ELECTRICITY AUTHORITY

EVALUATION OF EMERGENCY RESERVE SCHEME OPTIONS

9 May 2025

Prepared by: Sue Paul, Tim Robinson, Eija Samson

Document version: FINAL 2.1

Robinson Bowmaker Paul
Level 17, Grand Plimmer Tower
2-6 Gilmer Terrace
Wellington 6011
New Zealand
rbp.consulting

EXECUTIVE SUMMARY

The Electricity Authority (the Authority) has engaged Robinson Bowmaker Paul (RBP) to evaluate potential options for an Emergency Reserves Scheme like the Reliability Emergency Reserves Trading (RERT) scheme implemented in the National Electricity Market (NEM) in Eastern Australia.

CONTEXT

The Energy Competition Taskforce has proposed two work packages to strengthen the performance of the electricity market in the short to medium-term. The second of these packages relates to providing more options for consumers and includes four initiatives or tasks. Task 2D is part of this package and pertains to rewarding industrial consumers for providing short-term demand flexibility.

To support Task 2D, the Authority has engaged RBP to evaluate potential design options for an Emergency Reserves Scheme (ERS) but with a focus on utilising demand side resources instead of in- or out-of-market generation resources.

The Authority will consult on the desirability and potential options for an ERS, informed by the outputs of this project. This report is intended to inform a consultation paper the Authority will develop.

This project is part of a wider project around demand response initiatives.

ROLE AND PURPOSE OF THE ERS

The New Zealand electricity spot market is designed to provide accurate price signals to guide efficient long-, medium- and short-term decisions by participants. Most of the time, price signals and other information provided to market participants are accurate enough such that the market is able to ensure that demand and supply are balanced.

However, there have been historical incidents in which the market has been unable to balance demand and supply resulting in involuntary load shedding such as the low residual event on 9 August 2021 which occurred due to a confluence of adverse events or “perfect storm”, which included wind output being lower than what was forecast, demand increasing and loss of generation from a hydro plant.

These perfect storm events are rare, and it is not reasonable to expect the market to be able to balance supply and demand when such events occur. When such events occur, involuntary load shedding is the last resort mechanism to balance supply and demand. If an ERS were implemented, it could be used to limit the quantity of this involuntary load shedding in the rare instances that the market is unable to balance supply and demand (while ensuring price-led demand response due to spot market price signals are not compromised). Implementing an ERS would not completely remove the need for involuntary load shedding. In the context of power system reliability, an economic level of load shedding represents the quantity of unserved energy above which the net benefits associated with additional generation investment in the spot market would yield a positive net benefit. Hence, using Emergency Reserves to remove all involuntary load shedding can result in in efficient outcomes.

We therefore recommend the purpose of a potential ERS is to be a penultimate resort service that sits between the market and involuntary load shedding to enable the System Operator to limit involuntary load shedding to economic levels during times of extreme stress.

Given the purpose of the ERS is to be a penultimate resort used in rare circumstances, it seems prudent to integrate the ERS into the System Operator's existing low residual monitoring and management processes as opposed to re-inventing the wheel. For this reason, we further recommend that the ERS be viewed an additional resource in the System Operator's market toolkit when managing low residual conditions.

EVALUATION APPROACH

We have used multi-criteria analysis to evaluate potential design options as follows:

- We first identified the various design attributes or design decisions associated with implementing a potential ERS. These attributes are summarised in Figure 2.
- We then specified different options for addressing each attribute. These options and their implications are described in Chapter 2.
- Next, we specified six different "Strawman" scheme designs with each strawman possessing one more different feature for the various design attributes. Detailed descriptions of the strawmen schemes are provided in Table 12 and Table 16.
- Each strawman was then evaluated against the competition, reliability, and efficiency components of the Authority's statutory objective. The evaluation framework is described in

Section 3.1 and identifies outcomes associated with each component of the Authority's statutory objective. Each strawman scheme was evaluated with respect to how well it performs in terms of achieving these outcomes. Each strawman scheme is also evaluated against cost and complexity criteria. The performance of each strawman scheme against the Authority's statutory objective and cost and complexity criteria is summarised in Chapter 3, while the detailed evaluation against the outcomes associated with the statutory objective is provided in 3.3Appendix A.

- Finally, we recommend preferred policy settings for a potential ERS scheme in Section 3.3.

KEY FINDINGS

Procurement timeframes and triggers

Limited value in implementing a long-notice scheme in New Zealand

Our evaluation indicated that there is limited value in implementing a long-notice scheme (12 months to 3 months ahead) as any ERS requirements determined many months is unlikely to accurately represent real- or near-real time shortfall risk.

Medium- (3 months to 4 weeks ahead) and short-notice (4 weeks to 1 week ahead) schemes are likely to produce more accurate forecasts of potential shortfall risks. The System Operator is unlikely to incur material additional costs by operating medium- and short-notice procurement versus operating one or the other. Nevertheless, the quantity of procured Emergency Reserves is likely to be more accurate for short-notice procurement than medium notice. For this reason, we recommend a short-notice scheme only; however, given the short timeframe between triggering procurement and the forecast shortfall (which could be as little as a week), we recommend that System Operator operate a pre-qualified panel of providers to ensure timely procurement.

As with its existing Grid Support Contracts, we recommend that the System Operator run competitive tenders to establish their panel of pre-qualified providers.

Reliability indicators used to trigger procurement and activation can be developed using existing indicators; however, changes will be required

As the purpose of the ERS is to enable the System Operator to limit involuntary load shedding to economic levels during rare occasions of extreme system stress, one or more indicators or reliability standards are needed to enable the System Operator to determine:

- Whether a credible risk of involuntary load shedding exists over procurement and activation timeframes
- The quantity of energy or capacity needed to close the "reliability gap".

The New Zealand spot market successfully balances supply and demand the vast majority of the time and that involuntary load shedding tends to occur during extremely rare "perfect storm" events when wind and demand forecasts are "wildly out" from actual system conditions or when supply availability is reduced due to multiple unplanned outages.

This means that a reliability standard used to trigger an ERS should incorporate both the inherent and forecast uncertainty of the drivers of involuntary load shedding. Continuing with the example of wind generation, a reliability standard should reflect:

- Uncertainty in wind generation due to variations in weather conditions; and,
- Forecast uncertainty in wind generation forecasts which may result in wind forecasts being materially under-forecasted.

There are three sets of reliability indicators that are currently published to monitor security of supply in New Zealand:

- The Security Standards[1] developed by the Authority are published annually by the System Operator as part of its Security of Supply Annual Assessment (SOSA). Even in the absence of a long-notice scheme, it would be useful to use these standards to inform the System Operator's market testing and "pre-procurement" activities. However, we recommend:
  - Redeveloping the EUE threshold to estimate what percentage of annual demand it is economic to serve.
  - Annually calculating EUE for the year-ahead by using energy market modelling techniques in which uncertain variables such as planned outages, wind output and demand are made stochastic. This will yield a probability distribution for unserved energy, enabling the Authority (or System Operator as relevant) to calculate the average level of involuntary load shedding that is expected. Additionally, the model should be able to inform the quantum, location, and duration of potential outages.

[1] This includes a ten-year ahead forecast of the NZ WEM, SI WEM, and NI CM under various scenarios.

- The New Zealand Generation Balance (NZGB) is published by the System Operator forecasting the capacity margin over the next 200 days. The N-1 balance[2] is a reasonable indicator to trigger medium- or short-notice procurement. However, this balance is based on a specific set of system conditions and does not considered uncertainty due to unplanned outages, wind output and demand. To ensure that uncertainty is adequately incorporated, we recommend the use of Probability of Exceedance (POE) wind and demand forecasts[3] when calculating the N-1 NZGB. This will capture inherent uncertainty, but not forecast uncertainty (to capture forecasts being "wildly out". As discussed in Section 2.3.2 (Amended New Zealand Generation Balance), the time-frame in which the NZGB is calculated means that it is not practical to incorporate a forecast uncertainty measure.
- Residuals produced by the Weekly Dispatch Schedule (WDS) and Non-Responsive Schedule (NRS)[4] indicate the forecast capacity gap between 1 week ahead to 30 minutes ahead of real-time dispatch. These residuals can be used to trigger pre-activation of Emergency Reserves[5]. However, these residuals are calculated based on a specific set of system and market conditions with no consideration of real-time outcomes being materially different. As such, we recommend the use of a forecast uncertainty measure using historical residual errors and their drivers to develop a probability distribution of residuals. This distribution can then be used to determine POE forecast of residuals (at the time of pre-activation).
- We recommend residuals produced by the Non-Price Responsive Schedule Short (NRSS) or Real-Time Dispatch (RTD) schedule (as relevant) be used to activate Emergency Reserves. As

[2] This is the system's capacity cover, over the peak, the loss of the largest risk-setter (generator (AC risk) or HVDC pole (DC risk)).

[3] For example, a 90% POE wind and demand forecast would represent forecasts that would only be exceeded 10% of the time. Note that such a forecast could be provided by the new Intermittent Generation Forecasting Market Operations Service Provider (MOSP) role that the Authority is in the process of procuring. It is also important to note that demand is correlated with wind output (e.g., a windy winter day would be accompanied by lower temperatures and higher demand for temperature dependent loads); this means that the POE forecasts of wind and demand would need to be correlated.

[4] The NRSL schedule is run every two hours for the next 72 trading periods (36 hours). The schedule does not include difference bids or non-conforming bids. The NRSS schedule is produced every 30 minutes for the next 8 trading periods (4 hours).

[5] The NRSS schedule could be used to pre-activate reserves with a short lead-time of 2 hours or less while the NRSL schedule could be used to pre-activate reserves with a lead-time between 2 and 36 hours. Reserves with a longer lead-time could be pre-activated using the WDS (which is more uncertain).

these schedules are calculated within the hour prior to real-time, they are likely to accurately represent real-time conditions.

Better notification of price-led demand response will enable better forecasts of residuals used to pre-activate/activate Emergency Reserves

A potential ERS should not replace the price-led demand response by loads (e.g., retailers and large users) that currently occurs in response to forecast scarcity prices and/or the System Operator issuing a CAN in response to low NZGB values or residuals. Instead, Emergency Reserves should be pre-activated (or armed) and activated (or dispatched) assuming some level of price-led response will occur. If the System Operator can determine the level of price-led response, this will result in more accurate demand forecasts and therefore more accurate residuals. There may, therefore, be benefit in strengthening the low residual condition management process by requiring loads to provide information on any curtailment they have planned. Alternatively, the System Operator could use information from the Price-Responsive Schedules (PRS) to estimate the impact of price on demand.

Scheme Participation

Additionality is a critical criterion for participation

Additionality is the term for ensuring a demand-side (or other) resource's paid service is additional to what the resource would have done in response to price or other incentives. Paying a resource for something it would have done anyway will erode value delivered to consumers. For this reason, verification of resource additionality will be a critical component of determining whether a resource is eligible to participate in an ERS. We recommend that a pre-requisite for any resource participating in an ERS is that its response must be additional to what it would have done in response to price-led or other incentives.

Demand-side participation should be open to all demand sectors

While the flexibility potential is greatest amongst industrial customers, there is limited value in excluding the commercial and residential sectors.

We have considered the potential costs and complexities of incorporating residential aggregations in a potential ERS:

- Residential aggregations will likely comprise hundreds and possibly thousands of individual loads and/or devices. Performance verification may require the System Operator to implement systems to interface with the aggregator's Distributed Energy Resource Management System (DERMS) to so it can use each load's meter data to verify if the portfolio has curtailed from their estimated baseline consumption. This might require the System Operator to implement new systems that can interface with the aggregator's Distributed Energy Resource Management System (DERMS). This could potentially add significant cost to implementing an ERS. However, a lower cost and complexity alternative would be for the aggregator to provide aggregated meter data to the System Operator to enable performance verification. The System Operator may need to undertake occasional audits or spot checks of the raw meter data to ensure the aggregator is submitting accurate data.

- Residential loads will have volatile consumption patterns which may make it challenging to baseline them accurately. Inaccurate baselining may result in consumers paying for services that were partially or not delivered. However, System Operator could conduct due diligence checks on an aggregator's meter data to check whether reasonably accurate baselining is possible.

In the short-term, it is unlikely that most residential customers will meet the System Operator's additionality requirements (see Section 2.2.2 (Demand side providers), technical requirements (e.g., firmness and/or duration of response requirements.). However, as aggregator capability increases and technological developments result in more smart controllable devices being available to residential and commercial customers, the technical barriers to entry may reduce. For this reason, we recommend that a potential ERS be open to industrial, commercial, and residential customers.

Spot market generator participation could erode efficiency and competition objectives

Our evaluation noted that allowing spot-market generators to participate could have materially adverse impacts on efficiency and competition. This is because spot market generators could withhold capacity from the spot market to instead offer through the ERS and receive a higher price. Given the recent increase to the Scarcity Price Limits, generators should have sufficient incentive to offer energy through the spot market during periods of scarcity. We therefore recommend that generators registered in the spot market be excluded from providing Emergency Reserves.

Allowing unregistered generators to participate will increase potential pool of providers

Including unregistered generators is unlikely to affect efficiency objectives as their participation is unlikely to distort spot market prices or result in capacity being withheld from the spot market. Allowing these generators to participate will improve competition as there will be a greater pool of providers.

Compensation and relationship with VoLL

Fee structures should enable providers to represent their costs accurately

Providers should be allowed to specify preparation, availability, pre-activation and activation fees to ensure the fee structure closely matches their cost structure. This will increase participation and ensure that fee structures are efficient as providers can represent their costs more accurately. Limiting fee structures to only activation or availability and activation payments may result in providers inflating those fee categories to cover other costs that they are not able to explicitly include in their fee structure. This may result in consumers paying more than they would have had the provider been able to represent its costs more accurately.

ERS costs should be restricted to the average VoLL

To ensure involuntary load shedding is restricted to economic levels, System Operator should make reasonable endeavours to ensure that at the time of procurement, the forecast cost of ERS provision (based on forecast shortfalls) is not greater than the VoLL.

VoLL values will affect ERS pricing, Scarcity Price Limits and then EUE threshold for the amended Security Standards. As such, it will be important to ensure that VoLL estimates are accurate. We therefore further recommend that VoLL reviews should be conducted once every three years.

Preserving scarcity signals are key to maintaining investment signals but may result in minor inefficiencies

Scarcity pricing signals may be distorted unless activated Emergency Reserves are treated as instructed load shedding

Scarcity prices indicate that there is shortfall in the spot market and is a critical signal for new investment. Emergency Reserves being activated during scarcity periods could result in scarcity prices not binding when they should[6], thereby removing the investment signal. The RTDP schedule addresses the impact of load shedding on scarcity prices by adding instructed load shedding back onto the nodal load forecasts to restore the scarcity prices.

We recommend a similar approach be adopted for activated Emergency Reserves. That is, the activated reserves should be added back on to the relevant nodal load forecasts to ensure scarcity pricing signals are maintained.

[6] This occurs because curtailment or increased generation during periods of scarcity will decrease demand so that there is no longer insufficient generation.

Treating activated Emergency Reserves as instructed load shedding could result in "double payment" by loads

Adding activated reserves back onto the load forecast and requiring loads to pay ERS costs (see below) will result in loads paying both the scarcity price and the ERS activation cost. The quantum of this double payment is likely to be relatively minor (see discussion in Section 2.8.1). This double payment can be addressed by requiring System Operator to restrict the market costs (function of scarcity price) plus pre-activation/activation costs to be less than VoLL. As indicated above, the System Operator will be required to make reasonable endeavours to ensure the forecast cost of ERS provision (based on forecast shortfalls) is not greater than the VoLL. If the additional cost of implementing a further restriction limit market and pre-activation/activation costs is minor, then it would be prudent to implement this feature to prevent double payment. If, however, the additional cost and complexity is material, it may be prudent to accept a minor amount of double payment and inefficiency.

Cost allocation considerations

Our evaluation indicated that there is limited value in allocating the ERS costs to the generators as it is likely that these costs will be ultimately passed down to retailers and eventually consumers. Allocating to loads or purchasers directly on the other hand gives retailers more information about their exposure to these costs and may result in retailers proactively managing and reducing the consumption of their loads to avoid or reduce charges. We therefore recommend:

- Non-event costs (preparation and availability fees) be allocated nationally to loads based on share of monthly metered consumption.
- Event costs (pre-activation and activation fees) be allocated nationally to loads based on share of metered consumption during activation events.

System Operator discretionary parameters can be monitored through information provision

We have recommended leaving the following design attributes to the System Operator's discretion.

- The approach to selecting providers from offers during the procurement process (which will involve a competitive tender)
- Selecting providers to pre-activate and activate.

This is to provide System Operator with flexibility to make decisions that will enable them to meet the desired outcomes of the scheme. Instead of placing regulatory obligations on the System Operator, we recommend the following:

- To monitor whether the System Operator is selecting the highest value combination of providers during procurement and is making reasonable endeavours to limit the overall costs to VoLL, we recommend that the System Operator:
  - Report procurement activities to the Authority in detail including:
    i. Number of providers and their offer details
    ii. Providers selected and rationale behind the choice
    iii. Forecast ERS cost based on selection.
  - Publish redacted or aggregated information on its procurement activities (removing commercially sensitive information).
- To monitor whether the System Operator is pre-activating and activating providers efficiently, we recommend that the System Operator:
  - Report activation events to the Authority in detail including:
    i. Number of pre-activated providers and their offer costs
    ii. Number of activated providers and their offer costs
    iii. Providers selected and rationale behind the choice
    iv. Total availability and activation costs.
  - Published redacted or aggregated information on its pre-activation/activation events (removing commercially sensitive information).

RECOMMENDATIONS

Our recommendations for policy settings for a potential ERS is summarised in the table below.

Table 1: Recommended policy settings for ERS

Policy setting	Recommended settings
Participant pool	Only demand-side and unregistered generators that can meet System Operator additionality requirements allowed to participate
Procurement timeframe	Short-notice only: 4 weeks to 1 week ahead. We have assumed that the System Operator would run a competitive tender process to appoint a panel of pre-qualified providers (possibly with prices negotiated beforehand) that it can efficiently procure from.
Procurement trigger	• The NZGB would be used as a procurement trigger. ERS will be procured once N-1 balance falls below zero. Inherent uncertainty is incorporated into the N-1 balance forecast by using a POE wind forecast and correlated demand. • Security Standards (Expected Unserved Energy greater economic threshold for load shedding) are used to inform System Operator market testing activities
Pre-activation and activation triggers	• Pre-activate if NRS/WDS residuals fall below zero. Forecast uncertainty is incorporated into the NRS/WDS residuals using historical errors and their drivers to construct a probability distribution of residuals. • Activate if NRSS/RTD residuals (calculated in hour prior to real-time) falls below zero
Compensation mechanism and relationship to VoLL	• Providers can specify preparation, availability, pre-activation and activation fees to ensure the fee structure closely matches their cost structure. • System Operator should make reasonable endeavours to ensure that at the time of procurement, the forecast cost of ERS provision (based on forecast shortfalls) is not greater than the VoLL. • VoLL reviews should be conducted once every three years.
Scarcity pricing signals and potential double payment issue	Instructed ERS activation is added back onto nodal loads to maintain scarcity price signals. This will result in some double payment by loads who pay both scarcity prices and fund activation costs. As indicated above, this can be addressed by requiring the System Operator to restrict market and pre-activation/activation costs to be less than VoLL. We recommend adopting this approach as long as it does not result in material additional costs.
Cost allocation	• Allocate non-event costs (preparation and availability fees) nationally to loads based on share of monthly metered consumption • Allocate event costs (pre-activation/activation fees) nationally to loads based on share of metered consumption during activation events.

