ORNL/CON-426

ENERGY DIVISION

ELECTRIC-POWER ANCILLARY SERVICES

ERIC HIRST and BRENDAN KIRBY

February 1996

Sponsored by
Competitive Resource Strategies Program
Office of Energy Efficiency and Renewable Energy
U.S. Department of Energy

OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee 37831
managed by
LOCKHEED MARTIN ENERGY RESEARCH CORPORATION
for the
U.S. DEPARTMENT OF ENERGY
under contract No. DE-AC05-96OR22464

CONTENTS

SUMMARY

Since the Federal Energy Regulatory Commission (FERC) published its March 1995 proposed rule on open-access transmission, ancillary services have been an important and often controversial topic. Ancillary services are those functions performed by the equipment and people that generate, control, transmit, and distribute electricity to support the basic services of generating capacity, energy supply, and power delivery. FERC defined ancillary services as those "necessary to support the transmission of electric power from seller to purchaser given the obligations of control areas and transmitting utilities within those control areas to maintain reliable operations of the interconnected transmission system."

During the past year, several organizations have developed lists of ancillary services. These lists include from 6 to more than 40 services. The definitions of and distinctions among some of these services are often unclear. Load following and energy imbalance are particularly controversial.

Based on our earlier research on ancillary services, as well as that conducted by many others, we developed a revised set of seven services (Table S-1). Our set is similar to the one suggested by FERC (we separate FERC's system protection service into reliability reserve and supplemental operating reserve). We also provide extended discussions of each service, with an explanation of the complexities in shifting from the perspective of relationships among control areas in an interconnection to the perspective of customers and suppliers dealing with a system operator.

We consider load following to include four elements. The two control-area functions include maintenance of interconnection frequency at 60 Hz and maintenance of generation/load balance within the control area. The two customer functions include following the moment-to-moment fluctuations in load and following the longer-term (e.g., hourly) changes in load.

We consider energy imbalance to include three elements. Energy imbalance would include only the discrepancies within a defined deadband measured over a defined time interval (e.g., ±1.5% for 10 minutes). Standby service would be contractually arranged beforehand between the customer and a supplier (not necessarily the local control area). Unauthorized use would result in penalty charges imposed by the local control area if (1) the customer's load fell outside the deadband and (2) the customer had not arranged for standby service or the standby service failed to perform. Because unauthorized use is not a service, its charge would not be based on costs. Rather, its price would be designed to encourage customers to obtain standby service and to discourage them from leaning on the local control area.

Table S-1. Proposed set of electric generation and transmission ancillary services

Service	Unbundle to: Suppliers	Unbundle to: Customers	Controlled by system operator?	Can resource be provided competi-tively?	Must be inside local control area?
Scheduling and dispatch (1)	Y	Y	Y	N	Y
Generating reserves					
Load following (2)	Y	Y	Y[a]	Y[a]	N[a]
Reliability (3)	Y	Y	Y[a]	Y[a]	N[a]
Supplemental operating (4)	Y	Y	Y	Y	N
Energy imbalance (5)	Y	Y	Y[a]	Y	N
Real-power loss replacement (6)	Y	Y	Y	Y	N
Voltage control (7)					
Generation	Y	N	Y	?[b]	Y
Transmission	N	N	Y	N	Y

[a]If dynamic scheduling is feasible, these generating-reserve services can be provided and controlled by another supplier.

[b]Whether the market for generator voltage control is competitive depends on the specifics of each situation.

Additional work is needed to reach agreement among control-area utilities, other utilities (primarily municipalities and rural cooperatives), power marketers and brokers, independent power producers, electricity consumers, and regulators on various issues related to ancillary services. These issues include determining the definition and amount of each service required to maintain system reliability, whether or not that service can be provided competitively or can only be provided under regulation, whether the supplier can be outside the customer's control area, and whether the amount and timing of each customer's use of the service can be cost-effectively metered and billed. Suggestions for additional research include:

- Explain clearly the current reliability requirements and their rationales. How do utilities today derive their requirements for load-following, reliability, and supplemental reserves? Do the required amounts of these generating reserves vary from day to day and from season to season? What determines the allowable range of voltages throughout the transmission system and, therefore, the reactive-power requirements?

- What are the tradeoffs between cost and reliability associated with possible changes in today's rules and practices?

- Examine how these requirements might change in the future. Develop mechanisms to establish definitions, provision, and pricing of ancillary services for different industry structures (e.g., one with an hourly spot market and an independent system operator).

LIST OF ACRONYMS

ACE	Area-control error
AGC	Automatic generation control
EEI	Edison Electric Institute
EPRI	Electric Power Research Institute
FERC	Federal Energy Regulatory Commission
ISO	Independent system operator
MAIN	Mid-American Interconnected Network
NERC	North American Electric Reliability Council
NOPR	Notice of Proposed Rulemaking
ORNL	Oak Ridge National Laboratory

INTRODUCTION

BACKGROUND

Until a few years ago, only a few electrical engineers and system operators in electric utilities knew what ancillary services were. Because of increasing competition in, and changing regulation of, the U.S. electricity industry, ancillary services are now an important issue. The topic is important for the vertically integrated utilities that generally supply these services; for the utilities (such as municipalities and cooperatives), large industrial firms, and power marketers that must buy these services to effect power purchases; for independent power producers that might be able to provide some of these services; for state and federal regulators that must decide whether and how to unbundle and price these services; and especially for electricity consumers, whose continued reliability of service depends on ancillary services.

Ancillary services are those functions performed by the electrical generating, transmission, system-control, and distribution-system equipment and people that support the basic services of generating capacity, energy supply, and power delivery (Kirby, Hirst, and VanCoevering 1995). The Federal Energy Regulatory Commission (FERC 1995), in its Notice of Proposed Rulemaking (NOPR), defined ancillary services as "those services necessary to support the transmission of electric power from seller to purchaser given the obligations of control areas and transmitting utilities within those control areas to maintain reliable operations of the interconnected transmission system."

Unbundling generation and transmission services is likely to promote economic efficiency. Such unbundling will allow customers to choose those services that they need and will require customers to pay for the services that they use. Similarly, unbundling will encourage competition among suppliers and thereby lower the costs of these services. Overall, unbundling should increase customer choice, promote competition, and provide more accurate price signals to both suppliers and consumers of these services. The overall cost of ancillary services is roughly 6 to 20% of total generation and transmission costs, equivalent to almost $14 billion a year (Kirby and Hirst 1996).

Unbundling for its own sake has no value (Hogan 1995a). Unbundling should proceed only to the point at which the value of further unbundling no longer exceeds the cost of unbundling (Fig. 1). At some point, increasing transaction costs and the loss of economies of scope will outweigh the benefits of further unbundling. Also, some functions may be purchased separately from suppliers (e.g., dynamic reactive support and black-start capability) but remain bundled to customers.

