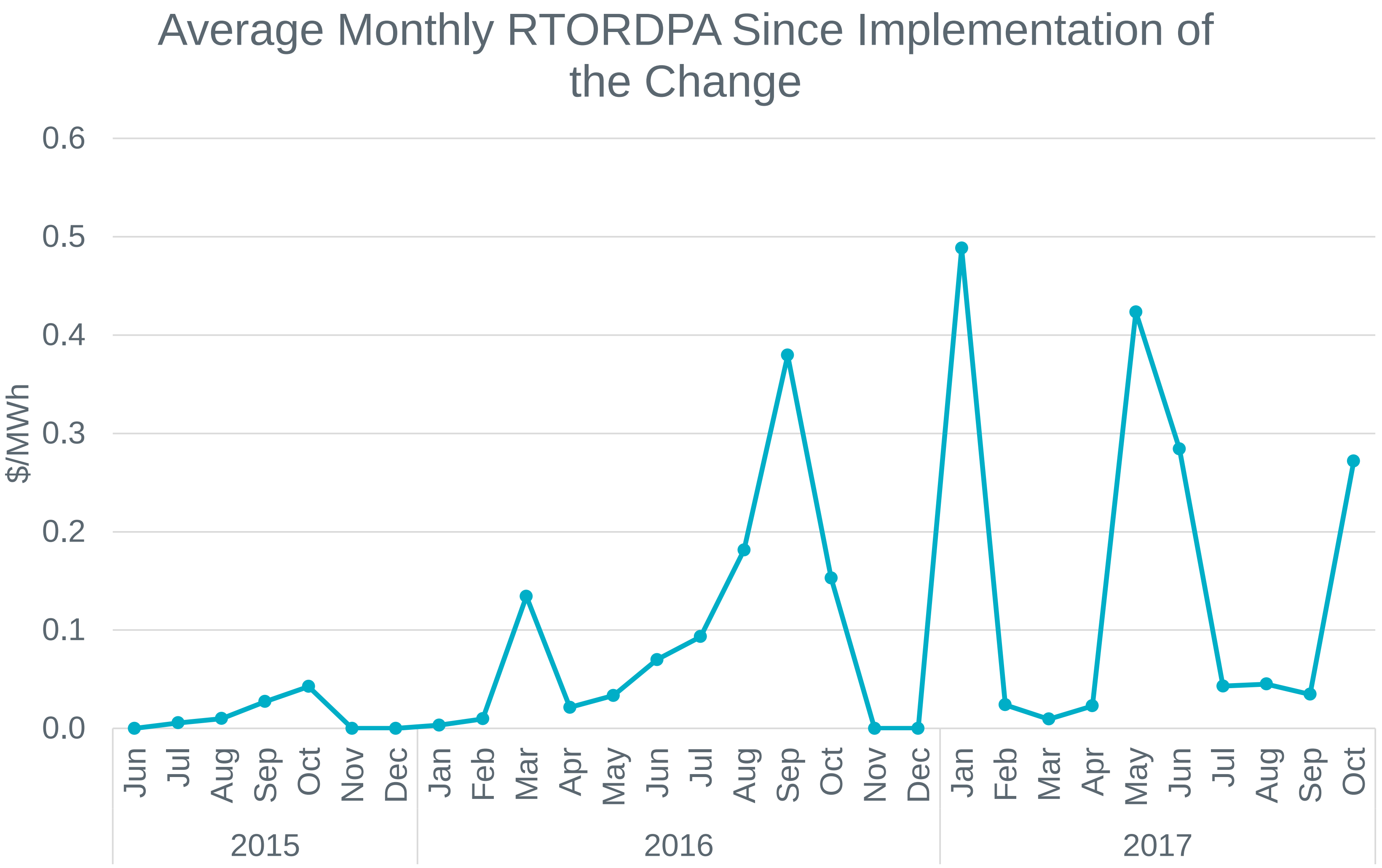
Summer 2017 On-Line Reserve Price Adders