Policy setting	Recommended settings
Information provided to market	• Annual EUE assessment setting out quantum, location and duration of potential shortfalls • Publication of NZGB N-1 balance • Quarterly updates of procurement activities • Quarterly updates of activation activities • Standardised contracts on System Operator website • Expressions of interest

Contents

1 INTRODUCTION

1.1 CONTEXT

The Energy Competition Taskforce has proposed two work packages to strengthen the performance of the electricity market in the short to medium-term:

- Package 1 relates to measures to enable new generators and independent retailers to enter, and better compete in the market.
- Package 2 relates to providing more options for consumers and includes four initiatives or tasks. Task 2D is part of this package and pertains to rewarding industrial consumers for providing short-term demand flexibility. The Authority notes:

  "*One of the ways to help manage our electricity supply is to lower demand at peak times. Industrial plants that use a lot of electricity can make a meaningful contribution to this by using less electricity when it's scarce and expensive.*

  The Task Force is considering measures that would enable industrials to be appropriately rewarded for the benefit their flexible electricity use brings to the system, freeing up more supply and reducing the need for more expensive electricity generation to manage peaks. This could also provide industrials with an additional revenue stream."

To support Task 2D, the Authority has engaged Robinson Bowmaker Paul (RBP) to evaluate potential design options for an Emergency Reserves Scheme (ERS) like the Reliability and Emergency Reserve Trader (RERT) scheme in the Eastern Australian National Electricity Market (NEM), but with a focus on utilising demand side resources instead of in- or out-of-market generation resources.

Decisions on whether to progress with an ERS will be made as part of the consultation process. The output of this project will inform the Authority's Task 2D Consultation Paper.

This project is part of a wider project around demand response initiatives.

1.2 The role and purpose of a potential Emergency Reserves Scheme

1.2.1 The purpose of an ERS

The New Zealand electricity spot market is designed to provide accurate price signals and other information to market participants to enable efficient decision making. This includes:

- Guiding long-term decisions on generation and transmission investment.
- Guiding shorter term decisions on whether to offer generation, go on plant outage or reduce consumption.

For the vast majority of the time, price signals and other information provided to market participants are accurate enough such that the market is able to ensure that demand and supply are balanced. For example:

- Pre-dispatch schedules indicating low wind or hydro availability should deter thermal generators (particularly those with long start-up times) to go on maintenance outages and instead, offer in as they are more likely to be dispatched.
- Pre-dispatch schedules indicating scarcity prices should incentivise retailers to trigger demand response schemes they have in place to reduce their exposure to the high spot prices (e.g., the load curtailment arrangement Meridian has with the Tiwai smelter).

However, there have been historical incidents in which the market has been unable to balance demand and supply resulting in involuntary load shedding. The low residual event on 9 August 2021 is one such example. Involuntary load shedding of 3% of demand was required to balance supply and demand due to a confluence of adverse events or "perfect storm", which included wind output being lower than what was forecast (meaning thermal plant that may have otherwise been available did not have time to start up), demand increasing due to weather conditions and loss of generation from Tokaanu (due to lake weed blockage).

These perfect storm events are rare, and it is not reasonable to expect the market to be able to balance supply and demand when such events occur. When such events occur, involuntary load shedding is the last resort mechanism to balance supply and demand. If an ERS were implemented, it could be used to limit the quantity of this involuntary load shedding in the rare instances that the market is unable to balance supply and demand.

In the remainder of this document, we assume that *the purpose of a potential ERS is to be a penultimate resort service that sits between the market and involuntary load shedding to enable the*

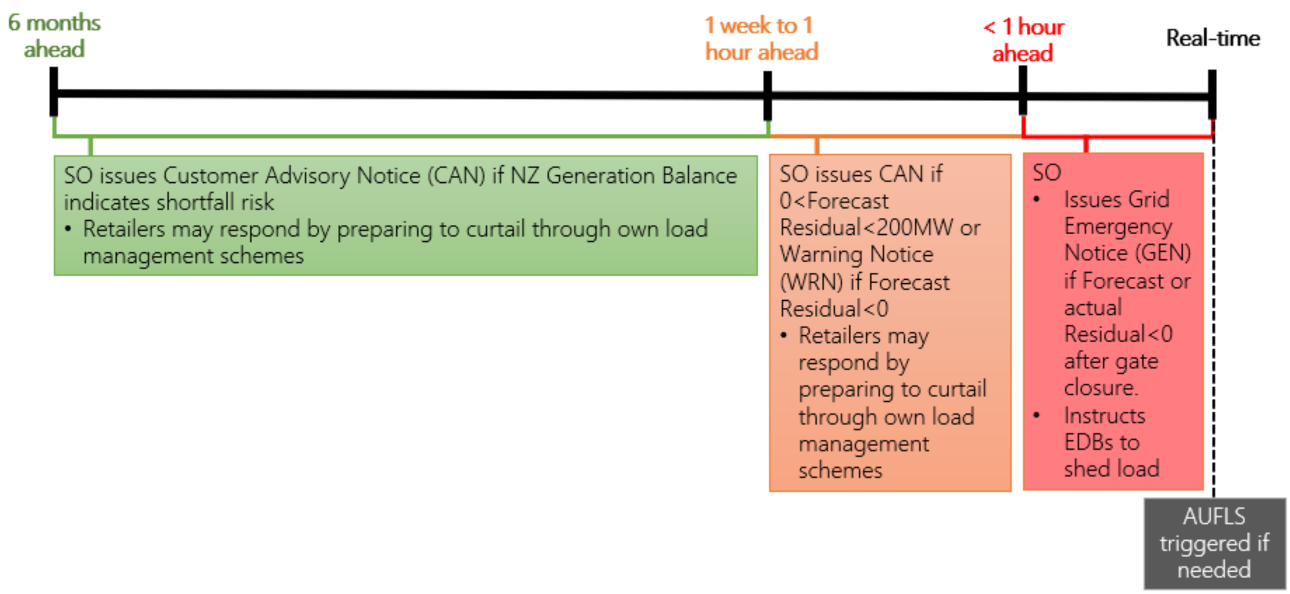
System Operator to limit involuntary load shedding to economic levels during times of extreme stress.

1.2.2 Integration of ERS with existing System Operator processes

Given that an ERS would sit between the market and involuntary load shedding mechanisms, it is useful to consider how an ERS would fit into the System Operator's existing processes for managing low residual conditions.

The System Operator has well-defined process for monitoring and managing low residual conditions using market mechanisms. Figure 1 summarises the current approach to monitoring and managing security of supply from six months ahead of real-time, to weeks and hours before real-time.

Figure 1: Overview of how market currently handles shortfall situations



The System Operator starts monitoring potential shortfalls six months ahead of real-time:

- If the New Zealand Generation Balance (NZGB)[7] indicates a margin of less than 200MW, then the System Operator issues a CAN to the market. This may result in retailers preparing to curtail through their own load management schemes they have contracted with their customers.
- Between one-week to one-hour ahead of real-time, if forecast residuals[8] indicate shortfall risk, the System Operators will issue a CAN and, if necessary, a WRN. At this stage:

[7] See Section 2.3.1 for a discussion on the NZGB.

[8] See Section 2.3.1 for a discussion on residuals.

  - Pre-dispatch schedules will show scarcity prices which will incentivise retailers to manage their load to reduce exposure to scarcity prices. This is price-responsive behaviour which indicates that the market is functioning well – *any ERS should not dilute incentives for such price-response*.
  - Generators who have not offered into the market may make generation available as they are likely to be dispatched – *any ERS should not displace such resources*.

- One-hour ahead of real-time (post gate-closure), if forecast residuals are still below zero, the System Operator will issue a GEN and instruct EDBs to shed load. This will result in involuntary load shedding.

- In real-time, if demand still exceeds supply and the frequency drops below 48 Hz, Automatic Underfrequency Load Shedding (AUFLS) will be triggered to balance supply and demand. This will result in further involuntary load shedding.

Given the purpose of the ERS is to be a penultimate last resort used in rare circumstances, it seems prudent to integrate the ERS into the System Operator's existing low residual monitoring and management processes as opposed to re-inventing the wheel. *Particularly, it is logical to activate Emergency Reserves (roughly) in the hour ahead of real-time before EDBs are instructed to shed load, as this would limit the level of involuntary load shedding, while ensuring that price-led demand response or late generator offers are not compromised*.

In proposing options for the design of the ERS, we have therefore assumed that *the ERS will be an additional resource in the System Operator's market toolkit when managing low residual conditions*. We have additionally assumed that:

- Activation (or dispatch) of Emergency Reserves would occur no earlier than one-hour prior to real-time to ensure all potential market response has been exhausted.

- The scheme would be analogous to how the System Operator currently manages and operates its Grid Support Contracts.

- The System Operator will be contract counterparty and responsible for procurement and activation.

- Procurement will occur through a competitive tender process.

- In terms of settlement, the System Operator will provide the costs to be collected from each market participant to the Clearing Manager, who will collect the payments and pay providers as part of the market clearing process.

1.3 Glossary

Activate	Refers to the System Operator dispatching an ERS resource to curtail their consumption or increase generation as relevant.
The Code	The Electricity Industry Participation Code
Emergency Reserves	Reserves procured, pre-activated, or activated under the Emergency Reserves Scheme (ERS)
Lead-time	The window of time between an ERS provider being pre-activated and activated.
NRSL	Non-price responsive schedule long. This pre-dispatch schedule is produced every two hours for the next 72 trading periods (36 hours). The schedule does not include difference bids or non-conforming bids.
NRSS	Non-price responsive schedule short. Similar to NRSL but produced every 30 minutes for the next 8 trading periods (4 hours).
NZGB	Non-price responsive schedule long. This pre-dispatch schedule is produced every two hours for the next 72 trading periods (36 hours). The schedule does not include difference bids or non-conforming bids.
Pre-activate	Non-price responsive schedule short. Similar to NRSL but produced every 30 minutes for the next 8 trading periods (4 hours).
Procure	Denotes the process used to identify, negotiate with, select, and contract resources to provide Emergency Reserves. For avoidance of doubt, procurement excludes pre-activation and activation.
RERT	Refers to the Reliability Emergency Reserve Trader scheme operated by Australian Energy Market Operator (AEMO) in the National Energy Market (NEM) of Eastern Australia.
Residuals	Denotes the capacity margin as calculated by SPD and given by the minimum of: • Total energy offered less energy cleared; and • Maximum offered capacity less energy cleared less the maximum of cleared Fast Instantaneous Reserves (FIR) and Sustained Instantaneous Reserves (SIR).
RTD/RTDP	Real-Time Dispatch/Real-Time Dispatch (Pricing) Schedule. The RTD schedule is the basis for dispatch and is produced every five-minutes for the upcoming five-minute dispatch interval. RTDP is run when the System Operator notes insufficient generation and Scarcity Prices occurring during pre-dispatch schedules and instructs load shedding. This load shedding is added back onto the relevant nodal loads in the RTDP schedule so that the Scarcity Price signals remain despite the load shedding.

Scarcity Price Limits	The prices associated with the three Energy Scarcity tranches modelled in the System Operator's SPD software. The Scarcity Price Limits are as follows: • 5% of load tranche valued at $21,000/MWh • 15% of load valued at $31,000/MWh • 80% of load valued at $50,000/MWh.
Security Standards	The set of margin indicators developed by the Authority under Part 7 of the Code, which includes: • New Zealand Winter Energy Margin (NZ WEM) • South Island Winter Energy Margin (SI WEM) • North Island Capacity Margin (NI CM)
SPD	The System Operator's Scheduling, Pricing and Dispatch software that calculates energy and some Ancillary Services dispatch schedules and prices in accordance with Schedule 13.2 of the Code
WDS	Weekly dispatch schedule. This pre-dispatch schedule is produced once a day for the next 7 days.
Value of Lost Load (VoLL)	The value different customers place on consuming electricity. This represents the $/MWh value above which customers have a financial incentive to cease consuming.

1.4 Structure of this report

The remainder of this report is structured as follows:

- In Chapter 2, we discuss different policy options and their potential implications with respect to the various design attributes of the ERS.
- Policy evaluation using multi-criteria analysis and the resulting preferred policy settings are provided in Chapter 3.
- Detailed evaluation of the strawmen schemes is included in 3.3Appendix A.

2 POLICY ISSUES

In this section we set out policy questions/issues that need to be considered if an Emergency Reserves Scheme were to be implemented in New Zealand. For each issue/question, we explore different design settings and briefly discuss the implications of the design setting.

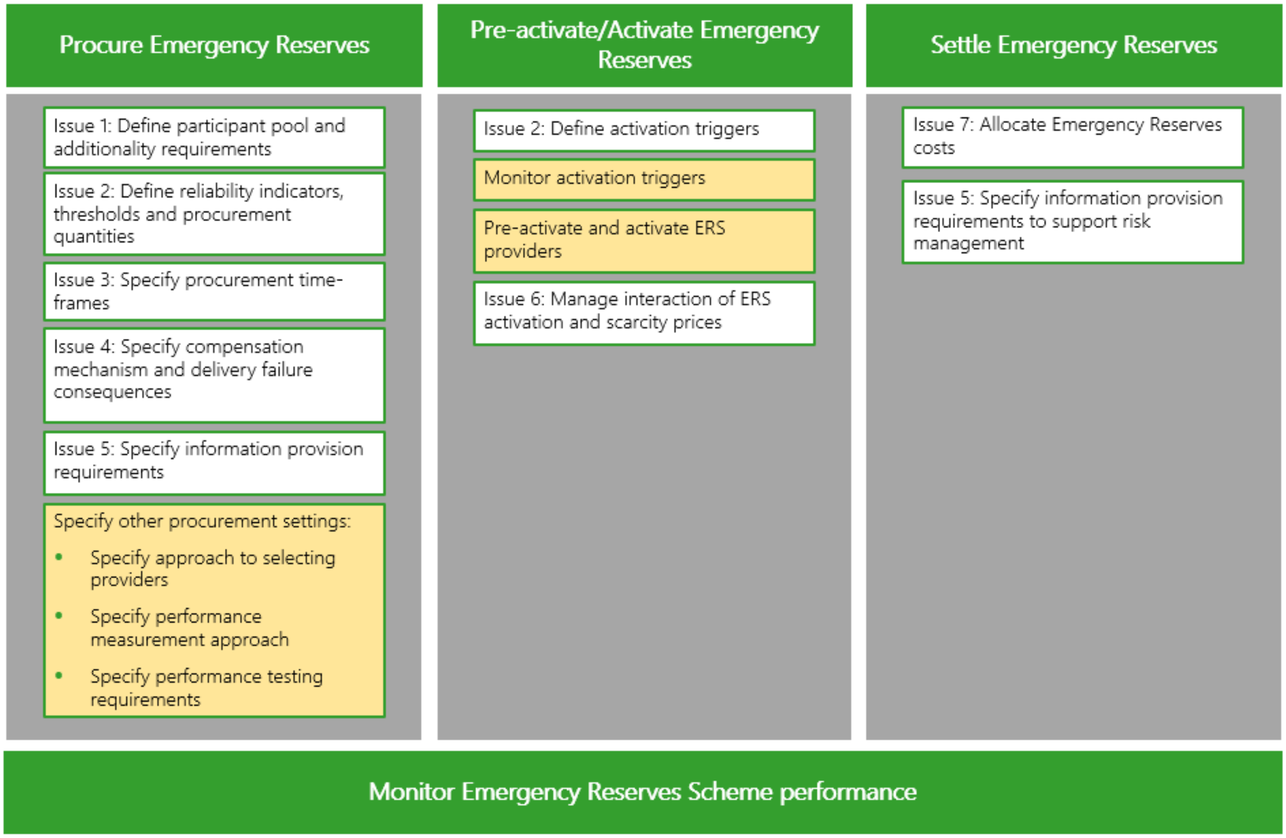
Selected design settings from this chapter are then used to create "Strawman" schemes which we evaluate in Chapter 3.

2.1 DESIGN ATTRIBUTES

Figure 2 groups policy issues into the various stages of the ERS: from procurement, through to pre-activation and activation, cost allocation and provider payment, and monitoring and oversight.

We recommend that attributes in yellow boxes should be left to the System Operator's discretion, with transparency around its decision making achieved through information provision settings (see Issue 5).

Figure 2: Design attributes associated with an Emergency Reserves Scheme (ERS)



In the remainder of this chapter, we discuss each policy issue separately.

2.2 Issue 1: Participant pool

This section is divided into two parts:

- Section 2.2.1 introduces the concept of additionality and discusses why it is instrumental in ensuring a potential ERS delivers value to New Zealand consumers.
- Section 2.2.2 explores the pros and cons of including different types of demand and supply-side providers to participate in a potential ERS.

2.2.1 Additionality of resource

Additionality is the term for ensuring a demand-side (or other) resource's paid service is additional to what the resource would have done in response to price or other incentives. In the context of an ERS, this means that a load that would have curtailed anyway (either in response to prices or because it is part of a load reduction programme) should not be allowed to provide Emergency

Reserves. Likewise, a generator that would have been exporting at the time Emergency Reserves is needed should not be allowed to participate either.

Paying a resource for something it would have done anyway will erode value delivered to consumers. For this reason, verification of resource additionality will be a critical component of determining whether a resource is eligible to participate in an ERS.

The System Operator requires resources to prove additionality to be eligible to provide Grid Support Contracts and has existing processes and tools to verify additionality. This includes:

- Verifying that loads providing demand response are not part of a load reduction program.
- Analysing historical data to verify whether a load has historically decreased consumption in response to price (irrespective of whether it is part of a load reduction program).

We recommend that a pre-requisite for any resource participating in an ERS is that its response must be additional to what it would have done in response to price-led or other incentives.

In the remainder of this document, we assume that only resources that can provide "additional" response would be allowed to participate in an ERS.

2.2.2 Types of providers

In this section we explore the impacts of different types of providers participating in an ERS.

Options include:

- Demand side:
  - Industrial customers
  - Commercial customers
  - Residential customers
- Supply side:
  - Generators (including Battery Energy Storage Systems (BESS)) that are not cleared through the spot market (i.e., unregistered generators)
  - Generators (including BESS) that cleared through the spot market but have not offered any capacity into the market at the time the Emergency Reserves are needed.

Demand side providers

Industrial sector

The industrial sector accounts for almost a third of New Zealand's electricity consumption but only 2% of Installation Control Points (ICP)[9]. A recent analysis of demand response potential in the industrial sector[10] noted:

- Large- and small-scale commercial and industrial consumers account for over half of demand response potential, with the highest potential being in cement and metal manufacturing, due to the flexibility of electricity demand for crushers, mills and smelting pots, and the large inventories typically present in those industries.
- Industrial demand response could potentially provide 6% of the typical daily peak demand and 2% of winter peak demand if relied upon at short notice. The report conservatively estimated that the scale of potential industrial DR ranges from 190MW to 300MW.
- While the technological costs of making a load demand-response capable is reasonably low, commercial considerations (i.e., costs incurred due to curtailing demand) can make demand response unattractive to industrial customers. Particularly, the cost of sustained demand response is high. Hence, substantial incentives are required to procure material quantities of demand response to cover such losses which can range from $1,000 to $5,000 per MWh (if load reductions are sustained). The report noted that strong financial incentives are therefore required to harness industrial demand response with financial incentives on the high end of the scale being $250,000 - $300,000 per MW (food manufacturing, wood, pulp & paper). Financial incentives for other industrial sub-sectors ranges $90,000 - $125,000.