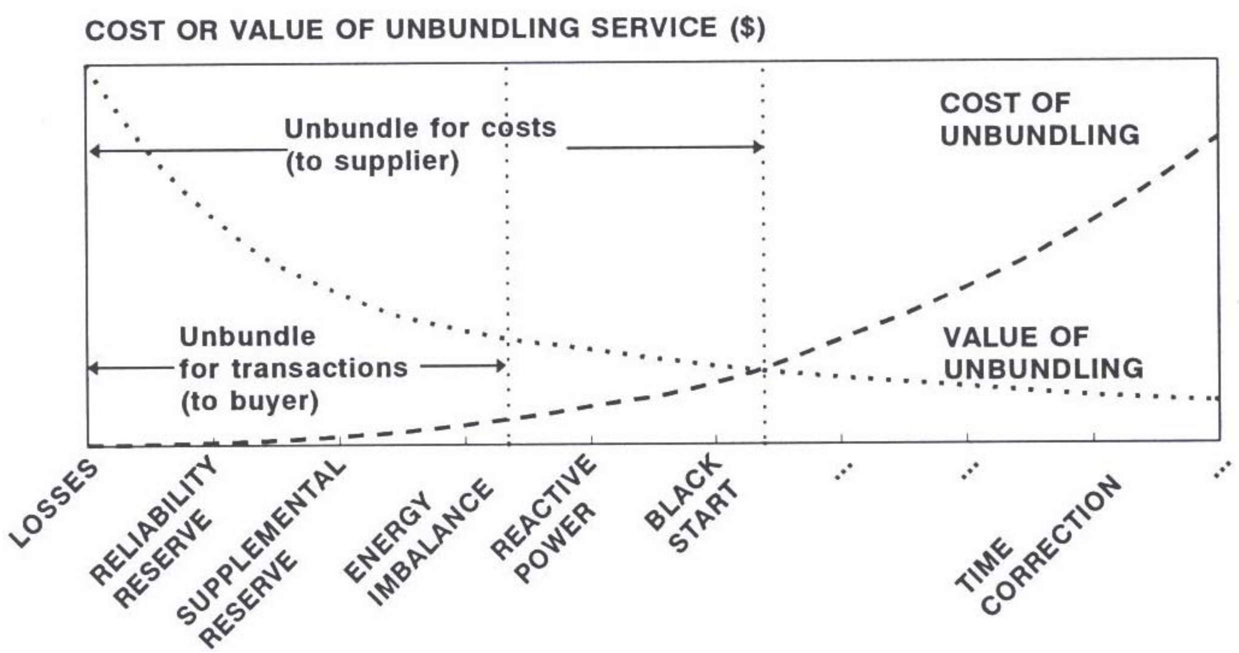

Fig. 1. The degree of unbundling generation and transmission services should reflect the costs and benefits of doing so.

This report, the second in a series of studies conducted by Oak Ridge National Laboratory (ORNL) for the U.S. Department of Energy, covers four topics:

- Recent events, in particular the comments filed with FERC on its NOPR;
- Alternative sets of ancillary services proposed by utility, regulatory, and research organizations;
- An improved set of ancillary services (fewer in number than the 19 included in our March 1995 report) with sharper definitions for each service;
- Identification of critical issues that remain to be resolved, including the relationships between ancillary services and reliability, competitive markets, and the structure of the electricity industry.

The remainder of this chapter reviews recent events. Chapter 2 presents our current thinking on a reasonable set of services and their definitions. Chapter 3 discusses unresolved issues and suggests several lines of research to pursue. The Appendix compares various sets of ancillary services and their definitions.

RECENT EVENTS

In March 1995, FERC issued a major proposed rule on open-access nondiscriminatory transmission service. FERC's notice identified six ancillary services: reactive power and voltage control, loss compensation, scheduling and dispatch, load following, system protection, and energy imbalance. Because of FERC's jurisdiction over wholesale electricity markets, its notice (often called the Mega-NOPR because of its importance and length) and discussion of ancillary services received considerable attention during the following months.

All the parties that responded to FERC's notice agreed that maintaining reliability in the U.S. electric system is important [Edison Electric Institute (EEI) 1995]. The parties generally agreed that some unbundling of generation and transmission services is desirable and that markets should be used to define and price these services where markets are competitive.

Many of the investor-owned utilities, which generally are vertically integrated, emphasized reliability and the importance of the control-area operator. They want transmission customers to be required to purchase all the ancillary services needed to maintain reliability, with the control-area operator having the final say in what is required. For example, EEI stated:

> The fundamental framework for 'control area' operations, or some further evolution of control area operations, must be maintained. Control areas ensure the stability and reliability of the transmission network through a regimen of scheduling, confirming and implementing all interchange transactions, and through monitoring frequency and controlling [generating] resources to maintain frequency. Certain operating procedures may need to be amended, deleted, or added to accommodate a broader wholesale power market with open access, as indicated by NERC [the North American Electric Reliability Council]. However, these changes must be consistent with the objectives of control area operations and NERC operating guidelines and reliability criteria. They should not result in any adverse economic consequences for utilities' native load customers, particularly retail customers.

To ensure reliability, some utilities wanted the control-area operator to have the right to specify the transaction-specific requirements and to verify that the service was provided in an adequate manner if provided by another entity.

Investor-owned utilities often commented on the apparent asymmetry between their obligation to offer services and the lack of obligation to take these services on the part of transmission customers. They also noted that most of the ancillary services are provided by generation, not transmission, and therefore, some of these services could be obtained in competitive markets. Some utilities indicated that the FERC proposals would not fully cover the costs of providing ancillary services; as examples, the 3% loss factor was deemed too low, and the 1.5% deadband for energy imbalance was deemed too high.

Transmission-dependent utilities, power marketers, and others that purchase ancillary services worried that transmission-owning utilities would use reliability as a way to charge too much for ancillary services or to charge twice for the same service. For example, separate charges for scheduling and dispatch or for reactive power may be unfair if the costs of these services are already included in the basic transmission tariffs. These parties worry that utilities will refunctionalize generation costs (which will be competitive) to transmission (which will remain regulated) and that utilities will try to load as much of the generation cost (e.g., turbines as well as generators, and as much of the exciter as possible) in the charge for VAR support. In essence, these entities want to be sure that the sum of the costs for unbundled services is no more than the cost of the bundled service.

