Opinion on

Flexible Ramping Product Refinements

by

James Bushnell, Member

Scott M. Harvey, Member

Benjamin F. Hobbs, Chair

Members of the Market Surveillance Committee of the California ISO

Final, September 9, 2020

I. Introduction

The Market Surveillance Committee (MSC) of the California Independent System Operator (CAISO) has been asked to comment on elements of the ISO's proposal for flexible ramping product refinements, in particular, the implementation of a nodal delivery test.[1] The initiative leading to this proposal has been discussed during MSC meetings on October 11, 2019 and May 29, 2020, while the issues motivating this initiative have been discussed at a variety of MSC meetings over the 2017 to 2019 period.

Summary of our recommendations: The CAISO proposal to implement a nodal delivery test is a reasonable approach to improving the deliverability of flexiramp and thereby enabling the CAISO to balance variations in variable energy resource output at lower cost. However, we have a few cautions regarding the development of this design. First, it will be important to evaluate the impact of the ultimate nodal delivery test design on FMM and particularly RTD solution time before proceeding to implementation. Second, it will also be important to assess how well the nodal delivery test design will perform before proceeding to implementation. This testing will be particularly important to the extent that the current design is modified in order to reduce solution time impacts. Third, the CAISO's software implementation should provide the flexibility to adjust key parameters on an ongoing basis as it is very unlikely that the CAISO will be able to specify the best values in advance, and the ideal values are also likely to change over time with changes in the CAISO and Western EIM resource mix and with the further expansion of the Western EIM. Fourth, stakeholders and the CAISO should both recognize that the

[1] California ISO, "Flexible Ramping Product Refinements, Draft Final Proposal," May 8, 2020,

successful implementation of a nodal delivery ability test will be one step in realizing the full potential of the flexiramp design concept and that it is likely that improvements will be needed to other elements of the current design. Fifth, the kind of changes in the scheduling of flexiramp in forward periods in STUC that DMM has suggested may ultimately be desirable, but it does not appear possible to assess the value of such changes, or even to develop a design for changes of this type until an improved test for deliverability is implemented and works at least roughly as intended and the CAISO implements the improved methods for determining future ramp needs that are also being developed as part of this initiative.

II. Background

An upward flexiramp constraint was implemented in HASP and RTPD in December 2011 in order to provide unloaded rampable capacity that could be used to balance variations in net load. When binding, this constraint could cause HASP to schedule additional interchange or cause RTPD to commit additional generation. This flexible ramping constraint was enforced in HASP and RTPD but was not enforced in RTD.

This design resulted in phantom ramp being included in the HASP and RTPD evaluation of available ramp capability. Phantom ramp is ramp that is counted as being available in RTPD (and after May 2014 in the FMM) but did not exist in the real-time dispatch (RTD). Phantom ramp is created when resources are scheduled to operate out-of-merit below their upper limit so as to create additional ramp capability in RTPD but are not actually dispatched this way in RTD. This issue was discussed in a number of Market Surveillance Committee Meetings over 2012-2015,[2] and in the Market Surveillance Committee's January 26, 2016 opinion on the Flexible Ramping Product.

Subsequently, the flexible ramping product was implemented in RTD on November 1, 2016, as well as a downward product in both RTPD and RTD. The flexiramp product design was intended to eliminate phantom ramp and make more ramp available for use in real-time dispatch. However, it was observed almost immediately after implementation of the flexiramp product design that the price of flexiramp was consistently $0/MW in RTD, reflecting the fact that it had a zero shadow price in the real-time dispatch. This zero price indicated that the ramp constraint was not binding and therefore was having no impact on the actual real-time dispatch. In other words, the design was not changing system dispatch in a way that affected costs. In contrast, a significant positive price would indicate that some out-of-merit dispatch was necessary to make the target level of ramp available, and that the amount provided was less than the full target amount (the full target amount would only be scheduled at a zero cost based on the ramp product demand curve).

[2] See for example, A. Hundiwale, D. Tretheway and L. Xu "Order 764 Implementation – Flexible ramping constraint performance," March 11, 2014;

There are at least two broad reasons why this could occur. One reason could be that the amount of ramp that the RTD flexiramp requirement was intended to maintain was always available in RTD at no cost,. However, this should not be the case if the design is operating as intended because the flexiramp product design was implemented precisely because this was not the case historically. The point of the flexiramp design was to make more ramp available in RTD than would otherwise be the case. The second broad reason is that there was something wrong with the implementation of the flexiramp design. As we explain below, the CAISO has identified a number of flaws in the original flexiramp implementation. Many of these flaws contributed to some degree to this outcome in which the constraint virtually never bound nor had any impact on the dispatch or on the amount of ramp available in RTD.

A number of implementation flaws have been identified and a series of changes have been made in the design to address these flaws through early 2019.[3] Additional implementation flaws were discussed in the price and performance report and in past Department of Market Monitoring market reports. First, the FRP implementation in FMM apparently has not enforced the FRP requirement in the buffer interval in the RTPD FMM solution.[4] Interchange schedules are fixed in the buffer interval but it is possible that the failure to enforce the FRP requirement in the buffer could potentially at times have resulted in no start instructions being given to very quick starting resources that were scheduled to come on-line to provide ramp capability during the period covered by the buffer interval, with the consequence that they were neither on-line nor coming on-line when they were needed to provide upward ramp capability.[5] It is not clear how often this has occurred over the past few years nor has the CAISO reported the typical magnitude of the effects. The CAISO Energy Markets Price Performance Report identified one interval in which 200 megawatts of capacity were apparently not on-line because the flexiramp constraint was not enforced in the buffer interval, but we do not know how often impacts of this magnitude

---

[3] A number of flaws in the way the ramp target was calculated were corrected beginning on February 18, 2018, with additional modifications implemented during March 2018. See Rahul Kalaskar, California ISO, "Flexible ramping product performance discussion," Market Surveillance Committee Meeting, February 2, 2018; Kyle Westendorf, Department of Market Monitoring, California ISO, "Flexible Ramping Product Uncertainty Calculation and Implementation Issues, April 18, 2018, California ISO, Market Performance and Planning Forum, April 19, 2018, California ISO, Department of Market Monitoring, 2018 Annual Report on Market Issues and Performance, May 2019 pp. 92-93; Amber Motley, "Flexible ramping product requirements and load forecast discussion," June 7, 2018.

