

Frequency Response Phase 2

Issue Paper

December 15, 2016

Table of Contents

1.	Executive Summary 3
2.	Introduction
3.	Background5
3.1. F	requency Response Mechanisms
3.2. F	requency Response Requirements10
3.2.1.	NERC Frequency Response Standard10
3.2.2.	WECC Regional Criterion16
3.2.3.	FERC Notice of Proposed Rulemaking16
3.3. 1	SO Markets Role
3.3.1.	Market Mechanisms or Requirements18
3.3.2.	Settlement Mechanisms20
4.	Identified issues
4.1. C respor	Does the current ancillary services paradigm position the system to be able to sufficiently and to meet reliability requirements?
4.1.1.	Market design might not ensure sufficient frequency response
4.1.2.	Frequency response performance might not be sufficient
4.2. D resour	Does the current market design produce price signals that incent capital investments on ces to be capable of primary frequency response?
4.2.1.	Ensure no barriers to entry29
4.2.2.	Produce price signals to incent capability or provision
Append	ix A: Stakeholder Engagement Plan

1. Executive Summary

This initiative will complete the second phase of the ISO's frequency response initiative. In this second phase of the stakeholder initiative, the ISO seeks to examine a market structure for primary frequency response procurement and compensation. The goal of introducing a primary frequency service would be in the short-term to continue to support compliance with NERC's frequency response requirement, which, without changes, will be more difficult in the long-term as the generation mix changes to accommodate a renewable portfolio standard of 50% renewables by 2050.

This issue paper will describe market design limitations identified with the ISO's ability to (1) position its fleet to provide sufficient primary frequency response that maintains grid reliability during the largest contingency events and (2) incentivize and compensate resources for frequency response capability and provision.

The ISO asks stakeholders to consider whether the current market design or tariff requirements produces effective price signals and compensation.

Does the current ancillary services paradigm position the system to be able to sufficiently respond to meet reliability requirements?

Does the current market design produce price signals that incent capital investments on resources to be capable of primary frequency response?

The issue paper will focus on evaluating whether the ISO should compensate resources for any or all of the following:

- Capital expenses associated with having frequency response capability such as having equipment installed to provide frequency response service
- Opportunity costs associated with holding the frequency responsive capacity in reserve
- Operating expenses associated with providing frequency response when a frequency response event happens.

2. Introduction

On January 16, 2014 in Order No. 794,¹ FERC approved NERC Reliability Standard BAL-003-1, which placed a new frequency response requirement on balancing authority areas (BAAs) (including the ISO). BAL-003-1 requires balancing authorities (BAs) to demonstrate sufficient primary frequency response to disturbances in system frequency. Primary frequency response is a service that provides an actual response to a frequency change where additional power is provided to the grid to arrest and stabilize frequency within 52 seconds by automatic,

autonomous response either through control devices or ISO-signals based on an algorithm that matches product specifications.

The new standard obligates BAs to achieve a specific performance level called its Frequency Response Measure (FRM) that meets the required response called its Frequency Response Obligation (FRO). Each FRO is calculated as a proportion of the Interconnection Frequency Response Obligation (IFRO), which is the minimum required frequency response needed to arrest a decline in frequency from the loss of the two largest generators (i.e., Category D Event) before tripping load at the under frequency load shedding (UFLS) threshold of 59.5 Hz.

For the ISO, this is conservatively estimated as 30% of WECC's IFRO, based on the ISO's share of WECC's annual generation and load. NERC has calculated WECC's 2017 IFRO at 858 MW/0.1Hz. The ISO's estimated FRO is 258 MW/0.1Hz using the 30% share assumption is shown in Table 1 below. Directly following is the actual 2017 FRO based on 2014 annual generation and load amounts at 23%. Compliance with BAL-003-1 will be measured by converting actual response (MW) into MW/0.1Hz. Of course in practice the frequency response in MW will vary by event depending on the frequency change.

California ISO's Estimated Requirement	2017 FRO	Resources
Western IFRO	858	MW/0.1Hz
Estimated ISO FRO	258	MW/0.1Hz
Actual ISO FRO	196	MW/0.1Hz

Table 1: ISO FRO Estimate

Each balancing authority area must meet its BAA's FRO as of December 1, 2016 or risk being assessed penalties. The ISO initiated a stakeholder initiative in August 2015 to adopt short-term solutions to mitigate non-compliance risks.

Under this effort the ISO assessed its current frequency response capabilities and historical frequency response performance rates measured in MW/0.1Hz and compared response levels to the new NERC requirements. The analysis showed that the ISO could, at times, be short of its required share of frequency response. In response to this expected frequency response deficiency, the ISO filed tariff revisions to ensure compliance with BAL-003-1 for 2017 through procuring Transferred Frequency Response and strengthening requirements for conventional resources at FERC on April 21, 2016.² The Transferred Frequency Response is an annual contract to allow a transfer of frequency response performance between BAAs. FERC approved the filing September 16, 2016.³

In its filings before FERC, the ISO committed to evaluate whether a market mechanism should be designed that encourages frequency response capabilities of all participating resources, enables the diverse mix of resources to provide services, and ensures ISO meets applicable reliability criteria.

CAISO/M&IP/Cathleen Colbert

Phase 2 of this initiative will constitute the ISO's commitment to FERC. Here, the ISO will evaluate the need for and merit of introducing long-term market design measures that could incentivize sufficient capability and performance levels to maintain grid reliability, especially as the resource mix evolves to include diverse technology types and increased levels of non-conventional resources.

The ISO proposes the following market design principles be adopted in the design of a primary frequency response procurement mechanism. The market design should:

- Produce market outcomes that enable the ISO to position its fleet to respond sufficiently to frequency disturbances in the post-event measurement period;
- Allow all technology types to participate in the procurement mechanism through ensuring there are no barriers to entry;
- Produce price signals that incentivize capital investments on resources to be capable of primary frequency response; and
- Ensure compensation of capital investments made to meet the required capability if frequency response capabilities become an interconnection requirement.

3. Background

The purpose of the background section is to provide context to support the evaluation of a primary frequency response procurement and compensation mechanism.

The remainder of the section will discuss:

- Frequency Response Mechanisms: Present the reliability need underpinning frequency response service. The reliability need is to control frequency to stable levels to support a well-functioning grid. Of note, the frequency response product under consideration in this initiative is tied to the time horizon based on NERC's definition of its post-event measurement period from 20 to 52 seconds after the event not to NERC's definition of primary frequency response.
- Frequency Response Requirements: Describe how NERC's Frequency Response and Frequency Bias Setting Standard, BAL-003-1 approved on January 16, 2014, imposed additional requirements to demonstrate sufficient frequency response on the system as the BAA. The requirements are designed to avoid hitting this UFLS during the loss of the two largest resources in the Western Interconnection. Of note, the post-event measurement period between 20 to 52 seconds includes performance from both primary and secondary response types.
- ISO Markets Role: Provide high level background on its market mechanisms, the product or services that they procure, and how the settlement mechanisms function to

compensate for this procurement. Additionally, this section will discuss how requirements are used to gain primary frequency response instead of a procurement mechanism.

3.1. Frequency Response Mechanisms

The purpose of this section is to present the reliability need underpinning frequency response service. The reliability need is to control frequency to stable levels to support a well-functioning grid. An interconnection needs to have a system frequency that is on average near the scheduled frequency value at 60 Hz. If frequency increases far above the scheduled value due to over-generation relative to demand it can lead to grid instability. If frequency decreases well below the scheduled value due to insufficient generation relative to demand it can lead to grid instability. The under-frequency events introduce a high grid reliability risk since if an under-frequency event persists it could cause cascading black outs.

NERC divides the frequency control mechanisms into three separate response types.⁴ The three frequency response types relative to the timing during a frequency excursion event.

The three response types are defined as:5

- Primary: actions provided by the Interconnection to arrest and stabilize frequency in response to frequency deviations. Primary control comes automatic generator response, load response (typically from motors), and other devices that provide an immediate response based on local (device-level) control systems.
- 2) Secondary: Actions provided by an individual Balancing Authority or its Reserve Sharing Group to correct the resource-load unbalance that created the original frequency deviation, which will restore both Scheduled Frequency and Primary frequency response. Secondary Control comes from either manual or automated dispatch from a centralized control system.
- 3) Tertiary: Actions provided by Balancing Authorities on a balanced basis that are coordinated so there is a net-zero effect on area control error (ACE). Examples of Tertiary Control include dispatching generation to serve native load, economic dispatch, dispatching generation to affect interchange, and re-dispatching generation. Tertiary Control actions are intended to replace Secondary Control Response by reconfiguring reserves.

Figure 1 below shows an illustration of an under-frequency event where the horizontal axis shows the minutes before and after the event and the arrows show where in time the response types would be expected to deploy. As can be seen, primary response type is generally deployed seconds following and maintained through 10 minutes after the event or when frequency has been fully restored. Secondary response type is generally observed 30 seconds following up to 30 minutes after the event. Finally, tertiary would replace secondary response so would be observed beginning around 10 minutes after until tertiary response is concluded.



Figure 1: Frequency Response Types

Expanding on this relationship, Figure 2 shows the post-event measurement period, Value B, shown in the green area in more granularity. This is the time period from 20 to 52 seconds after the event for which the CAISO performance will be measured for sufficient frequency response as defined by NERC. This chart introduces arrows that illustrate at what time during the first minute after the event that primary frequency response and depending on event location some types of secondary frequency response begin to appear on the grid.