Given the large potential from the industrial sector, the participation of industrial loads will be key to the procuring sufficient quantities of response to meet winter shortages. However, it is important to note that whether an individual load is incentivised to curtail consumption in return for a payment depends on the payment being greater than its individual Value of Lost Load (VoLL); i.e., the cost it would incur if it did not run its plant. At the same time, the load's curtailment would only benefit the market as a whole if the ERS payments are less than the average VoLL (averaged over different customer segments). Refer to Section 2.5.1 for a more detailed discussion on the interaction between VoLL, ERS remuneration and the energy price faced by a load.

[9] New Zealand Electricity Data Tables (data for the 2023 calendar year)

[10] Sense Partners: *Industrial demand flexibility: Sizing the potential for useful demand response* (March 2025).

Industrial customers are more likely to have controllable load components and reasonably predictable load patterns which will provide the System Operator greater certainty with respect to the quantum of service provided (as opposed to a volatile behind-the-meter (BTM) load that increases consumption, while other controllable components are curtailed resulting in a lower-than-expected level of curtailment). Transmission-connected loads will also have SCADA systems enabling performance monitoring at a lower granularity than 30 minutes. This will be important where an industrial load cannot provide a sustained response for an entire trading period or where the load can sustain a response over a trading period but where the System Operator needs to verify whether the response was sustained consistently (e.g., a steady ramp down) as opposed to bouncing up and down.

Commercial sector

The commercial sector accounts for a quarter of industrial consumption and 9% of ICPs[11]. The flexibility potential of commercial loads will be different to their industrial counterparts. The demand response study commissioned by the Authority (see footnote 9) cited a study by William & Bishop (2024) noting:

"*Commercial electricity demand is typically less flexible than industrial demand due to technical and economic constraints … But inherent thermal storage and occupants' thermal comfort preferences mean space heating and cooling of commercial buildings is considered a good candidate.*"

That is, the technical aspects of becoming demand-response capable (in terms of control and communications equipment) is more challenging for commercial customers than their industrial counterparts.

Residential sector

Residential loads are the most challenging to harness for demand response[12]. This is because:

- Residential loads are significantly more volatile than industrial and commercial loads, making their demand response less firm. This has two implications:
  - Demand response providers that are paid for curtailment typically have their service provision measured using baselining technologies which estimate a customer's demand in

[11] New Zealand Electricity Data Tables (data for the 2023 calendar year)

[12] A notable exception is hot water control. For example, distributors use ripple control to curtail hot water loads during peak times.

the absence of curtailment so that the quantity curtailed can be measured from meter data. A commonly used algorithm is the California Independent System Operator (CAISO) X/Y days approach in which the demand for a given trading period is calculated by averaging the demand during that period over the previous X days, where X is a member of a historical window comprising Y trading days. The average can be adjusted up or down if demand on the day of dispatch is higher or lower than the average. This algorithm works well for loads with reasonably stable profiles but can result in inaccuracies if the load is too volatile making the quantum of demand response less firm and making residential aggregation less attractive for aggregators.

  - Most consumer appliances do not have the smarts to communicate with a third party. While battery storage systems and smart EV chargers can be controlled by a third party, the volatility of the uncontrollable component of the BTM load means that any curtailment delivered through batteries or chargers could be cannibalised by the customer increasing consumption elsewhere. Aggregators in Australia have addressed this issue by recruiting high levels of redundancy into their aggregations to ensure they can meet their dispatch instructions.

- It may be challenging to find residential customers that provide additional value to the power system (over and above their usual consumption patterns). Residential customers will need to be aggregated by an aggregator (e.g., Enel X) to enable provision of Emergency Reserves. Such aggregators are unlikely to recruit load solely to provide Emergency Reserves. Instead, they would value-stack to allocate their aggregated resources to their highest value use. This may mean that aggregations that are capable of curtailing load during peak periods may already be curtailing to provide other flexibility services (e.g., to shift network load from peak periods to off-peak periods or providing interruptible load in spot market). This means that it may be challenging to find residential aggregations that provide additional value to the power system (see Section 2.2.1 for a discussion on additionality).

- There may be significant complexity in incorporating residential aggregations into the System Operator's existing low residual management processes. Residential aggregations will likely comprise hundreds and possibly thousands of individual loads and/or devices. Performance verification may require the System Operator to have visibility of all loads in the aggregation so it can use each load's meter data to verify if the portfolio has curtailed from their estimated baseline consumption. This might require the System Operator to implement new systems that can interface with the aggregator's Distributed Energy Resource Management System (DERMS).

This could potentially add significant cost to implementing an ERS. A lower cost and complexity alternative would be for the aggregator to provide aggregated meter data to the System Operator to enable performance verification. The System Operator may need to undertake occasional audits or spot checks of the raw meter data to ensure the aggregator is submitting accurate data.

Supply-side providers

Unregistered generators with spare capacity are a potential source of Emergency Reserves. As these generators do not participate in spot market trading (and are modelled as a decrease to demand for dispatch purposes), their participation is unlikely to distort spot market prices or result in capacity being withheld from the spot market. If such generators are allowed to participate, however, it will be important to only compensate them for the additional value provided, as opposed to compensating them for generating energy that they would have generated anyway. This means that the performance of such generators would need to be measured using a similar baseline method that is used for loads. That is, unregistered generators would be compensated based on the additional quantity they have generated over and above what they would have generated had there been no dispatch event.

Registered generators with unoffered capacity are another potential source for Emergency Reserves. In New Zealand, generators are not compelled to offer in their full availability; nor is the System Operator allowed to intervene and direct generators to run. This means that there may well be unoffered generation when the System Operator has issued an insufficient generation notice. Generators who have availability but do not offer will do so due to commercial reasons. As long as the scarcity price limits in the spot market at sufficiently high levels, available generators should be incentivised to offer in their capacity. Note that the Authority has recently increased the scarcity price limits as follows:

- The 5% load tranche previously valued at $5,000/MWh is now $21,000/MWh.
- The 15% load tranche previously valued at $10,000/MWh is now $31,000/MWh.
- The 80% load tranche previously valued at $20,000/MW is now $50,000/MWh.

Allowing registered generators to participate in the ERS could result in these generators withholding capacity in the spot market to be cleared via the ERS. This would not only distort the spot price but would mean consumers pay a spot market resource a price higher than the highest scarcity price limit (as a generator would only opt to be cleared via the ERS if it expected its ERS revenues to be greater than spot market revenues).

Given the recent increases to the scarcity price limits, spot market registered generators should be incentivised to make their capacity available during scarcity situations. For this reason, and to avoid distortions to spot prices, we recommend that registered generators not be allowed to participate in the ERS.

2.2.3 Participation options for evaluation

For the purposes of evaluation, we will consider the following options around participation:

- Restricting participation to the demand side only (including industrial, commercial, and residential sectors).
- Restricting participation to the demand side, but excluding residential customers.
- Restricting participation to the demand-side and unregistered generators (that are not cleared through the spot market) only.
- No participation restriction so that in addition to demand-side providers and unregistered generation, unoffered generation registered in the spot market can also participate.

For all four options, we assume that resources would need to meet additionality requirements to participate.

2.3 Issue 2: Reliability standards

As discussed in Section 1.2.1, the purpose of the ERS is to be a penultimate resort service that sits between the market and involuntary load shedding to enable the System Operator to limit involuntary load shedding to economic levels[13] at times of extreme system stress. This means that one or more indicators or reliability standards are needed to enable the System Operator to determine:

- Whether a credible risk of involuntary load shedding exists over procurement and activation timeframes.

[13] In the context of power system reliability, an economic level of load shedding represents the quantity of unserved energy above which the net benefits associated with additional generation investment in the spot market would yield a positive net benefit.

- The quantity of energy or capacity needed to close the “reliability gap”.

Determining the above requires developing a quantitative indicator that indicates involuntary load shedding is credibly likely to occur. As discussed in Section 1.2.1, the New Zealand spot market successfully balances supply and demand most of the time and that involuntary load shedding tends to occur during “perfect storm” events. It is further important to note that low wind output or high demand themselves are not the exacerbating factors – if wind is forecast to be low and/or demand high, then the System Operator’s existing processes will ensure that thermal generators receive the information and price signals they need to make themselves available. The problem is when wind is forecast to be available but then the wind speeds drop near or at real-time. At this stage, thermal plants who may be on planned outage may not have sufficient time to start up.

This means that a reliability standard used to trigger an ERS should incorporate both the inherent and forecast uncertainty of the drivers of involuntary load shedding. Continuing with the example of wind generation, a reliability standard should reflect:

- Uncertainty in wind generation due to variations in weather conditions; and,
- Forecast uncertainty in wind generation forecasts which may result in wind forecasts being materially under-forecasted.

The remainder of this section is organised as follows:

- We first discuss the various reliability indicators that are currently used for security of supply monitoring.
- We then discuss how these indicators could be used to trigger the procurement and activation of Emergency Reserves.

2.3.1 Existing Reliability Standards

There are three sets of reliability standards that are currently used for monitoring security of supply:

- The Security Standards developed by the Authority include the winter energy and capacity margins that the System Operator uses in the Security of Supply Annual Assessment (SOSA) to assess the ability of the electricity system to meet New Zealand’s needs over the decade ahead.
- The New Zealand Generation Balance (NZGB) is calculated by the System Operator and forecasts (200 days ahead) whether there will be sufficient generation capacity meet forecast demand securely.

- Forecast residuals are calculated by the System Operator seven-days ahead of dispatch to monitor whether there will be sufficient generation capacity meet forecast demand securely.

Each standard is described in further detail below.

Security Standards

The Security Standards are developed by the Authority under Part 7 of the Electricity Industry Participation Code. They include three margin indicators which measure system reliability over the decade ahead:

- The New Zealand Winter Energy Margin (NZ WEM) which is calculated as the ratio of expected available generation to expected demand.
- The South Island Winter Energy Margin (SI WEM) which is calculated as the ratio of the sum of expected available generation in the South Island plus HVDC contribution from the North Island to expected South Island demand.
- The North Island Capacity Margin (NI CM) which is the difference between the expected available capacity and expected demand in the North Island plus a the HVDC capacity contribution from the South Island. The HVDC contribution is calculated as a monotonically increasing function of the South Island capacity surplus.

The Authority developed these standards in 2012 and determined threshold quantities for each indicator as summarised below. The thresholds indicate the value below which the ratio of the benefit of additional generation investment to cost is greater than one. That is, if the margin falls below the threshold, then there is a net positive benefit in adding generation.

Table 2: New Zealand Security Standards as published in 2012

Security Standard	Threshold	Benefit to cost ratio	Expected Unserved Energy (EUE) as % of demand	Expected hours of shortfall
NI CM	630 – 780 MW	At margin of **690MW** the benefit to cost ratio is one.	N/A	At a margin of **690MW, 22** hours of shortfall can be expected
NZ WEM	14% - 16%	At a margin of **15%** the benefit to cost ratio is 1	At a margin of **15%** the EUE is **0.06%**	N/A
SI WEM	25.5% - 30%	Not published	Not published	Not published

Source: New Zealand Security Standards Assumption Document, 2012.

The standards above were developed in 2012 and are based on a specific set of assumptions around thermal outages, hydro inflows, wind availability, transmission losses, embedded generation, and the impact of price-led demand response. The benefit to cost ratio will have been developed based on generation costs and the value of lost load at the time the standards were developed. The Authority has advised that it plans to update the standards and thresholds later this year.

Transpower publishes the standards above for their annual SOSA under a range of different demand growth and generation investment scenarios. The purpose of the SOSA is informative as it provides a forecast of how different assumptions affect the margin values.

New Zealand Generation Balance

The System Operator publishes the NZGB daily, looking 200 days ahead to detect risks of low residual (generation-demand gap) and insufficient generation situations. The NZGB is used for capacity and outage planning.

There are two balances or margins that the System Operator publishes:

- N-1 balance. This is the system's capacity cover, over the peak, the loss of the largest risk-setter (generator (AC risk) or HVDC pole (DC risk)). This is calculated as
  - the available North Island generation capacity[14] plus
  - the HVDC contribution from the South Island less
  - the North Island demand less
  - the largest risk.
- N-1-G balance. This is the system's capacity to cover, over the peak, the loss of the largest risk setter if the second largest risk setter were also to become unavailable. This is calculated as:
  - the available North Island generation capacity plus
  - the HVDC contribution from the South Island less
  - the North Island demand less
  - 200MW of interruptible load less
  - the size of the largest risk less
  - the size of the second largest risk less
  - the North Island frequency keeping requirement (15MW).

[14] The NZGB assumes that any generation not notified through the Planned Outage and Capacity Planning (POCP) tool as being on outage is available for generation.

The system operator publishes a Planned Outage and Capacity Planning (POCP) assessment, highlighting potential shortfalls for instances where the base scenario with base load assumptions indicates an N-1-G shortfall and the shortfall is projected to occur within the next four weeks.

If the N-1-G balance falls below 200MW, then the System Operator issues Customer Advisory Notices.

As with the Security Standards, the NZGB is calculated under a range of scenarios with specific assumptions around wind and hydro availability. The NZGB does not consider the impact of forced outages on the generation balance,

Forecast residuals

From one-week ahead of real-time dispatch, the System Operator uses forecast residuals to monitor the generation balance. These forecast residuals are calculated as part of the pre-dispatch schedules as follows:

- The Week-Ahead Schedule (WDS) forecasts residuals a week ahead
- The Non-Price Responsive Schedule Long (NRSL) produces residuals for the next 36 hours (72 trading periods) every two hours.
- The Non-Price Responsive Schedule Short (NRSS) produces residuals for the next 4 hours (8 trading periods) every 30 minutes.

The residual[15] for a given trading period is defined as the minimum of:

- Offered energy balance (total energy offered less energy cleared); and
- Offered capacity balance (maximum offered capacity less energy cleared less the maximum of cleared Fast Instantaneous Reserves (FIR) and Sustained Instantaneous Reserves (SIR). This is analogous to the N-1 balance discussed in the section above.

If the residual falls below 200MW (but is greater than 0MW), then the System Operator issues Customer Advisory Notices (CAN) recommending load curtailment/management by customers. A Warning Notice (WRN) is issued if the residual is negative; if the residual is negative after gate closure, then the System Operator issues a Grid Emergency Notice (GEN).

---

[15] Unlike the NZGB calculations that assume any generator not on a planned outage is available, the residuals use offers submitted by generators to calculate availability. This means that unoffered generation is not included in residual calculation.

The residuals are calculated using the pre-dispatch schedules produced by the System Operator's Scheduling Pricing and Dispatch software. As such, their accuracy will depend on the offers submitted by generators; the further out from real-time, the less accurate the balance. The System Operator advises that generation offer changes are unlikely 12 hours out from real-time (reflecting the start-up time of the slowest generator); at this point the main drivers of uncertainty include wind, demand (which is correlated with wind) and unplanned outages.

2.3.2 Reliability standards for the Emergency Reserves Scheme

Reliability standards used to trigger procurement and activation of similar schemes have a single threshold value associated with them. If the indicator falls below this threshold, then additional reserves are procured or activated (depending on the timeframe). The difference between the forecast value of the indicator and the threshold is the reliability gap.

One of the key challenges in developing a trigger is measuring the uncertainty associated with the trigger value. As previously discussed, the reliability indicator or trigger must account for both inherent and forecast uncertainty, the latter accounts for uncertainty in forecast error. The New Zealand standards we have discussed above will have a specific value for a specific set of assumptions and system conditions; however, there are no probability distributions associated with these measures to enable specifying a threshold such that the probability of the indicator exceeding that threshold is 5% (or similar).

Additionally, procurement and activation triggers may vary depending on the procurement timeframe and even the lead-time needed to activate a provider.

By way of example, the procurement and activation triggers for the NEM RERT scheme is summarised in the table below:

Table 3: Procurement and activation triggers used in the RERT

	Procurement timeframe	Procurement trigger and threshold	Activation threshold and trigger
Long-notice	12 months to 10 weeks ahead of forecast shortfall		Lack of reserve condition (LOR2 and LOR3)[16] indicates

[16] The short-notice RERT is triggered by AEMO forecasting a LOR2 or LOR3 situation:

	Procurement timeframe	Procurement trigger and threshold	Activation threshold and trigger
Medium-notice	10 weeks to 5 weeks ahead of forecast shortfall	Expected Unserved Energy must not exceed 0.002% of demand	N-1 generation balance or generation balance will fall below zero.
Short-notice	5 weeks – 1 week ahead of forecast shortfall	Lack of reserve condition (LOR2 and LOR3) indicates N-1 generation balance or generation balance will fall below zero	

- Long- and medium-notice procurement is triggered in a region if the Expected Unserved Energy (EUE) is forecast to be greater than 0.002% of the region's demand. The quantity procured must be sufficient to close the reliability gap so that EUE equals 0.002% if the reserves are taken into account. The 0.002% threshold represents an "economic" level of involuntary load shedding such that if EUE exceeds this threshold, then additional generation becomes economic. The EUE is published annually for the next ten years in the Electricity Statement of Opportunities report and is calculated using Monte Carlo simulation and fundamental dispatch models to create a probability distribution of unserved energy. The published EUE is the sample mean of this distribution.
- Short-notice procurement, on the other hand, is triggered if AEMO determines a lack of reserve condition (LOR2, LOR3) is likely to occur; that is the generation balance, or the N-1 generation balance is forecast to fall below zero.
- All three types of emergency reserves are activated if AEMO forecasts a LOR2 or LOR3 condition. The LOR indicators are calculated using pre-dispatch schedules; hence it does not have uncertainty baked into it the way the EUE indicator does. That is, the LOR indicator represents the forecast balance for a specific set of system conditions, as opposed to the mean or some

---

- LOR2 is the equivalent of the N-1 balance published as part of the NZGB; LOR3 is declared if the difference between available generation and demand and the largest risk is forecast to fall below zero. Available generation takes interconnector support into account.
- LOR3 is declared if available generation capacity less demand is forecast to fall below zero.

quantile of a probability distribution. To address this, AEMO determines, in addition to the LOR2 and LOR3 balances, a Forecast Uncertainty Measure (FUM). AEMO has implemented a quantile regression model that uses forecast error (actual balance less balance forecasted via pre-dispatch schedules) as the response variable, and temperature, solar irradiance, forecast output of intermittent generators, demand forecast error, supply mix, current demand forecasting error and time at which the forecast is undertaken (ranges from 36 hours ahead to 30 minutes ahead) as predictor variables. The model predicts the 95th percentile of errors for a given set of system conditions. AEMO uses this error to create a 95% POE estimate of the N-1 generation balances. Hence, the FUM represents a value which will be exceeded only 5% of the time (given a set of system conditions). AEMO then calculates the generation and N-1 generation balance using the maximum of the forecast balance from pre-dispatch schedules (ST PASA) and the FUM[17].

As noted above, none of the New Zealand reliability standards incorporate uncertainty. As such, applying them to an ERS would be problematic as each standard would represent a specific set of assumptions that are uncertain. Below, we discuss potential options for amending the Security Standards, NZGB and residuals calculations to make them fit for purpose with respect to the ERS.