[4] The buffer interval is the first 15 minute interval in the RTPD look-ahead horizon. The prices, interchange schedules, and financial schedules for FMM settlements of this interval would have been determined in the prior RTPD run. The buffer interval is the interval prior to the "binding" interval in each RTDP run and is therefore the binding interval in the prior RTPD run. This temporal structure exists because RTPD must post interchange schedules more than 20 minutes in advance of the operating period in order to allow interchange transactions to be tagged to meet WECC requirements. The schedules for the buffer interval are resolved based on an updated load forecast, with interchange schedules fixed, to provide a more accurate starting point for units in determining schedules for the binding interval.

[5] See Guillermo Bautista-Alderete, "Price Performance in CAISO Energy Markets," Market Surveillance Meeting, October 11, 2019, p. 9; California ISO, "CAISO Energy Markets Price Performance Report," September 23, 2019, pp. 76-78.

have occurred.[6] This issue is intended to be addressed by changes being implemented in RTPD during fall 2020.[7]

Another issue identified by Department of Market Monitoring in past market reports is that proxy demand response resources have been bid into the CAISO market as 5 minute dispatchable resources during this period, despite the proxy demand response resources typically not being offered for dispatch on a 5 minute basis. [8] The CAISO determined in the analysis underlying the price performance study that these proxy demand response offers were being scheduled to provide flexiramp, both in the FMM and in RTD. Treating these offers as capable of providing dispatchable capacity in FMM and RTD would also have caused the FMM and RTD software to calculate that there was more flexiramp available to meet the target than was actually case, since almost none of this demand response was actually available to balance load in the real-time dispatch and only some was even available in FMM. [9]

The overall magnitude impact of the inclusion of non-dispatchable demand response in the flexiramp supply is not clear but the CAISO price performance report identified an interval where 75 megawatts of demand response was included in the flexiramp supply.[10] The Department of Market Monitoring has compiled data showing that in some months of 2019 demand response accounted for an average of about 20 megawatts of the flexiramp supply in intervals prior to power balance violations.[11] We understand that this issue has largely been addressed by changes in the registration of PDR resources during 2020 to reflect that fact that only a very few of these resources are actually dispatchable on a 5 minute basis.[12]

There appear at present to be two critical issues limiting the effectiveness of the flexiramp constraint that have yet to be addressed. The first of these issues is a weakness in the methods currently used to determine the target for flexiramp procurement that may result in the target for

[6] California ISO, "CAISO Energy Markets Price Performance Report," September 23, 2019, p. 80.

[7] See California ISO, "Flexible Ramping Product Refinements, Draft Final Proposal," May 8, 2020, p.3..

[8] The Department of Market Monitoring pointed out in its 2016 and 2017 market reports that proxy demand response resources that could not be dispatched on a five minute basis were being included in the real-time dispatch and sent dispatch instructions that they could not respond to in that time frame. See California ISO Department of Market Monitoring, 2016 Annual Report on Market Issues and Performance, May 2017, pp. 259-262; and California ISO Department of Market Monitoring, 2017 Annual Report on Market Issues and Performance, June 2018, pp. 260. The CAISO subsequently observed in the price formation report that these resources were also being scheduled, and presumably paid, to provide flexiramp in RTD, despite being unable to respond to 5 minute dispatch instructions.

[9] See Guillermo Bautista-Alderete, "Price Performance in CAISO Energy Markets," Market Surveillance Meeting, October 11, 2019, p. 12; California ISO, "CAISO Energy Markets Price Performance Report," September 23, 2019, ,pp.79-80.

[10] California ISO, "CAISO Energy Markets Price Performance Report," September 23, 2019, pp.79-80.

[11] California ISO, Department of Market Monitoring, 2019, Annual Report on Market Issues and Performance, June 2020, pp. 113-115.

[12] See California ISO, "Flexible Ramping Product Refinements, Draft Final Proposal," May 8, 2020, p. 8; California ISO, Department of Market Monitoring, 2019, Annual Report on Market Issues and Performance, June 2020, p. 115.

flexiramp procurement being set higher or lower than would be efficient.[13] The second of these issues appears to be the scheduling of flexiramp at locations at which it cannot be dispatched due to transmission congestion, which is included in the second broad reason for zero RTD prices we suggested above. The scheduling of flexiramp will appear to be low cost at these locations, precisely because the resources scheduled to provide flexiramp are constrained down in the energy market dispatch due to transmission congestion and therefore cannot be used to meet load in HASP, FMM or RTD. However, this transmission congestion also means that these resources often cannot be dispatched to provide ramp to balance variations in net load when needed.[14]

There were indications of flexible capacity being stranded behind transmission constraints within the CAISO during 2012. But because similar patterns did not appear to be present in subsequent years, the initial California ISO flexiramp product design did not apply a deliverability test within the CAISO. Rather, the design allowed flexiramp scheduled in one balancing area to meet the target in another balancing area if the sink balancing area had enough unscheduled import capacity to accommodate the import of flexible capacity.[15] There have been a number of indications over the past several years that this design has been ineffective in providing reasonable assurance that flexiramp schedule to meet the flexiramp requirement of the CAISO or another EIM entity can actually be dispatched to meet variations in net load within the target balancing area. The CAISO has made a number of software and market design changes intended to correct this problem (summarized below) but it appears that none of changes has materially improved outcomes.