Figure 2 shows the pre-event period (Value A) from 16 seconds prior to the contingency event at T_0 shown in the blue shaded area where the system is functioning near scheduled frequency. Once the contingency event occurs frequency begins to drop and resistance is provided solely through inertia and load damping until it is arrested at the Nadir point (Point C, arrested point) through additional power output.

When primary frequency response initially begins it can be sluggish in speed and vary in triggering and deploying the automatic response depending on the control technology used. During this period, called the transient period shown in yellow shaded area, the frequency fluctuates but continues to move in a direction that balances frequency. Importantly it shows the post-event measurement period from 20 to 52 seconds following the frequency event in the green shaded area.



Figure 2: Frequency Response Services (60 seconds after event)

Following the event through the post-event measurement period, Figure 2 shows that the response needs to arrest the frequency decay providing the additional power output that stops the frequency deviation and continue to provide additional power output until the stabilizing period. Equation 1 below shows the formula for a frequency excursion's slope and illustrates the role of inertial response and load damping to control the steepness of the curve. The speed at which a frequency response drops from the pre-event value at Value A before the drop is stopped at the arrested point at Point C is a function of size of the loss of generation, inertia constant of the system and the load damping.⁶

Equation 1: Slope of Frequency Excursion

	$Slope = \frac{\Delta P}{D + 2H}$
	Where: $\Delta P = P_{t+1} - P_t$ and $t = T_0$
Slope	Frequency Excursion Slope.
ΔP	Magnitude of the imbalance.

D	Load damping factor ranging from 0 to 2 where 2 represents all-motor load.
Н	Inertia Constant of system ranging from 2.5 to 6.5.
T_0	First SCADA scan showing change in frequency

Prior to deployment of primary frequency response, inertial response and load damping is provided to flatten the frequency excursion slope so that the deviation's slope is limited in its extremity until the frequency response is provided. Inertial response is provided automatically by large motorized loads and conventional resources with rotating mass, called synchronous inertial response. Theoretically, the more inertia on the system the quicker the arrest will occur thus requiring less primary response to restore frequency.

The concept of adding another type of frequency response has been raised in market design efforts in ERCOT as well as academic studies on renewable energy technologies. Fast frequency response could be provided by inertia-less technologies that can be responsive sooner than conventional resources – a matter of cycles after an event. Fast frequency response service could complement synchronous inertial response in minimizing the frequency excursion.⁷ Fast frequency response would act within the first few seconds to reduce the overall amount of frequency response needed to restore the system.

Primary frequency response begins several seconds after the event when generators or responsive demand with control devices read the frequency deviation on the system outsides their deadband (shown by the red arrows). Currently, the fast frequency response would fall under this type. The stabilized frequency in the post-event period remains less than the scheduled value but is at a stable and recovering level that will be fully restored through secondary and tertiary response.

Secondary and tertiary response services are provided in the ISO market by regulation and operating reserves in the ISO market as shown by the orange arrows and this provision will overlap with primary response as shown in the first chart. Following the deployment of primary response roughly 20 or 30 seconds after the event, secondary response begins to deploy through the actions of the grid operator sending a signal to adjust power output to continue to recover to near scheduled levels.

The concept of adding another type of frequency response, fast secondary response, has been raised and implemented in other market design. PJM and ERCOT have such products called fast regulation that receive a second signal that is tuned to a faster deployment of their obligation. For example, ERCOT defines its fast regulation as:

"A subset of Reg-Up in which the participating Resource provides Reg-Up capacity to ERCOT within 1 second of either its receipt of an ERCOT Dispatch Instruction or its detection of a trigger frequency independent of an ERCOT Dispatch Instruction. Except where otherwise specified, all requirements that apply to Reg-Up also apply to FRRS-Up."⁸

At around 10 minutes tertiary response will replace the secondary response if not fully restored. The primary frequency response type from any unloaded frequency responsive capacity continues beyond the post-event period (Value B) until frequency levels move to within a deadband (i.e. less than 59.964). If these actions are still insufficient to recover frequency and the frequency drops below the under frequency load shedding (UFLS) point at 59.95 Hz (shown as a black dashed line at the base of the chart) then it can shed load to restore frequency.

3.2. Frequency Response Requirements

The purpose of this section is to introduce the federal or region level requirements that NERC, WECC, or FERC has either imposed or is considering to impose on the ISO or generators to support frequency response.

The remainder of this section will discuss:

- NERC Frequency Response Standard
- WECC Regional Criterion
- FERC Notice of Proposed Rulemaking

3.2.1. NERC Frequency Response Standard

The purpose of this section is to describe the first requirement of NERC's frequency response and frequency bias setting standard, BAL-003-1.1, so that stakeholders and the ISO can be aligned on the obligation.⁹

To achieve this the ISO will discuss in the remainder of this section the following:

- Determining IFRO and FRO
- Measuring BAA performance
- Results of IFRO and FRO by year

3.2.1.1. Determining IFRO and FRO

The ISO's new frequency response obligation is the ISO's portion of a new obligation for the entire Western Interconnect, the Interconnection Frequency Response Obligation (IFRO). Federal Energy Regulatory Commission approved the Petition of the North American Electric Reliability Corporation (NERC) for approval of BAL-003-1, Frequency Response and Frequency Bias Setting on March 29, 2013 where this new standard requires the ISO to meet a BAA level Frequency Response Obligation (FRO) effective April 1, 2016.

The primary objective of Requirement 1 (R1) from BAL-003-1 is to determine whether a BA has sufficient frequency response for reliable operations.¹⁰ The standard requires each interconnection to maintain a minimum amount of frequency response. The IFRO is the

CAISO/M&IP/Cathleen Colbert

minimum MW/0.1 Hz frequency response to protect against a loss of generation event that would cause the system frequency encroaching on the trip setting of the Under Frequency Load Shedding (UFLS) relays. An important goal of this standard is to avoid tripping the first block of the UFLS relays for a frequency event resulting in the actual system frequency dipping to 59.5 Hz.¹¹

In the following equation, Equation 2, the formula for determining the IFRO is shown. It shows that the IFROs are determined by selected a protection criteria and distributes it across an estimated maximum delta frequency. If the frequency excursion's slope stays above the delta frequency the BAA can hold a 95% confidence level that the Interconnection frequency should stay above its UFLS threshold.

First on the protection criteria, FERC approved NERC's recommended target resource loss protection criteria (RLPC) reflecting the simultaneous loss of resources without system adjustments for the largest reasonably expected contingency. The Western Interconnection must protect against the Category D event, which is the loss of 2 Palo Verde resources (2,626 MW). The Interconnection will be credited for 120 MW of load that trips by a Remedial Action Scheme (RAS) following this Category D event. The Adjusted Resource Loss Protection Criteria (ARLPC) is 2,506 MW.

Next in setting the estimated maximum delta frequency, NERC is attempting to capture a maximum delta frequency that would support reliability at the 95% confidence level. NERC's adjustments used in determining the MDF that estimate the statistical "C-to-C" adjustment and the statistical "B-to-C" adjustments is effectively determining Point C by using phasor measurement resources (PMU) sub-second frequency data and makes an adjustment to compensate for slower EMS scans of Point C in the calculation of the calculate an IFRO.

Equation 2: Interconnection Frequency Response Obligation

$$DF_{BASE} = F_{START} - UFLS$$
$$DF_{CC} = DF_{BASE} - CC_{ADJ}$$
$$DF_{CBR} = \frac{DF_{CC}}{CB_R}$$
$$MDF = DF_{CBR} - BC'_{ADJ}$$
$$ARLPC = RLPC - CLR$$
$$IFRO = \frac{ARLPC}{MDF}$$

 F_{START} Starting Frequency is the 5% of the lower tail samples over three-year window representing 95% change frequency will be at or above value at start of an event.

UFLS Prevailing UFLS First Step

DF _{BASE}	Base Delta Frequency from F _{START} to UFLS First Step
CC _{ADJ}	95% confidence interval adjustment for differences between Point C when comparing 1-second and sub-second measurements of Point C.
DF _{CC}	Delta frequency adjusted for differences between 1-second and sub-second Point C observations
CB _R	Statistically determined ratio of arrested frequency response (Point C) to settled frequency response (Value B) where CB_R is defined as expected Value A – Point C / Value A – Value B plus a 95% confidence adjustment.
DF _{CBR}	Delta frequency adjusted for the ratio of Point C to Value B
BC' _{ADJ}	Statistically determined adjustment in the event nadir occurs below settled frequency response only applying to the Eastern Interconnect.
MDF	Maximum allowable delta frequency
RLPC	Resource Loss Protection Criteria (Largest G-2 Event)
CLR	Credit for load resources
ARLPC	Adjusted Resource Loss Protection Criteria
IFRO	Interconnection Frequency Response Obligation

The IFRO value will change from year to year primarily as the result of the estimated maximum delta frequency estimate changing due to changes in statistically derived factors used to determine the MDF. The three varying factors are:

- Statistical frequency variability over a three-year window of 1-second frequency measurements affecting F_{START}.
- Statistical "C-to-C" adjustment over a three-year window comparing 1-second and subsecond measurements of Point C.
- Statistical "C-to-B" ratio adjustment over a three-year window comparing the Value A Point C and Value A – Value B ratio.