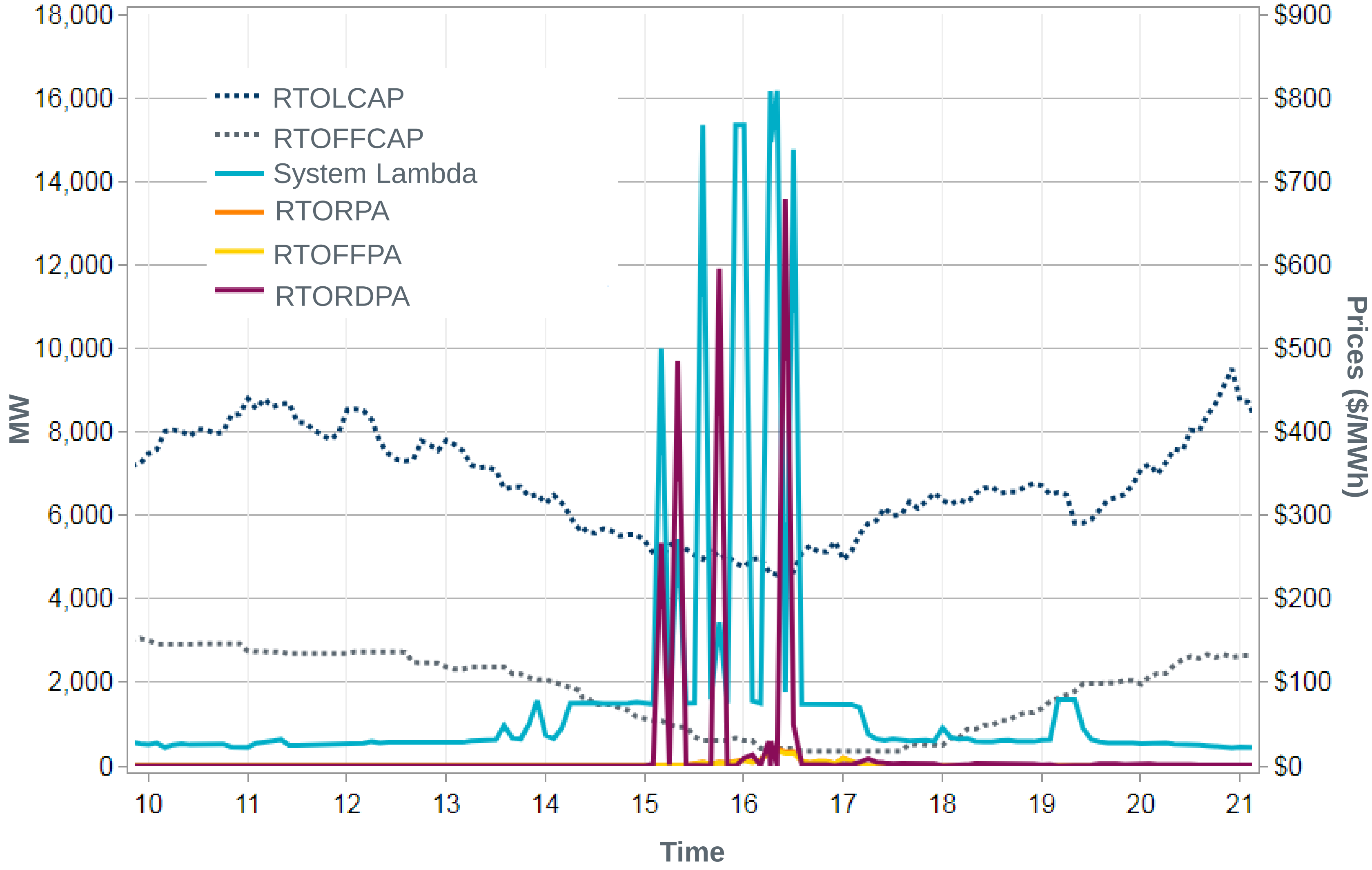
A Closer Look at August 9th, 2017



Historic Results – Reliability Deployment Price Adders



A Closer Look at October 4th, 2017



Upcoming Changes to ORDC

- The PUCT has directed ERCOT to address the impact of Resources committed by RUC (out-of-market action) on the Real-Time reserve price adders.
- The RUC capacity will be removed from RTOLCAP and may be added to RTOFFCAP, depending on the start time of the Resource(s) committed by RUC.
- ERCOT plans to implement this change by the summer of 2018.