After the flexiramp product design was approved in early 2016, the expansion of the EIM caused the CAISO to change elements of the flexiramp design as part of the year 1 enhancements. These changes expanded the set of BAs from which flexiramp could be procured to meet load in CAISO, while eliminating the evaluation of flexiramp requirements for all combinations of balancing areas, instead only testing whether the sink balancing areas had sufficient import capability (on all ties in combination) to accommodate a specified level of flexiramp imports.[16] Unfortunately, the application of this test did not turn out as intended. In particular, it did not provide a reasonable likelihood that flexiramp capacity could be dispatched to balance net load variations in the CAISO because the total import capability of a balancing area would not be available to support imports from a particular EIM balancing areas.

[13] This issue is also to be addressed in the California ISO's flexible ramping product refinements initiative but is not discussed in this opinion.

[14] These resources could be dispatched to meet variations in net load within the transmission constrained region in which they are located but could not be dispatched to meet variations in net load outside this constrained region.

[15] See James Bushnell, Scott Harvey and Benjamin Hobbs, "Opinion on Flexible Ramping Product," January 20, 2016 p. 9 and California ISO, Department of Market Monitoring, 2012 Annual Report on Market Issues and Performance, April 2013, p. 84. California ISO, Business Requirements Specification, Flexible Ramping Product, Version 1.4, March 23, 2016, pp. 60-61, Appendix 3.

[16] See California ISO filing letter in docket ER15-1919 June 2, 2015 p.8 and California ISO, Energy Imbalance Market Year 1 Improvements, Draft Final Proposal, February 11, 2015 pp. 29-31.

Internal CAISO analyses led the CAISO to conclude during 2017 that the amount of imports and exports of flexiramp being scheduled in the FMM between balancing areas exceeded the available transfer capability.[17] In April 2019 the CAISO constrained the amount of flexiramp that could be scheduled within a balancing area based on the available transfer capability out of the balancing area.[18] However, this test only assured that the flexiramp could be delivered to one other balancing area, it did not assure that the flexiramp would be generally deliverable within the EIM.[19]

This outcome in which flexiramp was scheduled at locations from which it could be delivered to other EIM entities but not to the CAISO, or more generally to some EIM entities but not to others, may in part be an inevitable result of the expansion of the Western EIM to include additional balancing areas within which flexible capacity could be scheduled. As the number of balancing areas within the EIM increases, the probability that some balancing areas will be located in a generation pocket in which generation cannot be dispatched up to meet load in the CASIO would also increase.

In addition, however, some EIM balancing areas made very small amounts of transfer capability available to support exports to the CAISO during 2018 and 2019 while making more transfer capability available for transfers to other EIM balancing areas.[20] The small amounts of transfer capability made available by some of these balancing areas for deliveries to the CAISO may have been, at least in some cases, a result of the potential for market power mitigation to be triggered when an EIM BAA was import constrained, with the consequence that the offer prices for hydro generation resources would be mitigated to levels that would result in net exports being scheduled from a balancing area that wanted to import power to allow it to refill reservoirs.[21] These patterns in the transfer capability made available could result in outcomes in which substantial amounts of flexiramp could be scheduled within a group of balancing areas that could be dispatched to meet variations in net load within that group, passing the deliverability test in RTPD, but the flexiramp would not be dispatchable to meet load within the CAISO.

In any case, The CAISO's 2019 analysis found that despite the multiple software and market design changes intended to provide greater assurance that flexiramp is scheduled where it can be

[17] See, for example, Lin Xu, California ISO, "Discussion on flexible ramping product," Market Surveillance Committee Meeting, September 8, 2017 and Rahul Kalaskar, California ISO, "Flexible ramping product performance discussion," Market Surveillance Committee Meeting, February 2, 2018.

[18] See California ISO, "CAISO Energy Markets Price Performance Report," September 23, 2019, p. 78 and California ISO, Department of Market Monitoring, 2019, Annual Report on Market Issues and Performance, June 2020, p. 115.

[19] See California ISO, "CAISO Energy Markets Price Performance Report," September 23, 2019, pp. 78-79.

[20] See California ISO, Department of Market Monitoring, 2019, Annual Report on Market Issues and Performance, June 2020, pp. 128-129; California ISO, Department of Market Monitoring, 2018 Annual Report on Market Issues and Performance, May 2019, pp. 113-114.

[21] See the discussion in James Bushnell, Scott Harvey and Benjamin Hobbs, "Opinion on Local Market Power Mitigation Enhancements," March 6, 2019 and in CAISO filing letter in Docket ER19-2347 July 2, 2019.

dispatched to meet unpredictable variations in net load, in some months *a considerable proportion of the flexiramp-up capability scheduled and needed to meet variations in CAISO net load could not be delivered*.[22] In addition to flexiramp being scheduled in balancing areas from which it could not be dispatched to meet variations in net load within the CAISO, the CAISO's analysis appears to show that a material amount of flexiramp was scheduled within constrained-down regions within the CAISO from which it also could not be dispatched to meet variations in net load.[23]

The California ISO's price performance report showed that during 2017, 2018 and 2019 the effective requirement for flexiramp located within the CAISO BAA (the gross requirement for the CAISO BA less the portion of the requirement that could be met with flexiramp in other EIM balancing areas) was almost always zero in both FMM and RTD.[24] While the requirement to schedule flexiramp within the CAISO was almost always zero, the price performance analysis also showed that a substantial proportion of the gross CAISO flexiramp requirement was in practice scheduled within the CAISO in both FMM and RTD, although this was not required by the deliverability test.[25] In addition, the amount of flexiramp scheduled within the CAISO in FMM was considerably larger than indicated by the data in the price performance report because the price performance report data does not include the additional flexiramp that was scheduled in FMM as a result of operator adjustments to the load forecast that created additional undispatched capacity within the CAISO that was cleared as energy rather than as flexiramp. Operator upward adjustments to the load forecast for both binding and advisory intervals that were in excess of actual net load forecast error cleared energy in FMM that exceeded the actual load forecast that in effect scheduled additional capacity that could be dispatched to meet load at the distributed load bus. While these operator actions had an effect similar to scheduling additional flexiramp, they were not the same and could fail to have the intended effect, as well as raising costs as discussed below.