These parties complained about the asymmetry between the treatment that control areas offer each other (e.g., inadvertent interchange is compensated with return of energy in-kind and involves no financial payments) and that which they offer transmission customers (e.g., energy imbalance outside the deadband requires cash payments that may include a penalty provision). Some of these entities complained that FERC's assumed loss factor is too high and that the deadband for energy imbalance is too low. They are concerned that local control areas will hold them to stricter standards than the control areas themselves meet in complying with reliability-council rules; such a double standard, they believe, will penalize their transactions. In general, transmission-dependent entities were concerned that transmission owners would find ways to overcharge for ancillary services and thereby limit access to transmission networks.

Independent power producers want the opportunity to sell ancillary services. They worry that control-area operators will erect barriers to competitive markets for generation-related services.

This brief review shows that, although most commenters agreed on basic principles concerning reliability and the use of markets to provide and price ancillary services, they did not agree on the specifics. Indeed, as discussed in the Appendix, the number and definitions of ancillary services were often in dispute. And the relationship between the control area's responsibility to maintain reliability and the customer's requirement to purchase ancillary services was hotly debated. Similarly, no agreement was reached on which services could be provided by competitive markets and which services must remain regulated.

At about the same time that FERC issued its NOPR, ORNL and the Electric Power Research Institute (EPRI) issued reports on transmission and ancillary services (Kirby, Hirst, and VanCoevering 1995; Wakefield et al. 1995). The ORNL study defined 19 generation- and transmission-related ancillary services and compared these definitions with those developed by five utilities. ORNL analyzed each service and indicated whether it was required or optional, had to be under the control of the system operator, and could be metered or controlled.

The EPRI study, although it did not deal explicitly with ancillary services, provided valuable background on the functions and operations of electric transmission networks. The study identified the capital and operating costs of transmission service and discussed different cost concepts (e.g., embedded costs, short-run marginal costs, and long-run marginal costs). The study also developed a framework to define transmission services (amount, firmness, and duration), to identify transmission-service costs (administrative, provision of facilities, and delivery), and to calculate these costs.

In June, the Michigan Public Service Commission (1995) issued a decision in a retail-wheeling case that began in 1992. This decision defined various ancillary services and set rates for each service for Detroit Edison and Consumers Power. This decision marked the first official definition of ancillary services and accompanying tariffs.

In July, EPRI sponsored a workshop on ancillary services (Zadeh and Meyer 1995). One full day of the workshop was devoted to the six services that FERC identified in its NOPR.

As part of an investigation into transmission access and pricing, the Texas Public Utility Commission (1995) sponsored several workshops on ancillary services during the summer of 1995. The Staff's report to the Commission recommended that the Commission develop a clear and comprehensive list of services (*IRP Report* 1995). The Commission could let markets determine the price for those services that can be provided competitively; for other services, the Commission could set prices on the basis of a utility's embedded costs. During a transition period, the Commission could allow a range of pricing, with a floor of short-run marginal costs and a ceiling of embedded costs.

In October, FERC conducted a Technical Conference on Ancillary Services with about 25 presentations representing all segments of the industry (EEI 1995). Oral comments made at the conference echoed the written comments submitted earlier to FERC.

Thus, the past year has seen considerable activity and progress on ancillary services. While the details and definitions of ancillary services remain elusive, the key concepts and purposes of these services are now widely understood and appreciated. As discussed in Chapter 3, much work remains to define a set of services that is mutually exclusive and exhaustive, to identify which services can be provided competitively, to identify which services must be under the direct control of the system operator, to denote which services can be obtained from outside the local control area, to characterize the relationship between reliability and these services, and to analyze the effects of electric-industry structure on the existence and pricing of these services.

The six sets of services discussed in the Appendix illustrate well the progress made during the past year as well as the work that remains to be done. Although similar in many ways, the sets differ in number, nomenclature, definitions, and other characteristics. Developing a workable set of ancillary services is difficult, we believe, because of the complexities in shifting from the perspective of relationships among control areas in an interconnection to the perspective of customers and suppliers dealing with a system operator. Today, control areas are operated by vertically integrated utilities with a long tradition of cooperation with the other control areas within the interconnection. In the future, system operators may own no generating resources, may have no retail customers, and may have to handle far more transactions than take place today.

Alvarado (1996) classifies ancillary services along three dimensions: real vs reactive power, time, and insurance. Real-power services include frequency regulation, ramping schedules, energy imbalance, loss compensation, and unit commitment. Reactive-power services include voltage regulation, capacitor switching, and generator scheduling. His time dimensions are instantaneous (up to a few seconds), fast (up to several minutes), and slow (up to several hours). The third dimension, insurance, recognizes that resources must be available to protect against contingencies.

Utility control-area operators use some generating units to follow moment-to-moment changes in load. More precisely, these units respond to changes in interconnection frequency and differences between actual and scheduled tie-line flows. This "service" is typically called regulation. Its customer-specific analog is load-following. Although utilities have considerable data on moment-to-moment fluctuations in loads, some of which include individual customers, these data have traditionally been used for operations and not kept for analysis. Thus, there is no foundation today for charging different customers different amounts for this service.

In a similar fashion, energy imbalance is the customer-specific analog to control-area inadvertent interchange. The NERC (1992) control-area performance criteria require control areas to maintain their area-control error (ACE) within tight limits. ACE, measured in MW, is the instantaneous difference between actual and scheduled interchange plus frequency bias. The first term reflects the control area's load balance with the rest of the interconnection and the second term reflects the interconnection's frequency deviation from the 60-Hz reference.

The first of the two criteria (A1) requires that, on an instantaneous power basis, the control area be in balance with the rest of the interconnection at least once every 10 minutes. The second criterion (A2) requires that the control area's energy imbalance (average ACE) be within a certain limit called L_d (roughly 0.2 to 0.5% of peak demand) every 10 minutes. Accumulated A2 discrepancies are called inadvertent interchange. Utilities generally compensate each other for such interchange through a return of energy in kind, with no cash payments. Energy imbalance requires customers to maintain a balance between their scheduled and actual deliveries. Within a defined deadband over a prescribed time (±1.5% and 1 hour in FERC's proposal), the energy can be returned in kind. Outside the deadband, customers must pay the local utility for imbalances; this payment may include a penalty provision.

Other problems relate to the definition of a utility. Although FERC's proposed rule requires transmission providers to offer ancillary services, most of these services are today managed by control-area operators. Most of the services are provided by generation, not transmission. In today's vertically integrated utilities, the control-area operator, the owner of generating equipment, and the transmission owner are usually the same, but that may not be true in the future.