Amended Security Standards

The current Security Standards were developed over ten years ago and represent energy and capacity margins for a specific set of assumptions. While the thresholds for the indicators are linked to a level of EUE below which additional generation investment is economic, the values of these indicators as reported in the SOSA likely no longer represent that level of unserved energy (given the thresholds were developed over ten years ago).

We understand the Authority will shortly update the definitions and thresholds for these standards. In doing so, we recommend that the Authority consider the following:

- Redeveloping the EUE threshold to estimate what percentage of annual demand it is economic to serve. This would be associated with a specific energy-margin; however, the metric that we are interested in is EUE (not the corresponding margin).
- Annually calculating EUE for the year-ahead by using energy market modelling techniques to create a probability distribution for unserved energy. This will enable the Authority (or System Operator as relevant) to calculate the average level of involuntary load shedding (unserved

[17] The FUM is only applied during the first 36 hours of the one-week ST-PASA horizon when deciding to activate.

energy) that is expected. Additionally, the model should be able to inform the quantum, location, and duration of potential outages.

Notwithstanding incorporating uncertainty into the modelling, an EUE value estimated 12 months ahead is still not a good predictor for what may happen in real-time. Particularly, a 12-month ahead forecast will not be able to capture the effect of forecast uncertainty on EUE. For this reason, triggering procurement using the EUE indicator may result in inefficient procurement. However, the EUE indicator and stochastic modelling would be a useful resource to inform the System Operator's procurement activities as it will indicate the quantity, location, and duration of potential shortfalls.

Amended New Zealand Generation Balance

The NZGB indicators are analogous to the LOR2 and LOR3 indicators used by AEMO to trigger short-term procurement under the RERT and to activate all RERT reserves. The key difference, however, is that AEMO applies a Forecast Uncertainty Measure when activating the reserves.

In the New Zealand context:

- Option 1: The N-1 balance could be used to trigger procurement as long as the window between procurement and forecast shortfall is reasonably short. For example, in the NEM, AEMO will use the LOR2 and LOR3 indicators (with no uncertainty incorporated) to trigger short-term procurement (1-5 weeks ahead); however, it uses the EUE indicator for medium-term procurement (5 weeks – 10 weeks ahead). Hence, the N-1 balance may not be appropriate for procuring on a medium-term basis unless uncertainty is incorporated.
- Option 2: The N-1 balance could also be revised to incorporate better assumptions for wind output. The Authority plans to tender for an Intermittent Generation Forecasting Service that could provide quantile estimates. For example, to err on the side of caution, the System Operator could use a 90% POE wind estimate. However, as demand is correlated with wind, the demand forecasts used in the generation balances may not be consistent with the 90% POE wind forecast. For example, high winds and high wind output would be correlated with lower temperatures which in turn is correlated with the demand of temperature-dependent loads. If this approach were to be used, demand forecasts would need to be correlated with the wind forecast. While this approach incorporates uncertainty due to wind and demand output, it does not capture forecast uncertainty. Incorporating forecast uncertainty into the NZGB will be challenging. The quantile regression model used by AEMO to estimate a FUM when pre-activating RERT resources uses predictor variables that would not be applicable in the context of

the NZGB. For example, the FUM quantile regression model takes weather conditions and supply mix at the time the forecast is taken. This is because market outcomes in real-time will be correlated to such variables when the lag is short (e.g., when the forecast is being done a few days to hours ahead). The NZGB is determined months and weeks ahead – hence the weather conditions and supply mix at the time of the forecast would not be correlated with the corresponding real-time market outcomes. For this reason, for the purposes of a procurement trigger, we recommend incorporating inherent uncertainty only.

Amended residuals

Residuals are calculated one-week ahead of real-time. The offered capacity component of the residual calculation is similar to the N-1 generation calculation, except that it is calculated using the pre-dispatch schedules. Residuals can be used for activation but not procurement (as residuals are not calculated for periods more than a week ahead).

The System Operator advises that residuals calculated between 1 week to 36 hours ahead can be unreliable due to generation offer changes. In the last 36 hours prior to dispatch, offer changes are less likely and 12 hours prior to dispatch offer changes are unlikely; however, uncertainty due to wind output and demand remain.

As discussed in Section 1.2.2, activation of Emergency Reserves should occur within the hour prior to real-time (to prevent Emergency Reserves displacing price-led demand response or late generation offers). This means that the System Operator could use the more accurate residuals from schedules run within the hour prior to real-time[18].

However, the System Operator may need to pre-activate or arm ERS resources early than an hour-ahead if those resources require a longer lead-time[19]. As with the NZGB, the residuals do not have an uncertainty measure. 12 hours ahead of dispatch, the key source of uncertainty is wind (noting that wind output is correlated with demand). We note the following options to use the residuals to trigger pre-activation of Emergency Reserves:

- Option 1: Use the residuals as calculated currently to pre-activate resources. The schedule used to calculate the residual will depend on the resource's lead-time. For example:

[18] This could be either the NRSS or RTD schedule, depending on system and market conditions at the time.

[19] The period between the System Operator pre-activating the service and the service being activated.

  - WDS residuals could be used to pre-activate resources with a lead time between 36 hours and one week
  - NRS residuals could be used to pre-activate resources with a lead-time less than or equal to 36 hours.

  As these residuals are calculated several hours to days before real-time, there may be insufficient reserves in real-time if the wind output decreases materially and vice versa.

- Option 2: A second option would be to substitute the intermittent generators' wind forecasts in the WDS and NRS schedules with a POE wind forecast (provided through the new Intermittent Generation Forecasting service provider). As above, the demand forecasts would need to be consistent with the wind forecast. While this would account for uncertainty in the wind output, it does not account for uncertainty in the wind forecast (to cover the potential of the wind forecast being materially over-estimated).
- Option 3: A third option is for the System Operator to incorporate a Forecast Uncertainty Measure (per the RERT trigger mechanism) into the residual calculations by developing a probability distribution of forecast errors and using the right tail of the distribution to quantify the impact of forecast uncertainty. As pre-activation would occur hours to days (but no more than a week) ahead of real-time, sophisticated techniques could be to predict the right tail quantiles of the forecast error probability distribution more accurately. For example, the System Operator could use the quantile regression approach used by AEMO to pre-activate and activate RERT resources.

2.3.3 Summary of options for procuring and activating ERS

The different options for triggering the procurement, pre-activation, and activation of ERS are recapped in the table below.

Table 4: Summary of options for triggering ERS procurement and activation

	Security Standards	NZGB	Residuals
Role in triggering procurement, and pre-activating/activating reserves	• Trigger long-notice procurement • Inform market-testing activities to help with shorter-term procurement	• Trigger medium-term and short-term procurement	• WDS and NRS residuals used to pre-activate (arm) Emergency Reserves. • NRSS or RTD residuals used to activate Emergency Reserves

	Security Standards	NZGB	Residuals
Threshold value	Modelled EUE exceeds EUE threshold which represents "economic levels" of unserved energy in the year-ahead	NZGB N-1 balance is below zero	Residuals are below zero
Reliability Gap	MW difference between Unserved Energy assuming EUE threshold and modelled Unserved Energy	Quantity of the shortfall[20]	Quantity of the shortfall
Changes needed for use in ERS	• EUE threshold to be recalculated. • Annual calculation of EUE using energy market modelling to create probability distribution of unserved energy	• Option 1: No/limited changes – use NZGB as it is currently calculated; or • Option 2: Incorporate POE wind forecast assumptions into calculation (requires demand forecast to be correlated with wind).	• Option 1: No/limited changes – use WDS/NRS residuals as currently calculated; or • Option 2: Replace wind generator forecasts in WDS/NRS residuals with a POE wind forecast and correlated demand forecast; or • Option 3: Incorporate an uncertainty measure into WDS/NRS residuals using historical forecast errors to develop a POE forecast of the residuals.

2.4 Issue 3: Procurement timeframes

Procurement timeframe refers to window of time ahead of the forecast shortfall during which the System Operator would enter into contracts for Emergency Reserves.

[20] The EUE threshold for long-notice procurement assumes that a certain level of load shedding is economic. The NZGB and Residuals threshold, however, does not allow for any load shedding. To ensure the System Operator procures and activates reserves while also allow for economic levels of load shedding, we recommend that the System Operator use reasonable endeavours to ensure the unit cost of ERS provision is less than the VoLL.

The RERT procures emergency reserves using three different timeframes as previously noted. For New Zealand, the options are to implement one, two or three procurement timeframes.

While longer procurement timeframes (a year to months ahead) provide the procurer with more time to arrange contracts for reserves, the reliability indicators used to trigger such procurement are less certain that indicators calculated closer to the time of shortfall. Hence, a long-notice procurement scheme may result in either under- or over-procurement.

Medium-term (months to weeks ahead) and shorter-term (weeks to one week ahead) timeframes increase the certainty of shortfall forecasts; however, the procurer has less time to arrange contracts with providers. For this reason, AEMO operates a pre-qualified panel from which to draw RERT providers from in medium- and short-notice procurement. Providers on the medium-notice panel can negotiate prices prior to finalising their contracts while providers on the short-notice panel have pre-negotiated offer prices.

In Chapter 3, we consider three options for the ERS:

- Long-notice procurement (12 months to 3 months ahead)
- Medium-notice procurement (3 months to 4 weeks ahead)
- Short-notice procurement (4 weeks to 1 week ahead).

2.5 Issue 4: Compensation and delivery failure consequences

A compensation regime for a potential ERS should meet the following criteria:

- Providers of Emergency Reserves must be able to recover their costs through the remuneration received and earn a return. Financial incentives for curtailment will depend on the relevant load's Value of Lost Load (VoLL) or the value it places on consuming electricity and reflects the price at which the customer is indifferent to whether they consume or not. A load will not be incentivised to curtail if the ERS remuneration is less than its VoLL. Different loads will have different VoLL.
- New Zealand consumers should receive value from the service being procured and should not "over-pay" for the service. That is, the remuneration provided to ERS providers should not

exceed the average VoLL or monetary value that different types of customers place on having a reliable electricity supply[21]. Additionally, consumer should not pay:

- For curtailment that would have occurred any way as part of price-led demand response or other incentives.
- For services that are likely to not be delivered.

In identifying compensation mechanisms for an ERS, it is therefore important to consider:

- The interaction between a provider's VoLL, the average VoLL, ERS remuneration and the energy price faced by the provider.
- The types of costs that may be incurred by providers.
- Delivery failure consequences.

2.5.1 Relationship between remuneration, VoLL and energy prices

As indicated above, there are some complexities with respect to incentivising demand response from in a way that benefits New Zealand customers as a whole. A customer's incentive to curtail its consumption depends on the relativity of four variables:

- The $/MWh electricity price (P) faced by the customer. If the customer is hedged, then it is insensitive to scarcity prices and will not have an incentive to curtail demand to avoid high spot prices.
- The customer's individual Value of Lost Load ($\mathbf{VoLL_{provider}}$, $/MWh) which reflects the price at which the customer is indifferent to whether they consume or not. For example, an industrial customer with customer VoLL of $80,000/MWh would want to keep consuming as long as the power price is below $80,000/MWh as to cease consumption would incur additional costs.
- The average VoLL ($\overline{\mathbf{VoLL}}$, $/MWh) that is used to guide procurement decisions. For example, in the NEM, AEMO is not allowed to pay more than the regulated average VoLL when procuring and activating reserves procured under the RERT. Each state has its own VoLL which ranges between $30,000 - $50,000[22].

[21] The average VoLL is an averaged value over different customer segment types and will not necessarily be equal to the VoLL faced by providers of ERS (each of whom may have different VoLL values).

[22] AER: 2024 review of Value of Customer Reliability

- The remuneration (R, \$/MWh) offered to the customer for curtailing their consumption (for simplicity, we have assumed only an activation payment is provided).

A customer will be incentivised to curtail its consumption for a reward as long as R> $\mathrm{VoLL}_{\mathrm{provider}}$. However, the market would only benefit if R< $\overline{\mathrm{VoLL}}$. This means that the remuneration price for Emergency Reserves needs to satisfy the following: $\mathrm{VoLL}_{\mathrm{provider}} < \mathrm{R} < \overline{\mathrm{VoLL}}$.

However, a customer may also be incentivised to curtail its consumption if $\mathrm{P} > \mathrm{VoLL}_{\mathrm{provider}}$ in response to spot price signals. This could occur where the load's VoLL is less than the scarcity price and where the load is not fully hedged. Remunerating such a load to provide Emergency Reserves may result in paying the load for consumption it may have curtailed anyway in response to spot prices.

This means:

- ERS compensation to a potential demand-side provider must be greater than their VoLL ($\mathrm{VoLL}_{\mathrm{provider}}$) but less than the average VoLL ($\overline{\mathrm{VoLL}}$).
- The System Operator should not procure from loads that have an incentive to curtail in response to spot prices. The issue of additionality is discussed in more detail in Section 2.2.1.

In terms of ERS design, we recommend:

- The System Operator use reasonable endeavours[23] to ensure that the unit cost of ERS provision does not exceed the average VoLL.
- A three-yearly review of VoLL across different customer segments and sectors. VoLL values will affect ERS pricing, Scarcity Price Limits and then EUE threshold for the amended Security Standards. As such, it will be important to ensure that VoLL estimates are accurate.

2.5.2 Provider remuneration

In the above section, we noted that ERS compensation (overall) to a demand-side provider must exceed their individual VoLL (to incentivise participation) while remaining under the average VoLL (to provide value to New Zealand consumers). This provides an upper bound to the overall compensation offered to potential providers.

[23] It is imprudent to place a blanket requirement on the System Operator to ensure the unit cost does not exceed VoLL. Forecast uncertainty means that despite the System Operator's best efforts, the cost of ERS could still exceed VoLL.

In this section, we discuss how that overall compensation should be structured. In doing so, it is useful to consider the types of costs a demand-side provider may incur. This may include:

- Availability costs which reflect capital costs incurred in setting up communication, control and measurement systems and other fixed costs associated with being a provider.
- Preparation costs incurred in meeting requirements to be on the System Operator's panel of pre-qualified providers. This could include the cost of testing to verify the load is capable of providing the contracted response.
- Pre-activation or arming costs the provider incurs as it prepares for activation. For example, an aggregator would need to determine which of its loads to activate to provide the contracted response. This may incur cost (e.g., where labour is required to undertake the resource selection process to meet contractual requirements, portfolio restrictions and optimise the aggregator's commercial position). These costs are likely to be fixed per pre-activation event.
- Activation costs or the cost to reduce consumption. This will be a variable cost and will be a function of the provider's VoLL.

In terms of options for structuring ERS remuneration we consider the following options in our evaluation:

- Allow the provider to propose fees for all four cost categories (availability ($/MW), preparation ($/MW), pre-activation ($/MW) and activation ($/MWh)).
- Allow activation ($/MWh) and availability payments ($/MW) only. This is like the structure of payments offered by the System Operator in some of its Ancillary Services contracts where the provider can specify an availability and an event (activation) fee. If a provider's cost structure is such that they also incur preparation and pre-activation costs, they will account for this by increasing their availability and/or activation fees to ensure they are able to cover costs. In doing so they will need to make assumptions around the number of potential activations. If the provider is conservative in its assumptions (i.e., assumes more activations than what actually eventuates), then this will inflate total costs relative to if the provider had been able to specify a $/MW pre-activation cost.
- Allow activation payments only ($/MWh). That is, a provider is only paid if they are activated, and they provide reserves. As above, if the provider has capital, pre-activation or preparation costs to cover, they will inflate their activation fee to be able to cover those costs. In doing so, they may over-estimate the actual cost incurred.

2.5.3 Delivery failure consequences

Demand response contracts typically have financial penalties that providers must pay if they fail to deliver the required response when activated. In its Grid Support Contracts, the System Operator uses a partial financial clawback approach in which the provider is not paid the activation and availability fees for periods where it has failed to deliver. This level of partial clawback does not provide a strong enough incentive for the provider to be available. The System Operator has concluded that financial penalties that reflected the actual cost of non-performance would be a barrier to entry resulting in few resources wanting to participate. For this reason, in addition to the financial clawback, the System Operator practices the following with respect to its Grid Support Contracts:

- Before entering into a contract, the System Operator performs technical and commercial due diligence to ensure that the service providers under the will be available and operate as required.
- Contracts include requirements for testing, with the System Operator reserving the right to make calls to test operational readiness.
- The System Operator considers resource fatigue risks[24] in evaluating individual or sets of proposals and in developing its call strategy.

We recommend that similar arrangements be incorporated into an ERS to address delivery failure risk.

2.6 ISSUE 5: INFORMATION PROVISION

Providing information to stakeholders would be an essential key in unlocking participation in the ERS. Participants of any market can only make decisions based on available information.

In the context of the ERS, potential providers may want to know the quantity of service needed by location (i.e., where service is relevant), the timeframe in which the service would be acquired or needed, along with the likelihood of being activated. The System Operator best placed to provide this information. We recommend that the System Operator provide the following key information:

- Annual EUE assessment setting out quantum, location, and duration of potential shortfalls. For long-notice procurement, the annual assessment would trigger long-notice procurement (if the

[24] This is where a resource's performance may be adversely impacted if it is activated too many times.

EUE threshold is violated). Even in the absence of long-notice procurement, an annual EUE assessment will still be useful as it can be used by the System Operator to plan its market testing activities in advance of procuring under medium- and short-notice schemes. Moreover, this assessment will also provide potential providers with information on future requirements. Such a publication should include an assessment of the quantity, location, and duration of shortfalls.

- Regular publication of the N-1 NZGB (the System Operator does this already).
- Quarterly updates of procurement activities setting out:
  - Quantities procured
  - Contracted compensation
- Quarterly updates of activation activities setting out:
  - Frequency and quantity of shortfalls
  - Quantities pre-activated by dispatch event
  - Quantities activated by dispatch event
  - Total costs by dispatch event
  - Unit cost by dispatch event.
- Standardised ERS contracts
- Expressions of Interest for market testing including (but not limited to) the service specification the System Operator is seeking and the technical requirements.

2.7 Issue 6: Manage interaction of ERS and Scarcity Pricing

The NRS schedules should reflect scarcity prices at the time Emergency Reserves are activated (as ERS would be activated if the residual falls below zero indicating insufficient generation).

If demand is decreased in real time (due to ERS activation), the scarcity prices may disappear as there may no longer be insufficient generation (depending on how much ERS is activated). This will distort the scarcity pricing signals in the market as the spot prices will no longer signal scarcity.

This issue also arises when distributors are requested to shed load in their local networks. This is addressed in the RTDP schedule by adding the instructed load shedding back onto the demand forecast at the relevant nodes. This restores the scarcity prices.

The same approach could be used for the ERS in such a way that activated Emergency Reserves are added back onto the demand forecast in the RTDP schedule to restore the scarcity prices. Note that this may be challenging for aggregated resources with multiple loads in different locations; here the System Operator would need designate the nodes at which the provider is deemed to be curtailing load.

For the purposes of the evaluation in Chapter 3, we will consider the following options:

- Do not include activated Emergency Reserves as instructed load shedding the RTDP schedule; or
- Include activated Emergency Reserves as instructed load shedding the RTDP schedule.

2.8 Issue 7: Cost recovery

ERS provider costs and remuneration can be broken down into two components:

- Event costs that are only incurred when a resource is pre-activated or activated. This includes pre-activation and activation fees.
- General fixed costs that are incurred irrespective of whether a resource is activated and will include preparation and availability fees.

We address each type of cost separately.