The CAISO's analysis in the price performance report suggests that deliverability constraints have been a critical factor limiting the effectiveness of the flexiramp dispatch, so implementation of the nodal deliverability test proposed in this initiative will be an important step towards improving the CAISO's ability to maintain reliability with its evolving resource mix. Nevertheless, precisely because the flexiramp design is a tool for managing real-time uncertainty, its real-time performance depends on many factors other than the performance of its deliverability test. Hence, we do not have assurance that this is the only remaining barrier to effective use of the design, and stakeholders should recognize the possibility that implementing

---

[22] See Guillermo Bautista-Alderete, "Price Performance in CAISO Energy Markets," Market Surveillance Meeting, October 11, 2019, pp. 10-11.

[23] California ISO, "CAISO Energy Markets Price Performance Report," September 23, 2019, p.80.

[24]See California ISO, "CAISO Energy Markets Price Performance Report," September 23, 2019, Figures 72 and 73, p. 74.

[25]Ibid.

these changes will provide some improvement in performance but also shed light on other limitations of the current design or data flow.

III. Operator Actions and the Market Software

The current ineffectiveness of the flexiramp product implementation has apparently caused operators to continue to take other actions in order to try to ensure that they generally have enough ramp capability available in real-time to balance variations in net load. These actions include committing long start units in RUC in the day-ahead time frame, and inflating the net load forecast used in HASP and FMM within the operating day. The extensive use of inflated load forecasts in HASP and FMM has the consequence of scheduling more imports in HASP and FMM, and committing more short-start generation, than would otherwise have been the case. Moreover, because the RUC, HASP and FMM software will model this extra load as being located at the distributed load reference bus, these operator actions have the impact of scheduling additional imports and committing more generation at locations where it can be dispatched to meet load at the CAISO reference bus.

While these out-of-market operator actions to increase the availability of ramp capability in real-time are likely necessary, at least in part, to meet reliability criteria and are therefore appropriate given the failed implementation of the flexiramp product design, the reliance on these actions has a number of limitations as a long run solution:

i. Ad hoc operator actions tend to result in over procurement and higher costs because of the lack of tools to fine tune and optimize operator procurement decisions;
ii. Out-of-market actions can be undone, at least in part, by market responses, requiring even larger out-of-market actions to achieve the intended impact;
iii. Operator actions to increase the supply of ramp that can be dispatched to meet CAISO load by increasing the load forecast, in effect purchases energy that is deliverable to the CAISO distributed load reference bus. But this is at best a high cost method of purchasing ramp capability and also may not operate as intended; and
iv. The scheduling of additional energy in FMM, rather than ramp capability, increases price volatility because the inflated load forecasts can result in power balance violations in FMM due to a lack of capacity while the scheduling of flexiamp would have resulted in a higher flexiramp and energy price but no power balance violation.

These four observations are discussed in greater detail below.

First, ad hoc operator actions are likely to be a high cost method of obtaining additional ramp capability for meeting variations in net load because it is unrealistic to expect operators to continually fine tune these ad hoc actions so as to balance the cost and value of additional ramp. Nor is it reasonable to expect them to make such cost effectiveness evaluations when relying on ad hoc adjustments. In addition, while increasing the load forecast for future intervals has effects that can be similar to the effect of scheduling more flexiramp, it needs to be understood that the effects are not necessarily the same. Modeling the flexiramp requirement in HASP and RTPD results in the software scheduling interchange and committing resources so that RTD **could** dispatch generation to meet load and maintain the target level of flexiramp in real-time. Raising the load forecast in HASP and FMM ensures that RTD **could** dispatch generation to meet that higher level of load over the RTD horizon, but will not necessarily provide a mix of generation nd interchange schedules that will enable RTD to meet a lower level of load while having a higher level of ramp available. The practical impact of this difference has not been studied and would be very complex to evaluate. Nevertheless, the potential for operator load forecast adjustments to result in a different set of interchange schedules and resource commitments than would be selected to provide a higher level of ramp capability is an inherent limitation of relying on ad hoc operator actions.

Second, ad hoc operator actions to increase ramp can lead to market responses that reduce the supply of ramp. For example, if operators increase the load forecast in HASP, causing more imports to be scheduled, these increased imports would be fixed in FMM. With increased price taking supply provided by hourly interchange schedules cleared in HASP, the FMM market would clear at a lower price, which could cause less 15 minute imports or more exports to be scheduled in FMM, offsetting the increased import supply in HASP.[26] This outcome can be avoided if the operators also increase the load forecast in FMM.