SUGGESTED ANCILLARY SERVICES

Building on the works discussed in Chapter 1 and the Appendix, we developed a revised set of seven ancillary services. In developing this set, we sought to identify those services that are essential to maintain electric-system reliability, are required to effect a transaction, or are a consequence of a transaction. We excluded various services that are optional, long-term in nature, too cheap to warrant the costs of metering and billing, naturally bundled with other services, or very location- and customer-specific. In reviewing our definitions and explanations, recognize that different utilities define and treat these services differently. For example, we adopted an expansive view of load following that includes intrahour ramping; many utilities would consider ramping a separate service.

The existence, definition, and pricing of ancillary services is a function of industry structure. We assumed the continued existence of control areas and the NERC control-area concepts and requirements. Based on several recent proposals, we anticipate creation of independent system operators (ISOs) that will replace today's utility operation of control areas. These entities will not be controlled by generation-owning organizations. Rather, they will be *independent* organizations with the primary goal of operating the electrical system in real time to maintain reliability. The ISO will control the short-term operation of enough generating capacity to maintain system reliability and stability. The ISO may (depending on the proposed industry structure) dispatch all of the generating units within the control area to minimize the cost of electricity production. Or the ISO may create a wholesale spot market for power, although in some proposals such a power exchange is a separate entity.

The ISO might unbundle some services that it purchases (i.e., when facing suppliers) and unbundle services differently that it sells (i.e., when facing customers); see Fig. 1 in Chapter 1. In addition, services should be unbundled only if the ISO can identify and measure the amount of service provided by suppliers and/or consumed by customers. (We consider data analysis, such as load-flow modeling, a form of measurement.) As computing and communication technologies improve, it may be possible to unbundle additional services. Finally, the incremental metering, accounting, billing, and auditing costs of unbundling must be less than the benefits of unbundling.

Our set of services includes scheduling and dispatch, which is a control-area-operator function requiring few resources (computers, meters, communications equipment, and people). The set also includes several generating services, such as load following, reliability reserve, and supplemental reserve, as well as loss replacement and energy imbalance. For various reasons, discussed below, there is considerable confusion about the definitions and purposes of these services. Figure 2 shows the generation-related ancillary services, their relative magnitudes, and the time scales in which they operate. Energy imbalance is shown with a question mark because it does not fit comfortably into these dimensions of magnitude and time. Voltage control, real-loss replacement, and load following are continuous services. Reliability reserves are usually idle but must be ready to respond fully within 10 minutes. Supplemental reserves are also usually idle but must be able to respond fully within 30 minutes. Finally, we include system voltage control, which requires both generating units and transmission-system equipment (Table 1). All of these services are required.

We present a rough estimate of the cost for each service, based on the data, analyses, and estimates we obtained from 10 utilities throughout the United States (Kirby and Hirst 1996). We present these estimates on a ¢/kWh basis, although the utilities charge for these services in different ways (e.g., $/kW-month and $/kVAR-month as well as ¢/kWh). Readers should view these estimates cautiously because the utilities defined the services in different ways, based their estimates on different cost concepts (e.g., marginal vs embedded cost), and often presented results based more on negotiations than on data and analysis. For example, utilities priced generating reserve based on: (1) a capacity charge only, (2) capacity and energy charges based on total consumption, or (3) a capacity charge plus an energy charge imposed only when the customer uses the reserve.

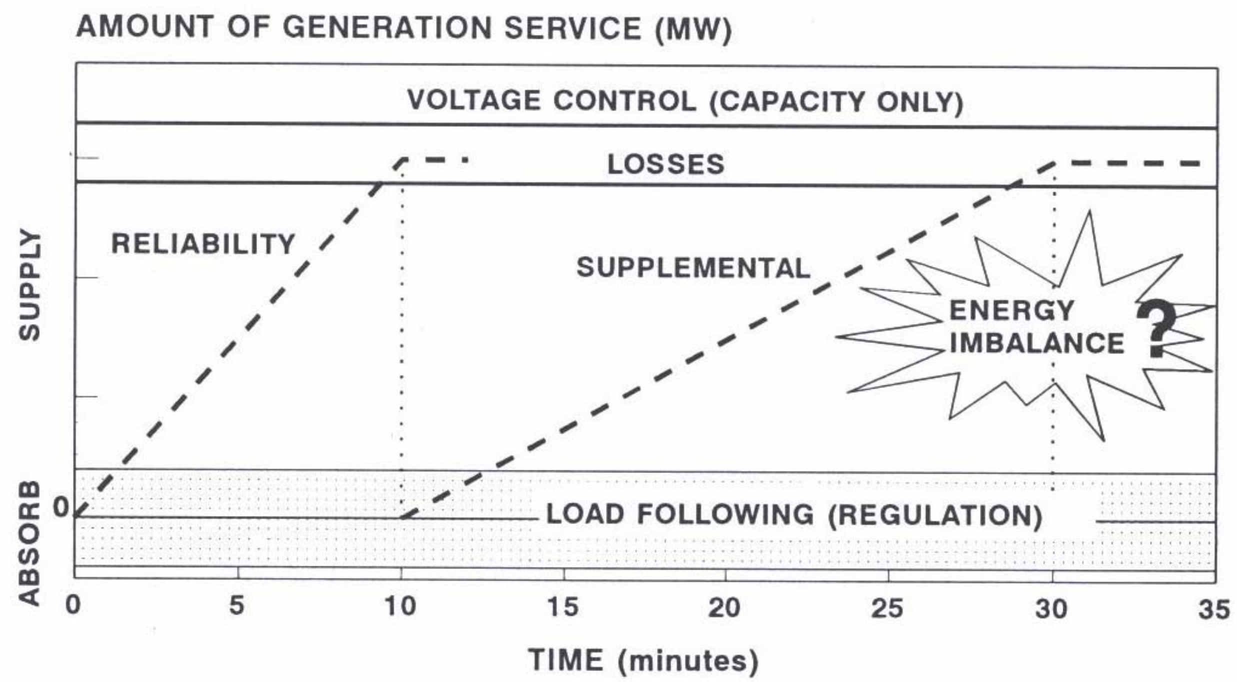

Fig. 2. Schematic showing approximate size and speed of response for different generation ancillary services.

Table 1. Proposed set of electric generation and transmission ancillary services

Service	Unbundle to: Suppliers	Unbundle to: Customers	Controlled by system operator?	Can resource be provided competitively?	Must be inside local control area?[a]
Scheduling and dispatch (1)	Y	Y	Y	N	Y
Generating reserves					
Load following[b] (2)	Y	Y	Y[c]	Y[c]	N[c]
Reliability (3)	Y	Y	Y[c]	Y[c]	N[c]
Supplemental operating (4)	Y	Y	Y	Y	N
Energy imbalance (5)	Y	Y	Y[c]	Y	N
Real-power loss replacement (6)	Y	Y	Y	Y	N
Voltage control (7)					
Generation	Y	N	Y	?[d]	Y
Transmission	N	N	Y	N	Y

[a] Whether these services must be located within the local control area will depend, in part, on the existence and location of transmission constraints.

