INDEX

TESTIMONY OF

JANET ROSS KLIPPSTEIN, JUERGEN M. BERMEJO, DANIEL H. FISHER,

MARK A. JACKSON, ERIC V. KING, PAUL T. KOSKI, and TIMOTHY C. MISLEY

Witnesses for Bonneville Power Administration

SUBJECT: BALANCING RESERVE CAPACITY COST ALLOCATION METHODOLOGY

Page

Section 1: Introduction and Purpose of Testimony 1

Section 2: Overview 2

Section 3: Embedded Cost Methodology 5

Section 3.1: Specific Aspects and Inputs of the Embedded Cost Methodology 7

Section 3.2: 120-Hour Peaking Capacity 10

Section 3.3: Embedded Cost Calculation 16

Section 3.4: Embedded Cost for Different Levels of Service 19

Section 4: Direct Assignment of the Wind Integration Team (WIT) Costs 20

Section 5: Variable Cost Pricing Methodology 24

Section 5.1: Specific Variable Costs Considered by the GARD Model 30

Section 5.2: Variable Cost of Reserves 49

Section 5.3: Alternative Variable Cost Forecast 50

Section 5.4: VERBS Credit Adjustment to Generation Inputs Revenue Forecast 51

Section 6: Types of Balancing Reserve Capacity Purchases 52

This page intentionally left blank.

TESTIMONY of

JANET ROSS KLIPPSTEIN, JUERGEN M. BERMEJO, DANIEL H. FISHER,

MARK A. JACKSON, ERIC V. KING, PAUL T. KOSKI, and TIMOTHY C. MISLEY

Witnesses for Bonneville Power Administration

SUBJECT: BALANCING RESERVE CAPACITY COST ALLOCATION METHODOLOGY

Section 1: Introduction and Purpose of Testimony

Q. Please state your names and qualifications.

A. My name is Janet Ross Klippstein, and my qualifications are contained in BP-14-Q-BPA-37.

A. My name is Juergen M. Bermejo, and my qualifications are contained in BP-14-Q-BPA-05.

A. My name is Daniel H. Fisher, and my qualifications are contained in BP-14-Q-BPA-19.

A. My name is Mark A. Jackson, and my qualifications are contained in BP-14-Q-BPA-28.

A. My name is Eric V. King, and my qualifications are contained in BP-14-Q-BPA-35.

A. My name is Paul T. Koski, and my qualifications are contained in BP-14-Q-BPA-38.

A. My name is Timothy C. Misley, and my qualifications are contained in BP-14-Q-BPA-49.

Q. What is the purpose of your testimony?

A. The purpose of our testimony is to sponsor section 3 of the Generation Inputs Study, BP-14-E-BPA-05 (Study) and Generation Inputs Study Documentation, BP-14-E-BPA-05A (Documentation), to explain the cost allocation methodology for Regulating Reserve, Dispatchable Energy Resource Balancing Service (DERBS) Reserve, and Variable Energy Resource Balancing Service (VERBS) Reserve.

Section 2: Overview

Q. What is the overarching principle of BPA's cost allocation methodology for generation inputs?

A. BPA's cost allocation methodology is based on the principle of cost causation. Those entities that are causing the costs should bear the responsibility for paying those costs. This principle avoids cost shifts between customer classes.

Q. Please explain how costs that are allocated to the various generation inputs are recovered and how these costs affect power rates.

A. Costs allocated to generation inputs are assigned to Transmission Services (TS) and recovered through transmission rates. The revenue received from providing generation inputs is a revenue credit to BPA power rates and thus reduces power rates by lowering the revenue requirement.

Q. Why are balancing reserve capacity costs assigned to TS?

A. TS uses balancing reserve capacity to provide Ancillary and Control Area Services to transmission customers. These services are required pursuant to reliability standards issued by the North American Electric Reliability Corporation (NERC) and the Western Electricity Coordinating Council (WECC). When Power Services (PS) supplies balancing reserve capacity to TS, PS no longer has such capacity available for other purposes.

Q. To which transmission rates are these costs assigned?

A. The costs for balancing reserve capacity are allocated to the proposed ACS-14 Regulation and Frequency Response, VERBS, and DERBS rates (balancing services). *See* Jackson *et al.*, BP-14-E-BPA-28. The other generation inputs that PS provides to TS for recovery in transmission rates are discussed in other testimonies. Messinger *et al.*, BP-14-E-BPA-25; Wellschlager *et al.*, BP-14-E-BPA-26; Salazar *et al.*, BP-12-E-BPA-27. The

variable cost portion of the cost allocation methodology addressed in this testimony also pertains to the spinning portion of Operating Reserve.

Q. What are the methods for allocating costs described in your testimony?

A. The three methods for allocating costs described in this testimony are embedded cost, direct assignment, and variable cost. Each method is explained in detail in section 3 of the Study and summarized below in this testimony.

Q. How are you proposing to allocate costs to the balancing services?

A. We are using a similar methodology to the methodology used in BP-12, with updated inputs, for the embedded cost component of the generation inputs rates. The direct assignment and variable cost components have changed slightly from BP-12, since BPA is no longer forecasting costs associated with the decremental balancing reserve capacity acquisition pilot. The direct assignment component includes the full direct assignment amount and not an amount reduced by revenue received from BPA's Green Energy Premium.

Q. Are there any new features you are addressing in this case regarding the allocation of costs associated with these balancing services?

A. Yes. In this rate case we are forecasting for the first time that the Federal Columbia River Power System (FCRPS) will not be the sole source for meeting the balancing reserve capacity needs for BPA's balancing authority area. Instead, BPA will be purchasing balancing reserve capacity from third parties when FCRPS capacity is either insufficient (on a forecast basis) or unavailable (on a real-time basis). In section 6 of this testimony, we describe BPA's proposed methodology for assigning these purchase costs to the customer groups that will be using these types of balancing reserve capacity.

Q. Why are you proposing to allocate an embedded cost component and a variable cost component to the balancing services?

A. We propose to allocate a share of the embedded cost of the FCRPS to the generation inputs that support VERBS, Regulation and Frequency Response Service, DERBS, and Operating Reserve because providing these services requires use of FCRPS capability. Study section 3.2. Users of the capability of the FCRPS should share in paying for a portion of the embedded costs of the FCRPS. Such an allocation ensures that the embedded costs of the FCRPS are allocated among those that benefit from the use of the FCRPS and avoids cost shifts among BPA's customers. If VERBS, Regulation and Frequency Response Service, DERBS, and Operating Reserve were not provided from the capability of the FCRPS, then the balancing reserve capacity of the FCRPS could be used for other system purposes. Moreover, if no embedded costs of the FCRPS were allocated to these services, BPA's other customers (primarily PF power customers) would bear the full costs of paying for the embedded cost of the FCRPS, even though a portion of the capability of the FCRPS was being used by transmission customers.

We also propose to allocate to the generation inputs that support VERBS, Regulation and Frequency Response Service, DERBS, and Operating Reserve certain variable energy costs that the system incurs by operating the FCRPS to provide balancing reserve capacity. Providing balancing reserve capacity requires Power Services to change the operation of the FCRPS from an efficient power operation to an operation that: (1) is less efficient in converting water to electricity (less efficiency); and (2) produces electricity at times in which it is less valuable than during efficient power operation (energy shift). Study section 3.4. Allocating these costs to the users that are creating the need to make these changes to the FCRPS is consistent with the ratemaking principle of cost causation.

Q. If BPA was not proposing to provide balancing reserve capacity, how would these variable costs be recovered?

A. If BPA provided less balancing reserve capacity, BPA would not incur the variable costs associated with the additional amounts of these generation inputs. Rather, power production would be more efficient, and BPA's revenues could be assumed to be increased by BPA selling additional amounts of power into the market. In addition, if less balancing reserve capacity were required, BPA could make additional power sales at higher prices because BPA would not be required to shift power production from Heavy Load Hours (HLH) (when power is more valuable) to Light Load Hours (LLH) (when power is less valuable).

Section 3: Embedded Cost Methodology

Q. What is meant by embedded cost?

A. Embedded cost refers to the actual depreciated cost of an electrical system, such as the cost of generation facilities used to provide balancing reserve capacity, operation and maintenance costs, and other associated costs. For purposes of BPA's embedded cost methodology, the embedded cost is a specifically defined portion of BPA's annual revenue requirement associated with the generation projects that are used to provide generation inputs for these balancing reserve capacity services.

Q. What is balancing reserve capacity?

A. Balancing reserve capacity is the generation capability ready to increase or decrease generation to provide load-resource balance in the balancing authority area in order to meet reliability standards. See Fisher *et al.*, BP-14-E-BPA-21, section 2, for further explanation.

Q. Please describe the Regulation and Frequency Response Service.

A. Regulation and Frequency Response Service is an Ancillary Service that provides for the continuous balancing of loads in the BPA balancing authority area to maintain frequency at 60 cycles per second. Study section 10. TS has the obligation to maintain this balance in accordance with NERC control performance criteria. *Id.* In order to provide this service, BPA requires balancing reserve capacity from generation inputs. *Id.* Balancing reserve capacity is provided from the FCRPS through the immediate responsiveness of automatic generation control (AGC) to maintain balance in moment-to-moment changes. *Id.*

Q. Please describe VERBS and explain how balancing reserve capacity from the FCRPS is made available to TS for this service.

A. VERBS is a Control Area Service that provides balancing reserve capacity to integrate wind and solar generation projects into the BPA balancing authority area. Study section 10. It includes three components: regulating reserve for moment-to-moment variation, following reserve to account for within-hour variation, and imbalance reserve to account for the difference between actual generation and schedules within the scheduling period. *Id.* VERBS is provided by increasing or decreasing committed online FCRPS generation through the use of AGC equipment as necessary to follow the moment-by-moment changes, the within-hour variations, and the differences in variable energy resources' actual generation relative to the schedule. *Id.*

Q. Please describe DERBS and explain how balancing reserve capacity from the FCRPS is made available to TS for this service.

A. DERBS is a Control Area Service that applies to all non-Federal thermal generators in the BPA balancing authority area. *Id.* This service is necessary to support the within-hour deviations of thermal generation from the hourly generation estimate *(i.e.,* generation

schedule). DERBS is the same service for non-Federal thermal generators that is provided by VERBS to wind and solar generators. *Id.*

Q. Briefly describe the proposed methodology for allocating embedded costs associated with the provision of FCRPS balancing reserve capacity for Regulation and Frequency Response Service, VERBS, and DERBS.

A. The embedded unit cost of Regulating, DERBS, and VERBS balancing reserve capacity is calculated by taking the costs associated with the hydro projects that are used to provide balancing reserve capacity for these services and dividing these costs by the average annual capacity amount of those same hydro projects (adjusted for other requirements). Study section 3.2. The embedded unit cost is then multiplied by the quantity of balancing reserve capacity forecast for each balancing reserve capacity service to yield the embedded cost allocation for that product. *Id.*

Section 3.1: Specific Aspects and Inputs of the Embedded Cost Methodology

Q. What FCRPS resources are used to provide the balancing reserve capacity for Regulation and Frequency Response Service, VERBS, and DERBS?