Finally Equation 3 shows the approach used to allocate the IFRO to the BAA-level based on the BA's share of Interconnection annual load¹² and annual generation¹³ from the most recently reported FERC Form 714 two year prior values or other representative data. FERC Form 714 is the Annual Electric Balancing Authority Area and Planning Area Report in which the ISO reports its BAA net energy for load and peak demand sources by month (Part II - Schedule C). These

values are reported for each month and an annual total. The IFRO allocation will be determined using the annual total reported on these forms or a similar calculation for BAAs not reporting with FERC Form 714.¹⁴

Equation	3:	BAA-level	Frequency	Response	Obligation
----------	----	------------------	-----------	----------	------------

	$FRO_{BA} = IFRO \times \frac{Annual \ Gen_{BA} + Annual \ Load_{BA}}{Annual \ Gen_{INT} + Annual \ Load_{INT}}$
FRO _{BA}	BAA frequency response obligation.
IFRO	Interconnection frequency response.
Annual Gen _{BA}	BAA net generation (MWH) annual total reported FERC Form 714, Part II – Schedule 3, Column C, line 13.
Annual Load _{BA}	BAA net energy for load (MWH) annual total reported FERC Form 714, Part II – Schedule 3, Column E, line 13 which is the sum of BAA net generation (MWH) and net h interchange (MWH).
Annual Gen _{INT}	The sum of all net generation (MWH) across all BAA in the Interconnection.
Annual Load _{INT}	The sum of all net energy for load (MWH) across all BAA in the Interconnection.

3.2.1.2. Measuring BAA performance

NERC determined that each BAA would report a sampling of events and the median frequency response reported in MW/0.1Hz would be used to estimate its performance on average over the year. The median frequency response value called the Frequency Response Measure (FRM), calculated in resources of MW/0.1Hz is the change in net actual interchange on the intertie lines between the pre-event period (Value A) and the stabilizing period after the event (Value B) per 0.1Hz drop of the frequency event measured between those two points. The ISO will submit its 2017 Frequency Response Measure (FRM) by March 7, 2018 and will be assessed for compliance based on the percentage difference between the FRM and FRO.¹⁵

NERC performed a field trial of SEFRD data to come to the recommendation for a sampling approach to the compliance of this standard instead of a single-event based compliance measure. The analysis showed a single-event based compliance measure is unsuitable for compliance evaluation when data has large degree of variability.¹⁶ The analysis further demonstrated that a sampling of at least 20 events is sufficient to stabilize the results and alleviate the problem associated with outliers in the measurement of BAA frequency response performance.¹⁷

One of the reasons NERC proposed using the median of a sampling of events when it established the standard because frequency response based on net actual interchange does not directly measure the additional power provided to the grid as flows may vary for many other reasons including an increase of internal demand requiring greater flows into the BAA instead of transfers in the export direction that would provide replacement power. Another reason is the challenges in capturing accurate intertie flow data depending on the type of telemetry technology used. The median measurement approach would reduce the influence of performance outliers and noise introduced by the various other factors to net actual interchange.

The median value of SEFRD data for actual frequency response performance described above expressed in MW/0.1Hz (FRM) will be determined from a

NERC will identify the sampling of 20-35 frequency excursion events, 2 or 3 events sampled from each calendar month,¹⁸ if the events fit the following criteria:¹⁹

- Value A should be relatively steady around the scheduled frequency of 60.000Hz.
- The change in frequency as defined by the difference between Pre-event period, Value A, and arrested frequency Point C is greater than 0.04 Hz and Point C for a frequency dip is less than 59.95 Hz.
- Typically, the time from the start of the rapid change in frequency until the point the frequency stabilizes within a narrow range should be less than 18 seconds.
- If any data point in Value B period recovers to the Value A level, the event will not be considered as a candidate.
- Events that include 2 or more events not stabilizing within 18 seconds, during large ramping or load changes, within 5 minutes of the top of the hour will be excluded from consideration if other acceptable events are available.

If 20 events cannot be identified during the period, similar acceptable events from the next year will be used and the compliance period will be extended to 24 months. NERC provides the sampling of events to each BAA on its Frequency Response Survey (FRS) Form1.²⁰ The ISO will submit to NERC its performance measurements for each event in FRS Form 2.²¹

For each event selected, the ISO must report its frequency response performance for each event on its Frequency Response Standard Forms 1 and 2. The performance is determined by calculating the change of a BA's net actual interchange on its tie lines with its adjacent balancing authorities adjusted for the loss of generation if the contingency occurred within the ISO BAA divided by the change in interconnection frequency using the SEFRD data.²² The performance will be measured by comparing the average of discrete scans of net actual interchange (Ni_A) across the defined sampling periods. The sampling periods define the windows for calculating the average of each SCADA scan for Value A and Value B as follows:

• Value A is the average of scans for 16 seconds prior to an event

• Value B is the average of scans for the 20 to 52 seconds after an event

The 20 – 52 second period following an event was selected because it measures response after transient period completely settles and show squelched response during the recovery period and would allow some secondary response from regulation services to improve compliance.²³

3.2.1.3. Results of IFRO and FRO by year

The ISO expects NERC to file an informational report with each year's IFRO by November effective December 1,²⁴ the updated Frequency Response Annual Analysis. NERC updates its statistical analysis determining three factors that contributes to the estimation of the MDF. The statistical analysis will be performed over a three year period prior to the upcoming compliance period. For example, the statistical analysis for the 2014 frequency response period (December 2013 – November 2014) uses data points across 2010-2012.

The NERC informational report serves as the vehicle for communicating recommended IFROs for a compliance period to FERC. The following are the formulae comprising the calculation of the IFRO and a table illustrating the results for the Western Interconnection for the 2013 – 2017 periods from those informational reports.²⁵ Note that the 2016 values will be used to set the 2017 compliance period's IFRO presented in the Introduction Section, Table 1.

Explana	ation of Calculations	2013 ²⁶	2014 ²⁷	2015 ²⁸	2016 ²⁹	2017	Units
F _{START}	Starting Frequency is the 5% of the lower tail samples over three-year window representing 95% change frequency will be at or above value at start of an event.	59.976	59.971	59.968	59.967	59.967	Hz
UFLS	Prevailing UFLS First Step	59.500	59.500	59.500	59.500	59.500	Hz
DF _{BASE}	Base Delta Frequency from F _{START} to UFLS First Step	0.476	0.471	0.468	0.467	0.467	Hz
CC _{ADJ}	95% confidence interval adjustment for differences between Point C when comparing 1-second and sub- second measurements of Point C.	0.004	0.008	0.011	0.000	0.000	Hz
DF _{CC}	Delta frequency adjusted for differences between 1-second and sub-second Point C observations	0.472	0.463	0.457	0.467	0.467	Hz
CB _R	Statistically determined ratio of arrested frequency response (Point C) to settled frequency response (Value B) where CB _R	1.625	1.774	1.672	1.598	1.566	Hz

Table 2: NERC recommended IFRO for the 2013 – 2017 evaluation periods

	is defined as expected Value A – Point C / Value A – Value B plus a 95% confidence adjustment.						
DF _{CBR}	Delta frequency adjusted for the ratio of Point C to Value B	0.291	0.261	0.273	0.292	0.298	Hz
BC' _{ADJ}	Statistically determined adjustment in the event nadir occurs below settled frequency response only applying to the Eastern Interconnect.	N/A	N/A	N/A	N/A	N/A	Hz
MDF	Maximum allowable delta frequency	0.291	0.261	0.273	0.292	0.298	Hz
RLPC	Resource Loss Protection Criteria (Largest G-2 Event)	2,740	2,626	2,626	2,626	2,626	MW
CLR	Credit for load resources	300	150	150	120	120	MW
ARLPC	Adjusted Resource Loss Protection Criteria	2,440	2,476	2,440	2,506	2,506	MW
IFRO	Interconnection Frequency Response Obligation	-840	-949	-906	-858	-841	MW/ 0.1Hz

3.2.2. WECC Regional Criterion

The ISO Tariff requires Participating Generators to follow the WECC regional criterion for generators with active governor controls that specifies minimum values for droop and deadband settings on all resources providing spinning reserve to aid in frequency response. Initially this requirement only applied to generators or demand awarded a spinning reserves obligation but was extended to all participating generators under Phase 1 in support and recognition of the WECC Regional Criterion for generators with active governor controls.

Further in the prior phase, the ISO also strengthened these Tariff requirements to align with NERC reliability guideline on primary frequency response specifications necessary to follow good utility practices that support grid reliability. For example, Participating Generators should ensure plant-level and governor-level controls are coordinated and frequency response is enabled at the highest level of governor controls.³⁰

3.2.3. FERC Notice of Proposed Rulemaking

The Notice of Proposed Rulemaking (NOPR)³¹ FERC issued on November 17, 2016 is the result of a Notice of Inquiry (NOI)³² FERC issued on February 8, 2016 that sought comment on the need for reforms to its rules and regulations regarding the provision and compensation of primary frequency response. The NOI examined the following issues:

- Whether newly-connecting resources should have the capability to provide frequency response;
- Whether existing resources should be required to ensure frequency response capability;

- What standards, if any, should be required for frequency responsive resources;
- To what extent the provision of frequency response should be mandatory; and
- What form of compensation mechanisms might be appropriate, if any, for those resources providing frequency response?