[b] Load following can be split into fast (roughly less than a second or two) and slow fluctuations. The response to fast fluctuations is distributed among generators based on machine characteristics (e.g., inertia and governor response) and on network impedances. The response to slow fluctuations can be directed to specific generators using AGC.

[c] If dynamic scheduling is feasible, these generating-reserve services can be provided and controlled by another supplier.

[d] Whether the market for generator voltage control (VARs) is competitive depends on the specifics of each situation.

SCHEDULING AND DISPATCH

Although scheduling and dispatch are two separate services, we lump them together because they are inexpensive and both are performed, or at least coordinated, by the ISO. Scheduling is the before-the-fact assignment of generation and transmission resources to meet anticipated loads. Because the ISO has the ultimate responsibility to maintain reliability within a control area, the ISO must coordinate the schedules. Scheduling can encompass different time periods: a week ahead (e.g., a utility will schedule its units on Thursday for each hour of the following week), a day ahead, and a few minutes before each hour. Scheduling generation occurs for flows out of a control area, flows into a control area, and flows through a control area. Monitoring transmission lines and equipment also occurs for flows through a control area.

Dispatch is the real-time control of all generation and transmission resources that are currently online and available to meet load and to maintain reliability within the control area. Dispatch can include decisions on which generating units to operate at what levels to minimize fuel and variable operating costs, but such least-cost dispatch is not necessary. That is, buyers and sellers, acting through bilateral contracts, can decide which units to operate at what levels. However, the system operator must have control of enough generation and transmission resources to maintain system reliability and stability, to minimize equipment damage, and to redispatch generating units because of transmission constraints. The ISO needs information from generators and customers concerning the value of transactions to economically redispatch (Hogan 1995b).

Scheduling and dispatch are very inexpensive, requiring only computers, metering and communication equipment, control-room operators, and accountants to manage the after-the-fact accounting and billing. Overall, this service costs less than 0.2 mills/kWh. Because only the system operator can perform these services, it cannot be provided competitively.

LOAD FOLLOWING

The various definitions of generation-reserve services that exist today are generally based on control-area concepts. These concepts start with the basic principle that loads and resources will maintain an instantaneous balance. In addition, frequency will be maintained close to 60 Hz.[*] Each control area in an interconnected system will maintain enough generating capacity online to provide for the area's loads, including the provision for contingencies, and to help maintain constant frequency (i.e., the A1 and A2 criteria discussed above).

Figure 3 shows ACE for a midwestern control area on its peak winter day in 1995. During the first 10 minutes of the peak hour, the system was undergenerating and did not achieve a zero crossing and therefore failed the A1 criterion. During the rest of the hour, however, the ACE crossed zero many times. Between 8 am and 9 pm, the system controlled its generation tightly enough to meet the A2 criterion (i.e., keeping the average load within ±48 MW of balance during each 10-minute period) almost all the time.

These concepts have given rise to various generation-related ancillary services, including frequency control, regulation, load following, energy imbalance, spinning reserve, supplemental reserve, nonoperating reserve, and standby service. Unfortunately, the definitions of, and boundaries among, these services are often unclear. The primary reasons for this lack of clarity,

[*]Many generators and some loads cannot operate at frequencies far from the standard 60 Hz. Frequency deviations of about a Hz can create destructive resonances in generators. In addition, system control is improved by maintenance of a standard frequency within narrow limits.

we believe, are (1) ambiguity about what can be purchased from another supplier and (2) differences between a control area and an individual customer’s load. In addition, the relationships among utilities and of the utilities with the regional reliability council complicate clear definitions for these services.