A. Of the entire FCRPS, only the "Big 10" hydro projects are capable of providing balancing reserve capacity for Regulation and Frequency Response Service, VERBS, and DERBS. Study section 3.2.1. The Big 10 hydro projects are Bonneville, Chief Joseph, Grand Coulee, Ice Harbor, John Day, Little Goose, Lower Granite, Lower Monumental, McNary, and The Dalles. *Id.*

Q. Why is provision of this balancing reserve capacity limited to these resources?

A. Balancing reserve capacity for those services must be provided by units that are connected to AGC, and the Big 10 projects are the only FCRPS hydro projects BPA connects to AGC. *Id.*

Q. How is the revenue requirement for the embedded cost methodology determined?

A. The embedded cost net revenue requirement associated with the Big 10 hydro projects is composed of (1) power-related costs of the Big 10 hydro projects on a project-specific basis; (2) an allocation of associated fish mitigation costs; (3) an allocation of administrative and general expense; and (4) three specific revenue credits. Study section 3.2.8.1; Documentation Table 3.5.

Q. Please describe the allocation of the fish mitigation and administrative and general costs.

A. The fish mitigation costs and the general and administrative costs are not set on a project-specific basis. Study Section 3.2.8.1. To allocate those costs to the Big 10 hydro projects, we determine the share of the Big 10 projects to total energy of the FCRPS. Using this method, the Big 10 projects are allocated 91 percent of the fish mitigation costs and the general and administrative costs because the Big 10 projects comprise 91 percent of the energy of the FCRPS. Study section 3.2.8.1.

Q. Please describe the revenue credits.

A. The three specific revenue credits are 4(h)(10)(C) (non-operations), Colville payment Treasury credit, and synchronous condensing. *Id.* The 4(h)(10)(C) (non-operations) revenue credit is applied because BPA receives a credit against its annual payments to the U.S. Treasury in recognition of funding the system wide mitigation of BPA's fish and wildlife Direct Program expense and capital costs. *Id.* The credit is based on the non-power portion of the FCRPS projects purposes, for which power ratepayers are not responsible. Power Revenue Requirement Study, BP-14-E-BPA-02, section 1.2.6. BPA makes annual payments under a settlement to the Colville Tribes tied to the generation at Grand Coulee Dam. The Colville payment Treasury credit recognizes that BPA receives a credit against its annual payments to the U.S. Treasury as the Federal government's contribution to the annual payment to the Colville Tribes. *Id*. section 1.2.3.

The 4(h)(10)(C) (non-operations) revenue credit and the Colville payment Treasury credit are part of the net revenue requirement for the Big 10 hydro projects. *Id.* section 2 and Table 2E. The synchronous condensing revenue credit is added to avoid double-counting, because the synchronous condensing costs are allocated to TS in a separate calculation in section 5 of the Study.

Q. How are the balancing reserve capacity amounts associated with each service quantified?

A. The balancing reserve capacity amounts for each service used in the embedded cost calculation are provided by the Balancing Reserve Capacity Quantity Forecast. Study section 2. The Study provides the quantities for Regulating Reserve, VERBS Reserve, and DERBS Reserve. In addition, the Study provides a quantity for Load Following Reserve, which includes the balancing reserve capacity for load, Columbia Generating Station, and non-AGC-controlled hydro generation (Federal and non-Federal hydro generation). Also, the Operating Reserve quantity forecast is an input to the embedded cost calculation. The Operating Reserve quantity is provided by the Operating Reserve Cost Allocation. *Id.* section 4.

Q. How is the amount of balancing reserve capacity that the Big 10 can produce quantified?

A. We forecast the amount of balancing reserve capacity that can be sourced from the FCRPS. *See* Kerns *et al.*, BP-14-E-BPA-23 (explaining the flexibility and operational limitations of the FCRPS). For purposes of allocating embedded costs and determining a unit cost for balancing reserve capacity, the amount of balancing reserve capacity for the Big 10 projects is quantified by determining the 120-hour peaking capacity of these projects. Study section 3.2.7.

Section 3.2: 120-Hour Peaking Capacity

Q. What is the 120-hour peaking capacity?

A. For the FY 2014–2015 rate period, we have identified 120-hour peaking capability as a critical measure to determine resource availability to meet sustained peak loads over an extended period of time and maintain operational reliability. Study section 3.2.6. Generally, the 120-hour hydro peaking capacity is defined as the averaged period of hourly Federal system hydro generation that is calculated from the highest six hours of generation for each of five weekdays of a four-week period. Study section 3.2.5. We propose to use the 120-hour peaking capability as a measure of the system's sustained capacity for allocating costs among all capacity uses of the system.

Q. Why do you propose to use 120-hour peaking capacity?

A. We use a 120-hour capacity estimate to represent the amount of available capacity on the Federal hydro system that is available to reliably serve Federal system load obligations while meeting balancing reserve capacity obligations and power and non-power requirements within any period or water condition. *Id.* BPA planning uses 120-hour capacity to simulate the typical operation of the Federal system to meet normal weather load obligations. The measurement of 120-hour capacity for use in the embedded cost methodology is reasonable because it represents the ability of the FCRPS to meet its capacity obligations while meeting the physical characteristics and limits placed on modeled Columbia River Basin projects, including hard project constraints (*e.g.*, flow limits, elevation limits), project outages (planned and forced outages), balancing reserve capacity requirements, one-percent efficiency restrictions, and non-power constraints (flood control, variable draft limits, fish operations pursuant to the Biological Opinions, and Canadian Treaty operations). Study section 3.2.3. It is not meant to represent a time of system stress to meet large weather deviations, additional reserve obligations, or other

extreme conditions. The 120-hour capacity quantification is the same capacity planning standard used in BPA short-term planning. Therefore, using 120-hour peaking capacity to calculate the unit cost of Regulating, VERBS, and DERBS balancing reserve capacity is an appropriate measure of the system capacity capability.

Q. What modeling tools do you use in the calculation of the 120-hour peaking capacity in this study?

A. We use the combination of HYDSIM and HOSS models to calculate the 120-hour peaking capacity of the FCRPS. HYDSIM is a computer model that simulates monthly energy hydro production under the physical characteristics and limits placed on modeled Columbia River Basin projects. HYDSIM simulates these regional hydro projects to meet system load while continuing to meet Pacific Northwest regional power and non-power requirements for the 80 water years of record (October 1928 through September 2008). HYDSIM provides monthly project flows, initial and ending conditions, and constraints that are used as inputs to the HOSS model. HYDSIM is described in the Power Loads and Resources Study, BP-14-E-BPA-03, section 3.1.2.1.

The HOSS model simulates the hourly operation of the Federal system to meet hourly loads for each period of the 80 water years for the study period. Both HOSS and HYDSIM use the same project constraints and reserve requirement estimates listed above; however, HOSS focuses on hourly constraints and HYDSIM focuses on monthly constraints. The HOSS model uses monthly project flows, initial and ending conditions, and constraints supplied by the HYDSIM model in the HOSS hourly simulation. Study section 3.2.4. The outputs of HOSS are not directly used for rate setting purposes. Rather, monthly Federal system regulated hydro generation energy relationships are developed to provide monthly HLH energy, LLH energy, and 120-hour hydro peaking capacity using outputs from HOSS.

Q. Please explain how the HOSS curves are derived and used.

A. The hourly output of HOSS is used to develop relationships between monthly average energy during each of the 14 periods of the year and the associated 120-hour hydro peaking capacity for each of the 80 historical water years. Study section 3.2.4. These relationships are created through curves that define peaking capacity as a function of monthly energy for each of the 80 hydro conditions. The data from HOSS are entered into an Excel spreadsheet, and the curve-fitting function in Excel is used to generate a peaking capacity equation for each month that reflects the 120-hour peaking capacity of the system for any given energy content for that period. Study section 3.2.6.

These equations (curves), one for each of the 14 periods of the year, are applied to the energy output of HYDSIM to produce the 120-hour hydro peaking capacity for each period. Study section 3.2.7. The resulting 120-hour hydro peaking capacity amounts represent the 120-hour peaking capacity of the regulated hydro projects.

Q. Why are curves applied to the energy output of HYDSIM, and how does this represent system capacity?

A. BPA uses the monthly energy output derived from HYDSIM to estimate the amount of regulated hydro energy available for the Federal system by period for each of the 80 water conditions of record. Study section 3.2.3. This HYDSIM analysis provides monthly regulated hydro generation forecasts for energy in average megawatts while incorporating power and non-power operating requirements for the study period. HYDSIM forecasts provide the base data for the energy analysis for the loads and resources studies used in ratemaking and long-range planning and are not intended to simulate hourly Federal system operations. The HOSS model does not provide the regulated hydro energy output; rather, HOSS provides the basis for the Federal system regulated hydro energy to 120-hour peaking capacity relationships. *Id.* section 3.2.4. In

order to provide consistency, the Power Loads and Resources Study, BP-14-E-BPA-03, the Power Risk and Market Price Study, BP-14-E-BPA-04, and the Generation Inputs Study all use the same HYDSIM regulated hydro energy forecasts as the base for the 120-hour peaking capacity curves.

The curves derived from HOSS are only the energy-to-capacity relationships per period. To obtain the capacity values for each period, the curves must be applied to a monthly energy value. The model that produces the energy values by month is HYDSIM. Because the curves take into account the power and non-power constraints on the system, the resulting capacity values after applying the curves to HYDSIM amounts represent the regulated hydro system capacity.

Q. Are you proposing to use average water conditions from 1958 to determine the 120-hour peaking capacity?

A. We propose to use 1958 water conditions to forecast the 120-hour peaking capacity, because it represents the ability of the Federal system to provide peaking capacity under average water and weather conditions. The use of an average water assumption to calculate the unit embedded cost of peaking capacity continues the same assumption that was used to develop the current balancing reserve capacity rates. By using average water BPA has taken into account that secondary sales are a use of the FCRPS. Products and services that include a benefit from secondary sales are also subject to risk mitigation tools. See Mandell *et al.*, BP-14-E-BPA-15, section 3.3, for a more detailed description of the proposal regarding application of risk mitigation tools to Ancillary and Control Area Service rates.

Q. Are the generation input uses of balancing reserve capacity accounted for in the 120-hour peaking capacity forecast?

A. Yes, for all balancing reserve capacity uses except Operating Reserve, the HOSS model incorporates balancing reserve capacity forecasts using both the *inc* and *dec* balancing reserve capacity amounts, which are provided by the Balancing Reserve Capacity Quantity Forecast. Study section 2.11. The balancing reserve capacity includes: (1) *inc* balancing reserve capacity, which is capacity available to ramp up generation to meet increasing within-hour load and/or decreasing within-hour generation: and (2) *dec* balancing reserve capacity, which is generating capacity available to ramp down to meet increasing within-hour generation and/or decreasing within-hour load. *Inc* and *dec* balancing reserve capacity are explained in greater detail in section 5 below. In HOSS, the *inc* requirement is treated as a reduction to available capacity to generate power, and the *dec* requirement is treated as an increase in the minimum generation requirement at Grand Coulee, Chief Joseph, McNary, John Day, and The Dalles. The treatment of Operating Reserve in the HOSS model is described in section 3.3 below.