2.8.1 Allocating Event Costs

In terms of allocating ERS event costs, there are three broad options:

- Allocate event costs to all generators registered in the spot market in proportion to metered generation during event.
- Allocate event costs to generators and loads in proportion to the absolute values of their metered generation and consumption respectively during event.
- Allocate event costs to loads only in proportion to metered consumption during event.

Each option is discussed in further detail in the table below. The above options are used in evaluating cost allocation settings for event-based costs in Chapter 3.

Table 5: Options for allocating ERS event costs

	Comments
Allocate event costs to generators only	Generators could be considered "causers" of the shortfall, if we were to assume that the shortfalls are occurring as a result of generators not offering available capacity or building new

	Comments
	generation. The rationale behind a causer pays allocation would be that generators would offer available capacity or build new generation to avoid these costs. However: • The decision to not offer capacity into the market may be due to genuine availability issues as opposed to commercial reasons; moreover, the increase in the scarcity price limits should incentivise generators withholding capacity for commercial reasons to offer that capacity into the spot market. • Decisions to build new plant are unlikely to occur simply to avoid ERS charges – the spot market price signals are the more important factor here. Moreover, delays in building new generation (that may result in insufficient generation) can be outside the generator's control. Given the above, generators are more likely to pass the costs down via contracts with loads than they are to build a new plant in response to avoid ERS cost exposure. These costs are likely to be passed by retailers down to their consumers. However, because the retailers are not directly being allocated the cost, they will not have the information they need to trade-off the cost of implementing load management schemes versus paying the charge. That is, because the cost is not explicitly linked to the consumption of their consumers, it will be challenging to model how to reduce that cost.
Allocate event costs to generators and load	Under this approach the cost of ERS would be roughly split between generators and loads. However, as above, it is likely that generators will pass their costs down to their contracted loads so that consumers end up paying for the ERS anyway. This approach does, however, provide retailers with better information to manage the consumption of their loads to avoid their share of the charges.
Allocate event costs to loads only	Under this approach costs are allocated only to loads as the beneficiaries of the scheme. This would provide retailers better information with respect to facilitating load management amongst its customers.

Double payment issue

As discussed in Section 2.7, failure to include activated Emergency Reserves as instructed load shedding in the RTDP schedule may result in scarcity prices no longer binding. This will distort scarcity pricing signals and adversely affect investment signals. This issue can be addressed by adding the Emergency Reserves back onto the relevant nodal forecasts to maintain the Scarcity Prices in real time. Adding activated reserves back on as instructed load shedding and allocating ERS costs to loads will result in double payment. That is, loads will pay the scarcity price for demand cleared through the spot market plus ERS activation costs. One way to mitigate this issue is to require the System Operator to restrict activation levels such that the total market costs paid by loads (scarcity prices times load served) and ERS activation costs as a proportion of load served is less than or equal to VoLL. This will ensure that consumers are not paying more than VoLL at times of scarcity.

It is worth noting, however, that the quantum of this "double payment" may be relatively minor, especially if event costs are similar to scarcity prices. For example, the price observed by loads paying event costs during scarcity conditions, in which Emergency Reserves are activated to reduce overall consumption by $x\%$, would be:

$$Price\ paid\ by\ loads\ (\$/MWh) = \frac{Scarcity\ Price\ (\$/MWh) \times Consumption + Activation\ Fee\ (\$/MWh) \times x\% \times Consumption}{Consumption}$$

$$= Scarcity\ Price\ (\$/MWh) + \ x\% \times \ Activation\ Fee\ (\$/MWh).$$

If the scarcity price and activation are similar in magnitude, and if the ERS is activated rarely, then this additional cost is likely to x % of the scarcity price. x is likely to be in the order of 1%-3%[25] and therefore relatively minor.

For the evaluation in Chapter 3, we consider two options with respect to addressing the double payment issue:

- Requiring System Operator to restrict activation levels so that total market and ERS activation costs as a proportion of load served is less than or equal to VoLL.
- Not requiring the above.

National versus island allocation

Costs can be allocated nationally or by island; so that reserves procured/activated to cover shortages in the North Island will only be recovered from North Island loads. This will result in the costs being spread across a smaller customer base. Given the lower demand in the South Island, this may result in very high charges for any reserves procured or activated in the South Island.

Moreover, the NZGB and residuals assume interconnector support; that is, energy sent from the South Island is included in the North Island margin. For example, let us say that South Island is projected to have shortfalls, but the HVDC is flowing north so they have no interconnector support. The shortage in the South Island could be abated by reducing the interconnector support to the

[25] For example, see the Authority's review into the 9 August 2021 incident.

North Island. Hence, the North Island benefits from the interconnector support while the South Island is worse off.

For the above reasons, we recommend considering allocating costs nationally as opposed to by island.

2.8.2 Allocating non-Event costs

As non-event costs are incurred irrespective of activation, we recommend that costs like availability and preparation costs be allocated as part of the monthly settlement process in proportion to metered generation/consumption. For the purposes of the evaluation in Chapter 3, we consider the following options:

- Allocate non-event costs to all generators registered in the spot market in proportion to monthly metered generation.
- Allocate non-event costs to generators and loads in proportion to the absolute values of their monthly metered generation and consumption respectively.
- Allocate non-event costs to loads only in proportion to monthly metered consumption.

2.9 Monitoring the ERS

Monitoring the performance of a potential ERS will be important to ensure that the scheme meets its objectives. This is particularly as we have recommended some critical design attributes be left to the System Operator's discretion. This includes:

- The approach to selecting providers from offers during the procurement process.
- Performance verification and measurement including the baseline methodology used to measure the quantity curtailed.
- Selecting providers to pre-activate and activate.

Each bullet is discussed in more detail below.

2.9.1 Provider selection during procurement

While we are proposing a requirement for SO to ensure forecast costs at the time of procurement are not greater than VoLL, we are not proposing a requirement for System Operator to select the least cost combination of compliant providers. Some providers may be "better" than others due to short lead-times and predictable load patterns – it is reasonable for the System Operator to select such a provider over a provider with less firm response and a longer-lead time, even if the latter has lower offer costs.

Ideally, the System Operator should select the combination of offers that minimise overall cost, while giving preference to providers with short-lead times and firmer responses.

Rather than making this a regulatory requirement, we recommend that the System Operator be required to report procurement activities to the Authority in detail including:

- Number of providers and their offer details
- Providers selected and rationale behind the choice
- Forecast ERS cost based on selection.

Reporting could be required under the Code or enforced through the SOSPA.

2.9.2 The baselining methodology used to measure service delivery will be left to SO discretion

Both demand and supply-side providers should be compensated based on the additional value they add. That is:

- A demand-side provider should be compensated based on load curtailment relative to what they would have otherwise consumed.
- A supply-side provider should be compensated on the additional generation output relative to what they would have otherwise generated.

This means that the System Operator requires a methodology or algorithm for estimating what a provider's baseline demand or generation levels are (i.e., what they would have consumed or generated had there been no activation event). We refer to this methodology as a baseline methodology.

Poorly chosen algorithms can result in:

- Providers gaming the methodology to artificially raise their baseline demand to get higher payments.
- Inaccurate baseline forecasts can result in providers being under- or over-compensated.

We recommend allowing the System Operator discretion to select a baselining methodology as this is an evolving space. Discretion will enable the System Operator to leverage off improvements in this area in the future. Moreover, the System Operator can negotiate different methods for different providers based on their load profile.

The System Operator could also use historical meter data from their providers to test how well their baseline methodology estimates demand.

2.9.3 Selecting providers to pre-activate and activate

While the System Operator must ensure that total market and ERS costs do not exceed VoLL, the combination of providers they select to meet this requirement is discretionary.

Again, rather than placing a regulatory requirement to minimise costs, we recommend an alternative to require the System Operator to report activation events to the Authority in detail including:

- Number of pre-activated providers and their offer costs
- Number of activated providers and their offer costs
- Providers selected and rationale behind the choice
- Total availability and activation costs.

The above could be required under the Code or enforced through the SOSPA.

3 STRAWMAN EVALUATION

In Chapter 2 we discussed a number of policy issues and specified different options for addressing those issues. In this chapter, we develop "strawmen schemes", each representing a potential design for the ERS. Each strawman is then evaluated against the Authority's statutory objectives using multi-criteria analysis and scored on how well they meet each objective.

The remainder of this chapter is structured as follows:

- The Evaluation Framework we have used to evaluate each strawman scheme is described in Section 3.1.
- Strawmen schemes are defined and evaluated against evaluation objectives in Section 3.2. Detailed evaluation against each outcome associated with each evaluation objective is provided in 3.3Appendix A.
- Recommended policy settings are summarised in Section 3.3.

3.1 EVALUATION FRAMEWORK

In this section we set out the framework and approach used to evaluate the strawmen designs against the Authority's statutory objectives.

The Electricity Industry Act (2010) [26] sets out the following statutory objective:

> *To promote **competition** in, **reliable supply** by, and the **efficient operation** of, the electricity industry for the **long-term benefit of consumers**.*

The statutory objective has three limbs: competition, reliability, and efficiency (each of which affects long-term benefits to consumers).

The Government Policy Statement further interprets the competition and efficiency limbs as follows:

> *An efficient wholesale electricity market with many different wholesale buyers and sellers of electricity, managing their own risks, responding to competitive pressures and accurate price signals, continually looking for ways to serve their current and potential customers more effectively than their competitors.*

[26] Interpretation of the Authority's statutory objective (www.ea.govt.nz)

To evaluate each strawman design against these objectives, it is important to specify the types of outcomes that would be observed if that objective were met.

To this end, we have developed outcomes associated with each objective. Each objective and its associated outcomes are described below.

3.1.1 Statutory Objectives and associated outcomes

A. Competition

From an economic perspective, more competition results in more benefits accruing to the consumer in the long-term. The more diverse the market, the likelier it is that competitive pressures will drive innovation up and price down.

To ensure a diverse market exists:

- there should be limited barriers to entry,
- providers should have access to high quality information to drive decision-making,
- providers should be able to earn a reasonable return on investment and
- providers should not face discrimination based on technology.

Additionally, policy settings designed to increase competition in the ERS should not have unintended adverse consequences on competition in the spot market.

The outcomes associated with the competition objective are summarised below.

Table 6: Competition outcomes

Objective	Desired Outcomes
Competition	• Providers have access to information that enables them to develop business models to provide emergency reserves; • Pricing incentivises entry by covering both operating and capital costs and enabling a reasonable return; • The ERS rules do not favour one technology type over another; • Competition in the ERS is not distorted by market power; and • Wider competition in the NZ electricity market is not distorted. For instance, ERS providers would not bypass participation in the spot market to participate in the ERS.

B. Reliability

In the context of the ERS, meeting the reliability objective well will require the System Operator to forecast Emergency Reserves requirements accurately so that involuntary load shedding can be avoided. Additionally, firmness of response will be important so that the System Operator can rely on the provider providing the required curtailment relative to the provider's baseline demand.

The outcomes associated with the reliability objective are summarised below.

Table 7: Reliability outcomes

Objective	Desired Outcomes
Reliability	• System Operator triggers procurement of Emergency Reserves appropriately when credible risk of capacity shortfall exists; • System Operator procures sufficient reserves to ensure unserved energy is restricted to economic levels; • ERS incentivises providers to be available. In other words, ERS must disincentive participant non-performance; • ERS service delivery or performance can be measured accurately so that System Operator has assurance of delivery; and • Providers are given sufficient notice for resources to be available.

C. Efficiency

Meeting the efficiency objective will require:

- Limiting unintended adverse interactions with spot market price signals. The accuracy of spot market price signals is critical to driving efficiency decision making with respect to operations and investment.
- Ensuring consumers are protected and benefit in the long-term.

The outcomes associated with the efficiency objective are summarised below.

Table 8: Efficiency outcomes

Objective	Desired Outcomes
Efficiency	• ERS does not distort wholesale market pricing by incentivising wholesale market participants to withdraw capacity • ERS does not distort scarcity pricing signals in the wholesale market

Objective	Desired Outcomes
	• ERS does not displace cheaper energy resources in the spot market • Loads/retailers have access to information to accurately identify, evaluate and manage their risks • ERS provides long-term value to consumers • Small customer interests are protected where consumer resources are used to provide ERS

3.1.2 Cost and complexity considerations.

Different designs will have varying levels of complexity and cost associated with them. To enable a robust evaluation, the performance of a design must be evaluated against the Authority's statutory objectives and potential implementation requirements, so that benefits can be traded off against cost and complexity.

Cost and complexity are evaluated based on where changes are likely to occur due to implementation of a specific design. Sources of cost and complexity are summarised below.

Table 9: Objectives and Desired Outcomes

Objective	Desired Outcomes
Cost and Complexity of Implementation	• Code changes • Market Information System changes • Telemetry Requirements • Scheduling and dispatch tools and process changes • Settlement and reconciliation tools and process changes • Other process and tool changes

3.1.3 Evaluation Approach

To measure how well each strawman design does against the statutory objectives and cost and complexity criteria, we measure how well each design achieves the outcomes associated with each objective.

Table 10 summarises how each strawman's performance is scored against competition, reliability, and efficiency outcomes.

Table 10: Evaluation scoring for Competition, Reliability and Security, and Efficiency Objectives

Score	Criteria
○	No objective outcomes achieved
◔	Some objective outcomes achieved partially
◑	Some objective outcomes achieved substantially or most met partially
◕	Most objective outcomes achieved substantially
●	All objective outcomes achieved substantially

Table 11 summarises how each strawman's performance is scored against cost and complexity outcomes.

Table 11: Evaluation scoring for Cost and Complexity Objective

Score	Criteria
○	Wide ranging change / Significantly high costs
◔	Many changes / High Costs
◑	Moderate Changes / Costs
◕	Few changes / low costs
●	Negligible changes / costs

3.2 Strawman Evaluation

In defining strawmen, we have used an iterative approach to eliminate policy settings that drive poor performance against the evaluation objectives.

In this section, we summarise how each strawman performs against the evaluation objectives at a high level. Detailed evaluation of each strawman against each objective outcome is included in 1.1.11.a.Appendix A.

3.2.1 Initial evaluation

Strawmen definitions

In the initial evaluation, we have defined three strawmen schemes as summarised in the table below.

- All three strawmen assume there will be three procurement timeframes: long-, medium- and short-notice.
- The structure of payments offered to provider varies across the three schemes.
  - Delivery failure consequences are handled through partial clawbacks and procurement due diligence checks as discussed in Section 2.5.3.
- The participant pool is varied across the strawmen to determine participation impacts on competition and efficiency.
  - In all three schemes, we assume additionality is a requirement and that the System Operator will not procure from a provider who does not meet additionality requirements.
- The approach to calculating the procurement and pre-activation triggers is varied across the strawmen.
- ERS activation is not included as instructed load shedding in the RTDP schedule for any of the strawmen.
- Each strawman has a different cost allocation methodology.

Table 12: Policy settings for Strawmen 1a, 1b and 1c.

	Strawman 1a	Strawman 1b	Strawman 1c
Procurement timeframe	• Long-notice: 12 months to 3 months ahead Medium-notice: 3 months to 4 weeks ahead Short-notice: 4 weeks to 1 week ahead	Same as Strawman 1a	Same as Strawman 1a
Procurement trigger	• Long-notice: Security Standards (Expected Unserved Energy greater economic threshold for load shedding) • Medium- & Short-notice: NZGB, N-1 balance under base scenario falls below zero (no adjustment for uncertainty)	• Long-notice: Same as Strawman 1a • Medium- & Short-notice: Same as Strawman 1a, however, uncertainty is incorporated into the N-1 balance forecast by using a POE wind forecast and correlated demand	• Long-notice: Same as Strawman 1a • Medium- & Short-notice: Same as Strawman 1b
Pre-activation and activation triggers	Pre-activate if NRS/WDS residuals fall below zero (no adjustment for uncertainty) Activate if NRSS/RTD residual falls below zero	Same as Strawman 1a, however, uncertainty is incorporated into the NRS/WDS residuals by using a POE wind forecast and correlated demand	Same as Strawman 1a, however, forecast uncertainty is also incorporated into the NRS/WDS residuals using historical forecast errors to develop a POE forecast of the residuals
Compensation mechanism	Providers are able to specify: • Activation payment ($/MWh)	Providers are able to specify: • Availability payments ($/MW) • Activation payments ($/MWh)	Providers are able to specify: • Preparation payments ($/MW) • Availability payments ($/MW) • Pre-activation payment ($/MW) • Activation payments ($/MWh)
Pricing restrictions	System Operator makes reasonable endeavours to limit cost of ERS provision to be less than or equal to VoLL	Same as Strawman 1a	Same as Strawman 1a

	Strawman 1a	Strawman 1b	Strawman 1c
Interaction with RTDP schedule	None	Same as Strawman 1a	Same as Strawman 1a
Participant pool	No restrictions – scheme is open to all demand side (including loads participating in Dispatchable Demand) and all generators (including registered generators with unoffered generation in the spot market)	Only demand side allowed to participate (all sectors); Dispatchable Demand loads are not allowed to participate.	Only demand-side (excluding Dispatchable Demand loads) and unregistered generators allowed to participate
Information provided to market	• Annual EUE assessment setting out quantum, location and duration of potential shortfalls • Publication of NZGB N-1 balance Quarterly updates of procurement activities • Quarterly updates of procurement and activation activities • Standardised contracts on System Operator website • Expressions of Interest	Same as Strawman 1a	Same as Strawman 1a
Cost allocation	• Allocate non-event costs nationally to generators based on share of monthly metered generation • Allocate event costs nationally to generators based on share of metered generation during activation events	• Allocate non-event costs nationally to generators and loads based on share of absolute value of monthly metered generation and consumption • Allocate event costs nationally to generators and based on share of absolute value of metered generation	• Allocate non-event costs nationally to loads based on share of monthly metered consumption • Allocate event costs nationally to loads based on share of metered consumption during activation events

	Strawman 1a	Strawman 1b	Strawman 1c
		and consumption during activation events	
VoLL review frequency	Once every three years	Same as Strawman 1a	Same as Strawman 1a

Strawman evaluation

Conclusions

Our evaluation of Strawman 1a, 1b, and 1c indicates the following:

- Allowing spot-market generators to participate could have materially adverse impacts on efficiency and competition – we recommend their exclusion.
- Excluding activated Emergency Reserves as instructed load shedding in the RTDP schedule results in distortion of scarcity pricing signals , which are critical for signalling investment. In the second evaluation (Section 3.2.2), we evaluate the impact of including activated Emergency Reserves as instructed load shedding on efficiency outcomes.
- Providers should be allowed to specify fees to cover their preparation, availability, pre-activation and activation costs to ensure they are able to reflect their cost structure in their fees as closely as possible. Limiting the category of fees will result in providers making conservative assumptions to estimate other costs they will incur and incorporating those estimates into the available fee categories. This may result in ERS costs being higher than what it would have been if the provider had been able to specify its cost structure more accurately.
- There is limited value in allocating the ERS costs to the generators as it is likely that these costs will be ultimately passed down to retailers and eventually consumers. Allocating to loads directly on the other hand gives retailers more information about their exposure to these costs and may result in retailers proactively managing and reducing the consumption of their loads to avoid or reduce charges.
- Implementing a long-notice scheme is likely to result in inaccurate forecasts of Emergency Reserves requirements due to the long time-period between the forecast and the potential shortfall manifesting. This may result in over- or under-procurement; the former would adversely affect efficiency outcomes while the latter will adversely affect reliability outcomes (absent a medium- or short-term procurement scheme). In the second evaluation (Section 3.2.2), we evaluate the impact of excluding a long-notice scheme on reliability and efficiency outcomes.
- Given the purpose of the ERS is to be a penultimate resort service that is used rarely during "perfect storm" events in which actual wind output and demand diverges materially from forecasts and/or where multiple unplanned outages occur in real-time, it is critical for the procurement and pre-activation triggers to capture inherent and forecast uncertainty. Using the

NZGB and residuals as currently calculated to trigger procurement and pre-activation is therefore not appropriate:

- The procurement trigger (N-1 NZGB) should incorporate the inherent uncertainty in wind output and demand by using POE wind and (correlated) demand forecasts (Strawman 1b). Incorporating forecast uncertainty into the NZGB is not practical nor prudent due to the reasons outlined in Section 2.3.2.
- When pre-activating Emergency Reserves, consideration of forecast uncertainty is critical given the ERS's role is to limit involuntary load shedding in the rare circumstances when forecasts are so "wildly out" from actuals that the market is unable to balance supply and demand. For this reason, we recommend the approach used in Strawman 1c of using historical residual errors (and their drivers) to construct a probability distribution of residuals.