Third, if operators obtain more ramp capability by increasing the net load forecast in FMM as well as in HASP, they buy additional energy in the FMM that is deliverable to the distributed load reference bus. This is an effective method of purchasing ramp that is likely to be deliverable but it is more expensive to buy energy rather than ramp in FMM, and then sell the excess energy back (dispatching resources below their FMM schedule) when the extra ramp is not needed in RTD. In addition, this method of buying deliverable ramp would be more expensive than necessary to the extent that not all net load uncertainty is located at the distributed load reference bus within transmission constrained regions. If instead some the net load uncertainty is located outside these constrained regions, this net load uncertainty could be met with the output of lower cost resources. Moreover if the load forecast is increased in FMM but not RTD, some of the ramp that is counted in FMM will be phantom ramp that will not be available in RTD because resources will not be dispatched up and down to create the ramp in RTD. Furthermore, while the CAISO operators can in effect obtain more upward ramp-up by increasing the load forecast in FMM and acquire more downward ramp by decreasing the load forecast, the operators cannot do both at the same time. In addition, purchases of additional energy in FMM to provide ramp up can reduce the ramp down that is available in RTD and

[26] Thus, if the load forecast was increased in HASP so that hourly imports offered at $80 cleared in the market and were scheduled to flow, this fixed supply could cause the price in FMM to fall to $50, with fewer imports scheduled.

reducing purchases of energy in FMM to provide ramp-down can reduce the ramp-up that is available in RTD.[27]

Fourth, the scheduling of additional energy in FMM, rather than ramp capability, increases price volatility because it can result in power balance violations in FMM due to a lack of capacity while the scheduling of flexiramp would have resulted in a higher flexiramp and energy price but no power balance violation. If there is not enough capacity available to meet the inflated load forecast in FMM, this will result in a price set by the power balance violation penalty while the flexiramp design was intended to have the price gradually raise with the flexiramp penalty price as capacity became tight. The use of operator load forecast adjustments to manage ramp could therefore also result in more extreme prices than would be the case if the flexiramp constraint were operating as intended.

IV. Nodal deliverability

The proposed nodal deliverability test will add two transmission related constraints to the RTPD and RTD optimization. The FMM and RTD formulations currently impose a constraint that the energy dispatched to meet load in the FMM or RTD solution would not overload any pre- or post-contingency transmission constraint. The proposed additional constraints will test 1) whether resources scheduled to provide flexiramp-up can indeed be dispatched up to meet increased net load modeled as located at a combination of the distributed load bus and the location of variable energy resource output without overloading any pre- or post-contingency transmission constraint; and 2) whether resources scheduled to provide flexiramp-down can be dispatched down to offset reduced net load modeled as located at a combination of the distributed load bus and the location of variable energy resource output without overloading any pre- or post-contingency transmission constraint.[28]

The scheduling of energy and flexiramp in the FMM and RTD optimization in the nodal delivery design will therefore need to satisfy three distinct sets of transmission constraints. The CAISO proposes to reduce the solution time impact of enforcing these additional constraints by simplifying some elements of the transmission flow calculations. While these simplifications may in the abstract reduce the accuracy of the transmission flow calculations to some degree, small inaccuracies in these calculations would not be a consequential concern since the flow calculations are based on only rough assumptions regarding the location of potential changes in net load. The small inaccuracies in these transmission flow calculations will likely be dwarfed by the inaccuracies in the assumptions on which the calculations are based.

---

[27] For example, increasing the load forecast could cause more imports to be scheduled or another unit to be committed, allowing other resources to be dispatched down to provide more ramp up, but these actions would generally reduce the ramp down that is available.

[28] See George Angelidis, California ISO, "Flexible Ramping Product Refinements: Appendix B, Procurement and Deployment Scenarios, Draft Technical Description."

The goal of the nodal delivery test design would be to improve deliverability of flexiramp enough that operators would not need to routinely take out-of-market actions in order to increase ramp and the price of flexiramp would generally reflect system conditions. However, the flexiramp design has a number of elements that need to be optimized in order for the design to achieve its intended benefits. It is possible that improving the deliverability test will achieve some benefits but may also highlight other elements of the flexiramp design that will need to be improved before the design can achieve the goal of reducing the need for ad hoc operator actions to maintain the level of ramp capability that is required to meet reliability targets with the evolving resource mix. Some of these other elements would be the target procurement of flexiramp (which is being also being addressed in the flexiramp product refinements initiative), the relative level of the targets for ramp procurement in FMM and RTD, the targets set for future intervals of STUC and RTPD, and the offer price of resources scheduled to provide flexiramp, as well as issues that may only become apparent after these improvements are implemented.

As indicated by our recommendation on p. 1, above, there are three elements of the proposed nodal deliverability test design that should be discussed. First, the design and implementation of the nodal deliverability test needs to consider solution time impacts. Second, it needs to be understood that the nodal deliverability test will not test the actual deliverability of flexiramp-up or down, which will depend on the actual realization of variations in net load in each interval. It would be impossible to test the deliverability of flexiramp to meet all possible variations in net load, and even if it were possible, imposing such a constraint might yield an inefficiently high cost solution. Because there is no perfect solution, but some solutions could be much better than others, it is important that the CAISO try to assess the performance of the design it utilizes for modeling the location of net load uncertainty before implementing the design. Third, because there is no clearly defined best way to model the location of net load uncertainty, we think it is important that the CAISO's implementation of the nodal deliverability test be designed to accommodate ongoing adjustments to the way the location of variations in net load is modeled in applying the test.

These considerations are discussed further below.

Solution Time Impact

The driver for this initiative is the goal of improving the locational distribution of flexiramp to increase the likelihood that resources scheduled to provide flexiramp can generally be dispatched to balance variations in net load. However, more accurate modeling of flexiramp deliverability can compromise other CAISO goals by increasing the solution time of RTPD and RTD. Increases in RTD solution time in particular can hinder the ability of RTD to balance variations in net load by increasing the latency of the load forecast used in RTD. The CAISO nodal deliverability test therefore needs to designed with the goal of avoiding material solution time impacts, and the ultimate decision to implement the design must consider these effects. As the CAISO comes to understand the solution time impact of each element of the nodal delivery test it will need to consider the cost benefit trade-offs between increased RTD solution time and the associated increase in net load forecast latency against improved flexiramp deliverability. Simplifications may be desirable. For example, while it is symmetric to test nodal deliverability for both upward and downward flexiramp, the CAISO might evaluate the solution time impact of the downward flexibility test and consider whether the downward test is necessary or whether it is necessary in all hours.[29] The CAISO could develop the software with the ability to apply both the upward and downward nodal deliverability test but also with the ability for CAISO operators to only enable one or the other test during appropriate system conditions.