The HYDSIM model incorporates the *inc* and *dec* balancing reserve capacity by reducing the availability of turbines that can be used to produce power on a monthly basis. The HYDSIM generation estimate of Federal hydro energy therefore reflects power set aside for balancing reserve capacity and Operating Reserve.

Q. Does the 120-hour peaking capacity take into account the impacts associated with providing balancing reserve capacity for the various services?

A. The 120-hour peaking capacity takes into account the generation input uses of balancing reserve capacity, because those products and services diminish the usable capacity of the Federal system and reduce its capability to serve firm load. The 120-hour peaking capacity amount is intended to reflect the capacity of the system for producing energy

after all constraints, limitations, and requirements are taken into account. Study section 3.2.5. Among these requirements is the capacity for balancing reserve capacity and Operating Reserve. In the embedded cost methodology, the net revenue requirement is spread across all system capacity uses. The embedded cost calculation adds the balancing reserve capacity and Operating Reserve requirement to the 120-hour peaking capacity to yield all system capacity uses for purpose of pricing and allocating costs. *Id.* section 3.2.8.

Q. Please explain the adjustments that are made to the 120-hour peaking capacity for the regulated hydro projects to derive the capacity of the Big 10 hydro projects.

A. Two adjustments are made to the 120-hour peaking capacity of regulated hydro projects to calculate the 120-hour peaking capacity of the Big 10 hydro projects. The regulated hydro projects are reduced by 3.35 percent for capacity transmission losses. *Id.* section 3.2.7. Then the transmission loss-adjusted regulated hydro projects are multiplied by 94 percent to reflect the Big 10 hydro projects as a portion of the regulated hydro projects. *Id.*

Q. Why is the adjustment for transmission losses made?

A. The 120-hour peaking capacity is adjusted for transmission losses because the FCRPS has inherent capacity losses to transmit generation to Federal system load obligations. These capacity transmission losses total 3.35 percent of the available 120-hour generating capacity. Power Loads and Resources Study, BP-14-E-BPA-03, section 3.1.5.

Q. How is the 94 percent adjustment of regulated hydro to represent the Big 10 derived?

A. The Federal system regulated hydro represents 14 projects whose hydro operations are simulated in HYDSIM and HOSS, and thus the 120-hour peaking capacity produced from the HOSS curves reflects all 14 projects. As described above, only the Big 10 projects are controlled in real time by AGC and provide balancing reserve

capacity. To determine the necessary adjustment to the regulated hydro 120-hour peaking capacity, we take a proportion of the regulated hydro annual average energy amounts. The proportion was determined by taking the total annual average energy production of the regulated hydro projects and dividing by the annual average energy of the Big 10 hydro projects. The proportion calculated on energy was applied to the 120-hour peaking capacity. The portion of the total capacity that is associated with the Big 10 projects is 94 percent of the regulated hydro projects. Documentation Table 3.2. This 94 percent adjustment for the 120-hour peaking capacity of the Big 10 hydro projects is a measure of the Big 10 hydro projects as compared to the 14 regulated hydro projects.

Q. Have you considered using any other measurements for system capacity in the embedded cost methodology, and if so, why is BPA proposing to continue to use the 120-hour peaking capacity?

A. We considered using critical water rather than average water for the 120-hour peak for the embedded cost methodology. We are proposing to continue with the 120-hour peaking capacity under average water as the allocation factor for embedded cost because the 120-hour peaking capacity takes into account non-power constraints and is representative of the available capacity for system planning purposes. It also represents an allocation factor that balances all firm uses of the hydro system.

Section 3.3: Embedded Cost Calculation

Q. Please describe the components of the hydro project system uses that are used to calculate the embedded cost.

A. The hydro project system use is the sum of: (1) the 120-hour peaking capacity of the Big 10 hydro projects; (2) balancing reserve capacity quantities for Regulating Reserve,

Load Following Reserve, DERBS Reserve, and VERBS Reserve; and (3) capacity reserve quantity for Operating Reserve adjusted to account for non-spinning Operating Reserve that may be provided by Federal hydro projects other than the Big 10.

Q. Are the balancing reserve capacity amounts added back to the HYDSIM/HOSS-developed 120-hour peaking capacity the same as the balancing reserve capacity assumptions within HYDSIM and HOSS themselves?

A. The data inputs for HYDSIM and HOSS for Regulating Reserve, Load Following Reserve, DERBS Reserve, and VERBS Reserve are the same as the balancing reserve capacity quantities forecast in section 2 of the Study, Documentation Table 3.4, with two exceptions. The monthly total *inc* and *dec* amounts used as data inputs to the hydro models are between two and five megawatts lower than the amounts forecast in Documentation Table 2.18. These numbers will be revised to match the forecast in section 2 in the Final Study. The other exception is that the monthly total *inc* balancing reserve capacity amount was capped at 900 MW, and the monthly total *dec* balancing reserve capacity amount was capped at 1100 MW. The data inputs for HYDSIM and HOSS include *inc* and *dec* amounts, while the embedded cost methodology only uses the *inc* amount for the calculation.

The Operating Reserve amounts input into HYDSIM and HOSS are calculated in a manner that is consistent with the reserve forecast in the Operating Reserve Cost Allocation in section 4 of the Study (*i.e.*, 3 percent of resources and 3 percent of load). However, instead of using the average Operating Reserve requirements, the Operating Reserve input for HOSS and HYDSIM is based on historical peak balancing authority area generation at the 95th percentile by month.

Q. Why is the Operating Reserve amount used as an input to the HYDSIM and HOSS models different from the amount used in the hydro project system uses?

A. The peak 95th percentile is used instead of the average, because it is at high levels of generation that Operating Reserve constrains the system. If the study were to assume the average Operating Reserve, then during periods of high generation, the HYDSIM and HOSS models would not adequately cover the required Operating Reserve.

Q. Please explain the slight adjustment made to the Operating Reserve to account for Operating Reserve not supplied by the Big 10.

A. The Operating Reserve quantity used is adjusted to account for the fact that Supplemental (non-spinning) Operating Reserve can be carried on projects in addition to the Big 10 hydro projects. Since the Big 10 hydro projects comprise 91 percent of the hydro projects in the BPA balancing authority area capable of providing Operating Reserve (Documentation Table 3.6, line 7 and Table 4.4, line 17), the Supplemental Operating Reserve quantity is reduced by nine percent to account for the portion that may be carried on other projects.

Q. What is the proposed embedded cost allocation for Regulation and Frequency Response Service, VERBS, and DERBS?

A. The proposed embedded unit cost of balancing reserve capacity need for Regulation and Frequency Response Service, VERBS, and DERBS is $6.93 per kilowatt per month of balancing reserve capacity need. Documentation Table 3.6, line 13. The embedded cost allocation is $5,072,760 for Regulation and Frequency Response Service; $43,409,520 for VERBS; and $5,072,760 for DERBS. Study section 3.2.8.3; Documentation Table 3.6, lines 14-16.

Q. How does the embedded cost forecast impact power rates?

A. The embedded cost forecast is included in the Generation Inputs for Ancillary, Control Area and Other Services revenue credit for the power rates. The revenue PS collects through providing generation inputs lowers the revenue requirement for power rates.

Q. How are the embedded costs used to calculate Ancillary and Control Area Services rates?

A. The embedded cost revenue forecast is recovered through the appropriate Ancillary and Control Area Services rates. *See* Study section 10; Jackson *et al.*, BP-14-E-BPA-28.

Section 3.4: Embedded Cost for Different Levels of Service

Q. What is the level of service assumption used in your base case for the Initial Proposal?

A. We assume a 99.5 percent level of service for the base case in the Initial Proposal. Fisher *et al.*, BP-14-E-BPA-21, section 5.1. We also assume an amount of customer self-supply of the imbalance component of VERBS, which affected the balancing reserve capacity quantity. See Study section 2.7.4 for further explanation of this assumption.

Q. What is the reason for providing the embedded cost calculations for a scenario where no customer elects to self-supply the imbalance component in VERBS?

A. We are providing cost calculations for one scenario in addition to the base case in order to inform the parties of the cost impacts. The base case assumes a level of self-supply based on the Customer-Supplied Generation Imbalance Pilot Program, which began in September 2010. Customers must make elections by April 1, 2013, whether or not to self-supply for the FY 2014–2015 rate period. This amount will be used for the self-supply amount for the Final Proposal. There is the possibility that no customers will elect to self-supply for the rate period.

Section 4: Direct Assignment of the Wind Integration Team (WIT) Costs

Q. What costs are proposed to be directly assigned to the VERBS rate?

A. We are proposing to directly assign Wind Integration Team costs to the VERBS rate. Study section 3.3.

Q. Please briefly describe why BPA formed the WIT.

A. As part of the WI-09 Settlement, BPA assembled the internal cross-agency WIT to explore technical solutions to address the challenge of balancing loads and resources to preserve system reliability while accommodating the rapid development of wind energy in the BPA balancing authority area. The mission of the WIT is to clearly define and execute a plan for integrating wind generation in a manner that allows for the continued highly reliable operation of the Federal power and transmission system at the lowest cost consistent with sound business and operations practices.

Q. Please describe the various initiatives the WIT has undertaken since its formation.

A. The WIT has developed and implemented numerous initiatives that have helped allow for a steady increase in the amount of wind interconnected to BPA's balancing authority area. These initiatives will continue in the FY 2014–2015 rate period. These initiatives include:

(1) Dispatcher Standing Order 216 (DSO 216): DSO 216 is BPA's primary reliability tool for managing variable generation. The purpose behind BPA's operating procedure known as DSO 216 is to fix the limits of BPA's balancing reserve capacity obligations and ensure the reliability of the BPA balancing authority. DSO216 has also been deployed as a way to reduce wind integration costs by substituting infrequent wind curtailments for additional balancing reserve capacity. This operating protocol limits the amount of Federal hydro capacity BPA must set aside to back wind while maintaining reliable power service. BPA developed automated tools and communication protocols to

either limit actual wind generation according to the set schedule or curtail e-Tags to actual generation in response to the amount of balancing reserve capacity deployed.

(2) Dynamic transfer capability (DTC): Methodology developed for determining dynamic transfer limits and application of that methodology to nine transmission paths. BPA then developed and implemented a process to allocate and award this DTC to requesting utilities. That process, collaboratively developed with customers, resulted in new DTC offers, awards, and Dynamic Transfer Operating Agreements that are now in effect.

(3) Forecasting, state awareness tools: BPA has created a wind speed and wind generation forecasting system in order to achieve better reserve management through more informed operational decisions and to enhance hydroelectric generation reliability as more wind generation is integrated. BPA will also be making the wind generation forecast for each wind project available to the project owner/operators to improve scheduling accuracy.

(4) Intra-hour power scheduling: The purpose of the Intra-hour Scheduling initiatives is to reduce the error between actual performance and schedules to allow for a more efficient use of the existing electric system without sacrificing reliability and to assist with the integration of variable energy resources.

- ***Committed Intra-Hour Scheduling***

  Committed Intra-Hour Scheduling is designed to reduce the amount of balancing reserve capacity held by BPA by reducing the difference between scheduled and actual output. In exchange for a rate discount, participating parties commit to schedule, with a scheduling accuracy equal to or better than 30-minute persistence for each scheduling period.