The NOI led to FERC issuing the NOPR. In the NOPR, FERC seeks to effect the following proposed policy changes:

- Require all (synchronous and non-synchronous) newly interconnecting large and small generating facilities, as well as all existing large and small generating facilities that take any action that requires the submission of a new interconnection request that results in the filing of an executed or unexecuted interconnection agreement, to adhere to the proposed frequency response operating requirements, on or after the effective date of any Final Rule issued in the proceeding;
- Require newly interconnecting resources to have droop and deadband settings of 5 percent and ±0.036 Hz respectively;
- Prohibit all newly connecting large and small generating facilities from taking any action that would inhibit the provision of primary frequency response, except under certain conditions;
- Require droop settings to be based on nameplate capability with a linear operating range of 59 to 61 Hz; and
- Exempt new nuclear resources from these requirements.

Moreover, FERC explicitly noted that it would *not* (a) imposed frequency response requirements on existing units, (b) impose a generic headroom requirement on resources, or (c) mandate or prohibit compensation related to frequency response. FERC clarifies that nothing in the rulemaking is meant to prohibit a public utility from filing a proposal for primary frequency response compensation under Federal Power Act (FPA) section 205, if it so chooses.

Further, FERC is requesting comment on whether additional primary frequency response performance or capability requirements for existing resources are needed, and if so, whether the Commission should impose those requirements by either (1) directing the development or modification of a reliability standard pursuant to section 215(d)(5) of the FPA or (2) acting pursuant to section 206 of the FPA to require changes to the pro forma open access transmission tariff.

Comments on FERC's NOPR are due January 24, 2017. The ISO plans to participate in the NOPR process. Once a final rule is issued, it will evaluate any impacts to the design effort in this initiative at that time.

3.3.ISO Markets Role

The purpose of this section is to describe the ISO's market mechanisms for various services and their compensation mechanisms. Additionally, in the section discussing market mechanisms the reliance on primary frequency response requirements to provide this service will be discussed. It is notable that for the other frequency response types – secondary and tertiary – the ISO procures these services but does not for the primary frequency response type.

To provide this background information, the remainder of this section will discuss:

- Market Mechanisms or Requirements
- Settlement Mechanisms

3.3.1. Market Mechanisms or Requirements

This section provides an overview of the competitive market mechanisms the ISO operates to facilitate the exchange of energy to serve demand and procurement of grid services to support grid reliability. The ISO believes the procurement of grid services such as frequency response capabilities and provision would be done through the ISO's market mechanisms to support grid reliability not through bilateral capacity markets for sufficient capacity to deliver energy.

To participate in the energy and ancillary service markets a market participant would submit either a supply offer representing its willingness to provide power or an ancillary service offer representing the service it is willing to provide at a given price. The market finds the optimal solution through an energy and cascaded AS procurement in a co-optimization process between ancillary service and energy offers.

A cascaded ancillary service procurement allows substituting a higher quality ancillary service for a lower quality ancillary service to ensure sufficient operating reserves are available since this substitution results in the least-cost, security constrained solution and with a higher reliability valued ancillary service obligation. An optimal solution.

For energy, the supply offer includes up to four components that represent the total production cost of the resource:

- startup costs associated with bringing a resource online from being shut down³³,
- transition costs associated with moving from one configuration to another for multi-stage suppliers (MSG),
- minimum load costs associated with operating the resource at the minimum operating level (Pmin) where a resource cannot drop below without compromising the resource's operation, and
- incremental energy costs associated with producing energy above Pmin.

For ancillary services, the ISO procures ancillary services for its frequency response services in the secondary and tertiary response types to meet its minimum reserves procurement requirements. For the primary response type, the ISO currently relies on primary frequency response requirements on resources with active governor controls to meet droop setting, deadband and control coordination tariff requirements.

Table 3 lists descriptions of the ISO frequency control ancillary services ranked in order of the highest reliability value to the lowest reliability value from primary frequency response to non-spinning reserves.

Requirement or Product Definition	Synchr onized to grid	Deployment Method	Deploy Deadline	Full AS Delivery Deadline	Sustain Provision
Primary Frequency Response Requirements	Yes	Frequency event triggers automatic, autonomous governor or under- frequency relay device response	Frequency change exceeds 0.0036 0.1Hz	N/A	N/A
Regulation (Up and Down)	Yes	AGC signal sends 4 second signals where resource will provide its entire AS obligation by the full delivery deadline based on its ramp rate.	4 seconds to meet the most recently sent signal	10 minutes	N/A
Spinning Reserves	Yes	Contingency Dispatch Operating Target (DOT) Replaces Real-time Market DOT	Immediately	10 minutes	2 hours
Non-Spinning Reserves	No	Contingency Dispatch Operating Target (DOT) Replaces Real-time Market DOT	Immediately	10 minutes	2 hours

Table 3: Current Ancillary Service Requirement or Products³⁴

As shown in the column titled 'Deployment Method' depending on the service the ISO will deploy the ancillary service differently.

For the primary frequency response service, the ISO does not deploy this service but instead relies on automatic, autonomous response directly at the generating resource after reading a frequency change of 36 mHz.

For the secondary frequency response service procured through the regulation products, when the ISO's area control error equation fluctuates to signal an under-frequency event, the ISO sends automatic generation control (AGC) signals to the generating resources with regulation up obligations. These signals are updated every four seconds moving the regulation capacity up based largely on the resource's ramp rate and tuned to deliver its full regulation award to the grid by 10 minutes.

For the tertiary frequency response service procured through the ISO spinning reserves and non-spinning reserves products, the ISO determines whether this obligation should be deployed using the real-time contingency dispatch (RTCD) run. This RTCD runs outside of the normal market runs in response to a significant contingency event when Operations expects the Operating Reserves are likely needed to resolve the system condition³⁵. The RTCD sends generating resources a set of dispatch operating targets for a single 10-minute interval after the real-time market run. These contingency dispatches must be responded to as soon as possible.

3.3.2. Settlement Mechanisms

The California ISO settles a resource's market award so that they are compensated at the price submitted in their supply offer through their market revenues and uplift payments. For the incremental energy produced, the supplier will receive payment for this energy at the energy price (market revenues). To the extent a resource has a market revenue shortfall where its market revenues do not exceed its supply offer, the resource is compensated for the difference between its supply offer and its market revenues through uplift mechanisms.

For settling energy awards, the market will produce prices that reflect the marginal cost of serving an additional resource of demand, which generally is set by the energy offers and does not include commitment cost offers. Energy prices are not intended to reflect the impact of start-up costs, transition costs, no load or minimum load costs. As noted above, these costs will influence which resources are committed so there is an indirect impact to the energy price. Instead, the energy price is intended to reflect the marginal cost of energy given commitment decisions.

For settling ancillary service awards, the market will produce prices that generally reflect the opportunity cost to the generator of the ISO procuring the ancillary service instead of procuring that capacity for energy. The market values the ancillary service price as the combined reduction of the total procurement costs of the combined energy and ancillary service market due to relaxing the ancillary service constraints. The price will represent the combined reduction from the expanded system, system, and region constraints depending on which of these constraints are binding. Note, in this event there are insufficient ancillary service offers to meet the minimum

reserves requirement, the ISO triggers an ancillary services scarcity event and reserves will be settled based on a scarcity reserve demand curve instead³⁶.

When the market revenues from both the energy and ancillary service awards are insufficient to cover the costs submitted in the supply and ancillary service offers, the ISO will provide the remainder of the resource's compensation through uplift mechanisms. Uplift mechanisms provide make-whole payments to suppliers who had a market revenue shortfall. The California ISO will generally pay this make-whole payment either through its bid cost recovery mechanism for market awards or excess cost payments for out-of-merit exceptional dispatches³⁷ (uplift payments). The need for uplift payments tends to occur more when energy demand is lower or when the ISO dispatches a resource to operate at or near its minimum load.

4. Identified issues

The ISO has identified two principal issues with its average frequency response performance in trying to meet the NERC Frequency Response and Frequency Bias Setting Standard: (1) The ISO's ability to position its fleet to provide sufficient primary frequency response that maintains grid reliability during the largest contingency events, and (2) its inability to adequately incentivize resources for frequency response capability and provision.

Recall the four market design principles that should be achieved in a well-functioning market to ensure sufficient frequency response capability and provision to meet reliability standards:

- 1. Produce market outcomes that enable the ISO to position its fleet to respond sufficiently to frequency disturbances in the post-event measurement period;
- 2. Allow all technology types to participate in the procurement mechanism by ensuring there are no barriers to entry;
- 3. Produce price signals that incentivize capital investments on resources to be capable of primary frequency response; and
- 4. Ensure compensation of capital investments made to meet the required capability if frequency response capabilities become an interconnection requirement.

In the remainder of this section, the ISO will elaborate on the questions introduced in the executive summary. The ISO asks stakeholders to consider whether the current market design or tariff requirements produce effective price signals and compensation for sufficient frequency response capability and provision.

Does the current ancillary services paradigm position the system to be able to sufficiently respond to meet reliability requirements?

Does the current market design produce price signals that incent capital investments on resources to be capable of primary frequency response?

4.1. Does the current ancillary services paradigm position the system to be able to sufficiently respond to meet reliability requirements?

The goal of this section is to evaluate whether the ISO's current ancillary service paradigm positions the system to be able to respond sufficiently to meet reliability requirements. This section defines "ancillary services" to also include the frequency response capability required to be provided by spinning reserves and provided by regulation. In the past, the ISO assumed that, because conventional resources have frequency response capability, as long as operating reserves are maintained there will be unloaded capacity for primary frequency response to be provided.