Future change? Real-Time Energy and AS Co-Optimization

- ERCOT Stakeholders have been discussing implementing energy and AS co-optimization (RTC) in the Real-Time Market since 2015

- Latest version of whitepaper: http://www.ercot.com/content/wcm/key_documents_lists/131797/RT_Co-optimization_Scope_UPDATED_09292017.docx

Future change? Real-Time Energy and AS Co-Optimization

- Main discussion items
  - Coordination of the VOLL, Power Balance Penalty cost, SWCAP, and the Maximum value on the respective AS Demand Curves

	Current Value	Example of Value with RT co-optimization
VOLL ($/MWh)	9000	9000 (no change)
Power Balance Penalty ($/MWh)	9001	9001 (no change)
System Wide Offer Cap ($/MWh)	9000	2000
Max Value on AS Demand Curve ($/MWh)	N/A	7000

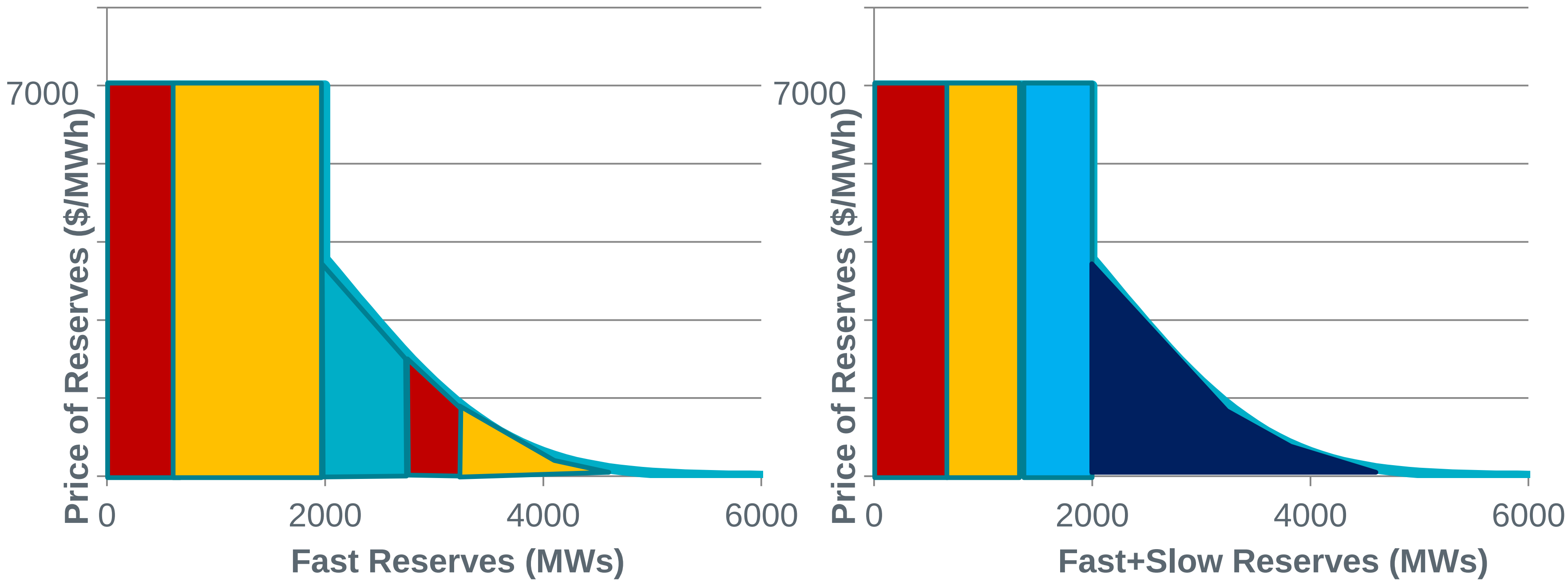
Real-Time Energy and AS Co-Optimization using ORDC

- Main discussion items - continued
  - Proposed AS Product Set:
    - Reg-Up/FRRS-Up ← No change
    - Reg-Down/FRRS-Down ← No change
    - RRS ← No change
    - Spinning Operating Reserve (SOR): provided by On-Line Resources that can convert capacity to energy in 30 minutes – same requirement as current Non-Spin
    - Non-Spin Operating Reserves (NSOR): provided by Off-Line Resources that can convert capacity to energy in 30 minutes – same requirement as current Non-Spin

Real-Time Energy and AS Co-Optimization using ORDC

- Main discussion items - continued
  - Determination of individual AS Demand Curves
    - ERCOT Staff proposal is to determine individual AS Demand Curves based on the based on the disaggregation of the two ORDCs (Fast, Fast+Slow).
    - The aggregation of the RegUp, RRS, SOR and NSOR AS Demand Curves should result in giving back the two ORDC curves.