Evaluation score and rationale

Table 13, Table 14 and Table 15 respectively summarise the performance of Strawman 1a, 1b, and 1c against the evaluation criteria.

Table 13: Strawman 1a evaluation

Evaluation criteria	Evaluation score	Comment
Competition		Performs poorly against the competition objective. The provision of activation payments only may deter participation as providers will not be able to recover fixed costs if they are procured but not activated. Additionally, allowing generators registered in the spot market may result in spot-market generators out-competing demand-side resources. Finally, the failure to account for instructed load shedding under the ERS in the RTDP schedule may dilute investment signals (if scarcity prices are suppressed) which may deter entry of new generators into the spot market.
Reliability		This strawman performs well against some reliability objectives as the System Operator has three opportunities to procure Emergency Reserves. Additionally, as the NRSS/RTD residuals (calculated one hour ahead) will be used for triggering activation, the activation quantity should be reasonably accurate. As this scheme design does not incorporate an uncertainty measure into its medium and short-term procurement triggers (N-1 NZGB under the base scenario), there is a risk that System Operator may under-procure reserves; this risk is less serious for providers procured under the short-notice scheme. As the pre-

Evaluation criteria	Evaluation score	Comment
		activation trigger also has no uncertainty measure incorporated into it, there is a risk that activation requirements may exceed what was pre-activated.
Efficiency		This scheme does poorly against the efficiency objective - this is largely driven by: • Spot market generators with unoffered generation being allowed to participate in the scheme. This may result in these generators withdrawing capacity from the spot market to receive a higher price in the ERS. • Scarcity price signals being distorted as a result of activated Emergency Reserves not being treated as instructed load shedding in the RTDP schedule. • Providers inflating their proposed activation fees to cover other (estimated) costs may result in higher costs than what would have been incurred if providers had been allowed to specify all costs incurred. • Potential over-procurement occurring as a result of the System Operator: – Procuring over a long-term timeframe where shortfall forecasts are highly uncertain. – Applying conservative assumptions in calculating the N-1 NZGB when triggering medium- and short-term procurement (to prevent under-procurement by using the current base scenario assumptions). Under this strawman design, the ERS costs are allocated to generators only. It could be said that this means that consumers are entirely shielded from the costs and will only derive benefits from the scheme if generators respond to these charges by either ensuring all their generation is offered or by investing in new build. However, it is more likely that generators will pass these costs down to their contracted retailers who will in turn pass that cost down to loads.
Cost and complexity		The scheme does well against cost and complexity objectives as the N-1 NZGB and residuals calculations should not require major changes when triggering the medium- and short-notice schemes. Moreover, the System Operator can leverage its existing Grid Support Services procurement processes and systems to conduct procurement under all three timeframes of the ERS. The inclusion of residential customers could potentially increase costs if the System Operator needs visibility of individual loads in a residential aggregation. However, these costs can be avoided if an aggregator is allowed to submit aggregated meter data to the System Operator for verification. In this case, the System Operator may incur some costs conducting spot-checks or audits on aggregator meter data to verify accuracy.

Table 14: Strawman 1b evaluation

Evaluation criteria	Evaluation score	Comment
Competition		Strawman 1b performs better against competition objectives compared to 1a by disallowing the participation of generators registered and cleared in the spot market. However, as with Strawman 1a, the failure to account for instructed load shedding under the ERS in the RTDP schedule may dilute investment signals and thereby deter entry of new generators into the spot market.
Reliability		Strawman 1b performs better against reliability objectives than 1a. The key difference in this scheme is that the System Operator incorporates a POE forecast of wind and (correlated) demand into the NZGB and residual indicators to respectively procure and pre-activate reserves; hence System Operator is less likely to under-procure or under-pre-activate under this scenario. However, forecast uncertainty is not explicitly modelled in pre-activation triggers. This could result in Emergency Reserves not being pre-activated in cases where wind and/or demand forecasts are significantly different to what occurs in real-time.
Efficiency		Strawman 1b performs better against efficiency objectives than 1a but still only partially meets efficiency objectives. It performs better than Strawman 1a because: • Registered generators and Dispatchable Demand loads in the spot market are not allowed to participate in the ERS – This will remove incentives to withhold capacity from the spot market. Moreover, spot market generation will not be displaced by more expensive (spot market unoffered) generation. • As providers can specify availability and activation payments, premiums added to these cost categories to cover other costs (e.g., pre-activation and preparation) are likely to be lower than offering an activation payment alone. Nevertheless, there is still some risk of inefficient pricing as providers will need to make conservative estimates on the number of activations and tests required by the System Operator. • Loads and generators are allocated a share of the ERS costs. The direct allocation of some of the cost to the retailers (in addition to the reporting by System Operator on procurement and activation activities) will provide retailers with information and incentives to manage the consumption of their loads to avoid or reduce these charges. There is some potential for over-procurement as forecast shortfalls under the long-term and medium-term timeframes are likely to be less accurate than short-term procurement.
Cost and complexity		Strawman 1b performs similarly to 1a. The main difference in terms of cost drivers is the use of a POE wind and (correlated) demand forecast when calculating the N-1 NZGB and pre-activation quantities (using the WDS or NRSL schedule). This will incur greater cost as System Operator will need to make changes to its NZGB application and to SPD's WDS/NRSL schedules. However, these costs are unlikely to be significant as the new

Evaluation criteria	Evaluation score	Comment
		Intermittent Generation Forecasting MOSP would likely be developing POE wind forecasts anyway.

Table 15: Strawman 1c evaluation

Evaluation criteria	Evaluation score	Comment
Competition	◕	See Strawman 1b. Strawman 1c also allows unregistered generators not cleared through the spot market to participate. Hence, Strawman 1c performs marginally better than Strawman 1b.
Reliability	●	As with Strawman 1b, this scheme performs well against reliability objectives. The key difference between Strawman 1b and 1c is that the approach to incorporating uncertainty into the pre-activation trigger (WDS or NRSL depending on lead time) is more sophisticated as it uses historical forecast residual errors to build a POE forecast of the residuals themselves. This means forecasts developed under this approach will consider the effects of forecasts being materially out from actuals.
Efficiency	◕	See Strawman 1b, except that under this scheme, all ERS costs are allocated to loads on a national basis in proportion to their metered consumption. In terms of cost allocation, this is the preferred approach as retailers have a greater incentive to manage their ERS costs by managing the consumption of their loads. As with Strawman 1a and 1b, there is some potential for over-procurement as forecast shortfalls under the long-term and medium-term timeframes are likely to be less accurate than short-term procurement
Cost and complexity	◑	Strawman 1c performs worst against cost and complexity objectives. This is because this design incorporates forecast uncertainty into forecast residuals using historical forecast residual errors to build a POE forecast of the residuals themselves. This will require significant changes to SPD/SPD adjacent tools to calculate and incorporate this measure. For this reason, this scheme is associated with moderate changes/costs.

3.2.2 Second evaluation

Strawmen definitions

In the second evaluation, we have defined three strawmen schemes as summarised in the table below.

- Strawmen 2a and 2b assume there will be two procurement timeframes: medium- and short-notice. Strawman 2c assumes there will only be short-term procurement.
- The participant pool is restricted to the demand (excluding loads participating in Dispatchable Demand in the spot market) and unregistered generators to improve performance against competition and efficiency objectives. In Strawman 2b, we restrict demand side participation to industrial and commercial customers only.
- For all three strawmen:
  - The procurement trigger (N-1 NZGB) will incorporate POE forecasts of wind and (correlated) demand to account for the inherent uncertainty in wind output and demand.
  - The pre-activation triggers (WDS/NRS residuals) will incorporate a forecast uncertainty measure per the approach in Strawman 1c.
- ERS activation is included as instructed load shedding in the RTDP schedule for all three strawmen.
- ERS costs are allocated to loads only for all three strawmen. However, Strawman 2b and 2c require System Operator to restrict activation levels so that total market and activation costs incurred by loads is less than or equal to VoLL.
- All three strawmen allow providers to specify fees in preparation, availability, pre-activation and activation cost categories.

Table 16: Policy settings for Strawmen 2a, 2b and 2c.

	Strawman 2a	Strawman 2b	Strawman 2c
Procurement timeframe	• Medium-notice: 3 months to 4 weeks ahead • Short-notice: 4 weeks to 1 week ahead	Same as Strawman 2a	Short-notice: 4 weeks to 1 week ahead
Procurement trigger	Same as Strawman 1b Uncertainty is incorporated into the N-1 NZGB forecast by using a POE wind and correlated demand forecast	Same as Strawman 1b	Same as Strawman 1b
Pre-activation and activation triggers	Same as Strawman 1c WDS/NRS residuals with forecast uncertainty incorporated using historical forecast errors to develop a POE forecast of the residuals.	Same as Strawman 1c	Same as Strawman 1c
Compensation mechanism	Same as Strawman 1c Providers are able to specify: • Preparation payments ($/MW) • Availability payments ($/MW) • Pre-activation payment ($/MW) • Activation payments ($/MWh)	Same as Strawman 1c	Same as Strawman 1c
Pricing restrictions	System Operator makes reasonable endeavours to limit cost of ERS provision to be less than or equal to VoLL	Same as Strawman 2a. Additionally, when pre-activating and activating reserves, System Operator must ensure that the sum of the market costs (Scarcity Price times	Same as Strawman 2b

	Strawman 2a	Strawman 2b	Strawman 2c
		load served) and activation costs as a proportion of load served ≤VoLL	
Interaction with RTDP schedule	Instructed ERS activation is added back onto nodal loads to maintain scarcity price signals	Same as Strawman 2a	Same as Strawman 2a
Participant pool	Only demand side allowed to participate (all sectors)	Same as Strawman 1a, but residential customers are excluded	Same as Strawman 1a, but unregistered generators are also allowed to participate
Information provided to market	Same as Strawman 1a, 1b, and 1c	Same as Strawman 1a, 1b, and 1c	Same as Strawman 1a, 1b, and 1c
Cost allocation	Same as Strawman 1c • Allocate non-event costs nationally to loads based on share of monthly metered consumption • Allocate event costs nationally to loads based on share of metered consumption during activation events	Same as Strawman 1c	Same as Strawman 1c
VoLL review frequency	Same as Strawman 1a, 1b, and 1c	Same as Strawman 1a, 1b, and 1c	Same as Strawman 1a, 1b, and 1c

Strawman evaluation

Conclusions

Our evaluation of Strawman 2a, 2b, and 2c indicates the following:

- Removing spot market generators materially improves performance against both competition and efficiency objectives.
- Including unregistered generators does not affect efficiency objectives but will improve competition as there will be a greater pool of providers.
- Excluding residential customers results in better performance against the efficiency objective but results in slightly poorer performance against the competition objective. Efficiency outcomes are improved by excluding residential aggregations as these loads may have volatile consumption patterns that makes baselining their demand challenging. This may result in inaccurate performance verification and System Operator paying for services that were not fully delivered. Additionally, System Operator will incur additional costs in verifying aggregator meter data if settlement occurs on the basis of aggregators submitting meter data aggregated across all their loads (as opposed to submitting meter data at each customer ICP). Nevertheless, allowing residential aggregations to participate does not affect performance against reliability and efficiency objectives materially. For this reason, we recommend that participation restrictions not be applied to residential aggregations.
- The inclusion of activated ERS as instructed load shedding also results in significant improvements to performance against the efficiency objectives for all three designs as scarcity pricing signals are preserved. However, maintaining these signals and allocating costs to loads will result in "double payment" which can be addressed by System Operator restricting the sum of market and activation costs to be less than VoLL (see Strawman 2b and 2c). Adding this requirement slightly improves Strawman 2b and 2c's efficiency performance relative to 2a. This is because the quantum of the double payment is likely to be low. At the same time, implementing a requirement to restrict total costs to be less than VoLL may increase System Operator costs. We recommend adopting the approach in Strawman 2b and 2c[27], if the cost of doing so is not materially higher to only requiring System Operator to restrict ERS costs to less than VoLL at

[27] i.e., removing the double payment issue by ensuring total market costs (scarcity price plus ERS activation costs) as a proportion of load is less than VoLL.

time of procurement and pre-activation/activation. For this reason, we recommend adopting the approach in Strawman 2b and 2c.

- Allowing providers to specify their fees across multiple cost categories improves competition and efficiency outcomes.
- Incorporating uncertainty measures into procurement and pre-activation triggers improves performance against reliability and efficiency objectives.
- Removing medium-term procurement (and only including short-term procurement) has a negligible impact on reliability and efficiency outcomes. System Operator is unlikely to incur material additional costs by operating medium- and short-notice procurement versus operating one or the other. Nevertheless, the quantum of Emergency Reserves is likely to be more accurate for short-notice procurement than medium notice. For this reason, a short-notice scheme is preferable as long as the System Operator conducts prudent market testing activities and maintains a pre-qualified panel of providers procurement logistics should not be an issue.

Evaluation scores and rationale

Table 17, Table 18 and Table 19 respectively summarise the performance of Strawman 2a, 2b, and 2c against the evaluation criteria.

Table 17: Strawman 2a evaluation

Evaluation criteria	Evaluation score	Comment
Competition	●	Strawman 2a substantially meets all competition criteria as: • Providers can specify fees aligned to their cost structures (preparation, availability, pre-activation and activation fees) • Competition in the ERS and wider market is not distorted by the participation of spot market generators.
Reliability	●	Strawman 2a performs similarly to 1c in terms of reliability. The respective incorporation of inherent and forecast uncertainty into procurement and pre-activation triggers will mitigate the risk of under procurement and/or under-activation.

Evaluation criteria	Evaluation score	Comment
Efficiency		This scheme does reasonably well against the efficiency objective: • Spot market generators with unoffered generation are not allowed to participate in the scheme, so there is no risk of generators withdrawing capacity from the spot market to participate in the ERS. • Scarcity price signals are preserved as activated Emergency Reserves are treated as instructed load shedding in the RTDP schedule • The risk of over-procurement is less than in Strawman 1a-1cas there is no long-notice procurement. However, loads double pay as they have to cover ERS costs and pay the scarcity price (as under this strawman activated reserves are added back onto the load forecast in the RTDP schedule). As noted in Section 2.8.1, the quantum of this double payment may be relatively minor.
Cost and complexity		The scheme performs similarly to Strawman 1c against cost and complexity objective as this option incorporates forecast uncertainty using historical forecast residual errors to build a POE forecast of the residuals themselves. This will require significant changes to SPD/SPD adjacent tools to calculate and incorporate this measure. Additionally, System Operator must also implement/amend dispatch tools to incorporate activated Emergency Reserves into the RTD load input.

Table 18: Strawman 2b evaluation

Evaluation criteria	Evaluation score	Comment
Competition		Strawman 2b is similar to Strawman 2a with the exception that residential customers are not allowed to participate in the ERS. For this reason, Strawman 2b performs slightly worse on competition outcomes.
Reliability		As with Strawman 2a, the respective incorporation of inherent and forecast uncertainty into procurement and pre-activation/activation triggers will mitigate the risk of under procurement and/or under-activation.
Efficiency		Strawman 2b does slightly better than 2a because: • The double payment issue is resolved by restricting activation levels to ensure total market and activation costs (as a proportion of load served) is restricted to VoLL. • The removal of residential aggregations from participating will remove the risk of service delivery from residential aggregations being measured inaccurately.

Evaluation criteria	Evaluation score	Comment
Cost and complexity	◔	Strawman 2b performs worse than 2a as System Operator will need to implement tools to ensure that it can restrict activation levels such that total market and activation costs (as a proportion of load served) is restricted to VoLL (to address the double payment issue).

Table 19: Strawman 2c evaluation

Evaluation criteria	Evaluation score	Comment
Competition	●	Strawman 2c perform the best in terms of competition outcomes as there are no restrictions on participation (residential aggregations are allowed (unlike 2b) as are unregistered generators.
Reliability	●	As with Strawman 2a, the respective incorporation of inherent and forecast uncertainty into procurement and pre-activation/activation triggers will mitigate the risk of under procurement and/or under-activation.
Efficiency	●	Strawman 2c performs similarly to Strawman 2b as System Operator will need to implement tools to ensure that it can restrict activation levels such that total market and activation costs (as a proportion of load served) is restricted to VoLL (to address the double payment issue). Additionally, unlike 2b, only short-term procurement is available which reduces the risk of over-procurement as shortfall calculations are likely to be more accurate than with procurement over longer timeframes.
Cost and complexity	◔	Strawman 2c performs similarly to Strawman 2b: • System Operator will need to implement tools to ensure that it can restrict activation levels such that total market and activation costs (as a proportion of load served) is restricted to VoLL (to address the double payment issue). • However, as with Strawman 2a, the System Operator will incur some costs verifying and auditing the meter data of residential aggregators

3.3 Recommended Policy Settings

Recommended policy settings are summarised in Table 20 and are briefly described below:

- Participation:
  - Only demand-side providers (excluding Dispatchable Demand) and unregistered generators should be allowed to participate to ensure there are no distortionary impacts on spot market prices and competition. In terms of residential aggregations, we have assumed that performance verification of services would involve the aggregator submitting aggregated meter data from its portfolio to facilitate settlement calculations (as opposed to the System Operator implementing a DERMS to gain visibility of loads in a residential aggregation). As previously noted, this will require the System Operator to conduct spot checks or audits on the quality of the aggregator's meter data. We further note that allowing residential aggregations with volatile consumption patterns may have accuracy issues when verifying performance, and that there is a risk that consumers pay for services that are not or partially delivered. For this reason, it will be important for the System Operator conduct due diligence checks on an aggregator's meter data to check whether reasonably accurate baselining is possible.
  - All providers must satisfy additionality requirements. For demand-side participants, the System Operator will require assurance that the relevant loads have not historically curtailed in response to price or other incentives.
- Procurement timeframe:
  - Short-notice provides the greatest certainty with respect to procurement requirements.
  - As long as the System Operator is prudent with market testing activities and maintains a pre-qualified panel, the short-time frame between procurement and activation requirements will not be an issue.
  - We also note that the market testing will be critical in terms of identifying the best resources (i.e., those with a short lead time and controllable/predictable load that can provide a firm response).
- Procurement trigger
  - The N-1 NZGB should be used to trigger the short-term procurement. The NZGB would have to be amended to captures inherent uncertainty in wind and demand by using POE of

forecasts of wind and correlated demand (we have assumed the new MOSP responsible for intermittent generation forecasts would be able to provide this data as part of their role).
  - The Security Standards (amended as discussed in Section 2.3.2) should be used to inform the System Operator's market testing and procurement planning activities.