This section will evaluate whether this assumption is supported by the market design, and then it will compare that to actual performance data in the following sections:

- Market design might not ensure sufficient frequency response
- Frequency response performance might not be sufficient

4.1.1. Market design might not ensure sufficient frequency response

The ISO posits that its market design may not ensure primary frequency response during certain periods. It is not a strong preventive position to rely solely on requirements for frequency responsive equipment without also ensuring frequency responsive unloaded capacity is maintained on the system to respond to contingency events.

The ISO's current ancillary service paradigm includes the indirect procurement of secondary or tertiary frequency response services through its regulation and operating reserves. For primary frequency response, the ISO and other organized markets have generally relied on the fact that conventional resources when interconnected are built with frequency response capability that provides automatic, autonomous, proportional response to a frequency change. Table 4 with descriptions of the ISO's ancillary services is shown again below for convenience.

Requirement or Product Definition	Synchr- onized to grid	Deployment Method	Deploy Deadline	Full AS Delivery Deadline	Sustain Provision
Primary Frequency	Yes	Frequency event triggers automatic, autonomous governor or under-	Frequency change exceeds	N/A	N/A

Table 4: Current Ancillary Service Requirement or Products³⁸

Response		frequency relay	0.0036		
Requirements		device response	0.1Hz		
Regulation (Up and Down)	Yes	Automatic Generation Control (AGC) signal sends 4 second signals where resource will provide its entire AS obligation by the full delivery deadline based on its ramp rate.	4 seconds to meet the most recently sent signal	10 minutes	N/A
Spinning Reserves	Yes	Contingency Dispatch Operating Target (DOT) Replaces Real-time Market DOT	Immediately	10 minutes	2 hours
Non-Spinning Reserves	No	Contingency Dispatch Operating Target (DOT) Replaces Real-time Market DOT	Immediately	10 minutes	2 hours

4.1.1.1. Primary Frequency Response Requirements

For the primary frequency response service, the ISO does not deploy this service but instead relies on automatic, autonomous response directly at the generating resource after reaching a frequency change of 36 mHz. In Phase 1 the ISO identified several reasons that this assumption could be threatened. First, if the resource does not have frequency responsive capacity unloaded when the governor triggers a response the resource will not move above its maximum operating limit. Resources dispatched to full load cannot provide primary frequency response. Second, resources that have multiple configurations, multi-stage generators, may have a different maximum capacity for each configuration. Third, if a resource does not have a frequency bias in its plant-level and generator-level load controls the frequency response triggered by the governor will be blocked since a load control algorithm is designed to bring the resources' operation in alignment with its dispatch operating target.

In addition to reliance on resources with active governor controls, the ISO expanded its design in Phase 1 to allow the ISO to enter into bilateral contracts to purchase a transfer of frequency response performance from external BAAs. The Transferred Frequency Response product mitigates risk of the ISO system not being positioned in a manner to respond sufficiently but it does not itself position the system. Therefore, this product does not contribute to improving grid operations ability to position its system but is a beneficial compliance instrument backed by existing frequency response performance surplus.

4.1.1.2. Secondary Response – Operating Reserves

Under the cascading ancillary service procurement process, the ISO procures its operating reserves and allows substitution of higher reliability values service, regulation, for the lower reliability service, spinning reserves, to meet its operating reserve requirements. First, the ISO posits that operating reserves, whether for regulation or spinning reserves, are significantly different than primary frequency responsive reserves since the primary purpose of regulation and spinning reserves are to be able to be deployed and provide their full ancillary service obligation by the specified deadline (e.g. 10 minutes). Conversely, primary frequency responsive reserves should be intended to deploy their full amount within several seconds as soon as frequency drops to below 59.964 Hz and their full amount provided is proportional to the actual frequency change.

Generally these services restore the system within longer time frames than needed for the rebound from primary frequency response, constrained by the ramp rate of the unit, and are for transforming reserves into energy. The costs for these services are based on these assumptions. On the other hand, resources on governor controls are constrained to providing a certain amount of primary frequency response proportional to the change in frequency. When deployed, they are operating outside the bounds of their normal operations ramp rate. These stressed levels can increase wear and tear in the long-term on the unit. For this service the frequency response reserves are translated into frequency response that provides additional power to stabilize with costs in excess of those for energy production.

Regulation reserves provide secondary response. Spinning reserves function to provide tertiary response by replacing regulation's secondary response. It functions more like tertiary since regulation does not have an obligation to maintain its response beyond 10 minutes and when this response type falls off spinning reserves would replace it through re-dispatch to fully restore frequency.

In addition to being maintained as operating reserves, whether regulation or spinning reserves, these reserve obligations will provide unloaded capacity on the system that can provide primary frequency response capability. Increased amount of operating reserve procurement from synchronized resources would not guarantee that with the additional operating reserves that the ISO performance would be more reliable. This is due to a few factors including that the ISO market typically awards regulation to a limited number of resources and that the relationship between unloaded capacity and frequency responsive capacity is not 1:1.

Given the finite amount of primary frequency response by a resource, if the ISO only increases its operating reserve levels or awards reserves to a limited number of resources it would have increased the headroom on the system but not increased the frequency response headroom. Without a market constraint to ensure the increased procurement was allocated to resources based on its frequency response capabilities and distribute the obligation across many

resources, the ISO cannot determine in advance if this intervention would improve its performance.

Further, while regulation awards support ability to provide additional reserves to contribute to the secondary frequency response type, the algorithm designed to deliver the full responsibility in 10 minutes does not necessarily begin to send signals to deploy the reserves into additional power output until about 30 seconds after an event. Note that this will allow regulation to contribute to the ISO performance to meet the reliability requirement. However, this contribution is limited unless the algorithm is altered to provide a faster signal such as those used in nonconventional response type of fast secondary or fast regulation.

Regulation awards do provide reliability value for the secondary frequency response service as currently designed. When the ISO's area control error equation fluctuates to signal an under-frequency event the ISO will send automatic generation control (AGC) signals to resources with regulation up obligations. The AGC signal is based largely on the Area Control Error equation which tracks interchange flow imbalances between actual flows and those scheduled. This formula includes a frequency bias term applied to the frequency change on the system which allows for AGC signals while primarily to support actual intertie flows to meet scheduled flows to move resources in a direction that supports system frequency even if it would be contrary to aligning flows. These signals will be updated every four seconds moving the regulation capacity up based largely on the resource's ramp rate and tuned to deliver its full regulation award to the grid by 10 minutes.

This market design to calculate the AGC signal to provide secondary response is rooted in the definition of secondary frequency control: "Operating signals sent by the BAA to signal response that will correct generation and load imbalances that created the original frequency excursion to restore scheduled frequency after primary response stabilized frequency levels." Conventional resources would have slower response times to an ISO signal. Today even with the current AGC calculation, regulation will bring additional power on the system overlapping with resources providing primary frequency response. Some other organized markets, such as PJM and ERCOT, are increasing levels of additional power in response to imbalances as result of operating signals through the introduction of fast regulation services. These services are triggered through ISO signal but there is a second AGC signal sent to resources that are fast responding and awarded regulation. ERCOT defines this service as:

"A subset of Reg-Up in which the participating Resource provides Reg-Up capacity to ERCOT within 1 second of either its receipt of an ERCOT Dispatch Instruction or its detection of a trigger frequency independent of an ERCOT Dispatch Instruction. Except where otherwise specified, all requirements that apply to Reg-Up also apply to FRRS-Up."³⁹

This market design change was largely triggered to the increased penetration of fast resources that have superior speed capability. These resources can respond to ISO signals even more quickly than the 30 seconds observed and are more dynamic than conventional regulation.

In summary from a design perspective, the ISO evaluated whether procuring additional reserves from its current ancillary service products or issuing exceptional dispatches to position the system with frequency responsive capacity that is unloaded. The ISO posits that from a design perspective this would either not be effective or would introduce unnecessary costs due to frequency responsive capacity being different than unloaded capacity. Further since there is a surplus of response on the Western Interconnection the ISO expanded its market design to include a bilateral contract between the BAA under the new standard to transfer a portion of frequency response – a more efficient solution.

4.1.2. Frequency response performance might not be sufficient

The ISO analyzed its primary frequency response performance to evaluate whether its hypothesis that the current design does not allow it to ensure it has sufficient frequency response capability positioned on its system. The ISO believes the analysis demonstrates that in practice the ISO is not operating its system to be assured of sufficient primary frequency response during certain periods. The ISO notes that it has mitigated this risk through the design of the Transferred Frequency Response product transacting between entities under the NERC compliance obligation.

The ISO analyzed its historical primary frequency response performance using the standard's performance measurement. The analysis uses the single event frequency response data from 2012 - 2016 and was analyzed relative to 2016's actual FRO of 198 MW/0.1Hz.⁴⁰ Recall single event frequency response performance is measured as the change in net actual interchange between Point B and Point A given the change in frequency between Point B and Point A (i.e. MW/0.1Hz). Further the ISO analyzed its annual performance on average using the median approach under the standard, which reduces outliers' influence.

For 146 frequency events across January 2012 through November 2016, the ISO's average performance was 214 MW/0.1Hz with a median of 203 MW/0.1Hz. With a 25th percentile of 138 MW/0.1Hz, the ISO could reasonably expect its median performance to exceed this percentile. However the ISO has observed a trend of deteriorating performance year over year due in part to changes in the resource mix. As Shown in Table 4, the ISO observed a deterioration of its performance year over year where its median performance has steadily decreased from 263 MW/0.1Hz in 2012 to 168 MW/0.1Hz for the 2016 compliance period.