Nodal Delivery Test Accuracy

The proposed nodal delivery test should avoid concentrating the procurement of flexiramp where there is no possibility for it to be dispatched to meet load in CAISO or other EIM regions, but it needs to be recognized that it is uncertain how well the proposed nodal delivery test will work. It is not possible to get this modeling precisely right but it is important that the CAISO try to understand how well or poorly this design is likely to perform, and what adjustments in the basic design may be needed before the CAISO designs the software. There will potentially be complex interactions between the way net load uncertainty is modeled locationally, the penalty prices in the flexiramp demand curve, and energy market prices that will lead to unintended outcomes if they are not tested in advance.

Moreover it is important that stakeholders recognize the limitations of the nodal design in testing deliverability. First, the nodal deliverability test only tests the deliverability of the flexiramp procured to meet the gross flexiramp requirement for each balancing area; it does not test whether any flexiramp in excess of this target would be deliverable. The gross requirement for each balancing area is calculated to be net of the diversity benefit, so it presumes that additional flexiramp would be deliverable. Second, the nodal deliverability test only tests whether the flexiramp would be deliverable to this single configuration of increased net load within each balancing area. It does not test deliverability for other locational distributions of net load uncertainty.

The CAISO should not wait until the new dispatch engine is built to assess how well this nodal delivery test will perform. It would be desirable to test the performance of the proposed design for the nodal delivery calculations prior to both implementation and development of the software to identify potential issues. It will be difficult to fully test these calculations because current operations are substantially impacted by the ad hoc operator actions that are taken to increase ramp capability because of the deliverability issues with scheduled flexiramp. However, we think it would be a useful diagnostic test to apply the proposed nodal test to procurement of flexiramp in selected historic FMM and RTD intervals in which bottled flexiramp was identified

---

[29] It appears that with increased amounts of variable energy resources being offered at prices at or above -$50, there have been few downward power balance violations over the past few years. See California ISO, Department of Market Monitoring, 2019, Annual Report on Market Issues and Performance, June 2019, pp. 104-106 .

during ramp-related price spikes in the analysis the CAISO undertook during 2019,[30] and assess whether the design would have worked as intended during these intervals or whether there are additional factors that need to be considered. The CAISO could also reduce HASP imports in historical hours with large load biases and test if the flexiramp design procures flexiramp where it can be dispatched to balance net load was well as the out-of-market operator actions.

Design Flexibility

While testing of the nodal delivery design will help ensure that the design achieves some of the intended benefits, stakeholders should anticipate that there will likely be issues with the performance of the nodal test. Some elements of the design will likely be found to not be optimal and it is likely that changes will be needed over time even in elements that appear to perform well initially because the CAISO market, resource mix and net load variability will also be changing over time. Managing variable energy resources involves modeling uncertainty, this is very different from calculating a transmission flow for a known generator to meet a known load.

There is no single scenario that can optimally model deliverability for all potential patterns in which net load might rise or fall in future intervals. More of the fluctuation in net load could be in one LAP or another when there are major variations in net load. Further, the variations could also be concentrated at a small set of large wind or solar generation resources, rather than being spread over all the variable energy resources in the CAISO footprint.

It is not presently possible to test a material number of potential deployment scenarios for the deliverability of energy from capacity scheduled to provide flexiramp to balance potential variations in net load at a variety of distinct locations on the CAISO transmission grid in the market software without unworkable performance impacts. The CAISO proposes to test the deliverability of energy from flexiramp-up with one scenario and to test the ability of the CAISO to dispatch flexiramp-down with one scenario. It is therefore important to conduct off-line tests of how flexiramp procured to meet the two chosen scenarios will perform across a range of conditions and assess if adjustments to these scenarios could the robustness of the deliverability test to variations in actual system conditions.

It is also important to recognize that it is unrealistic to expect the CAISO to be able to choose the best modeling approach in its initial implementation of the nodal delivery test design, nor to expect that the optimal modeling approach will not change over time with changes in the resource mix and expansion in the scope of the Western EIM. We therefore think it is important that the CAISO attempt to be forward looking in developing the software to implement the nodal deliverability test and where workable, design key elements to be based on values referenced in tables that can be updated, rather than fixed logic built into the software. For example, PG&E

[30] See California ISO, "CAISO Energy Markets Price Performance Report," September 23, 2019, pp. 83-89.

pointed out that the CAISO should not model all net load variability as located at the distributed load reference bus, that some should be modeled at the location of variable energy resources. The CAISO has made such a change in the proposed design and is studying how net load uncertainty should be distributed between the distributed load reference bus and VER locations.

We agree with this design choice, but believe the CAISO should recognize that it may not choose the optimal value for distributing load between the distributed load reference bus and generation in its initial implementation. Therefore, the values used to distribute net load uncertainty in this manner should be a parameter than can be adjusted as needed to improve performance. We also suggest that the CAISO consider building in the flexibility to distribute net load uncertainty in different proportions at least to DLAPS, if not further to CLAPs, rather than proportionately to the distributed load reference bus. Similarly, we suggest that the CAISO consider building in the flexibility to distribute net load uncertainty in different proportions to the VER output in DAP or CLAP regions or to weather zones. The CAISO may want to parameterize other elements so they can be adjusted as needed.

Conclusion

If the nodal test performs roughly as intended, operators will need to take fewer actions to inflate HASP and FMM load in order to create ramp capability. Instead, these commitment decisions will be made by the market software, HASP and FMM, in order to meet the flexiramp target and these commitment decisions should be more consistent with market prices than is the case under the current design and operating practices.