- Pre-activation and activation triggers:
  - WDS/NRS residuals should be used to pre-activate providers. Residuals will need to incorporate forecast uncertainty in wind and demand by using historical forecast errors to develop a POE of forecast of residuals. While this approach is more complex than only accounting for inherent uncertainty, it is more fit for purpose for a potential ERS.
  - NRSS/RTD residuals (calculated in the hour prior to real-time) should be used to activate providers. This schedule is close enough to real time to be reasonably accurate; additionally, it will give System Operator sufficient time to send activation instructions and prepare instructed load shedding inputs for the RTDP schedule.
  - As previously discussed, pre-activation and activation processes will be integrated into the System Operator's existing process for managing low residual conditions.

- Compensation and relationship with VoLL:
  - Providers should be allowed to specify preparation, availability, pre-activation and activation fees to ensure the fee structure closely matches their cost structure. This will increase participation and ensure that fee structures are efficient as providers can represent their costs more accurately.
  - To ensure involuntary load shedding is restricted to economic levels, System Operator should make reasonable endeavours to ensure that at the time of procurement, the forecast cost of ERS provision (based on forecast shortfalls) is not greater than the VoLL.
  - VoLL reviews should be conducted once every three years.

- Scarcity pricing signals and resulting double payment issue:
  - To preserve scarcity signals, activated ERS should be added back onto the relevant nodal load forecasts for the RTDP schedule so that scarcity prices still set the spot price.
  - As indicated below, ERS activation costs should be allocated to loads. However, adding activated reserves back onto the load forecast will result in loads paying both the scarcity price and the ERS activation cost. As previously discussed, the quantum of this double payment is likely to be relatively minor. At the same time, it can be addressed by requiring System Operator to make reasonable endeavours to ensure the market costs (function of

scarcity price) plus pre-activation/activation costs are not greater than VoLL. As indicated above, the System Operator will be required to make reasonable endeavours to ensure the forecast cost of ERS provision (based on forecast shortfalls) is not greater than the VoLL. If the additional cost of implementing a further restriction limit market and pre-activation/activation costs is minor, then it would be prudent to implement this feature to prevent double payment. If, however, the additional cost and complexity is material, it may be prudent to accept a minor amount if double payment and inefficiency.

- In terms of cost allocation:
  - Non-event costs (preparation and availability fees) should be allocated nationally to loads based on share of monthly metered consumption.
  - Event costs (pre-activation and activation fees) should be allocated nationally to loads based on share of metered consumption during activation events.

Table 20: Recommended policy settings for ERS

Policy setting	Recommended settings
Participant pool	Only demand-side and unregistered generators that can meet System Operator additionality requirements allowed to participate
Procurement timeframe	Short-notice only: 4 weeks to 1 week ahead. We have assumed that the System Operator would run a competitive tender process to appoint a panel of pre-qualified providers (possibly with prices negotiated beforehand) that it can efficiently procure from.
Procurement trigger	• The NZGB would be used as a procurement trigger. ERS will be procured once N-1 balance falls below zero. Inherent uncertainty is incorporated into the N-1 balance forecast by using a POE wind forecast and correlated demand. • Security Standards (Expected Unserved Energy greater economic threshold for load shedding) are used to inform System Operator market testing activities
Pre-activation and activation triggers	• Pre-activate if NRS/WDS residuals fall below zero. Forecast uncertainty is incorporated into the NRS/WDS residuals using historical errors and their drivers to construct a probability distribution of residuals. • Activate if NRSS/RTD residuals (calculated in hour prior to real-time) falls below zero
Compensation mechanism and relationship to VoLL	• Providers can specify preparation, availability, pre-activation and activation fees to ensure the fee structure closely matches their cost structure.

Policy setting	Recommended settings
	• System Operator should make reasonable endeavours to ensure that at the time of procurement, the forecast cost of ERS provision (based on forecast shortfalls) is not greater than the VoLL. • VoLL reviews should be conducted once every three years.
Scarcity pricing signals and potential double payment issue	Instructed ERS activation is added back onto nodal loads to maintain scarcity price signals. This will result in some double payment by loads who pay both scarcity prices and fund activation costs. As indicated above, this can be addressed by requiring the System Operator to restrict market and pre-activation/activation costs to be less than VoLL. We recommend adopting this approach as long as it does not result in material additional costs.
Cost allocation	• Allocate non-event costs (preparation and availability fees) nationally to loads based on share of monthly metered consumption • Allocate event costs (pre-activation/activation fees) nationally to loads based on share of metered consumption during activation events.
Information provided to market	• Annual EUE assessment setting out quantum, location and duration of potential shortfalls • Publication of NZGB N-1 balance • Quarterly updates of procurement activities • Quarterly updates of activation activities • Standardised contracts on System Operator website • Expressions of interest
VoLL review frequency	Once every three years

Appendix A DETAILED STRAWMAN EVALUATION

This appendix sets out the detailed evaluation of the eight different strawmen schemes against the outcomes associated with each of the three components of the Authority's statutory objective as well as cost and complexity criteria.

A.1 Strawman 1a, 1b, 1c

A.1.1 Summary of overall performance against objectives

Evaluation criteria	Strawman 1a score	Strawman 1b score	Strawman 1c score
Competition	Some objective outcomes achieved partially	Most objective outcomes achieved substantially	Most objective outcomes achieved substantially
Reliability	Some objective outcomes achieved substantially or most met partially	Most objective outcomes achieved substantially	All objective outcome achieved substantially
Efficiency	Some objective outcomes achieved partially	Some objective outcomes achieved substantially or most met partially	Most objective outcomes achieved substantially
Cost and complexity	Few changes/few costs	Few changes/few costs	Moderate changes/costs

A.1.2 Strawman 1a

A. Performance against competition criteria

Competition outcome	Strawman performance
Providers have access to information that enables them to develop business models to provide emergency reserves	Sufficient information provided for potential providers to have transparency of procurement potential and technical requirements. Outcome fully achieved
Pricing incentivises entry	Provision of activation payments only may limit interest in participation as providers may not be able to recover other costs accurately. In the absence of specifying other fee categories (e.g., availability, preparation, and pre-activation), providers may still choose to participate by inflating the activation fee to cover other costs. However, the provider still risks incurring unrecoverable costs if they are not activated. As such, participation may be limited under this strawman. Outcome not achieved
ERS rules do not favour one technology type over another	The scheme is open to both the demand and supply-side with no discrimination based on type of load. Outcome fully achieved.
Competition in the ERS is not distorted through the use of market power	Registered generators in the spot market could out-compete load-resources in the ERS. Outcome partially achieved
Wider competition in the NZ electricity market is not distorted	Allowing spot market registered generators to participate in the ERS could result in less capacity being offered into the spot market. Additionally, not accounting for load shedding instructed under the ERS into the RTDP nodal load inputs may result in scarcity prices not binding despite a scarcity situation existing. This may have an adverse impact on the entry of new generators. Outcome not achieved.

B. Performance against reliability criteria

Reliability outcome	Strawman performance
System Operator triggers procurement of Emergency Reserves appropriately when credible risk of capacity shortfall exists	The System Operator will use EUE falling below its economic threshold to trigger the long-notice scheme. While this trigger accounts for uncertainty in variables such as outages and potential wind and demand, the long period between procurement and forecast shortfall means that that the procured resources may not be needed closer to the time of the shortfall. Under-procurement is less of a concern as the System Operator still has an opportunity to procure under the medium- and short-notice schemes. However, as the NZGB indicator (under the base scenario) has no uncertainty built into it, it is likely to be an inaccurate predictor of procurement requirements. This may result in both under- and over-procurement. The former would affect reliability outcomes while the latter affects efficiency. Outcome partially achieved
System Operator procures sufficient reserves to ensure unserved energy is restricted to economic levels	Estimates of ERS requirements under the long-notice scheme are likely to be highly uncertain and could result in under-procurement (absent a medium and short-notice scheme). Estimate of ERS requirement under the medium- and short-notice schemes may also be inaccurate and unreliable as the procurement trigger does not have uncertainty incorporated into it. As above, this may result in either over- or under-procurement. Outcome partially achieved
ERS provides incentives to providers to be available or disincentives for non-performance	Financial clawbacks of fees will provide some incentive to providers to respond as contracted. In addition, System Operator's due diligence activities to ensure providers can comply with contractual requirements should largely ensure provider compliance. Outcome largely achieved

Reliability outcome	Strawman performance
Service delivery/performance can be measured accurately (so that System Operator has assurance of delivery)	The accuracy of the performance measurement will be a function of the baseline methodology used to calculate what the provider would have consumed had they not been curtailed. Baseline methodologies based on the CAISO X/Y days approach tend to work better for predictable and less volatile loads and generators – so would work better for some C&I customers than most residential households. The choice of baseline methodology or other performance measurement methodology will be left to System Operator discretion. As almost all customers in NZ have AMI metering with 30-minute interval capability, performance measurement at 30-minute granularity will not be an issue. However, if System Operator dispatches ERS based on five-minute dispatch intervals (per the RTD schedule) or requires a steady ramp to be maintained over a trading period, then metering will be required at lower granularity (e.g., SCADA or more granular AMI metering). The System Operator's market testing activities will also be critical in determining which loads (or generators) are likely to provide the firmest response so that System Operator can prioritise the recruitment for those resources with the firmest resource. Outcome largely achieved (assuming residential loads are allowed to participate on the demand-side; volatility of residential consumption may compromise the ability to verify service delivery accurately).
Providers given sufficient notice for resources to be available	The lead-time between pre-activation (notice) and activation should ideally be short as that will enable activation of reserves on a more certain trigger value. However, some loads may require hours, maybe days, to be ready for activation - requiring short lead times may therefore limit participation. In such cases, pre-activation would occur based on a less certain indicator which means activation may need to be cancelled or levels of activation changed. The System Operator must therefore balance the need for certainty with the need to incentivise participation. The market testing activities will be important in this context, as it will provide System Operator time to find resources with reasonable short lead-times. Outcome fully achieved

Reliability outcome	Strawman performance
System Operator triggers activation of Emergency Reserves appropriately when credible risk of capacity shortfall exists	As the NRSS or RTD schedule will be used for activation, the activation quantity is likely to be reasonably accurate (as these schedules would be run within the hour prior to real-time). However, note that pre-activation levels are less certain as pre-activation will occur based on the WDS or NRSL schedule with no adjustment for uncertainty; there is some risk that activation requirements may exceed or be less than the quantity pre-activated reserves. Outcome partially achieved

C. Performance against efficiency criteria

Efficiency outcome	Strawman performance
ERS does not distort wholesale market pricing by incentivising wholesale market participants to withdraw capacity	As registered spot market generators are allowed to participate under this strawman, they may be incentivised to withdraw capacity from the spot market to get better revenues through the ERS. Outcome not achieved.
ERS does not distort scarcity pricing signals in the wholesale market	Scarcity prices indicate that there is shortfall risk in the spot market and is a critical signal for new investment. Excessive ERS dispatch could result in scarcity prices not binding when they should, thereby removing the investment signal. Under Strawman 1a, ERS dispatch is not accounted for in the RTDP schedule; as a result, scarcity prices that were binding in the NRSS schedule may no longer exist after the RTDP run thereby removing the scarcity signal. Outcome not achieved.
ERS does not displace cheaper energy resources in the spot market	As long as the scarcity prices are set high enough, generators should be incentivised to offer their capacity into the spot market before ERS is triggered, so that ERS is a true penultimate-resort mechanism. For this reason, regular VoLL reviews will be required to ensure that the scarcity price caps are set at efficient levels. However, as spot market generators are allowed to participate under this strawman, such generators could bypass the spot market and provide ERS at a higher cost thereby displacing cheaper generation.

<table>
<tr><th>Efficiency outcome</th><th>Strawman performance</th></tr>
<tr><td></td><td> As ERS activation is triggered using the NRSS/RTD residuals, it is unlikely that there will be material generation offer changes between activation and real-time dispatch causing ERS resources to displace generators that would have otherwise been cleared in the RTD schedule. Outcome partially achieved </td></tr>
<tr><td>Loads/retailers have access to information to accurately identify, evaluate and manage their risks</td><td> Sufficient information provided for loads to determine potential ERS cost exposure. Outcome fully achieved </td></tr>
<tr><td>ERS provides long-term value to consumers</td><td>This strawman allows providers to specify activation payments only. However, as providers need to recover all their costs, they are likely to add those additional (fixed) costs to their proposed activation fee. In doing so, they will need to make assumptions around the number of activations. Conservative assumptions can result in overall ERS costs being higher than what they would have been had the provider been able to specify all its costs in separate fee categories. This will erode value to consumers.
As forecast shortfalls under the long-term and medium-term timeframes are likely to be less accurate than short-term procurement, there is some risk of over procurement which may erode value for consumers.
Meeting the efficiency outcome also depends on the extent to which the System Operator can ensure the unit cost of ERS provision does not exceed VoLL. The System Operator's ability to do this depends on the accuracy of VoLL estimates. VoLL estimates will affect the EUE threshold (defined via the Security Standards), the quantity of ERS pre-activated and subsequently activated and the Scarcity Price Limits. Setting the VoLL estimate too low (but higher than the largest scarcity price cap) could result in System Operator over-activating while the converse applies if VoLL is set too high. A three-yearly review will ensure VoLL is updated reasonably frequently.
As the cost is allocated to generators, it could be argued that loads will not face any cost. However, generators may well pass the ERS costs down to retailers (as opposed to making unoffered generation available or investing in new generation) so that consumers ultimately face the cost.
Including residential customers may result in service delivery being measured inaccurately, meaning a provider could be compensated for under-performing due to service delivery issues. Outcome not achieved </td></tr>
</table>

Efficiency outcome	Strawman performance
Small customer interests are protected where consumer resources are used to provide ERS	This outcome applies where customer assets are being aggregated. The presence of full retail competition, the energy ombudsman, utilities disputes means that small customers have the ability to switch retailers and lay complaints if needed. Outcome fully achieved

D. Performance against cost and complexity criteria

Source of cost	Strawman performance
Code changes	Some changes will be needed to incorporate ERS as a new Ancillary Service Few changes/low costs
Market information changes	Few changes/low cost changes to WITS
Telemetry requirements	Dependent on what types of loads and generators participate: • Transmission connected loads will have the required telemetry for 30 minutes response duration (AMI) or less (SCADA). • Some commercial and residential customers may incur additional costs measuring performance at a granularity lower than 30 minutes. Few changes/low costs (if commercial and residential customers are allowed to participate and do not meet lower granularity metering requirements)
Scheduling and dispatch tools and process changes	There will be some cost associated with modelling and publishing EUE annually (to inform the market and System Operator's procurement activities). However, as the NZGB and residuals are not being modified in anyway, and the activated ERS load is not added back onto the nodal loads for the RTDP schedule, the changes to the NZGB application and SPD should be minimal. Few changes/low costs
Settlement and reconciliation tools and process changes	Minor changes to clearing agent systems will be required to collect ERS costs and transmit to System Operator; this is similar to how other Ancillary Services costs are allocated and collected and as such changes should be minor.

Source of cost	Strawman performance
	Negligible changes/costs
Other process and tool changes	The System Operator has existing processes to manage its Grid Support Contracts and can leverage these process and tools to manage ERS procurement. Nevertheless, there may be moderate costs associated with procurement as the System Operator will need to create pre-qualified panels to ensure procurement can be done on a timely basis for medium- and short-term procurement. The cost of implementing the cost recovery calculation is also likely to be relatively minor as the cost allocation methodology is relatively simple and similar to how other Ancillary Services costs are recovered. The inclusion of residential customers could potentially increase costs if the System Operator needs visibility of individual loads in a residential aggregation. However, these costs can be avoided if an aggregator is allowed to submit aggregated meter data to the System Operator for verification. In this case, the System Operator may incur some costs conducting spot-checks or audits on aggregator meter data to verify accuracy. Moderate changes/medium costs

A.1.3 Strawman 1b

A. Performance against competition criteria

Competition outcome	Strawman performance
Providers have access to information that enables them to develop business models to provide emergency reserves	See Strawman 1a. Outcome fully achieved
Pricing incentivises entry	Strawman 1b allows providers to specify both availability and activation payments. This will incentivise participation better than Strawman 1a as the providers will have a separate category to specify fixed costs. However, pre-activation costs incurred by the provider will need to be estimated and incorporated into availability and/or activation fee components. Outcome largely achieved

Competition outcome	Strawman performance
ERS rules do not favour one technology type over another	The scheme is open to all demand-side participants and does not discriminate against type of load. However, all generators are excluded. Outcome largely achieved
Competition in the ERS is not distorted through the use of market power	As spot market generators and Dispatchable Demand loads are excluded, there should be no distortion to competition. Outcome fully achieved
Wider competition in the NZ electricity market is not distorted	Same as Strawman 1a, except that spot market generators are not allowed to participate; given sufficiently high scarcity prices, such generators should not be incentivised to withhold capacity resulting in more expensive ERS resources being used. For this reason, Strawman 1b performs slightly better than 1a. However, as with Strawman 1a, not accounting for load shedding instructed under the ERS into the RTDP nodal load inputs may result in scarcity prices not binding despite a scarcity situation existing. This may have an adverse impact on the entry of new generators. Outcome partially achieved

B. Performance against reliability criteria

Reliability outcome	Strawman performance
System Operator triggers procurement of Emergency Reserves appropriately when credible risk of capacity shortfall exists	Strawman 1b meets this criterion better than Strawman 1a as the NZGB N-1 balance will be more conservative with the use of a POE wind and correlated demand forecast. However, long-term procurement (and to a lesser extent, medium-term procurement) may result in highly uncertain shortfall forecasts and result in under-procurement. Outcome largely achieved
System Operator procures sufficient reserves to ensure unserved energy is restricted to economic levels	Strawman 1b meets this criterion slightly better than Strawman 1a as the NZGB N-1 balance will be more conservative with the use of a POE wind and correlated demand forecast. As above, however, there is a risk of under- and over-procurement due to the uncertain nature of the long- and medium-term shortfall forecasts. Outcome largely achieved

Reliability outcome	Strawman performance
ERS provides incentives to providers to be available or disincentives for non-performance	See Strawman 1a. Outcome largely achieved
Service delivery/performance can be measured accurately (so that System Operator has assurance of delivery)	See Strawman 1a. Outcome largely achieved
Providers given sufficient notice for resources to be available	See Strawman 1a. Outcome fully achieved
System Operator triggers activation of Emergency Reserves appropriately when credible risk of capacity shortfall exists	Strawman 1b meets this criterion slightly better than Strawman 1a as the residuals used to pre-activate reserves will be more conservative with the use of a POE wind and correlated demand forecast. However, as no forecast uncertainty is incorporated, the residuals will not capture shortfalls occurring due to forecasts being materially out. Outcome partially achieved

C. Performance against efficiency criteria

Efficiency outcome	Strawman performance
ERS does not distort wholesale market pricing by incentivising wholesale market participants to withdraw capacity	As only the demand-side is allowed to participate under this strawman, there should be no incentives for spot market generators to withdraw capacity. Outcome fully achieved
ERS does not distort scarcity pricing signals in the wholesale market	See Strawman 1a. Outcome not achieved
ERS does not displace cheaper energy resources in the spot market	Strawman 1b meets this criterion better than Strawman 1a as spot market generators are not allowed to participate under this design. Hence there is no risk of ERS resources displacing spot market generation. Outcome fully achieved

Efficiency outcome	Strawman performance
Loads/retailers have access to information to accurately identify, evaluate and manage their risks	See Strawman 1a. Outcome fully achieved
ERS provides long-term value to consumers	Strawman 1b performs better than 1a: • As providers can specify availability and activation payments, premiums added to these cost categories to cover other costs (e.g., pre-activation and preparation) are likely to be lower than offering an activation payment alone. Nevertheless, there is still some risk of inefficient pricing as providers will need to make conservative estimates on the number of activations and tests required by the System Operator. • However, ERS costs are shared between generators and loads. As with Strawman 1a, the share allocated to generators may end up being passed back down to consumers through retailers. As with Strawman 1a: • Since forecast shortfalls under the long-term and medium-term timeframes are likely to be less accurate than short-term procurement, there is some risk of over procurement which may erode value for consumers. • Including residential customers may result in service delivery being measured inaccurately, meaning a provider could be compensated for under-performing due to service delivery issues. Outcome partially achieved
Small customer interests are protected where consumer resources are used to provide ERS	See Strawman 1a. Outcome fully achieved

D. Performance against cost and complexity criteria

Source of cost	Strawman performance
Code changes	See Strawman 1a. Few changes/low costs
Market information changes	See Strawman 1a.