Compliance Period	Ν	2016 Frequency Response Obligation	Frequency Response Obligation (FRO)	Frequency Response Measure (FRM)	Frequency Response Shortfall (FRO-FRM)	Minimum Frequency Response Performance
2012	27	197	193	263	-71	56
2013	26	197	193	210	-35	95
2014	33	197	219	219	-6	60
2015	24	197	209	183	32	22
2016	36	197	198	168	17	-193

Table 4: ISO's Annual Performance Trend⁴¹

Table 4 shows the frequency response shortfall on average for each compliance period FRO and FRM for an event increased to roughly 32 MW/0.1Hz on average for 2015 relative to a surplus in 2012.⁴² The 2016 period continued the trend of on average providing insufficient levels of primary frequency response to meet the new NERC standard. Operations identified the main driver of this trend is largely the result of the increased proportion of renewable generation. Renewable generation amounts will continue to increase as California reaches its renewable goals.

To better project expected performance, the ISO isolated the 60 events during 2015 and 2016 compliance periods and assessed its single event performance. Only 19 of the 60 events met the FRO of 198 MW/0.1Hz. During this timeframe, the ISO's average performance dropped to 179 MW/0.1Hz with a median of 172 MW/0.1Hz. Figure 3 shows the ranked single event performance data (MW/0.1Hz) relative to the ISO's 2016 FRO (MW/0.1Hz).



Figure 3: ISO's Frequency Response Performance 2015 - 2016

Given these results, the ISO expects its rate of performance will likely fall short of the FRO.⁴³ In conclusion these results support the ISO hypothesis that its current design is not sufficient to ensure it has positioned the system in a manner that meets its frequency response obligation. This is evidenced by its need to procure a transfer of frequency response from external BAAs.

4.2. Does the current market design produce price signals that incent capital investments on resources to be capable of primary frequency response?

The goal of this section is to evaluate whether the ISO market design produces price signals that would incent investments on resources to be frequency responsive since requirements for

frequency responsive equipment alone have not been shown to produce sufficient frequency response performance.

The ISO will evaluate this through a discussion of whether its current market design accomplishes three of the four market design principles proposed in the introduction to the issues discussions. The ISO believes these principles are needed to incentivize and compensate resources for frequency response capability and provision.

This section will evaluate this question through discussing how the market design could:

- Ensure no barriers to entry
- Produce price signals to incent capability or provision

4.2.1. Ensure no barriers to entry

The ISO believes a competitive procurement mechanism should allow all technology types to participate in the procurement mechanism through ensuring there are no barriers to entry. Free entry and free exit are necessary conditions to be considered a perfectly competitive market. Generally speaking limiting as much as possible barriers to entry and exit will improve the efficiency of any market design.

The ISO believes that there are virtually no technological barriers to providing frequency response of any resource in operation today – even intermittent technologies are able to provide both synthetic inertial response and frequency responsive headroom. The only exception the ISO is aware of might be run-of-river hydro or tidal technology. These technologies would need storage to hold the "prime mover" of water in reserve or to store the energy generated when the water is flowing.

Various technology types have diverse frequency response capabilities. The various technologies can provide different levels of frequency response across the timeframe associated with frequency response standard. The ISO believes that while resources should be treated equivalently regardless of technology type from a requirement stance, the ISO will evaluate how frequency response characteristics vary by technology type. In the remainder of this section the ISO will discuss the potential technical capabilities of various technology types and concerns raised by NREL or others that compensation for capital investments would be necessary to ensuring there are no barriers to entry of any frequency response market.

In the initial review of its fleet or potential technologies, the ISO will introduce the various types grouped by whether the technology type primarily serves to contribute to power balance through (1) consuming a prime mover to generate power and might have augmentations to increase efficiencies of the primary process, (2) reducing its demand on the system, (3) providing power stores that were stored during periods of lower demand (i.e. lower price) and providing stores during higher periods of demand (i.e. higher price), or (4) harnessing intermittent natural resources to generate power during times the resource is available.

Regardless of the type of technology of the generating resource if it is fully loaded or has outer loop controls in place to address physical operational constraints for reasons that include ambient temperature limitations, outages of mechanical equipment or regulatory considerations the resource it will not have frequency responsive capacity available for reserves. These outer loop controls will have a significant impact on capability of resources who are sensitive to ambient temperature so will apply temperature loop controls maintaining its output at levels safe for operation given the temperature and on those who are under water management requirements.

4.2.1.1. Technologies that consume prime mover to generate power

The first category of technologies that consume a prime mover to generate power and might have augmentations to increase efficiencies of the primary process would include:

- Gas Turbine
- Steam Turbine
- Combined Cycle
- Combined Heat Plant
- Pumped hydro facilities
- Nuclear

Gas and steam turbines generally all come equipped with governor controls that provide automatic, autonomous frequency response as long as the frequency responsive reserves are available. Their capability varies depending on the control mode in which it is operating as long as its frequency responsive capacity is not fully loaded. Additionally the technology types that leverage these technologies but have increased the efficiency of the plant through combined cycle generation or capturing waste heat for other purposes experience the same limitations on their gas or steam turbines response depending on the control mode.

The characteristics for each control mode are as follows:44

- Boiler-follow mode: prompt and sustained increase in power followed by long delay from 10 seconds to minutes until full increase in power is achieved.
- Turbine-follow mode: slower response relative to boiler-follow mode.
- Coordinate control: compromise between fast response of boiler-follow mode and sluggish response of turbine-follow mode while maintaining better control of boiler.
- Sliding-pressure control: sluggish response to PFR since it's relying on controlling boiler and fuel which respond more slowly than turbine control valves left open in this mode.

Further, combined cycle gas turbines (CCGT) have additional complexity to their frequency response capability. This technology type builds off existing turbine technology with augmentations that after the gas turbine spins and produces heat waste that heat waste is recovered through heat recovery steam generators that can generate steam to feed into a

steam turbine. The heat waste from the first cycle generates the prime mover for the next cycle. The plants are generally only expected to have a percentage of their maximum capacity be frequency responsive even when operating at no load. This is because the ISO understands that for a CCGT plant with 3 turbines where 2 are gas and 1 is steam there are many instances where only the gas turbines came equipped with frequency responsive. Frequency response capability through governors. The steam turbines largely are not frequency responsive. Frequency response could be increased if there is an incentive to perform equipment upgrades on the steam turbines installing active governors – this would require expectation of compensation for capital expenses as well as perhaps additional opportunity costs from impacts to overall plant efficiency.

Combined Heat Plants are expected to also have the technical capabilities to provide frequency response capability. In the ISO's fleet the number of Combined Heat Plants are a very limited amount of the total nameplate capacity active in its market and largely fueled by gas, biogas, or biomass prime movers. This technology type really builds off existing turbine technology with augmentation to capture the heat waste from the electricity generation process to provide thermal energy for other processes. The main difference between this technology type and that of CCGTs is that the captured heat does not necessarily contribute to generating a prime move for power production in CHPs but instead used for processes such as space heating, hot water heating or other industrial processes. The fuel or steam turbines have the capability of providing frequency response as described in the discussion of gas and steam turbines.

Hydro facilities respond fairly predictably to a frequency excursion event since hydro facilities are generally not fully loaded but are instead used as peaking or spinning reserve resources⁴⁵. With reservoir of water on hand, additional stores of water can be flowed into the turbine area by moving the wicket gates controlling the flow of water. Their response is relatively slow with the beginning of their response expected by T_{30} and roughly 2/3 of their capacity by T_{60}^{46} . Since the most reliability value is gained from response early in the period following an event, this technology would have less reliability value than faster responding types. The slower power response rate is driven in part by the feedback signal used by the resources. Many older hydro resources use a mechanical signal that uses the wicket gate position as a proxy for the power signal. The wicket gate signal has a larger margin of error on the actual amount of water the gate is allowing into the turbine than electrical power signals. The response could be improved if electrical power signals were instead used for the feedback loop. This would require compensation of capital expenses for the equipment upgrades.

The possibility of large amounts of active nuclear technology in California appears unlikely at this time since California only has one active nuclear power plant, Diablo Canyon, which is being evaluated for retirement in 2024/2025 time frame. None the less a nuclear plant has the technical capabilities to provide primary frequency response but seldom provides this response since it is typically operated at full load – does not provide frequency responsive reserves. However it will act to limit the severity of a frequency event by providing inertia. An exception this is in New England where the Lepreau Generating Station in New Brunswick provides primary frequency response.⁴⁷

4.2.1.2. Technologies that reduce demand on system

The second category of technologies includes any demand responsive technology on the wholesale or perhaps even the distributed side of the electric system. These technologies reduce demand on the system when generation is at insufficient levels to meet demand.

Any participant that is demand responsive has technology with the technical capability to provide dependable frequency response constrained to the amount of demand it can off load at a given time based on its contracts. However if the load does not have an under frequency relay device installed or other demand responsive technologies such as electric vehicles or air conditioners do not have control devices such as enhanced inverters then the capacity is not frequency responsive. There would need to be compensation for the capital expenses associated with the equipment upgrades to install the control devices necessary to respond to a frequency event in the primary time horizon.