- ***California Independent System Operator (CAISO) Intra-Hour Scheduling Pilot***

CAISO Intra-Hour Scheduling Pilot is designed to move a portion of the balancing needed for wind that is serving load in California to the CAISO. This reduces the amount of balancing reserve capacity held by BPA by reducing the difference between scheduled and actual output.

(5) Customer-supplied generation imbalance: The purpose of the Customer-Supplied Generation Imbalance (CSGI) effort is to provide customers a choice of balancing reserve capacity suppliers and to reduce the overall dependence on the FCRPS for balancing capacity and energy.

(6) Supplemental Service: As the amount of wind on the BPA system has continued to increase and approaches the point of exhausting the ability of the FCRPS to provide balancing reserve capacity (*incs* and *decs*), BPA has begun exploring the purchase of balancing reserve capacity from third parties. BPA developed a Supplemental Service program to allow parties, including BPA, to procure additional balancing capacity and energy. BPA is now working on an "Enhanced Supplemental Service" to increase the ability of market participants to make shorter-term purchases of balancing capacity to meet the balancing needs of their variable energy resources.

(7) WebExchange (WebEx): Previously known as Intra-Hour Transaction Accelerator Platform (ITAP). The purpose of the WebEx is to provide a bulletin board market for hourly and intra-hour power products, with functionality to submit transmission service requests (TSRs) and e-Tags for confirmed transactions. This market will speed up the transaction process, making it more feasible to meet the short deadlines for intra-hour schedules.

Q. Please briefly explain how BPA funds the WIT.

A. The WIT is an internal cross-agency team made up of employees from PS and TS, with support from Corporate and Legal. For FY 2012–2013 the WIT was funded with funds that were directly assigned to VERBS as well as from the reinvestment of Green Energy Premium (GEP) revenues. For FY 2014–2015, we propose that the WIT budget be directly assigned to the VERBS rate.

Q. What changes from the BP-12 rate case are you proposing for costs directly assigned to the VERBS rate?

A. In BP-12, BPA assigned the cost of the TS employees' participation in the WIT along with the TS portion of Corporate and Legal employee costs for supporting WIT initiatives. The Power Services' share of WIT costs, including the PS portion of Corporate and Legal employee costs for supporting WIT, were covered by unspent GEP revenues.

In BP-14, we are proposing to directly assign the full WIT costs. It is assumed that the GEP revenues will be fully reinvested prior to the FY 2014–2015 rate period and will not be available to help fund WIT initiatives in that rate period. In addition, Staff is not proposing to include the Dec Acquisition Pilot Project in the direct assignment for this rate case. In place of the Dec Acquisition Pilot Project, Staff proposes to offer enhanced Supplemental Service. Staff proposes that the implementation costs of enhanced Supplemental Service, which will allow BPA to make *inc* and *dec* purchases, will be included in the WIT budget.

Q. What is the policy rationale for directly assigning these WIT costs to the VERBS rate?

A. These costs are for staffing BPA's WIT and associated projects described above that are for the benefit of variable energy resources in the BPA balancing authority area. Directly assigning these costs to the VERBS rate is consistent with BPA's principles of cost

causation and BPA's goal to ensure no cost shifts result from the integration of variable energy resources.

Section 5: Variable Cost Pricing Methodology

Q. Why are you proposing to include a variable cost component in the generation inputs cost allocation?

A. BPA operates an interconnected system of dams and reservoirs. Providing balancing reserve capacity affects the FCRPS output in MW, timing of energy generated, and revenues received. Losses of efficiency and value occur as the system is set up to allow balancing reserve capacity to be deployed, and additional losses occur as the balancing reserve capacity is actually deployed. The inclusion of this variable cost component in the cost allocation study allows BPA to appropriately allocate these losses to the parties who benefit from balancing reserve capacity services.

Q. Generally, how is the variable cost of balancing reserve capacity calculated?

A. The variable cost of balancing reserve capacity is calculated in two general steps. The first step calculates the cost of making the FCRPS capability ready and available should the need to deploy balancing reserve capacity arise. These costs are referred to throughout this testimony as stand-ready costs. The second step calculates the cost of actually deploying balancing reserve capacity as the need arises. These costs are referred to throughout this testimony as deployment costs. The tool used to calculate both of these costs is the Generation and Reserves Dispatch (GARD) Model. Study section 3.4.1.

Q. What is the GARD Model, and what does it do?

A. The GARD Model is an MS Excel 2003-based model developed by BPA. All inputs and outputs are based in Excel spreadsheets. The core of the model is written in Visual Basic

for Applications (VBA). The GARD Model analyzes variable costs in two general categories. The first category is the stand-ready costs, which are the costs associated with making a project capable of providing balancing reserve capacity. The next general cost category is the deployment costs, which are those costs incurred when the system uses its reserve capability to actually deliver in response to a reserve need.

Q *Please describe changes that have been made to the GARD Model inputs as compared to the inputs that were used in the FY 2012–2013 rates.*

A. The model has been changed from a 70-water-year study to an 80-water-year study to match the upgrade made to HYDSIM.

Q. Please describe the inputs for the GARD Model.

A. The GARD Model uses inputs from the HYDSIM model, a hydraulic model of coordinated river operations that calculates monthly average generation values by project for each of the 80 historical water years, and actual system data. The primary inputs into the GARD Model are tables of project-specific generation values calculated by HYDSIM.

Q. What are the tables of project-specific generation values used for?

A. These generation tables are used by the GARD Model to determine the generation request which, in turn, determines the project's unit commitment and dispatch. The generation request is the amount of Heavy Load Hour (HLH) or Light Load Hour (LLH) generation that a specific project is being asked to produce. The project's unit commitment consists of the number and type of units put online. The dispatch is the generation from the committed units, which must equal the generation request while at the same time meeting the balancing reserve capacity obligation.

Q. How do you determine a project's specific HLH and LLH generation request?

A. HYDSIM output tables are input into a pre-processing spreadsheet to calculate each project's HLH and LLH generation request prior to considering balancing reserve capacity needs. Determining a project's specific HLH and LLH generation request begins with monthly energy amounts for each of the 80 historical water years from HYDSIM. Monthly energy amounts are taken for Grand Coulee (GCL), Chief Joseph (CHJ), John Day (JDA), and The Dalles (TDA). These four projects are generally referred to as controller projects in this testimony. Additionally, the pre-processing spreadsheet calculates amounts of pre-existing *dec* capability for each project by month and historical water year based on the calculated LLH generation and the project's minimum flow and/or minimum generation requirements. The purpose of pre-existing *dec* capability input is to ensure that the GARD Model does not unnecessarily move energy out of HLH and into LLH when providing *dec* capability.

Q. Please describe the nature of the generation request in more detail and explain how it is applied to the GARD Model.

A. The generation request is a set of generation values, for each month of each water year for HLH and LLH, for each controller project. Given these generation values, the GARD Model finds the efficiency-maximizing unit commitment and dispatch. This process mimics the basepoint setting process, where the hydro duty scheduler submits requested generation amounts to each project and the project commits and dispatches its units in the most efficient manner possible in order to meet the requested generation.

Q. If the Big 10 projects are all capable of AGC response, why are costs calculated based on only four controller projects?

A. Although all of the Big 10 projects are capable of being, and at various times of the year are, armed for AGC response, GCL, CHJ, JDA, and TDA are the only projects analyzed,

because these four controller projects are most often armed by the hydro duty scheduler for AGC response. Study section 3.4.2.1. The projects used in the GARD Model are representative of the entire Big 10, because the variable costs calculated by the GARD Model are based on recovering the total amount of balancing reserve capacity needed for generation inputs provided to TS, and this cost calculation is unlikely to vary to any significant degree whether it is spread over the four controller projects or over the Big 10 projects. Using the four controller projects that are armed most often made the GARD Model manageable, as compared to trying to factor in all 10 projects that have varying degrees of use depending on the time of year. *Id.*

Q. What is the next step in calculating variable costs?

A. Supervisory Control and Data Acquisition (SCADA) data are used to develop relationships between average energy and HLH generation for each of the respective controller projects. This evaluation is constrained by unit availability, requirements to operate within 1 percent of peak efficient generation, and minimum turbine flow constraints. The results of the evaluation are functional relationships between average energy production and HLH generation.

Q. What time period does the SCADA data span?

A. Data are from the period January 1, 2002, through December 31, 2007.

Q. Why is the 2002–2007 data set used in developing the functional relationship between energy and HLH?

A. The 2002–2007 period is used to balance the need for a robust data set with the desire for operations that are similar to current practice and bound by similar constraints. Going back farther in time would include periods when operations were significantly different from current operations. Additionally, this period serves well as a base case because it predates the large wind fleet buildup.

Q. Please explain what inc *reserves and* dec *reserves are.*

A. *Inc* reserves, a component of balancing reserve capacity, are required to maintain load-resource balance when an undergeneration situation exists within the BPA balancing authority area. In an undergeneration situation, instantaneous loads are higher than planned and/or the instantaneous generation is lower than planned. Under these circumstances, FCRPS generation must automatically *inc* to maintain system balance. The quantity of reserve is set to cover 99.75 percent of all under-generation magnitudes calculated by the Balancing Reserve Capacity Quantity Forecast.

Conversely, *dec* reserves, also a component of balancing reserve capacity, are reserves required to maintain load-resource balance when an over-generation situation exists within the BPA balancing authority area. In an over-generation situation, instantaneous loads are lower than planned and/or the instantaneous generation is higher than planned. Under these circumstances, FCRPS generation must automatically *dec* to maintain system balance. The *dec* reserve is the amount of generation that the FCRPS must be capable of decreasing to maintain load-resource balance. The quantity of reserve is set to cover 99.75 percent of all over-generation magnitudes calculated by the Balancing Reserve Capacity Quantity Forecast.

Q. Please explain how inc *reserves and* dec *reserves are treated in the GARD Model.*

A. Reserves are input into the GARD Model by general reserve type. Specifically, the reserves are input into the model by quantity of *inc* and *dec* regulation, *inc* and *dec* following, *inc* and *dec* imbalance, and total Operating Reserve. Given these reserve classifications, the GARD Model determines the required amounts of spinning and non-spinning reserve to meet *inc* obligations and the amount of generation required to meet *dec* obligations. It is worth noting that all *dec* reserves must by definition be supplied from generation that is spinning.

Q. How are the quantities of spinning and non-spinning reserve determined?

A. The determination of the quantities of spinning reserve versus the quantities of non-spinning reserve is derived from the NERC requirements as well as system operator judgment. NERC requires that at least 50 percent of the balancing authority area Operating Reserve obligation be met with spinning capability responsive to AGC. NERC also requires that 100 percent of the balancing authority area Regulating Reserve must be carried on units with spinning capability responsive to AGC, because Regulating Reserve must respond on a moment-to-moment basis. In contrast, following and imbalance reserve do not have NERC-defined criteria, and therefore it is assumed that at least 50 percent of the *inc* following reserve must be carried as a spinning obligation and up to 50 percent as a non-spinning obligation. For imbalance reserve, up to 100 percent of the *inc* obligation may be met with non-spinning capability. Study section 3.4.2.2.

Q. Why does the GARD Model need to maintain the spinning portion of the Operating Reserve obligation at all times as compared to the balancing spinning obligation?