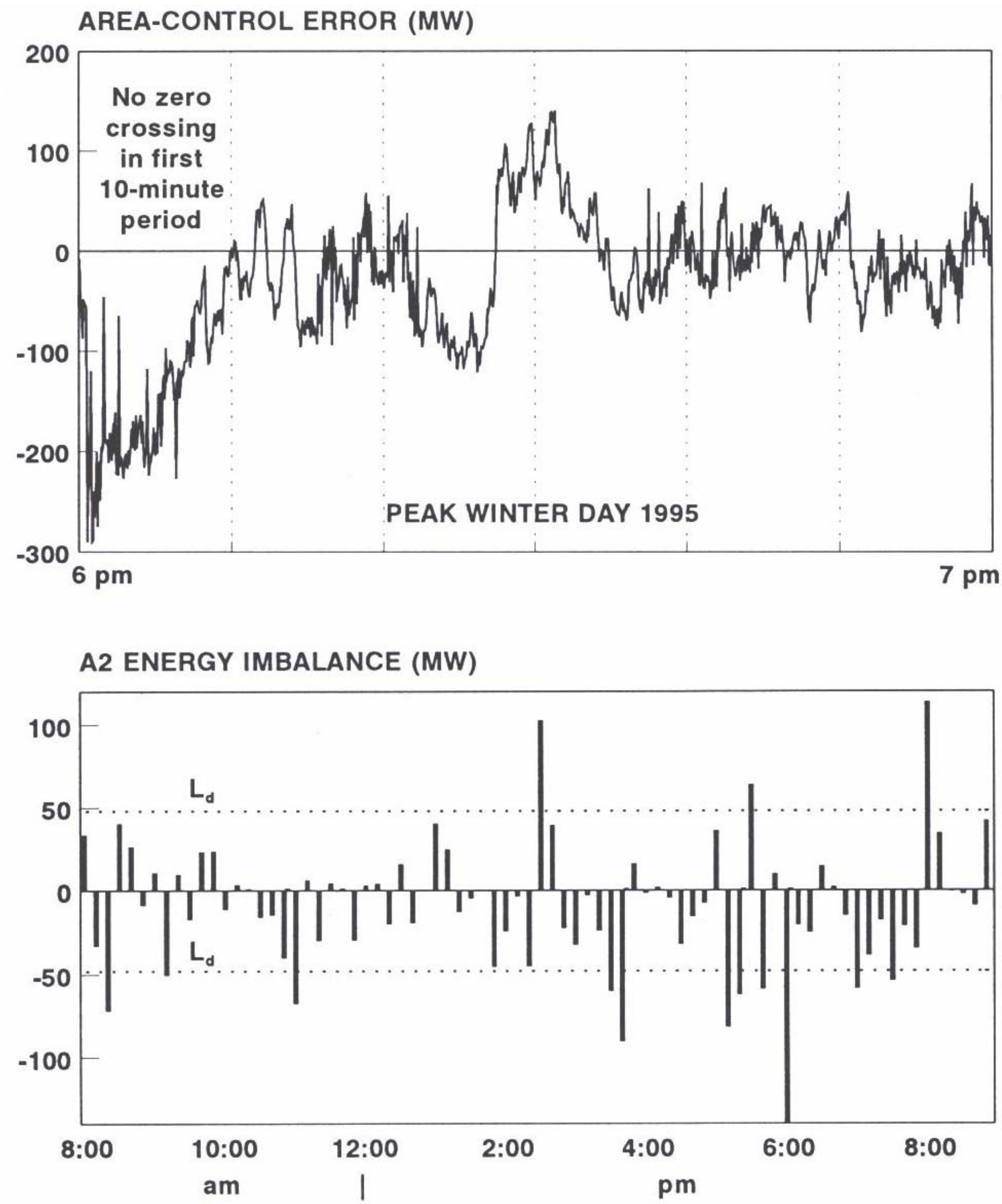

Fig. 3. **Area-control error for a Midwestern system during a winter weekday. The top figure shows ACE and compliance with the A1 criterion from 6 pm to 7pm, when the system peak occurred. The bottom figure shows compliance with the A2 criterion for each 10-minute period that day from 8 am to 9 pm.**

Figure 4 shows the load for a hypothetical customer from 7 am to 8 am on a weekday morning. The total load consists of three primary components. The first element is the minimum constant (base) load during the hour, 80 MW in this example. The second element is the trend during the hour (the morning pick-up in this case); here that element increases monotonically from 0 MW at 7 am to 14 MW at 8 am. The third element is the random fluctuations in load around the underlying trend; here the fluctuations range over ±2 MW.* Combined, the three elements yield a range of loads during this hour of 78 MW to 96MW, with a mean of 85 MW.

The third element, fluctuations, can be categorized by speed. Generators respond automatically (based on their inertia, governor control, impedance, and electrical proximity to the load) to fluctuations that occur faster than about a couple of seconds. Generators respond to slower fluctuations based on signals from the control-area operator's automatic generation control (AGC) system. The AGC system measures ACE every two to six seconds and sends signals to those generators that provide load following to increase or decrease output. Thus, generator response to fast fluctuations is automatic, based on the electrical properties of the generators and transmission system. Generator response to slower fluctuations, on the other hand, is managed by the control-area operator.

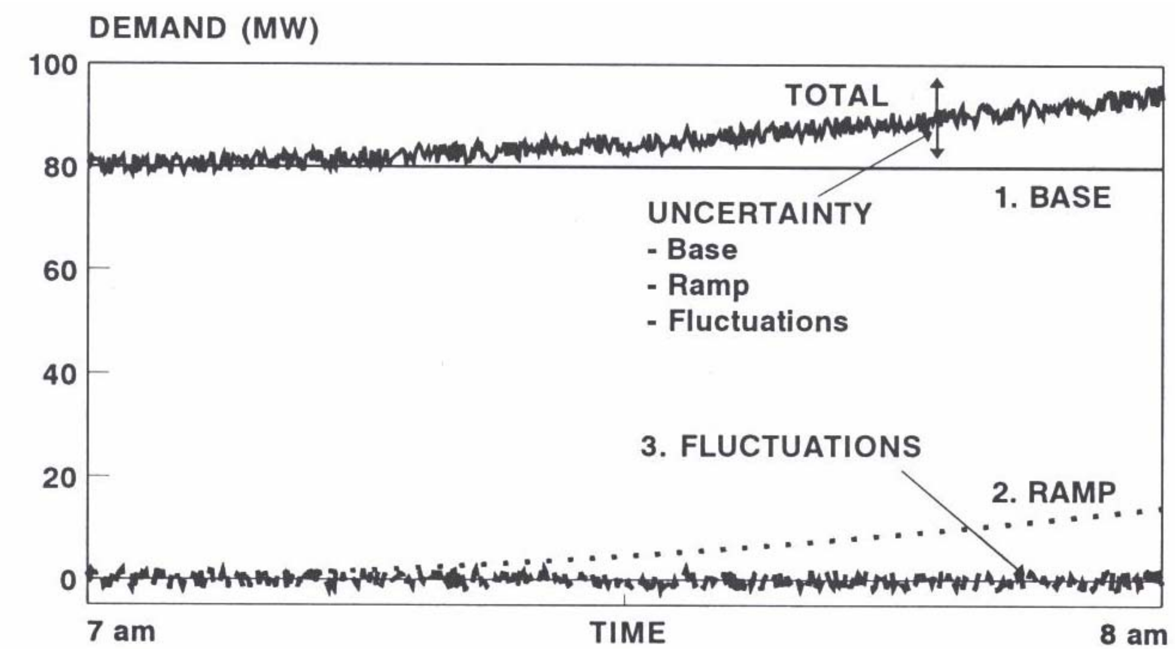

Fig. 4. Components of a hypothetical customer's time-varying load, from 7 am to 8 am on a weekday morning.

*Based on data from two utilities, one in the southwest and one in the midwest, the maximum change in system load from one hour to the next is five to ten times greater than the intrahour load fluctuations.

In addition to the three elements discussed above, uncertainty about loads further complicates the decision on how to meet load. All three elements are likely to vary from day to day with changes in weather, employment, and other factors unique to the particular customer. For example, a sudden change in wind direction might drive clouds over the service area, in response to which loads would increase as customers turn on lights. Under the traditional industry structure, the local utility must have available additional generation capacity to respond to these unanticipated changes in load. In the emerging structure, the customer may need to purchase additional services, either from its control area or from another supplier.

Given these different components of a customer's load, what can that customer reasonably purchase from a supplier? The customer can easily purchase a block of power consistent with its base demand. Even here, however, ramping requirements at the start and end of the hour complicate the situation.[*] Can a remote supplier provide ramping? Although dynamic scheduling is often technically feasible, it is not widely used today. Thus, today's protocols for scheduling transactions generally do not cover the intrahour ramping component. Can a remote supplier follow random fluctuations? Again, not under today's protocols. A remote supplier could fully meet this customer's time-varying load only if that load was telemetered to the supplier's control area and to the supplier's generating units, as well as to the customer's control area. In this process, called dynamic scheduling, the load would effectively be removed from the customer's control area and placed in the control area of the supplier. Dynamic scheduling is widely accepted in principle, but its implementation is rare.

Given this parsing of a customer's load, the various ancillary services discussed today clearly do not match the elements of the load. For example, is load following the same as momentary fluctuations or is it equal to the sum of these fluctuations and ramping? How does regulation differ from load following? Where does energy imbalance fit in?

The law of large numbers further complicates quantification of the load-following requirements for each transaction. The cost to follow system load is much less than the sum of the costs to follow the individual loads that, in aggregate, comprise the system load. This benefit occurs because the individual loads, especially the fluctuations, are generally not correlated with each other.