Source of cost	Strawman performance
	Few changes/low costs
Telemetry requirements	See Strawman 1a. Few changes/low costs
Scheduling and dispatch tools and process changes	Changes will be required to the NZGB application and SPD or SPD-adjacent tools to incorporate POE estimates into the N-1 balance and NRS/WDS residuals respectively. As the POE estimates will come from the new Intermittent Generation Forecasting MOSP, System Operator should incur minimal costs in incorporating those estimates into NZGB and SPD/SPD-adjacent tools. The new MOSP may also incur some additional costs (compared to them not having to provide the POE wind forecast along with the correlated demand estimate). Moderate changes/ costs
Settlement and reconciliation tools and process changes	See Strawman 1a. Negligible changes/costs
Other process and tool changes	See Strawman 1a. Moderate changes/medium costs

A.1.4 Strawman 1c

A. Performance against competition criteria

Competition outcome	Strawman performance
Providers have access to information that enables them to develop business models to provide emergency reserves	See Strawman 1a. Outcome fully achieved
Pricing incentivises entry	Strawman 1b allows providers to specify preparation, availability, pre-activation and activation payments. This will incentivise participation better than Strawman 1b as the providers will have separate categories to accurately reflect their cost structures.

Competition outcome	Strawman performance
	Outcome fully achieved
ERS rules do not favour one technology type over another	Same as Strawman 1b except unregistered generators are also allowed to participate. Outcome fully achieved
Competition in the ERS is not distorted through the use of market power	See Strawman 1b. Outcome fully achieved
Wider competition in the NZ electricity market is not distorted	See Strawman 1b. Outcome partially achieved

B. Performance against reliability criteria

Reliability outcome	Strawman performance
System Operator triggers procurement of Emergency Reserves appropriately when credible risk of capacity shortfall exists	Same as Strawman 1b. Outcome largely achieved
System Operator procures sufficient reserves to ensure unserved energy is restricted to economic levels	Same as Strawman 1b. Outcome largely achieved
ERS provides incentives to providers to be available or disincentives for non-performance	See Strawman 1a. Outcome largely achieved.
Service delivery/performance can be measured accurately (so that System Operator has assurance of delivery)	See Strawman 1a. Outcome largely achieved
Providers given sufficient notice for resources to be available	See Strawman 1a. Outcome fully achieved

Reliability outcome	Strawman performance
System Operator triggers activation of Emergency Reserves appropriately when credible risk of capacity shortfall exists	Strawman 1c meets this criterion better than Strawman 1b as the residuals used to pre-activate reserves incorporate forecast uncertainty and therefore capture shortfalls occurring due to forecasts being materially out. Outcome fully achieved

C. Performance against efficiency criteria

Efficiency outcome	Strawman performance
ERS does not distort wholesale market pricing by incentivising wholesale market participants to withdraw capacity	This strawman allows both the demand side and unregistered generators to participate. As unregistered generators are not cleared through the spot market, there is no risk of capacity being withheld from the spot market. Outcome fully achieved
ERS does not distort scarcity pricing signals in the wholesale market	See Strawman 1a. Outcome not achieved
ERS does not displace cheaper energy resources in the spot market	See Strawman 1b. Outcome fully achieved
Loads/retailers have access to information to accurately identify, evaluate and manage their risks	See Strawman 1a. Outcome fully achieved
ERS provides long-term value to consumers	Strawman 1c performs better than 1b for the following reasons: • ERS costs are allocated to consumers on a national basis in proportion to consumption. Allocating the cost directly to retailers may incentivise them to manage their loads to avoid these costs. • As providers can specify preparation, availability, pre-activation and activation payments, fees are likely to resemble actual provider cost structures. This means providers do not need to incorporate conservative estimates into other fee categories. This will result in better outcomes for the consumer. As with Strawman 1a:

Efficiency outcome	Strawman performance
	• Since forecast shortfalls under the long-term and medium-term timeframes are likely to be less accurate than short-term procurement, there is some risk of over procurement which may erode value for consumers. • Including residential customers may result in service delivery being measured inaccurately, meaning a provider could be compensated for under-performing due to service delivery issues. Outcome largely achieved
Small customer interests are protected where consumer resources are used to provide ERS	See Strawman 1a. Outcome fully achieved

D. Performance against cost and complexity criteria

Source of cost	Strawman performance
Code changes	See Strawman 1a. Few changes/low costs
Market information changes	See Strawman 1a. Few changes/low costs
Telemetry requirements	See Strawman 1a. Few changes/low costs
Scheduling and dispatch tools and process changes	Strawman 1c has the most complex approach to incorporating uncertainty into the NRS/WDS residuals; System Operator will likely incur significant costs to amend SPD/SPD adjacent tools to incorporate a forecast uncertainty measure using historical residual errors and their drivers. Many changes/high costs
Settlement and reconciliation tools and process changes	See Strawman 1a. Negligible changes/costs

Source of cost	Strawman performance
Other process and tool changes	See Strawman 1a. Moderate changes/medium costs

A.2 Strawman 2a, 2b, 2c

A.2.1 Summary of overall performance against objectives

Evaluation criteria	Strawman 2a score	Strawman 2b score	Strawman 2c score
Competition	All objective outcome achieved substantially	Most objective outcomes achieved substantially	All objective outcome achieved substantially
Reliability	All objective outcome achieved substantially	All objective outcome achieved substantially	All objective outcome achieved substantially
Efficiency	Most objective outcomes achieved substantially	All objective outcome achieved substantially	All objective outcome achieved substantially

Evaluation criteria	Strawman 2a score	Strawman 2b score	Strawman 2c score
Cost and complexity	Moderate changes/medium costs	Many changes/high costs	Many changes/high costs

A.2.2 Strawman 2a

A. Performance against competition criteria

Competition outcome	Strawman performance
Providers have access to information that enables them to develop business models to provide emergency reserves	See Strawman 1a. Outcome fully achieved
Pricing incentivises entry	See Strawman 1c. Outcome fully achieved
ERS rules do not favour one technology type over another	The scheme is open to all demand-side participants and does not discriminate against type of load. However, all generators are excluded. Outcome largely achieved
Competition in the ERS is not distorted through the use of market power	As spot market generators and Dispatchable Demand loads are excluded, there should be no distortion to competition. Outcome fully achieved
Wider competition in the NZ electricity market is not distorted	As spot market generators are not allowed to participate, ERS generation is less likely to displace cheaper spot market generation. Furthermore, as this design adds the load shedding instructed under the ERS into the RTDP nodal load inputs, scarcity pricing signals will be maintained ensuring investment signals are not diluted. Outcome fully achieved

B. Performance against reliability criteria

Reliability outcome	Strawman performance
System Operator triggers procurement of Emergency Reserves appropriately when credible risk of capacity shortfall exists	See Strawman 1c. Additionally, this strawman includes medium- and short-term procurement only, meaning that the procurement levels are likely to be more accurate than those forecast under a long-term timeframe. Outcome fully achieved
System Operator procures sufficient reserves to ensure unserved energy is restricted to economic levels	See Strawman 1c. As above, this strawman includes medium- and short-term procurement only, meaning that the procurement levels are likely to be more accurate than those forecast under a long-term timeframe. Outcome fully achieved
ERS provides incentives to providers to be available or disincentives for non-performance	See Strawman 1a. Outcome largely achieved
Service delivery/performance can be measured accurately (so that System Operator has assurance of delivery)	See Strawman 1a. Outcome largely achieved
Providers given sufficient notice for resources to be available	See Strawman 1a. Outcome fully achieved
System Operator triggers activation of Emergency Reserves appropriately when credible risk of capacity shortfall exists	The residuals used to pre-activate reserves incorporate forecast uncertainty and therefore capture shortfalls occurring due to forecasts being materially out. Outcome fully achieved

C. Performance against efficiency criteria

Efficiency outcome	Strawman performance
ERS does not distort wholesale market pricing by incentivising wholesale market participants to withdraw capacity	See Strawman 1b – spot market generators are excluded. Outcome fully achieved

Efficiency outcome	Strawman performance
ERS does not distort scarcity pricing signals in the wholesale market	Under this scheme, instructed load shedding under the ERS is added back onto the nodal load inputs for the RTDP schedule. Adding the activated reserves back onto the load forecast will ensure that scarcity prices set the spot market price thereby preserving the scarcity pricing signals. Outcome fully achieved
ERS does not displace cheaper energy resources in the spot market	As spot market generators are not allowed to participate under this design, there is no risk of ERS resources displacing spot market generation. Additionally, as ERS activation is triggered using the NRSS/RTD residuals (within the hour prior to real-time), it is unlikely that there will be material generation offer changes between activation and real-time dispatch causing ERS resources to displace generators that would have otherwise been cleared in the RTD schedule. Outcome fully achieved
Loads/retailers have access to information to accurately identify, evaluate and manage their risks	See Strawman 1a. Outcome fully achieved
ERS provides long-term value to consumers	As with Strawman 1c: • ERS costs are allocated to consumers on a national basis in proportion to consumption. Allocating the cost directly to retailers may incentivise them to manage their loads to avoid these costs. • As providers can specify preparation, availability, pre-activation and activation payments, fees are likely to resemble actual provider cost structures. This means providers do not need to incorporate conservative estimates into other fee categories. This will result in better outcomes for the consumer. However, as this strawman includes activated ERS as instructed load shedding in the RTDP schedule, loads will end up paying both the scarcity price and activation costs., thereby diluting the value received by consumers. As noted in Section 2.8.1, the quantum of this double payment is likely to be minor if ERS is activated rarely, results in a minimal proportion of load being curtailed (1%-3%) and the activation payments are similar in magnitude to the prevalent scarcity price. As with Strawman 1c:

Efficiency outcome	Strawman performance
	• There is also a minor risk of over-procurement when triggering procurement using the medium-term procurement option as forecast shortfalls are likely to be less accurate than those calculated under the short-term procurement option. • Including residential customers may result in service delivery being measured inaccurately, meaning a provider could be compensated for under-performing due to service delivery issues. Outcome partially achieved
Small customer interests are protected where consumer resources are used to provide ERS	See Strawman 1a. Outcome fully achieved

D. Performance against cost and complexity criteria

Source of cost	Strawman performance
Code changes	See Strawman 1a. Few changes/low costs
Market information changes	See Strawman 1a. Few changes/low costs
Telemetry requirements	See Strawman 1a. Few changes/low costs
Scheduling and dispatch tools and process changes	As with Strawman 1c, this strawman has a complex approach to incorporating uncertainty into the NRS/WDS residuals; System Operator will likely incur significant costs to amend SPD/SPD adjacent tools to incorporate a forecast uncertainty measure. Additionally, moderate changes will be required to dispatch systems to enable activated ERS to be added back onto the nodal loads in the RTDP schedule. Many changes/high costs

Source of cost	Strawman performance
Settlement and reconciliation tools and process changes	See Strawman 1a. Negligible changes/costs
Other process and tool changes	See Strawman 1a. Moderate changes/medium costs

A.2.3 Strawman 2b

A. Performance against competition criteria

Competition outcome	Strawman performance
Providers have access to information that enables them to develop business models to provide emergency reserves	See Strawman 1a. Outcome fully achieved
Pricing incentivises entry	See Strawman 1c. Outcome fully achieved
ERS rules do not favour one technology type over another	The scheme is open to most demand-side participants but excludes residential customers and all generators. Outcome partially achieved
Competition in the ERS is not distorted through the use of market power	See Strawman 2a. Outcome fully achieved
Wider competition in the NZ electricity market is not distorted	See Strawman 2a. Outcome fully achieved

B. Performance against reliability criteria

Reliability outcome	Strawman performance
System Operator triggers procurement of Emergency Reserves appropriately when credible risk of capacity shortfall exists	Same as Strawman 2a Outcome fully achieved
System Operator procures sufficient reserves to ensure unserved energy is restricted to economic levels	Same as Strawman 2a Outcome fully achieved
ERS provides incentives to providers to be available or disincentives for non-performance	See Strawman 1a. Outcome largely achieved
Service delivery/performance can be measured accurately (so that System Operator has assurance of delivery)	Similar to Strawman 1a. However, the exclusion of residential customers will remove the issue of potentially inaccurate performance verification due to load volatility. Outcome fully achieved
Providers given sufficient notice for resources to be available	See Strawman 1a. Outcome fully achieved
System Operator triggers activation of Emergency Reserves appropriately when credible risk of capacity shortfall exists	See Strawman 2a. Outcome fully achieved

C. Performance against efficiency criteria

Efficiency outcome	Strawman performance
ERS does not distort wholesale market pricing by incentivising wholesale market participants to withdraw capacity	See Strawman 1c. Outcome fully achieved
ERS does not distort scarcity pricing signals in the wholesale market	See Strawman 2a.

Efficiency outcome	Strawman performance
	Outcome fully achieved
ERS does not displace cheaper energy resources in the spot market	See Strawman 2a. Outcome fully achieved
Loads/retailers have access to information to accurately identify, evaluate and manage their risks	See Strawman 1a. Outcome fully achieved
ERS provides long-term value to consumers	Performs similarly to Strawman 2a with the following differences: • Under this design, System Operator must also restrict its activation quantities to ensure both market costs (scarcity price times demand served) and activation costs as a proportion of load served is less than or equal to VoLL. Consumers therefore derive greater value under this scheme than Strawman 2a (as the double payment issue in 2a will have been addressed). • Excluding residential customers will remove the risk of service delivery from residential aggregations being measured inaccurately. As with Strawman 2a, there is also a minor risk of over-procurement when triggering procurement using the medium-term procurement option as forecast shortfalls are likely to be less accurate than those calculated under the short-term procurement option. Outcome largely achieved
Small customer interests are protected where consumer resources are used to provide ERS	See Strawman 1a. Outcome fully achieved

D. Performance against cost and complexity criteria

Source of cost	Strawman performance
Code changes	See Strawman 1a. Few changes/low costs

Source of cost	Strawman performance
Market information changes	See Strawman 1a. Few changes/low costs
Telemetry requirements	See Strawman 1a. Few changes/low costs
Scheduling and dispatch tools and process changes	See Strawman 2a. Additionally, System Operator will have to implement a tool that restricts activation levels so that customers do not pay more than VoLL for both market costs (scarcity prices) and activation costs. Wide ranging change / Significantly high costs
Settlement and reconciliation tools and process changes	See Strawman 1a. Negligible changes/costs
Other process and tool changes	See Strawman 1a. However, unlike the other strawmen evaluated, under Strawman 1b, System Operator will not incur additional costs associated with verifying residential aggregation performance. Moderate changes/medium costs

A.2.4 Strawman 2c

A. Performance against competition criteria

Competition outcome	Strawman performance
Providers have access to information that enables them to develop business models to provide emergency reserves	See Strawman 1a. Outcome fully achieved
Pricing incentivises entry	See Strawman 1c. Outcome fully achieved

Competition outcome	Strawman performance
ERS rules do not favour one technology type over another	The scheme is open to all demand-side participants and does not discriminate against type of load. Additionally, unregistered generators are allowed to participate. Outcome fully achieved
Competition in the ERS is not distorted through the use of market power	See Strawman 2a. Outcome fully achieved
Wider competition in the NZ electricity market is not distorted	See Strawman 2a. Outcome fully achieved

B. Performance against reliability criteria

c	Strawman performance
System Operator triggers procurement of Emergency Reserves appropriately when credible risk of capacity shortfall exists	Same as Strawman 2a. Outcome fully achieved
System Operator procures sufficient reserves to ensure unserved energy is restricted to economic levels	Same as Strawman 2a. Outcome fully achieved
ERS provides incentives to providers to be available or disincentives for non-performance	See Strawman 1a. Outcome largely achieved
Service delivery/performance can be measured accurately (so that System Operator has assurance of delivery)	See Strawman 2a. Outcome largely achieved
Providers given sufficient notice for resources to be available	See Strawman 1a. Outcome fully achieved

c	Strawman performance
System Operator triggers activation of Emergency Reserves appropriately when credible risk of capacity shortfall exists	See Strawman 2a. Outcome fully achieved

C. Performance against efficiency criteria

Efficiency outcome	Strawman performance
ERS does not distort wholesale market pricing by incentivising wholesale market participants to withdraw capacity	See Strawman 1c. Outcome fully achieved
ERS does not distort scarcity pricing signals in the wholesale market	See Strawman 2a. Outcome fully achieved
ERS does not displace cheaper energy resources in the spot market	See Strawman 2a. Outcome fully achieved
Loads/retailers have access to information to accurately identify, evaluate and manage their risks	See Strawman 1a. Outcome fully achieved
ERS provides long-term value to consumers	• As with Strawman 2b, System Operator must also restrict its activation quantities to ensure both market costs (scarcity price times demand served) and activation costs as a proportion of load served is less than or equal to VoLL. Consumers therefore derive greater value under this scheme than Strawman 2a (as the double payment issue in 2a will have been addressed). • Unlike Strawman 2c, residential customers are allowed to participate, so there is some risk of service delivery from residential aggregations being measured inaccurately. • Unlike Strawman 2a and 2b, 2c only has short-term procurement which is likely to result in more accurate procurement quantities than with a medium-term timeframe. Outcome fully achieved

Efficiency outcome	Strawman performance
Small customer interests are protected where consumer resources are used to provide ERS	See Strawman 1a. Outcome fully achieved

D. Performance against cost and complexity criteria

Source of cost	Strawman performance
Code changes	See Strawman 1a. Few changes/low costs
Market information changes	See Strawman 1a. Few changes/low costs
Telemetry requirements	See Strawman 1a. Few changes/low costs
Scheduling and dispatch tools and process changes	As with Strawman 2a, System Operator will have to implement a tool that restricts activation levels so that customers do not pay more than VoLL for both market costs (scarcity prices) and activation costs. Wide ranging change / Significantly high costs
Settlement and reconciliation tools and process changes	See Strawman 1a. Negligible changes/costs
Other process and tool changes	See Strawman 1a. Moderate changes/medium costs