A. Operating Reserve is not used for balancing purposes and may be deployed only for qualifying contingencies. WECC's currently proposed standard, which is expected to go into effect during the rate period, requires Operating Reserve to be maintained within the balancing authority area equal to three percent of hourly integrated load plus three percent of hourly integrated generation. Because the Operating Reserve requirement is an instantaneously calculated requirement, even momentarily using Operating Reserve capability for balancing purposes is a violation of WECC/NERC disturbance criteria.

Q. What amount of regulation inc *and* dec, *following* inc *and* dec, *imbalance* inc *and* dec, *and Operating Reserve are assumed for purposes of running the GARD Model?*

A. On average, the regulation *inc* and *dec* are 126 MW and -125 MW, the following *inc* and *dec* are 377 MW and -386 MW, and the imbalance *inc* and *dec*, adjusted for self-supply

and committed intra-hour scheduling, are 399 MW and -559 MW. The Operating Reserve obligation is 559 MW. These amounts are derived from the Balancing Reserve Capacity Quantity Forecast, Table 2.18 of the Documentation, and the Operating Reserve Cost Allocation, section 4.3 of the Study.

Section 5.1: Specific Variable Costs Considered by the GARD Model

Q. What are the basic categories of variable costs used in the GARD Model?

A. There are two broad categories: stand-ready costs and deployment costs.

Q. Please describe the stand-ready costs.

A. In order to meet the potential reserve requirements on any given hour, BPA's system must be set up to respond to these reserve needs going into the operational hour. Stand-ready costs are those variable costs associated with making the FCRPS capable of providing the required reserve. In short, they are the costs that arise from making the FCRPS ready to deploy reserves as needed. There are four general subcategories of stand-ready costs: energy shift, efficiency changes, cycling losses, and spill.

Q. What are deployment costs and how do they differ from stand-ready costs?

A. Deployment costs are those variable costs realized when the FCRPS actually deploys the stand-ready reserve to automatically *inc* or *dec* generation in order to meet a within-hour reserve need that changes due to variations in loads and resources. The costs of meeting the within-hour variations in loads and resources are referred to as "deployment costs." There are three general cost subcategories for deployment costs: response losses, cycling loss, and spill.

Q. How does the act of providing balancing reserve capacity result in the type of stand-ready cost quantified as energy shift cost?

A. Energy may shift out of the HLH period in order to make *dec* capability available during the LLH period and/or to make available sufficient non-spinning and/or spinning *inc* capability during the HLH period. In the first instance, fuel normally used to meet peak generation needs is consumed during periods of lowest demand to make sufficient generation capability available on the FCRPS to fully deploy available *dec* reserves without violating minimum generation requirements. The need to shift energy is typically driven by the need to generate during the graveyard period. Depending on water conditions, energy may also be shaped into the shoulder LLH period to make available *dec* capability. In making available non-spinning and spinning *inc* capability, energy shift impacts typically manifest as a reduction first in super-peak generating capability followed by a bleeding into the shoulder HLH period. Should additional *inc* capability be required after completely flattening generation across the HLH period, such as in high flow scenarios, energy is shifted into the shoulder LLH period and, eventually, into the graveyard period.

Q. How are the energy shift costs calculated?

A. Energy shift costs are calculated by determining how much energy is moved into or out of the four blocking periods of super-peak (top eight demand hours of the day), shoulder HLH (this varies, but typically consists of clock hours 07:00 through 12:00 and 21:00 through 22:00), shoulder LLH (clock hours 23:00 through 00:00 and 05:00 through 06:00), and graveyard (clock hours 01:00 through 04:00). Energy shifted into or out of a given time block is multiplied by the price for that time block, resulting in a revenue or expense. If the energy is shifted into a time block with a higher price than the time block BPA would have otherwise sold the energy in, then the net result is a revenue for BPA.

If the energy is shifted into a time block with a lower price than the time block BPA would have otherwise sold the energy in, then the net result is an expense for BPA. Totaling the revenues from energy shifted into a time period and the expenses from energy shifted out of other time periods yields the energy shift cost. The greater the differential in price between time periods, the greater the cost of energy shift. Prices for the time periods are taken from the Power Risk and Market Price Study, BP-14-E-BPA-04, section 2.4.

To calculate the energy shift cost associated with a quantity of balancing reserve capacity, the GARD Model compares generation values for the no-reserves case to the test case where balancing reserve capacity must be carried. Study section 3.4.3.1. If energy is moved out of HLH and into LLH due to the need to increase LLH generation and/or graveyard generation and/or to make sufficient super-peak or HLH *inc* capability, the HLH generation is reduced and the LLH generation is increased relative to the base case. Because HLH generation is more valuable than LLH generation, an economic impact is realized as HLH generation is reduced and LLH generation is increased.

Q. Does an energy shift cost always arise as a consequence of providing balancing reserve capacity?

A. No. To the extent the generation request plus the *inc* balancing reserve capacity is less than maximum generation, energy will not be shifted into the LLH period, because there is sufficient generating capability available to meet the generation request as well as the full *inc* balancing reserve capacity.

Additionally, to the extent that the shoulder LLH and graveyard generation is already above minimum generation, there is no need to pull energy out of the HLH period. In these instances, "pre-existing *dec*" capability is said to exist. If the pre-

existing *dec* capability does not fully meet the *dec* requirement, energy is shifted out of the HLH period and into the shoulder LLH and/or graveyard periods.

Q. *Is the energy shift cost dependent on the amount of spinning reserves versus non-spinning reserves?*

A. No. Energy shift costs are the same whether the *inc* reserve is carried as spinning or non-spinning. Energy shift due to *inc* reserve obligations is a function of the difference between maximum generating capability and the requested generation from a controller project. If the difference between the maximum and requested generation is less than the *inc* reserve need, energy must be shifted. Whether the reserve is carried as spinning, non-spinning, or some combination thereof is then a function of the unit commitment and dispatch. The mere need for reserve capability is the cause of energy shift costs, regardless of whether that capability is available as spinning or non-spinning.

Q. *In determining energy shift costs, why are the super-peak and graveyard time periods explicitly taken into account?*

A. The super-peak and graveyard time blocks are explicitly considered for the energy shift calculation because these periods are first impacted by the need to shift energy when making available *inc* and/or *dec* capability. Not considering these periods would understate the energy shift impact. When providing *dec* capability, the first period where energy shift may be required is the relatively low-value graveyard time block. Explicitly calculating how much energy is put into the graveyard period for the *dec* provides greater accuracy in the cost calculation. Likewise, should energy be put into the LLH period to provide *inc* capability, accounting for the quantity of energy going into the graveyard time block versus the shoulder LLH block improves the accuracy of the cost calculation.

Q. Does shifting energy to make incs *available provide an offset to the cost of providing* dec *balancing reserve capacity?*

A. No. The typical scenario where energy is shifted to provide *inc* balancing reserve capability is periods of high river flows. In these instances, sufficient LLH generation exists to provide the required *dec* balancing reserve capacity. Any additional energy shifted into the LLH period due to increasing the HLH *inc* capability results in having *dec* capability greater than required.

Q. Based on the amount of balancing reserve capacity forecast in the Balancing Reserve Capacity Quantity Forecast, what is the proposed forecast cost for energy shift?

A. The total average energy shift included in the Initial Proposal for the FY 2014–2015 period is 1,845,055 MWh, worth $11,984,612. Documentation Table 3.11, lines 1-3.

Q. Are energy shift costs reflected in any other rates study?

A. Yes. The impact of the energy shift is implicit in the RevSim studies, described in the Power Risk and Market Price Study, BP-14-E-BPA-04, section 2.5.

Q. What is RevSim, and does RevSim accurately reflect the energy shift cost?

A. RevSim is a revenue simulation model used to calculate the net revenues and net revenue at risk. *Id.* RevSim underestimates the energy shift cost, because it does not consider monthly time periods any more refined than the HLH and LLH blocks. As a result, an adjustment is passed to RevSim based on the GARD super-peak graveyard adjustment in order to fully reflect the cost of the energy shift. The MWh of energy shift sent to RevSim from the GARD Model are incremental amounts above and beyond the hydro generation amounts supplied to RevSim by the Loads and Resources study, BP-14-E-BPA-03. Power Risk and Market Price Study, BP-14-E-BPA-04, sections 2.5 and 2.6.

Q. Besides the adjustment for energy shift, are any of the other results from the Variable Cost Pricing Methodology provided to RevSim?

A. Yes. The full impact of all remaining cost categories are sent to RevSim. These impacts (represented as MWh adjustments) include efficiency changes, cycling costs, spill costs, response losses, deployment cycling costs, and deployment spill costs. *Id.* section 2.5. Each of these cost categories will later be explained in further detail.

Q. How are the MWh amounts taken into account in RevSim?

A. RevSim receives a data table containing the incremental amounts of energy shift. The impacts are characterized as MWh of gains and losses by diurnal time period, month, and water year.

Q. Please describe the subcategory of stand-ready costs known as efficiency change.

A. Efficiency changes are increases or decreases in the controller projects' ability to generate power given a flow of water.

Q. Why does the act of providing balancing reserve capacity change the efficiency of the FCRPS?

A. Efficiency may change when providing reserves because the generation for a given time period may change. For example, providing *dec* capability for the LLH time period may require the shifting of energy from the HLH period into the LLH period. This shifting of energy results in increased LLH generation. Altering the generation dispatch often results in a change in the efficiency of a given controller project, because project efficiency changes with respect to generation. In order to make sufficient reserve capability, the dispatch and resulting efficiency from the controller projects may be altered.

Q. How does the GARD Model use the inputs described above to calculate the efficiency losses?

A. The GARD Model uses the above inputs to calculate the change in efficiency between a base case, where balancing reserve capacity is not required, and the test case requiring defined quantities and components of balancing reserve capacity. For each month of each water year for HLH and LLH, the GARD Model finds the efficiency-maximizing unit commitment and dispatch for GCL, CHJ, JDA, and TDA. The efficiency-maximizing unit commitment and dispatch are found for the base case and the case in which balancing reserve capacity is being carried. Any observed difference in the efficiency between the base case and balancing reserve capacity case is the change in efficiency. More precisely, the GARD Model models efficiency by minimizing the amount of water consumed per MW generated. Efficiency changes may occur when having to change generation patterns for energy shift or when making available additional spinning reserve capacity *(inc)* by maintaining generation at the same level but putting additional units online.

When holding balancing reserve capacity, it is often necessary to shift energy, with the consequence of changing the requested generation from a controller project for a given time period. Changing the generation level typically results in a change in efficiency, because generator units at a given controller project will be dispatched to an alternate loading, and the number and type of unit used to meet the generation request will often differ from the base (no balancing reserve capacity) case.

Similarly, should additional spinning reserve capacity be required, the generation request remains unchanged; however, requiring additional spinning reserve capacity causes changes to the unit commitment and dispatch. Changing the unit commitment and

unit dispatch at a controller project often adversely changes the ability of the controller project to minimize flow per unit of generation.

Q. Is there any interplay between the cost of energy shift and efficiency changes, and if so, how does the GARD Model account for this interaction?