[*]Operators agree on certain protocols to use in scheduling transactions. For example, the control area in which the supply is located may ramp up its generation from zero at 5 minutes before the hour to full load at 5 minutes after the hour, with the receiving control area doing just the opposite. At 55 minutes, the first control area might begin to ramp down its generation to reach zero at 5 minutes after the next hour.

The generating units that provide ramping are chosen because they can respond to controls and will fit into optimal (i.e., least-cost) dispatch when they are loaded. The units that follow fluctuations need to respond more rapidly to control signals (in MW/minute) and, because they oscillate throughout the day, may not need to fit into the least-cost dispatch. Both types of generating unit respond to unscheduled changes in load.

We believe that ramping and fluctuations could be combined into a single service. One could define the base component of demand at the average power level during an hour (e.g., at 85 MW for the example in Fig. 4). Such a definition of the base would require the units that respond to time-varying demands (elements 2 and 3) to “absorb” energy (i.e., reduce their output) during the first part of each hour and to provide energy during the later part of each hour during the morning pickup; the reverse would occur during the afternoon dropoff. Such a definition of the base component would require accurate prediction of the time-varying load during each hour to ensure that elements 2 and 3 included only capacity and no net energy. Units that provide these services would also have to provide enough base energy to allow for under-delivery at certain times. A more practical approach is to define the base as the lowest likely level of demand during a particular time period (80 MW in our example) and to have the service that meets elements 2 and 3 provide both capacity and energy.

The volatility of electric-arc-furnace steel mills provides a vivid example of the complications in defining load following (Fig. 5). Although this steel-mill load averaged 114 MW during this 3-hour period, its load varied from 2 MW to 206 MW, with a standard deviation equal to 54% of its mean load. By comparison, the total system load (net of this steel mill) had a standard deviation equal to only 2% of its mean during the same 3-hour period.

This discussion suggests that some of the generation-related ancillary services can be combined into one customer load-following service (the sum of ramping and fluctuations in Fig. 4). If this service accurately follows the time-varying loads of customers, the control-area requirements traditionally met by frequency control and tie-line regulation will be automatically satisfied except for outages and losses. Also, by definition, energy imbalance will be zero under these conditions. Operating reserves will still be needed to protect against generator and transmission contingencies.

Thus, load-following service, in our view, includes four separate components. The two control-area functions are maintenance of interconnection frequency at 60 Hz and maintenance of generation/load balance within the control area. The two customer functions include following the moment-to-moment fluctuations in load and following the longer-term (e.g., hourly) changes in load.

The output of the generating units used to provide load-following service is adjusted continuously and automatically to compensate for changes in system load. These generating units have governors, which automatically adjust unit output in response to frequency changes. The units also have AGC equipment that responds to signals from the system operator’s computer to change output in response to changes in ACE. Typically, utilities assign up to 1% of their generating capacity to load following.

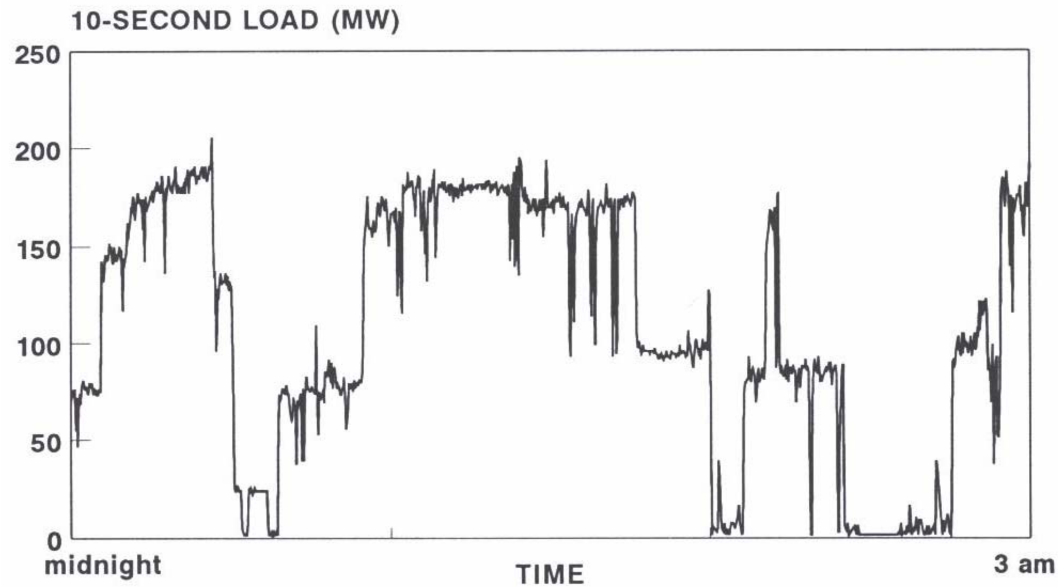

Fig. 5. Variation in the load of an electric-arc steel mill from midnight to 3 am.

Load following includes the fixed costs of the particular generating units used to follow load plus the variable costs associated with increased O&M, higher heat rates, and the requirement to operate these units out of merit order. Because these units are constantly increasing or decreasing output, their operating lifetimes may be shorter than if they were operated at a more nearly constant output; this lifetime loss adds to the fixed cost for this service. Finally, some of the capital costs of governors and AGC equipment should be assigned to load following. Overall, load following costs about 0.5 mills/kWh.

In principle, customers should be charged for load following on the basis of their contribution to the volatility of the aggregate load (e.g., the covariance between an individual load and the system load). In practice, it may be simpler to charge on the basis of the volatility of individual loads (e.g., its standard deviation). An interesting question concerns the appropriate time period over which to determine such temporal variations. The usual 30- or 60-minute interval may be too long to capture the effects of load volatility on utility costs. As shown in Fig. 5, some customers impose substantial load-following costs on the utility, while other customers (such as paper mills) may impose near-constant loads that require little of this load-following service.

Customers with fluctuating loads that impose costs on generators should be charged for those costs. Generating units located within or electrically close to the control area can provide load following and should be compensated for their response to these fluctuations. Compensation for fast fluctuations (roughly within two seconds) should be based on the inherent characteristics of generators and the transmission grid. Compensation for slow fluctuations, which could be competitively procured, should be based on the unit's response to AGC signals.

OPERATING RESERVES

Operating reserves are, in some respects, the supply-side analog of load-following service. While load following matches generation to load based on the time-varying nature of demand, operating reserves balance generation to load in response to unexpected generation or transmission outages. (In practice, the spinning portion of reliability reserve also includes the load-following service.) Generating reserves used to meet generating and transmission outages are split in two:

- Reliability reserves, which include spinning reserves and other generating units that can be started quickly, all of which must be fully available within 10 minutes; and

- Supplemental-operating reserves, which include generating units that can begin to provide power within 10 minutes and are fully available within 30 minutes. These reserves are intended to replace the reliability reserves and stand ready to meet additional contingencies.