These modeling issues are complex and the CAISO changes are unlikely to eliminate all of the deliverability issues. The system conditions that make ramp undeliverable might be most associated with system conditions in which net load is much higher in a few areas in the network but not in others, causing more congestion. Such deployment scenarios could in theory be included in addition to the proposed scenarios with more uniform system-wide increase or decrease in net load, but would significantly increase computational costs.

As we have explained above, however, it needs to be understood in assessing the CAISO's design proposals that it will be highly challenging to develop a deliverability test for capacity that would be dispatched for energy to balance unpredictable variations in net load, and to sett ramp targets for the amount of such capacity that would be cost effective to procure. These very difficult problems will require continuing modeling and design enhancements.

V. DMM concerns

DMM has expressed concerns in a number of reports and presentations that the current flexiramp design does not procure enough ramp capability to balance variations in net load over periods longer than the time from the initialization of the binding FMM run to the binding RTD interval.[31] While FMM procures the target amount of flexiramp for each future interval and enough energy to cover the then current load forecast for the interval, the amount of flexiramp procured for each interval is only intended to cover FMM binding interval to RTD net load forecast uncertainty, i.e., changes in the net load forecast over 40 minutes or so. There is a potential for larger changes in load forecast to occur over a period of several hours with a total change in the net load forecast that is much larger than the amount of flexiramp procured in FMM to meet net load uncertainty in the binding RTD interval.

The potential for these larger net load forecast errors is indirectly addressed by the current flexiramp design which updates the net load forecast in each run of STUC/HASP/RTPD, with the consequence that the net load modeled in these software engines gradually moves up or down with changes in the net load forecast. RTPD is therefore updating the net load forecast on a continuous basis and would commit resources to meet flexiramp targets if net load forecast rose above day-ahead market forecasts and schedules. These commitment decisions will be inadequate to meet ramp needs in future intervals only if the CAISO does not have enough resources that can be committed in the relevant time frame in which the change in the load forecast is identified. Because this net load uncertainty will be realized gradually over time rather than appearing all at once at 37.5 minutes prior to the RTD interval, the current design will potentially address these larger forecast errors through the routine operation of STUC, HASP, and RTPD. However, this is not assured if too much uncertainty is realized too close to real-time and not enough resources are available that can be committed within the required time frame.

The Department of Market Monitoring points to operator actions taken to inflate the load forecast used in HASP as evidence that the current flexiramp design does not take adequate account of uncertainty further out in time.[32] However, it appears to us that the timing of the operator actions identified by the Department of Market Monitoring is also consistent with these

[31] See Department of Market Monitoring, California ISO, “2019 Annual Report on Market Issues and Performance, June 2020, p. 117; Department of Market Monitoring, California ISO, “2018 Annual Report on Market Issues and Performance, June 2020, p. 270 , Ryan Kurlinski, Department of Market Monitoring, California ISO, “Enhancing the flexible ramping product to better address net load uncertainty,” Western EIM Body of State Regulators, webinar, June 12, 2020, Ryan Kurlinski, Department of Market Monitoring, California ISO, “Day-ahead market enhancements discussion, Market Surveillance Committee Meeting, July 30, 2020, California ISO, Department of Market Monitoring, “Comments on Flexible Ramping Product Refinements,” April 7, 2020. California ISO, Department of Market Monitoring, “Comments on Issue Paper on Extending the Day-Ahead Market to EIM Entities,” November 22, 2019.

[32]See Department of Market Monitoring, California ISO, “2019 Annual Report on Market Issues and Performance, June 2020, p. 117; Department of Market Monitoring, California ISO, “2018 Annual Report on Market Issues and Performance, June 2020, p. 270 , Ryan Kurlinski, Department of Market Monitoring, California ISO, “Enhancing the flexible ramping product to better address net load uncertainty,” Western EIM Body of State Regulators, webinar, June 12, 2020, Ryan Kurlinski, Department of Market Monitoring, California ISO, “Day-ahead market enhancements discussion, Market Surveillance Committee Meeting, July 30, 2020, California ISO, Department of Market Monitoring, “Comments on Flexible Ramping Product Refinements,” April 7, 2020. California ISO, Department of Market Monitoring, “Comments on Issue Paper on Extending the Day-Ahead Market to EIM Entities,” November 22, 2019.

actions being taken to address the known failure of the flexiramp design in RTD and FMM to provide ramp in the FMM and RTD time frame that is the focus of this initiative. It does not appear that any of the DMM analysis clearly shows that operators are taking actions to commit units out-of-market in the STUC time frame in order to acquire additional ramp from slow starting units, rather than simply inflating the load forecast in HASP and FMM to address the failures of the flexiramp implementation addressed by this initiative. Conversely, however, as we have pointed out above, it is not assured that addressing the various issues covered by this initiative will correct all of problems with the current implementation and we have not reviewed any California ISO analysis that would clarify the extent to which operators are taking actions in the STUC time frame to increase the ramp that will be available to cover load forecast error over the next several hours. Hence, we have little ability to provide an empirical assessment of the current situation.

When operators believe long-start units will need to be committed to provide ramp, it appears that this issue is currently addressed by operator commitments in RUC. The need for such long-start commitment in the day-ahead market to maintain availability of ramp capability will be incorporated in the day-ahead market with the enhancements to day-ahead market that are currently under development in the Day-Ahead Market Enhancements initiative. In particular, the procurement of flexible capacity (imbalance reserves) in the day-ahead market would support commitment of longer starting units in that time frame. As that initiative moves forward, thought will need to be given to how imbalance reserves scheduled in the day-ahead market will be accounted for in STUC unit commitment decisions for long-starting resources with day-ahead market schedules. It appears premature to develop a detailed design address these issues in the current flexiramp initiative since the design of the enhanced day-ahead market is still under discussion.