A. Yes, there is an interaction between energy shift and efficiency changes. Any time energy shift occurs, the generation request from a controller project changes. Generation changes due to shifting energy may result in changing the efficiency of the controller project. The GARD Model accounts for the interaction by separately calculating the costs associated with energy shift and the impact associated with efficiency changes. Study sections 3.4.3.1 and 3.4.3.2.

Q. Are efficiency losses always a consequence of providing balancing reserve capacity?

A. Changing a controller project's unit commitment and dispatch may result in an efficiency gain or loss. Gains may be observed if the generation request, inclusive of balancing reserve capacity, decreases the water flow per MW generated. For example, under a high flow condition, carrying *inc* balancing reserve capacity may push units operating beyond peak efficiency back toward peak efficiency. Another example may occur on LLH, during which increased *dec* balancing reserve capacity result in a higher generation level, requiring the commitment of additional online units. The additional online units may allow for an increased ability to fine-tune the optimization of the unit dispatch, with a resultant increase in overall controller project efficiency.

Q. Please explain how efficiency can increase when providing balancing reserve capacity when the baseline measurement is an optimal unit commitment and dispatch.

A. An optimal efficiency exists for any given level of generation. Adding balancing reserve capacity without changing generation almost always reduces efficiency. But adding balancing reserve capacity and changing generation may result in an increase in

efficiency, because the efficiency lost by adding balancing reserve capacity is offset by the efficiency gained by moving the generation to a generation level yielding a more efficient operation from the outset. Without balancing reserve capacity, the achieved efficiency gain from altering the generation level would have been even greater, because a more efficient unit commitment and dispatch could be achieved.

Q. How does a spinning reserve obligation versus a non-spinning reserve obligation affect efficiency losses?

A. For any given *inc* reserve capacity obligation, the greater the proportion of spinning capability required, the greater the efficiency loss. Conversely, the higher the proportion of the balancing reserve capacity carried as non-spinning, the lower the efficiency loss. This is true because the greater the spinning obligation, the greater the impact the obligation has on unit commitment and dispatch. Spinning reserve capacity is defined as unloaded turbine capability, so increasing the spinning obligation requires increasing the unloaded turbine capability. Unloading turbines and increasing the number of online turbines generally results in decreasing the efficiency of the controller project. Conversely, carrying more *inc* reserve capacity as non-spinning provides the controller project with increased flexibility to optimally commit and dispatch units, resulting in improved efficiency. The greater the non-spinning allowance, the more the *inc* obligation can be provided by idle turbines.

Q. How are efficiency losses priced?

A. Efficiency losses are priced using the market price forecast HLH price, as calculated in the Power Risk and Market Price Study, BP-14-E-BPA-04, section 2.4.

Q. Why is the HLH price used for all periods?

A. The HLH price is used for all periods because no matter whether losses occur during the HLH period or the LLH period, they both result in decreased HLH generation. For the

HLH period, HLH generation is directly lost during that period. For the LLH period, the same is true, because losses realized during the LLH period result in increased consumption of water that would otherwise have been used to generate during the HLH.

Q. Are efficiency gains also priced at the HLH price?

A. Yes.

Q. Are efficiency changes accounted for in any other rate studies?

A. Yes. All efficiency impacts calculated by the GARD Model are sent to RevSim, so the impact to secondary net revenue is correctly reflected. Power Risk and Market Price Study, BP-14-E-BPA-04, section 2.5.

Q. Based on the amount of reserves described in the Balancing Reserve Capacity Quantity Forecast, what is the proposed forecast cost for efficiency losses?

A. The total annual average efficiency change for the FY 2014–2015 period is a gain of 37,631 MWh, worth $929,237. Documentation Table 3.11, lines 4-6.

Q. Please describe the sub category of stand-ready costs known as cycling losses.

A. Cycling losses originate from the additional synchronization and ramping of units in order to have reserve capability standing ready to respond.

Q. How are the cycling losses calculated?

A. Unit cycling losses originate from the additional synchronization and ramping of units. For cycling, the number of units cycled online or offline is calculated by comparing the online units for each unit family at a given controller project in the base, no reserves, case to the online units in the case where the reserve requirement is being met. To the extent that more or fewer units are online, a cycling cost is realized. Study section 3.4.3.3.

Q. How does the GARD Model use the inputs described above to calculate the cycling loss?

A. The loss calculations are controller project-specific and are functions of the individual unit family efficiency curves and the level of generation required from the individual

units to meet the generation request for the amount of reserves required to stand ready. For each unit cycle, synchronization and ramping losses are calculated. During synchronization, water is lost as the unit is spun to synchronize to grid frequency. Water losses during synchronization are equal to 10 percent of full gate flow for three minutes. Ramping losses occur as the unit ramps up to its required generation level. Losses associated with ramping are calculated by evaluating the integral of the specific unit efficiency function from minimum generation to requested generation. The GARD Model fully ramps units to their requested generation level over seven minutes, which is the typical time to fully ramp up to a unit's requested loading.

Q. Why is the cycling loss assumed to occur on each HLH or LLH period?

A. Cycling losses are assumed to occur in each HLH and LLH period for a given month, because the generation level for each HLH and LLH period in a given month is considered an average generation representative of all HLH and LLH periods in the month. That is, in the GARD Model, all HLH periods for a month are considered identical to each other, and all LLH periods for a given month are considered identical to each other. An observed unit cycle during any HLH or LLH period is said to occur for each HLH or LLH period in the month. For example, if one additional unit is online during the HLH period relative to a case without a reserve requirement, that one additional unit is assumed to be needed for each HLH period in the month. That is, for a month with 31 HLH periods, 31 additional unit cycles would occur relative to the case without reserves. These additional unit cycles occur for all time periods, because the balancing reserve capacity must be available at all times.

Q. How are the cycling losses priced?

A. All base cycling losses are priced at the monthly HLH price from the market price forecast, as calculated in the Power Risk and Market Price Study, BP-14-E-BPA-04, section 2.4.

Q. Why is the HLH price used for all periods?

A. The HLH price is used because the base cycling impact (that is, losses in energy) is taken out of the HLH period. In other words, since BPA seeks to maximize HLH sales, efficiency losses translate to lost HLH generation with resulting lost sales.

Q. Does the GARD Model or the Variable Cost Pricing Methodology include any additional maintenance costs associated with the additional cycling required by standing ready to provide reserves or deploying reserves?

A. The calculation of cycling losses, whether for stand-ready purposes or deployment purposes, does not attempt to account for any additional maintenance costs due to frequent cycling of the units. There are additional maintenance costs associated with the additional cycling of units to provide reserves, but for purposes of this rate proposal we have not attempted to evaluate these additional costs. All maintenance costs are included in the overall reserve requirement that is used in the embedded cost analysis.

Q. Are cycling losses accounted for in any other rate studies?

A. Yes. All cycling losses calculated by the GARD Model are sent to RevSim, so the impact to secondary net revenue is correctly reflected. Power Risk and Market Price Study, BP-14-E-BPA-04, section 2.5.

Q. Based on the amount of balancing reserve capacity described in the Balancing Reserve Capacity Quantity Forecast, what is the proposed forecast for cycling cost?

A. The total annual average cycling loss for the FY 2014–2015 period is 5,482 MWh, worth $170,803. Documentation Table 3.11, lines 7-9.

Q. Please describe the subcategory of stand-ready costs known as spill losses.

A. Spill losses may occur given the combination of a large *inc* reserve obligation and high river flows. In this instance, water may need to be spilled, because it cannot be shifted into LLH and must be moved past the controller project.

Q. How are the spill losses calculated?

A. Spill losses may be incurred when the GARD Model must flatten (that is, set equal) the HLH and LLH generation in order to provide balancing reserve capacity. The flattened generation profile maximizes the combined *inc* and *dec* capability across all hours. Should the GARD Model still fail to carry sufficient *inc* capability, it will begin spilling to achieve the joint objective of meeting the *inc* reserve obligation as well as the controller project flow requirements. Study section 3.4.3.4.

Q. How does the GARD Model use the inputs described above to calculate the spill loss?

A. The loss calculations are controller project-specific and are functions of the amount of response, the total *inc* reserves required, and the total quantity of energy that must pass the project. The GARD model calculates the amount of energy that cannot be passed through the turbines for power production due to the need to stand ready with *inc* capability.

Q. Are the spill losses a result of having to provide balancing reserve capacity, or are these costs a result of having too much water in the FCRPS?

A. The spill losses calculated in the GARD Model are the result of having to provide balancing reserve capacity. Any spill resulting from excess water in the FCRPS is handled by HYDSIM and is not part of the balancing reserve capacity cost. Incremental amounts of spill above and beyond that normally occurring in a high-water scenario may result from providing *inc* reserves. It is these *inc* amounts that are reflected in the GARD Model's spill costs.

Q. How are the spill losses priced?

A. Spill losses are valued in the diurnal period in which they occur. Spill occurring during the HLH period is valued at the HLH price, and spill occurring during the LLH period is valued at the LLH price.

Q. Are spill losses accounted for in any other rate studies?

A. Yes. All spill losses calculated by the GARD Model are sent to RevSim, so the impact to secondary net revenue is correctly reflected. Power Risk and Market Price Study, BP-14-E-BPA-04, section 2.5.

Q. Based on the amount of reserves described in the Balancing Reserve Capacity Quantity Forecast, what is the proposed forecast for spill cost?

A. The total annual average spill loss for the FY 2014–2015 period is 63,951 MWh, worth $1,444,655. Documentation Table 3.11, lines 10-11.

Q. Please describe deployment costs.

A. Deployment costs are those variable costs incurred when the FCRPS automatically increases or decreases generation in order to balance the system. These costs are distinct from the stand-ready costs. The cost subcategories for deployment costs are response losses, cycling loss, and spill loss.

Q. What inputs does the GARD Model use to determine deployment costs?

A. In addition to the same inputs used in calculating the stand-ready cost, the GARD Model uses a SCE file containing *inc* and *dec* signals for each minute of each month being studied. The SCE signal is the sum of the difference between actual and scheduled balancing area generation and the difference between actual and scheduled balancing area load. The SCE file contains the total difference between actual and scheduled for all load and generation on a one-minute time-step, resulting in an amount of *inc* or *dec* that must be provided by the controller projects.

Q. Why is this additional input used?

A. The SCE file is used to model the efficiency changes, unit cycles, and spill during the course of actually deploying reserves.

Q. Are there any overlaps between the deployment and the stand-ready costs?

A. No. Deployment costs are realized only when controller projects respond to a balancing reserve capacity need. Stand-ready costs, on the other hand, are realized when setting up the FCRPS with the capability to respond.

Q. Please describe the sub-category of deployment costs known as response losses.

A. Response losses are a type of efficiency loss experienced when committed units are deploying *inc* or *dec* reserves in response to a balancing need. The GARD Model responds to a balancing need on a minute-to-minute basis, as directed by the SCE file, by dispatching committed units with the objective of maintaining load-resource balance while continuing to minimize the water used to produce power at the requested level of generation. Study section 3.4.4.1.

Q. How are response losses different from efficiency losses?

A. Response losses are incurred when a controller project that is standing ready to respond actually increases or decreases generation in response to a reserve need.