AS Demand Curves for ERCOT style constraints

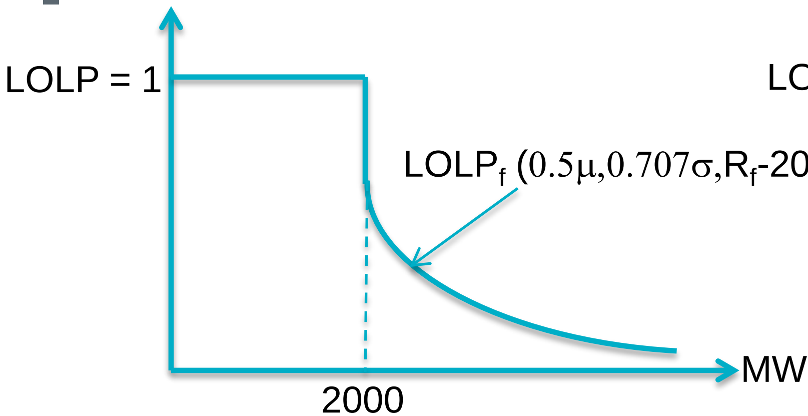


0.5 weightage applied to each ORDC during disaggregation to AS Demand Curve

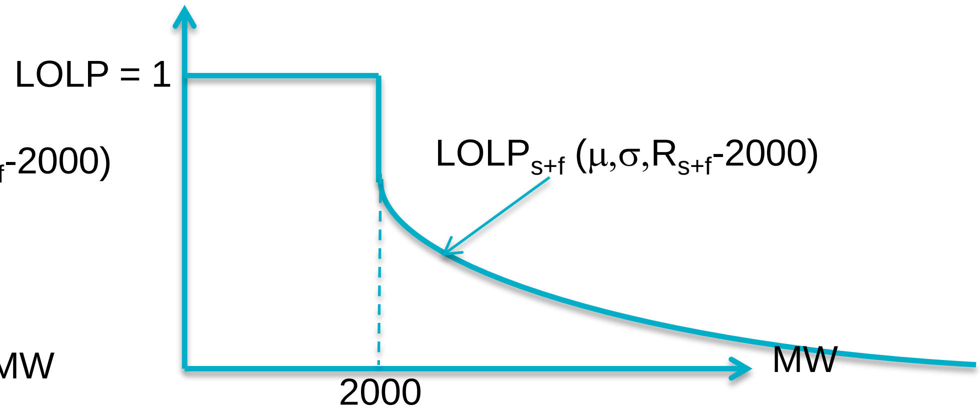
Questions?

Appendix

Calculating the Reserve Price Adders



Fast Reserves (R_f=RTOLCAP)



Fast + Slow Reserves (R_{s+f}=RTOLCAP+RTOFFCAP)

- Off-Line Reserve Price Adder (RTOFFPA)

  $\text{RTOFFPA} = 0.5*(\text{VOLL}-\lambda)*\text{LOLP}_{s+f}(\mu,\sigma,R_{s+f} - 2000)$

  $\text{LOLP}_{s+f} = 1-\text{CDF}(\mu,\sigma,R_{s+f} - 2000)$

- On-Line Reserve Price Adder (RTORPA)

$\text{RTORPA} = \text{RTOFFPA} + 0.5*(\text{VOLL}-\lambda)*\text{LOLP}_f(0.5\mu,0.707\sigma,R_f - 2000)$

$\text{LOLP}_f = 1-\text{CDF}(0.5\mu,0.707\sigma,R_f - 2000)$

λ is the Power Balance Constraint Shadow Price

Real-Time Pricing Methodology

- All energy prices (LMPs) computed by Real-Time SCED are increased by the amount of the On-Line Reserve Price Adder
- All unloaded On-Line Resource MW Capacity minus the On-Line Ancillary Service Responsibility (RegUp, RRS, On-Line Non-Spin) is paid or charged @ On-Line Reserve Price Adder
  - This incents the Resource to follow dispatch (Base Point) from Real-Time SCED
  - Referred to as the AS imbalance payment or charge for On-Line Resources
- All <u>eligible</u> (Resources that have cold starttime of 30 minutes or less) Off-Line available MW capacity minus the Off-Line Non-Spin Service Responsibility is paid or charged @ Off-Line Reserve Price Adder
  - Referred to as the AS imbalance payment or charge for Off-Line Resources