4.2.1.3. Technologies that provide power from storage

The third category includes an array of storage technologies on either the wholesale or distributed systems. These technology types are designed to store power during periods of lower demand (i.e. lower price) and then deliver its stores during periods of high demand (i.e. higher price). Like demand response, the ISO anticipates these technology types all have the natural capability to provide primary frequency response as long as the necessary control devices are installed that can read under frequency deviations and trigger an automatic, autonomous response. Further the ISO notes that since the power is stored these technology types have a higher reliability value due to its speed and potentially greater amounts per resource that could be provided to arrest and stabilize the frequency event.

4.2.1.4. Technologies that harness intermittent natural resources

The fourth and final category the ISO has identified includes variable energy resources. These resources harness intermittent natural resources to generate power during periods where the resource is available. This category includes:

- Wind
- Solar
- Run-of-river Hydro
- Tidal

The ISO understands that during periods where the intermittent resources are available most variable energy resources would have the technical capability to provide primary frequency response although it would likely require the resource to operate less efficiently by not producing as much active power output as possible when the conditions are favorable.

Wind and solar resources can vary the amount of power produced to support system power balance and for wind can even provide inertial response. However both wind and solar variable energy resources are generally decoupled from grid frequency based on the power electronics-

based converters.⁴⁸ Both wind and solar if equipped with active power control devices such as enhanced inverters could provide primary frequency response but only if operating below maximum output levels.

The mechanisms for operating with some frequency responsive reserves on a wind resource were discussed in a National Renewable Energy Lab study on wind power's ability to provide frequency response outlines a multitude of research and development efforts into wind's capability to provide inertial, primary or secondary response services. In regard to primary response NREL notes, "studies operated the turbine at higher than optimal tip-speed ratios to maintain a reserve of available wind power. With an overhead of un-utilized power, in an under frequency event, the controller can increase the generator torque to extract inertia from the turbine, causing the turbine to slow down to the optimal tip-speed ratio"⁴⁹.

The mechanisms for operating a solar resource at levels lower than its maximum output level, also called deloaded operations, were discussed in a paper by Rahmann and Castillo on "Fast Frequency Response Capability of Photovoltaic Power Plants: The Necessity of New Grid Requirements and Definitions". Rahmann briefly describes how a photovoltaic plant could be run under deloaded operations, "For a determined temperature and irradiance, PV units can be deloaded by operating them at reduced/increased DC voltage with respect to the optimal DC operation voltage (VMPPT). Both alternatives result in an output power reduction ΔP ".

Accepting the conclusion that both wind and solar resource can operate at levels lower than their maximum output levels, it is generally accepted that with additional controls both would be able to read frequency changes and trigger a response. These papers support that general belief.

From a theoretical as well as observed basis, wind and solar resources are being demonstrated to have the technical capabilities to respond but to provide the service it would require (1) upgrading equipment to have enhanced inverters that coordinate with their other controls, (2) operating at less efficient levels such as at higher than optimal tip-speeds for wind to maintain reserves, and (3) assumption of opportunity costs of not extracting and selling as much power as possible given wind levels. All of these factors drive a need to compensate for the capital expenses, operating expenses, and opportunity costs for the provision.

Finally, the ISO believes that run-of-river hydro or tidal technologies without either ability to store water for generation after an event or power storage to store surplus to provide power after an event may not have the technical capabilities to provide frequency response. The ISO seeks comments on whether this supposition is accurate. Regardless the ISO initial position is that hybrid technologies that pair an intermittent based technology with storage would be a feasible if there is sufficient incentive to make that capital investment.

4.2.2. Produce price signals to incent capability or provision

The ISO believes that there should be no restrictions on the type of resources that can participate in a frequency response market and that if incentives are aligned appropriately then the economic barriers to entry can be removed. For a well-functioning market, the should (1)

produce price signals that incentivize capital investments on resources to be capable of primary frequency response and (2) ensure compensation of capital investments made to meet the required capability if frequency response capabilities become an interconnection requirement.

The exact settlement mechanism for a frequency response market would vary depending on the design of the procurement mechanism. These proposed principles would be maintained in the later design.

In a report on Variable Renewable Generation Can Provide Balancing Control to the Electric Power System, NREL directly discussed the barriers to providing the technical capability. NREL points to economic reasons as the main barrier to entry:

"Active power controls are not widely used today by wind and solar generators because economic barriers—including upfront equipment costs, lost opportunity costs, and potential to increase loading impacts on the turbine components—discourage their use. Renewable generation operators are not able to justify such costs unless they are compensated for providing these services or are required to do so. In addition, wind and solar generators must operate below full generation output to provide certain balancing services. This has economic consequences. The cost for renewables to provide support services is the lost opportunity cost of generating energy that could be sold into market. Renewable generators have higher opportunity costs than conventional generators that save on fuel costs by not providing power. Thus, wind and solar generators require greater compensation to provide grid support services because otherwise they could provide power to the system from a free fuel supply—the wind or sun."⁵⁰

First, since the ISO does not have a market for primary frequency response and it has been established through this paper that there are additional costs associated with having the capability and providing the response to all the other markets – capacity, energy, regulation or spinning reserves – that there is limited to no price signals to incentive sufficient response.

The ISO is aware from its review of the Notice of Inquiry comments submitted by various stakeholders and its discussion of the various technology types in the previous section that there are multiple components to the provision of frequency response: maintaining the frequency responsive headroom; producing energy for frequency response; providing frequency response at faster speeds, whether in the form of increased wear and tear and future maintenance, or making capital investments to improve frequency response capabilities. In this initiative, the ISO will investigate what pricing and compensation mechanisms are best able to capture these costs as well as incent appropriate future investment into frequency response capability.

In analyzing what types of costs are incurred to be frequency response capable and to provide that response, the ISO reviewed comments submitted by other stakeholders in FERC's Notice of Inquiry process on the question of compensation. The answers informed the ISO's identification of issues impacting how to ensure price signals incent desired behaviors.

Specifically FERC sought comments on what form of compensation mechanisms might be appropriate, if any, for those resources providing frequency response?

First, the ISO agrees with many commenters' position that FERC should not enact a national standard for compensation since the generation portfolio and specific market rules would need to vary to reflect specific needs of each organized market. This initiative will evaluate the specific needs of the California ISO and what specifications would be necessary to ensure sufficient frequency response capability and performance that meets the ISO's needs.

Second, the ISO agrees that capability and responsive reserves are two of the potential costs associated with being frequency responsive as noted by commenters but adds that when response is needed and the frequency responsive reserves are deployed there are additional costs of providing that response not directly captured in energy markets pricing delivery of a five minute product⁵¹. Table 9 below shows where these types of costs would fall within three category of expenses – capital expenses, opportunity costs, and operating expenses.

Type of costs	Capital Expense	Opportunity Cost	Operating Expense	Payment Type
Maintaining the frequency responsive headroom		X		Opportunity Cost
Producing energy for frequency response			X	Energy Price
Providing frequency response at different speeds			Х	Capability Payment
Making capital investments to improve frequency response capabilities	X			Capability Payment

Table 9: Frequency Response Expenses

The ISO notes that in the NOI comments some stakeholders representing specific technology types with superior frequency response capabilities raised the diverse capability of the fleet based on technology types. These stakeholders seemed to support incentives for resources to have capabilities that provide greater reliability value for grid services.

For instance they requested consideration of compensation for capabilities for (1) inertial response,⁵² (2) speed, (3) percent of total capacity that is frequency responsive, and (4) ability to sustain response. As to the last three capabilities, energy storage stakeholders,⁵³ Western Interconnection Advisory Board, and some public interest organizations advocated for capability payment based on the speed of the response.

On the other hand, there were several commenters that argued that generators should be compensated for their operating costs of providing the energy only excluding the frequency response operating costs that impose wear and tear on the resource for operating at such stressed levels. Further others were strongly of the opinion that the frequency responsive capacity should clear in a market like other ancillary services and be compensated for its opportunity costs.

If the ISO elects to procure frequency response, it is reasonable to compensate resources for their frequency responsive capability, frequency responsive reserves, and frequency response provision. The principle to compensate resources for their service based on their willingness to provide the service at a given price under competitive market conditions is a long standing market design principle. ISO notes that this is analogous to the idea of a make-whole payment in the energy market where a supplier's total production costs are guaranteed if market revenues fail to cover the offered or mitigated price.

Appendix A: Stakeholder Engagement Plan

The ISO is targeting completion of this initiative and requesting board approval by September 2017 for fall 2018 implementation. The current schedule for this initiative is shown below.

Milestone	Date	
Issue paper posted	December 15, 2016	
Stakeholder call	December 22, 2016	
Stakeholder written comments due	January 11, 2017	
Straw Proposal Posted	April 6, 2017	
Stakeholder call	April 13, 2017	
Stakeholder written comments due	April 27, 2017	
Draft final proposal posted	July 25, 2017	
Stakeholder call	August 1, 2017	
Stakeholder written comments due	August 15, 2017	
Board of Governors meeting	September 2017	

The California ISO will discuss this issue paper with stakeholders during a call on Thursday, December 22, 2016. Stakeholders are asked to submit their written comments to <u>initiativecomments@caiso.com</u> by close of business on January 11, 2017.

Notes

³ Id.

¹¹ Frequency Response Initiative Report – October 2012 at 51 – 52.

¹² Annual Gen_{BA} is total annual BAA Net Generation (MWh) as reported on FERC Form 714, column c of Part 11 – Schedule 3, <u>http://www.ferc.gov/docs-filing/forms/form-714/view-soft.asp</u>.