Q. How are response losses calculated in the GARD Model?

A. The GARD Model calculates response losses by continually optimizing the unit dispatch, loading each online unit such that the marginal cost of each unit is identical while at the same time meeting the requested generation level and maintaining the Operating Reserve obligation. The efficiency changes are calculated on a minute-to-minute basis. All efficiency changes occurring during the LLH period are totaled by month and reported as a monthly LLH total. Likewise, all efficiency changes occurring during the HLH period are totaled by month and reported as a monthly HLH total. *Id.*

Q. *Does the energy used in providing the* inc *reserves and* dec *reserves balance over time?*

A. Yes. Over time, BPA assumes that the energy consumed during the deployment of balancing reserves nets to zero. Wellschlager *et al.*, BP-14-E-BPA-26, section 4.

Q. *If, over time, the expected average control error is zero MW, why are response losses not zero on average?*

A. Response losses are not zero because the gains and losses in efficiency incurred during reserve deployment are not symmetrical. That is, the change in efficiency for a 1 MW increase in generation is not equal to the change in efficiency for a 1 MW decrease in generation. The degree of this asymmetry is driven by the particular characteristics of a given unit family, the unit's generation level, the unit commitment, and corresponding efficiency prior to reserve deployment.

Q. *How are response losses priced?*

A. Response losses are priced at the monthly HLH price from the market price forecast, as calculated in the Power Risk and Market Price Study, BP-14-E-BPA-04, section 2.4.

Q. *Why is the HLH price used for all periods?*

A. The HLH price is used because response impacts, losses and gains in energy, are taken out of or put into the HLH period.

Q. *Are response gains also priced at the HLH price?*

A. Yes.

Q. *Are response losses accounted for in any other rate studies?*

A. Yes. All response losses calculated by the GARD Model are sent to RevSim, so the impact to secondary net revenue is correctly reflected. Power Risk and Market Price Study, BP-14-E-BPA-04, section 2.5.

Q. Based on the amount of balancing reserve capacity described in the Balancing Reserve Capacity Quantity Forecast, what is the proposed forecast cost for response losses?

A. The total annual average response loss for the FY 2014–2015 period is 42,397 MWh, worth $1,372,188. Documentation Table 3.11, lines 12-13.

Q. Please describe the subcategory of deployment cost known as deployment cycling losses.

A. Deployment cycling losses are realized during the course of reserve deployment when committed units responding to a balancing need cannot continue deploying *inc* or *dec* reserves while staying within unit-specific operating constraints, and/or additional units are needed to continually maintain the Operating Reserve obligation. When committed units have reached their limits, additional units are brought online in the event of continued *inc* deployment or taken offline in the event of continued *dec* deployment.

Q. How are deployment cycling losses different from cycling losses?

A. Cycling losses are associated with the additional unit cycles required to meet both a project's generation request and stand-ready reserve requirement. Deployment cycling losses are experienced during actual reserve usage as additional units are brought online or taken offline.

Q. Is there overlap between response losses and deployment cycling losses?

A. No. Response losses apply to only those units that are already committed and online when reserves are deployed. Deployment cycling losses are associated with non-spinning units being brought online to respond to the *inc* signal or units being taken offline to respond to the *dec* signal.

Q. How are deployment cycling losses modeled?

A. The GARD Model determines how many units from each unit family are cycled by re-optimizing the unit commitment and dispatch. As generating units are cycled on or off, water is lost to synchronization and/or ramping. Study section 3.4.4.2.

Q. How does the GARD Model use the inputs described above to calculate the deployment cycling losses?

A. The deployment cycling loss calculations are controller project-specific and are functions of the individual unit efficiency curves as well as the level of generation required from the individual units. For each unit on cycle, synchronization and ramping losses are calculated. For each unit off cycle, only ramp-down losses are calculated. Water lost during synchronization to grid frequency equals 10 percent of full gate flow for three minutes. Losses associated with ramping are determined by calculating the excess water used in ramping a unit from minimum loading up to its requested generation. The amount of this loss varies and is directly related to the characteristics of the unit being ramped. The GARD Model fully ramps units to their requested generation level over seven minutes. As with cycling losses for stand-ready cost, the calculation of cycling losses does not attempt to account for any additional maintenance costs that may be realized due to frequent cycling of the units.

Q. How are deployment cycling losses priced?

A. All deployment cycling losses are priced at the monthly HLH price from the market price forecast, as calculated in the Power Risk and Market Price Study, BP-14-E-BPA-04, section 2.4.

Q. Why is the HLH price used for all periods?

A. The HLH price is used because energy lost due to cycling is taken out of the HLH period.

Q. Are deployment cycling losses accounted for in any other rate studies?

A. Yes. All deployment cycling losses calculated by the GARD Model are sent to RevSim, so the impact to secondary net revenue is correctly reflected. Power Risk and Market Price Study, BP-14-E-BPA-04, section 2.5.

Q. Based on the amount of reserves described in the Balancing Reserve Capacity Quantity Forecast, what is the proposed forecast cost for deployment cycling losses?

A. The total annual average deployment cycling loss for the FY 2014–2015 period is 4,797 MWh, worth $157,629. Documentation Table 3.11, lines 14-15.

Q. Please describe the subcategory of deployment cost known as deployment spill losses.

A. Deployment spill arises if GCL receives a *dec* reserve deployment request requiring generation changes jeopardizing its dynamic tailwater limitations. GCL's ability to drop generation is limited because of tailwater bank stability concerns, which limits the rate of change in project outflow. More specifically, the limit is on how quickly the outflow may decrease. If GCL's tailwater elevation decreases too quickly, the downstream river banks begin to collapse.

Q. How are deployment spill losses calculated by the GARD Model?

A. Deployment spill losses are calculated using a functional relationship between GCL's generation and its minimum generation. This is a dynamic relationship serving to maintain downstream bank stability. Should violation of tailwater constraints become a risk, GCL will have to spill water during the course of the *dec* deployment to maintain acceptable rates of change in tailwater elevation as the generation is dropping.

Q. How does the GARD Model account for unique characteristics of the GCL project?

A. As stated above, spill may occur if the generation drop exceeds the drop rate allowed by the project. The drop rate constraint is a particular feature of GCL. If the *dec* request exceeds the limit, additional water must be spilled to slow the rate of change in the tailwater elevation. The tailwater constraint is represented in the GARD Model as a function of GCL LLH generation.

Q. How are deployment spill losses priced?

A. All deployment spill losses are priced at the LLH market price from the market price forecast, as calculated in the Power Risk and Market Price Study, BP-14-E-BPA-04, section 2.4. This is because the energy cannot be shaped into HLH. In these instances, the system must continue to move the water on LLH even if generation is decreasing during a *dec* reserve deployment. Because the water must continue to move, the opportunity to shape into HLH does not exist, and only LLH generation is forgone.

Q. Is deployment spill accounted for in any other rate studies?

A. Yes. All deployment spill calculated by the GARD Model is sent to RevSim, so the impact to secondary net revenue is correctly reflected. Power Risk and Market Price Study, BP-14-E-BPA-04, section 2.5.

Q. Based on the amount of reserves described in the Balancing Reserve Capacity Quantity Forecast, what is the proposed forecast cost for deployment spill?

A. For FY 2014–2015, the total annual average deployment spill loss incurred deploying *decs* is 79 MWh, worth $2,223. Documentation Table 3.11, line 16.

Section 5.2: Variable Cost of Reserves

Q. What is the total forecast variable cost of providing reserves?

A. The total variable cost of providing reserves for the FY 2014–2015 period is $14,202,873 annually, based on the average amount of reserves described in the Study, section 2, and the spinning portion of the Operating Reserve described in the Study, section 4. Documentation Table 3.12, line 6.

Q. What is the purpose of apportioning cost to the different types of reserves?

A. The total variable cost of providing balancing reserve capacity is apportioned according to type of reserve to ensure cost recovery from the appropriate rates. Specifically, the

total cost is apportioned into the cost of load Regulating Reserve, load following and energy imbalance for load, VERBS, DERBS, and Spinning Operating Reserve. The costs we allocate to load Regulating Reserve, VERBS, DERBS, and Operating Reserve are recovered through the Ancillary and Control Area Services rates. The remaining costs, those for load following and energy imbalance reserve, are recovered through the PF rate. Documentation Table 3.12.

Q. How is the balancing reserve capacity cost apportioned to reserve type?

A. The balancing reserve capacity costs are apportioned to reserve type in direct proportion to a given reserve type's composition of spinning (*inc*), non-spinning (*inc*), and *dec* reserves. Study section 3.4.5. From the GARD Model, the total cost for supplying all reserve products is totaled into the general categories of spinning, non-spinning and *dec*. The proportion of spinning reserve required for each of the reserve products is calculated relative to the total reserve. This proportion is used to allocate the total cost to the reserve product. Similarly, the proportion of non-spinning reserve required by each of the reserve products is calculated relative to the total reserve. This proportion is used to allocate the total cost of non-spinning reserve to the specific reserve products. *Dec* reserves are all spinning, and the process for allocating the *dec* cost is performed in the same fashion.

Section 5.3: Alternative Variable Cost Forecast

Q. Were any other variable cost forecasts generated?

A. Yes. In addition to calculating the cost of reserves based on the 99.5 percent level of service in the Balancing Reserve Capacity Quantity Forecast with self-supply of generation imbalance and a reduction for 30/30 committed scheduling, another study was run. The other study calculated the Variable Cost without self-supply of generation

imbalance but does include the reduction for 30/30 committed scheduling. Both studies capped the amount of balancing reserve capacity provided from the FCRPS at 900 MW *inc* and 1100 MW *dec*. *See* Documentation Table 3.16.

Section 5.4: VERBS Credit Adjustment to Generation Inputs Revenue Forecast

Q. Are there times of the year where BPA is not able to provide the full balancing reserve capacity?

A. Yes. As explained in Kerns *et al.*, BP-14-E-BPA-23, there are times at which the flexibility of the FCRPS is limited due to the need to meet non-power, high-priority, hydraulic objectives. During these times the full balancing reserve capacity may not be available. BPA staff (Kerns *et al.*, BP-14-E-BPA-23, section 4) project that the FCRPS will be able to supply the rate case forecast amount of balancing reserve capacity approximately 98 percent of the time on an average annual basis.

Q. How are you proposing to account for these reductions in the balancing reserve capacity of the FCRPS?

A. As described in Fisher *et al.*, BP-14-E-BPA-21, section 6, we are proposing to provide VERBS ratepayers a rate credit so that VERBS customers do not bear the costs of Federal balancing reserve capacity that BPA is unable to provide because of hydro system limitations.

Q. How is the value of the VERBS rate credit determined?

A. We propose to make an adjustment to the generation inputs revenue forecast based on an expected value of the percent of balancing reserve capacity availability from the FCRPS and the VERBS Credit rate. As noted above, BPA estimates balancing reserve capacity availability will be 98 percent on an annual average basis over the FY 2014–2015 rate period. For the estimated 2 percent of balancing reserve capacity unavailability from the

FCRPS, we applied the VERBS Credit *inc* rate of $7.30 per kW per month of balancing reserve capacity plus the VERBS Credit *dec* rate of $0.60 per kW per month of balancing reserve capacity to produce a result of approximately $1.0 million on an annual average basis for the rate period. Documentation Table 3.21.