We agree with DMM that evaluation of ramp needs and commitment of resources to meet this need in real-time in a time frame longer than covered by RTPD may be needed. This need could be addressed by modifying the flexiramp requirements that are modeled in STUC. It should be noted that operator actions to increase the load forecast and to increase the flexiramp requirement will have very similar impacts in the STUC time frame. The main difference would be that increasing the load forecast will cause STUC to commit capacity that can be dispatched to meet load at the distributed load bus while the location of capacity committed to meet the flexiramp requirement will need to meet the nodal deliverability test.

There are a number of complex issues in using STUC to evaluate ramp needs in this longer time frame so addressing them will not be a simple design nor a simple software implementation.

Some of the issues are:

i. Interchange schedules and offers are not fixed in the time frame that STUC is run. Any evaluation of ramp needs more than 2 hours in the future will require assumptions about future import supply offers.
ii. Resource offer prices for both internal CAISO resources and EIM resources are not fixed for periods more than 2 hours in the future in the time frame that STUC is run.
iii. Current CAISO tools do not enforce energy limits over the duration of the operating day in real-time, potentially resulting in inefficient use of storage resources, as we discuss in our ESDER 4 opinion.[33] Extending the look-ahead period to cover the remainder of the operating day would address this limitation but would have serious performance impacts.
iv. The CAISO may need to develop some form of a longer look-ahead tool as the role of storage increases. But the design of such a tool, or even if it will be developed, is not known at present.

The CAISO has had considerable difficulty developing even reasonably workable flexiramp procurement targets for FMM and RTD, and this will be even more complex in the STUC time frame. This difficulty tends to argue against delaying the implementation of the current flexiramp improvements while attempting to develop a design for defining flexiramp targets even farther in advance of the operating hour. Moreover, it is not clear that it is possible for the CAISO to accurately identify the key issues that will exist with the improved flexiramp design until the current round of changes is implemented and we are able to observe congestion patterns and the level of flexiramp shadow prices in RTD under the nodal delivery design.

Another factor complicating assessment of the need for this and other potential improvements to the flexiramp design is that the current design between FMM and RTD still has the feature that a portion of the FRU "procured" in FMM can be "phantom ramp" that is not actually available for dispatch in RTD. The core issue is that the ramp calculated in FMM may reflect an out-of-merit dispatch in FMM to create additional ramp in the CAISO or elsewhere in the EIM. FMM schedules do not affect the dispatch of resources in the CAISO or EIM, the actual dispatch is determined in RTD. Because the FRU requirement is set much lower in RTD than in FMM, out-of-merit dispatch that is required to meet the FMM FRU target will likely not occur in RTD and the associated flexiramp up will not be available in RTD.

Given these considerations, and absent clear evidence that operators need to take a material amount of out-of-market actions to commit long start resources during the operating day in order to maintain the ramp capability to balance load and generation, as well as uncertainty as to how long it would take the CAISO to develop a design that would improve on these ad hoc operator actions to the extent they are occurring, we agree with the CAISO's design choice to focus initially on correcting the deliverability problems with the flexiramp design and on improving the forecast of near term net load uncertainty (another element of this initiative but not discussed in this opinion). A near-term focus on implementing the nodal deliverability test would not

[33] J. Bushnell, S. Harvey, and B. Hobbs, "Opinion on the ESDER 4 Initiative," Draft, August. 2020.

foreclose the CAISO from making use of the flexibility of the current design by enabling operators to make ad hoc adjustments to flexiramp targets for future hours.

VI. Flexiramp and Shortage Pricing

Another consequence of the inadequate modeling of flexiramp deliverability is that the implementation of the current flexiramp design does not produce FMM or RTD energy prices that reflect growing shortages of ramp when the system becomes tight on ramp capability. This appears to be the case because there is never a shortage of ramp in the FMM or RTD because there is always more than enough ramp capability somewhere in the EIM to meet the RTD target. The problem is that there is often a shortage of ramp that can be dispatched to meet CAISO load in RTD.

It was precisely the apparent lack of shortage prices reflecting a lack of ramp capability that led the MSC to conclude in early 2017 that flexiramp was not working as intended when we observed that the shadow price of ramp was consistently $0/MW in real-time, even during power balance violations. Although the CAISO has, as discussed above, corrected a number of various implementation flaws in the flexiramp design, these improvements have not changed the statistic that the shadow price of ramp is zero in more than 99% of all RTD intervals.

With the enforcement of the nodal deliverability test, the flexiramp demand curve will have the effect that the price of ramp should gradually rise as supply gets tight and less extra ramp is available. This tightening in the supply of ramp should be better reflected in the energy price in RTD as generation is dispatched out-of-merit to meet ramp targets, and higher cost generation is dispatched up to meet load. If the improvements operate as intended, there should also be fewer power balance violations in RTD.

The application of the nodal deliverability test will also impact the supply of flexiramp in FMM, however the net impact on FMM prices is more complex to assess than for RTD. This is because the reduced need for operators to inflate the FMM load forecast should result in a lower energy dispatch in FMM which would tend to reduce FMM energy prices and reduce the frequency of power balance violations in FMM, while the enforcement of the nodal delivery constraint on the procurement of flexiramp in FMM would tend to increase energy prices. However, the procurement of flexiramp in FMM will be defined by the flexiramp demand curve and decline when flexiramp is expensive, while inflated load forecasts in FMM are enforced that the power balance violation price. Hence, overall the FMM price should in the future provide a better price signal for increased net imports into the CAISO or other BAAs as ramp supply gets tight, reducing the need for ad hoc operator actions to maintain ramp by inflating the load forecasts used in HASP, FMM and RTD and result in a more measured increase in FMM prices as supply tightens.