¹³ Annual Load_{BA} is total annual BAA Net Energy for Load (MWh) as reported on FERC Form 714, column c of Part 11 – Schedule 3 which is the sum of the BAA Net Generation and Net Actual Interchange, <u>http://www.ferc.gov/docs-filing/forms/form-714/view-soft.asp</u>.

¹⁴ BAL-003-1 Standard, Attachment A, http://www.nerc.com/pa/Stand/Reliability Standards/BAL-003-1.pdf.

¹⁵ NERC Petition at 13, http://www.nerc.com/FilingsOrders/us/NERC Filings to FERC DL/Petition and Exhibits A-J_FINAL.pdf.

¹⁶ Frequency Response Initiative Report, October 2012, Page 70, http://www.nerc.com/pa/rrm/bpsa/alerts dl/2015 alerts/frequency response initiative report.pdf.

¹⁷ Frequency Response Initiative Report, October 2012, Page 72, http://www.nerc.com/pa/rrm/bpsa/alerts dl/2015 alerts/frequency response initiative report.pdf.

¹⁸ Procedure for ERO Support of Frequency Response and Frequency Bias Setting, Pages 1-2,

http://www.nerc.com/comm/OC/RS Landing Page DL/Related Files/Procedure_Clean_20121130.pdf.

¹⁹ Procedure for ERO Support of Frequency Response and Frequency Bias Setting Standard, Pages 1 – 2, http://www.nerc.com/comm/OC/RS Landing Page DL/Related Files/Procedure_Clean_20121130.pdf.

²⁰ MyBA 2016 FRS Form 1.10 Western Interconnection, http://www.nerc.com/comm/OC/RS Landing Page

DL/Frequency Response Standard Resources/MyBA_2016_FRS_FORM_1 10_Western_Interconnection 2nd Qtr 2015.xlsm.

²¹ MyBA FRS Forms 2.10 MultiBAInterconnection, http://www.nerc.com/comm/OC/RS Landing Page DL/Frequency Response Standard Resources/MyBA_FRS_Form2.10_MultiBAInterconnection.xlsm.

²² Single Event Frequency Response Data (SEFRD) is the individual sample of the event selections and actual performance measurements data from a Balancing Authority which represents the change in Net Actual Interchange (NIA), divided by the change in frequency, expressed in MW/0.1Hz.

²³ Frequency Response Standard Background Document, November 2012, Page 13,

http://www.nerc.com/comm/OC/RS Landing Page DL/Related Files/Bal-003-

1_Background_Document_Clean_20121130.pdf.

²⁴ BAL-003-1 Detailed Implementation Timeline, North American Electric Reliability Corporation, August 20, 2014, https://www.wecc.biz/Reliability/WECC-0115 BAL-002-WECC-2 Request to Retire R2 - NERC BAL-003-1 Implementation Plan Timeline.pdf.

²⁵ Frequency Response Initiative Report, October 2012, Pages 56 – 61, http://www.nerc.com/pa/rrm/bpsa/alerts dl/2015 alerts/frequency response initiative report.pdf.

²⁶ Frequency Response Initiative Report, October 2012, Pages 56 – 61, http://www.nerc.com/pa/rrm/bpsa/alerts dl/2015 alerts/frequency response initiative report.pdf.

²⁷ 2013 Frequency Response Annual Analysis, http://www.nerc.com/FilingsOrders/us/NERC Filings to FERC DL/FR Annual Report 12-27-13 Final.pdf.

²⁸ 2014 Frequency Response Annual Analysis, http://www.nerc.com/FilingsOrders/us/NERC Filings to FERC DL/Final_Info_Filing_Freq_Resp_Annual_Report_03202015.pdf.

¹ Frequency Response and Frequency Bias Setting Reliability Standard, 146 FERC ¶ 61,024 (2014) (FERC Order No. 794).

² See California Independent System Operator Corp., 156 FERC ¶ 61,182 (2016).

⁴ Time control includes small offsets to scheduled frequency to keep long-term average frequency near the scheduled frequency at 60 Hz.

⁵ NERC RM13-11 filing, March 29, 2013.

⁶ Frequency Response Initiative Report, October 2012, 39, <u>http://www.nerc.com/pa/rrm/bpsa/alerts dl/2015</u> <u>alerts/frequency response initiative report.pdf</u>.

⁷ Claudia Rahmann and Alfredo Castillo, University of Chile, Fast Frequency Response Capability of Photovoltaic Power Plants: The Necessity of New Grid Requirements and Definitions, September30, 2014, Page 3.

http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=5&cad=rja&uact=8&ved=0ahUKEwjqqJDIn_XQ AhVR62MKHU4FCU4QFgg2MAQ&url=http%3A%2F%2Fwww.mdpi.com%2F1996-

^{1073%2}F7%2F10%2F6306%2Fpdf&usg=AFQjCNEIYNUwHCgavy4Yb2Rnt3F0C00Qzw.

⁸ Nodal Protocols, Section 2-24.

⁹ See FERC Order No. 794., <u>http://www.ferc.gov/whats-new/comm-meet/2014/011614/E-2.pdf</u>.

¹⁰ NERC Petition at 15, http://www.nerc.com/FilingsOrders/us/NERC Filings to FERC DL/Petition and Exhibits A-J_FINAL.pdf.

²⁹ 2015 Frequency Response Annual Analysis, http://www.nerc.com/comm/OC/RS Landing Page DL/Related Files/2015_FRAA_Report_Final.pdf.

³⁰ Reliability Guideline: Primary Frequency Control v1.0 Final,

http://www.nerc.com/comm/OC/Reliability%20Guideline%20DL/Primary_Frequency_Control_final.pdf.

³¹ Essential Reliability Services and the Evolving Bulk-Power System—Primary Frequency Response, 157 FERC ¶ 61,122 (2016).

³² Essential Reliability Services and the Evolving Bulk-Power System—Primary Frequency Response, 154 FERC ¶ 61,117 (2016).

³³ These costs will vary be the amount of time the unit has been shut down generally referred to as "hot",

"intermediate", or "cold" starts. "Cold" starts will be the most expensive of the three as it is likely to require the most fuel or auxiliary power to bring the unit from off to on.

³⁴ ISO Market Instruments Business Practice Manual.

³⁵ ISO Market Operations Business Practice Manual at Section 7.9.

³⁶ ISO Market Operations Business Practice Manual at Section 4.4.1.

³⁷ The California ISO settles the excess cost for exceptional dispatches used to mitigate or resolve congestion as a result of transmission related modeling limitations through exceptional dispatch uplift settlements (Charge Code Configuration Guide 6488). The California ISO settles the excess costs for system emergency exceptional dispatch energy types through the real-time excess cost uplift settlements (Charge Code Configuration Guide 6482). Both of these excess cost uplift settlements are made at the supplier's offer price or better.

³⁸ ISO Market Instruments Business Practice Manual.

³⁹ ERCOT Regulation Discussion Paper, April 2014,

http://www.ercot.com/content/meetings/fast/keydocs/2014/0418/Regulation%20Discussion%20Paper%20as%20of%2004172014_1140.doc

⁴⁰ NERC Balancing Authority Area Frequency Response Obligations,

http://www.nerc.com/comm/OC/RS%20Landing%20Page%20DL/Related%20Files/BA%20FRO%20Allocation%20for %20OY%202017.pdf.

⁴¹ This chart was updated to reflect the 2016 FRO published by NERC instead of the FRO based on a planning assumption of 30% of the IFRO. Additionally the 23% used by NERC was applied to the prior years' IFROs to adjust the compliance periods' FRO to a share of 23%.

⁴² Frequency response shortfall or surplus is respectively the MW/0.1Hz amount the frequency response performance rate either does not meet or exceeds the FRO during the post-event measurement period.

⁴³ Estimate subject to revision.

⁴⁴ Interconnected Power System Response to Generation Governing: Present Practice and Outstanding Concerns, IEEE, May 2007 at 1-5.

⁴⁵ This holds true unless units are needed for water flow control and then they will not be able to provide PFR.

⁴⁶ Interconnected Power System Response to Generation Governing: Present Practice and Outstanding Concerns, IEEE, May 2007, Pages 1-6.

⁴⁷ Interconnected Power System Response to Generation Governing: Present Practice and Outstanding Concerns, IEEE, May 2007, Pages 1-15.

⁴⁸ NREL, Page 8, http://www.nrel.gov/docs/fy12osti/54605.pdf.

NREL Variable Renewable Generation Can Provide Balancing Control to the Electric Power System, Page 1, <u>http://www.nrel.gov/docs/fy13osti/57820.pdf</u>.

⁴⁹ NREL, Page 10. <u>http://www.nrel.gov/docs/fy12osti/54605.pdf</u>

⁵⁰ NREL Variable Renewable Generation Can Provide Balancing Control to the Electric Power System, Page 1, http://www.nrel.gov/docs/fy13osti/57820.pdf.

⁵¹ For the purposes of comparing the cost of provision the ISO asserts the 5 Minute Market energy prices are the best comparison to compensation of any uninstructed imbalance energy due to primary response would be settled. ⁵² Nuclear Energy Institute requested that units with large turbines, such as nuclear units, should receive

compensation for the inertial response that they provide. Tacoma Power and the Tennessee Valley Authority strongly agreed with this statement.

⁵³ AES, the Energy Storage Association, and Grid Storage Consulting were some of these stakeholders.