Q. Was this rate credit reflected in the generation input revenue forecast included in the Initial Proposal?

A. No. The generation inputs revenue forecast was completed for the Initial Proposal before the calculation for the impact of the VERBS Credit was included in the revenue forecast. The VERBS Credit will be included in the generation inputs revenue forecast when the final rate case numbers are run for the Final Study, which will be based on the best information available at that time.

Section 6: Types of Balancing Reserve Capacity Purchases

Q. Have you identified the potential for BPA to purchase balancing reserve capacity during the FY 2014–2015 rate period to support its Balancing Services?

A. Yes. We have identified the potential for four types of balancing reserve capacity purchases during the FY 2014–2015 rate period. Study section 3.5.

Q. What are the four types of balancing reserve capacity acquisitions that you have identified?

A. As described in greater detail in the Study, section 3.5, we have identified four types of balancing reserve capacity purchases. Type 1 purchases are planned purchases needed to make up the shortfall between the planned Federal balancing reserve capability (900 MW) and the rate case planned balancing needs of the Base Service after adjusting for any self-supply of generation imbalance (CSGI or other). Type 2 purchases are needed when BPA is operationally unable to provide the planned Federal balancing

reserve capacity (up to 900 MW). Type 3 purchases are needed to provide the Full Service balancing service option. Type 4 purchases are needed to support an unplanned increase in balancing services needed by the balancing authority area.

Q. Why do you separate balancing reserve capacity acquisitions into four types?

A. Each type is made for a specific purpose and under a specific circumstance. These differences in purpose and circumstance justify a distinction and warrant different methods for allocating their costs.

Q. How do you propose to allocate costs for Type 1 purchases?

A. We propose to allocate Type 1 purchase costs based on a methodology that determines which balancing service use causes the need for BPA to acquire balancing reserve capacity. The methodology is applied equally and consistently across three identified use categories (Categories): load, dispatchable energy resources, and variable energy resources (which includes solar resources). Study section 3.5.3.

Q. Why do you apply the principle of cost causation to purchases to these three Categories?

A. We propose these three Categories because of the significant difference in operational characteristics between load, dispatchable energy resources, and variable energy resources, as demonstrated through the Balancing Reserves Capacity Quantity Forecast Study. *Id.* section 2. That said, we acknowledge that there can be a significant difference in operational characteristics among customers within each of these three use Categories. Given that the forecast need to acquire balancing reserve capacity is new to BP-14, we propose an approach that identifies cost causation at the first level of operational characteristic differentiation.

Q. What do you mean by "operational characteristic differentiation"?

A. We use operational characteristics to mean the amount of reserve requirement needed relative to a Category's size – a percentage that measures the amount of reserve

requirement a Category, a resource type, or a particular resources needs in proportion to its nameplate (or equivalent for load). As identified in the Balancing Reserves Capacity Quantity Forecast Study, this percentage varies significantly among the three Categories, but also within the Categories. For example, on aggregate, dispatchable energy resources have a balancing reserve capacity requirement equal to 1.1 percent of their nameplate while variable energy resources have a balancing reserve capacity requirement equal to 11.2 percent of their nameplate (given certain scheduling elections and self-supply assumptions). Documentation Table 3.19. However, within the variable energy resource Category, there are also significant differences between solar resources (2.0 percent) and wind resources participating in 30/30 committed scheduling (10.5 percent). Documentation Table 3.20.

Q. Explain how a lower level of operational characteristic differentiation could be applied to the methodology (i.e., defining more than three use categories).

A. While not part of our initial proposal, this cost causation methodology could be applied to more use categories (at a more granular level). A more granular approach could be used to identify and distinguish cost causation within the three use categories we propose (load, dispatchable energy resources, and variable energy resources). Subcategories of variable energy resources could be created to distinguish the causation of purchases between different variable energy resources and their elections. For example, a subcategory could be used to distinguish among: (1) a solar resource and a wind resource; (2) a resource with a self-supply commitment that reduces BPA's need to purchase balancing reserve capacity; and (3) a resource that uses scheduling practices that contribute more to (or cause) BPA's need to purchase balancing reserve capacity. While these types of subcategories are not part of our Initial Proposal, the expanded category examples listed above have been included to illustrate how the cost causation

methodology would work with more than our initially proposed three use Categories. Documentation Table 3.20.

Q. How do you propose to allocate costs for Type 2 purchases?

A. We propose to allocate Type 2 purchase costs to all resources (with the exception of AGC-controlled hydro resources) based on their proportional monthly balancing reserve capacity need. AGC-controlled hydro resources are excluded since they are resources providing balancing reserve capacity. Study section 3.5.4.

Q. Why do you use a different approach for Type 2 purchases than is proposed for Type 1 purchases?

A. Type 1 purchases are caused by demand, which means they can be linked to a particular customer or customer category (based on cost causation). In addition, the amount of Type 1 purchases needed during the rate period will be forecast in the final studies. This allows us to use a straightforward method for determining cost causation with regard to Type 1 purchases and set rates accordingly. Type 2 purchases are different. Type 2 purchases are caused by an unplanned reduction in BPA's supply of balancing reserve capacity. Furthermore, the amount of Type 2 purchases will not be known until after rates are set. Essentially, Type 2 purchases are the result of the actual capability of the FCRPS being less than the forecast used in the rate case (*i.e.*, rate case forecast error). This is just one of many actuals that will be different from the forecasts and assumptions used in the rate setting process. For these reasons, we propose that these purchase costs be allocated more broadly than the other three types of purchases and that BPA not reassess purchase responsibility used for Type 1 purchases in the middle of a rate period. Instead, we propose to allocate the costs of Type 2 purchases to all resources (with the exception of AGC-controlled hydro) and reevaluate the forecast planned FCRPS capability in the next rate setting process.

Q. Why does load not share in the cost of Type 2 purchases?

A. All balancing capacity uses are impacted if the FCRPS cannot provide the rate case forecast amount of balancing reserve capacity. While variable energy resources and dispatchable energy resources are exposed to the cost of Type 2 purchases, load is impacted by funding the VERBS Credit. Load is financially impacted in two different ways depending on the customer type – Slice or Non-Slice. Slice customers are affected through upward pressure (increased costs to Slice customers) on the Slice True-Up, since the Slice True-Up accounts for reduced revenue provided from the VERBS. Non-Slice customers are impacted through upward pressure on (or the use of) cash reserves, since the portion of BPA's revenue requirement that was forecast to be recovered from VERBS customers is no longer collected. In order to mitigate this particular source of Power Services' revenue risk, we propose to downward adjust the full generation inputs credit that is included in the Composite Cost Pool revenue credit by the expected value of the VERBS Credit. *See* Study section 3.6. If load were allocated Type 2 purchase costs, it would result in load paying twice for the same event when the FCRPS is unable to provide the rate case planned amount of balancing reserve capacity: once through the VERBS Rate Credit from Power financial reserves, and again through the allocation of Type 2 purchase costs.

Q. How do you propose to allocate Type 3 purchases?

A. Type 3 purchases are made for the Full Service option available to variable energy resources. Consistent with cost causation, we propose these costs be allocated to variable energy resources that take Full Service.

Q. How do you propose to allocate Type 4 purchases?

A. Type 4 purchases are made if BPA has an unplanned increase in balancing services. This can occur for several reasons: (1) a resource has elected to self-supply but is unable to

continue self-supplying one or more components of VERBS or DERBS; (2) a resource has a projected interconnection date after FY 2015 but interconnects during the FY 2014–2015 rate period; (3) a variable energy resource commits to a specific scheduling practice but does not maintain a scheduling performance consistent with or better than that committed scheduling practice; (4) a variable energy resource elected Base Service with a committed scheduling option but chooses to take Base Service with a longer scheduling period; or (5) a non-Federal thermal resource operating in another balancing authority area chooses to dynamically transfer into the BPA balancing authority area during the FY 2014–2015 rate period. We propose that the costs that result from the need for Type 4 purchases be directly assigned to the customer that received the expanded balancing service. This allocation would apply until BPA can reflect the increased balancing need in the rate setting process.

Q. Is your cost allocation proposal consistent with the principle of cost causation?

A. Yes. Our proposal is that demand (or need) caused purchases be allocated to the balancing capacity user that caused that cost. Type 1, Type 3, and Type 4 purchases are demand-caused purchases and can be linked to a particular customer or customer Category. As stated above, Type 2 purchases are different because they are supply-caused and are a result of forecast error. For this reason, we propose to allocate the costs to the users that are affected by the supply.

Q. Why is your proposed Type 1 purchases cost allocation methodology fair and equitable?

A. The methodology is fair and equitable because it is applied consistently across all users of balancing capacity. Simply stated, Type 1 purchase costs are allocated to the Categories with remaining balancing service need after all Categories are first provided equal access to the cost of the planned FCRPS balancing reserve capacity, which is an amount of FCRPS-sourced balancing reserve capacity equal to approximately 3.5 percent of a

Category's nameplate (or nameplate equivalent for load). As described above, the operational characteristics of users of balancing capacity vary considerably, with some users requiring more than 14 percent of their nameplate while others require less than 2 percent. Documentation Tables 3.19 and 3.20. The FCRPS is planned to be able to provide 3.5 percent of the total nameplate (or equivalent for load) of all balancing users. The cost allocation methodology provides equitable treatment among all balancing users by identifying the balancing reserve capacity uses that cause BPA's need to acquire Type 1 purchases and allocating those costs to those uses.

Q. Why is your proposed Type 2 purchases cost allocation methodology fair and equitable?

A. Given the circumstances under which Type 2 purchases occur (an unexpected decrease in the supply of balancing reserves capacity), we believe the proposed allocation is fair and equitable. Under our proposal, costs are allocated proportionally to the users that are affected by Type 2 purchases.

Q. Why is your proposed Type 3 purchases cost allocation methodology fair and equitable?

A. Our proposal for Type 3 purchases is to allocate these costs to the customers that take the service that creates the costs. If no customers take Full Service, BPA will not need to make any Type 3 purchases and will not incur any Type 3 purchase costs. Type 3 purchase costs are a direct result of a customer's use of balancing reserve capacity under the Full Service option and, thus, the customers that take Full Service should be allocated the costs of providing that service.

Q. Why is your proposed Type 4 purchases cost allocation methodology fair and equitable?

A. Our proposal for Type 4 purchases is to directly assign the costs of Type 4 purchases to the customers that cause the need for those purchases. For example, Type 4 purchases can be the result of customers doing something different than what they committed to do.

It would be unfair and inequitable to allocate these purchase costs to customers other than to the customers that were unable to meet their commitments.

Type 4 purchases can also result when a new customer that was not expected requests balancing service in the middle of a rate period. While this particular event is the result of forecast error, it is demand-caused, which justifies a different approach than what was used for Type 2 purchases. BPA has routinely required that unanticipated service be at the marginal cost of that service until BPA can properly account for that service the next time it sets rates. For example, this approach is demonstrated through the BPA's Priority Firm (PF) and New Resources (NR) Unanticipated Load Rate. *See* Power Rate Schedules, BP-14-E-BPA-09.

Q. Does this conclude your testimony?

A. Yes.