The spinning (synchronized) portion of reliability reserves is controlled in the same way as load-following service. Both detect and respond to ACE (i.e., discrepancies in actual and scheduled interchanges plus frequency deviations). An important difference is that load following is responding all the time to small changes in system load, while operating reserves respond to infrequent, but usually larger, failures of generation or transmission. We split these reserves into their two parts because they differ in the types of equipment used to provide the service, the number of potential providers (including loads for supplemental-operating reserve), the extent to which they must be controlled by the system operator, and their capital and operating costs.

The costs of reliability and supplemental operating reserves include both fixed and variable components. The fixed-cost component is the annualized cost of the generating units plus control equipment used to provide these reserves. This component might also include the fuel costs associated with the out-of-economic dispatch of generators required to provide these reserves. When the reserves are called upon (i.e., to respond to outages), additional fuel costs will be incurred. Some utilities impose both a fixed cost (in $/kW-month) and a variable cost (in ¢/kWh, when these reserves are used) for operating reserves. Overall, these reserves cost about 1.7 mill/kWh, with more than half the cost from reliability reserve.

Reliability Reserve

The nine regional reliability councils differ in the details of their requirements for reserves. The Mid-American Interconnected Network (MAIN) Guide No. 5A illustrates how the councils set the required levels of operating reserves (MAIN 1995). MAIN defines operating reserve as "that capability above system demand required to provide for regulation, load forecasting error, equipment forced outages and area protection."

With a 1994 peak demand of 42,200 MW, MAIN requires a total of 470 MW for regulation (what we call load following), 1.1% of peak demand. Regulating margin is the "portion of spinning reserve capacity required to respond to all normal, minute-to-minute demand changes in order to meet NERC Control Performance Criteria."

Spinning reserve is "that amount of unloaded generation which is synchronized, ready to serve additional demand and which can be fully applied within 10 minutes." MAIN's additional requirement for spinning reserve is 620 MW, 1.5% of peak demand. Thus, total spinning reserve equals 2.6% of peak demand.

Spinning reserve is spread over as many units as is practical because it is easier to get the required rapid response by adjusting several units a small amount rather than by adjusting a single unit a large amount. Any generating unit equipped with a governor and AGC can help provide this service.

Nonspinning reserve is "that amount of generating capability, or its equivalent, not connected to the system but capable of serving demand within 10 minutes, or interruptible load that can be removed within 10 minutes." The MAIN operating reserve requirement that need not be spinning is 620 MW, 1.9%.

Thus, MAIN requires generating reserves that total 4.5% of peak demand. This amount is based on the gross capability of the largest resource online that day in the region plus the sum of the individual systems' regulating margins. MAIN does not specify a minimum amount of supplemental-operating reserve; however it does require that operating reserves be restored "as soon as practicable."

Overall, reliability reserve costs about 0.9 mill/kWh.

Supplemental-Operating Reserve

Utilities maintain additional generation reserves to cover times when the spinning reserves are insufficient. Such reserves are expected to be available within 30 minutes. Each utility's supplemental operating reserves are set by its NERC region and generally total about 3% of the expected daily peak load. These reserves not only back up reliability reserves but are also used to restore the generating mix to a least-cost configuration. Any generating unit or interruptible load could help supply this service if it can be fully available within 30 minutes.

Supplemental reserves are less expensive than reliability reserves because they do not require governors or AGC. Also, supplemental reserves are not necessarily maintained in as ready a state as are reliability reserves. Overall, supplemental reserves cost about 0.8 mill/kWh.

As with load following, only the system operator knows when and how much operating reserves are needed. However, any generating unit or interruptible load within or near the control area can provide the service.

ENERGY IMBALANCE

Energy imbalance is unavoidable because it is impossible for each control area to exactly match its generation to load in an interconnected system. As defined in FERC's NOPR, energy imbalance is a confusing service. In some sense, it is the customer equivalent of a control area's inadvertent interchange. At both the customer and control-area levels, the service is intended to serve primarily as an accounting mechanism to ensure appropriate compensation (to the local control area for energy imbalance and to other control areas for inadvertent interchange) for the unavoidable small discrepancies between actual and scheduled flows.

FERC's definition of energy imbalance specifies a deadband of ±1.5%. If the deviation between actual and scheduled flows, measured over each 1-hour period, is within this deadband, the customer can return the imbalance "in kind" during a like time period (onpeak or offpeak) within a 30-day period. Within this deadband, over- and under-generation can offset each other within like time periods.

If the deviation falls outside the deadband, then FERC proposes to charge the customer 10¢/kWh for imbalances outside the deadband. FERC is not clear on whether these charges apply to both undergeneration (where the customer is taking unscheduled energy and power from the local control area) and overgeneration (where the customer is supplying unscheduled energy and power to the local control area). We assume that the customer would pay for undergeneration and would receive some compensation for overgeneration. The prices for under- or overgeneration would be set at some percentage above or below the local control area's marginal cost of production to discourage imbalances.

In our view, this definition of energy imbalance is too broad. It encompasses both an accounting service (intended to compensate the control area for minor discrepancies) as well as a penalty for substantial deviations from schedule, which are more akin to backup services. Figure 6 illustrates the reasons for our belief that FERC's definition is too relaxed. The four hypothetical loads have different frequencies and amplitudes but, averaged over the 60 minutes shown, each has a zero imbalance. However, the load-following burden and costs imposed on the local control area increase in going from A to B to C to D. Under FERC's definition of energy imbalance, the compensation to the local control area would be the same for the four loads-zero.

We believe that if a customer obtains load-following and reliability reserves from the local control area, then it is already obtaining the service that energy imbalance would offer. That is, the customer should not have to pay the local control area for load-following and reliability reserves and at the same time for energy imbalance. Presumably, if a customer purchased load-following and reliability reserves from another supplier, that supplier would be responsible for energy-imbalance charges.

We suggest a three-part split of energy imbalance:

- Energy imbalance would include only the discrepancies within the defined deadband as measured over the defined time interval. The deadband could be set at ±1.5%, ±3%, or some other number (perhaps based on the NERC control performance criterion L_d, which is typically ~0.5% of system peak). The appropriate time period for measuring energy imbalance could be set at the 10 minutes used in the NERC A1 and A2 criteria or the 60 minutes proposed by FERC. To prevent chronic abuse, it may help to set tight limits on the deadband and the reconciliation period and to allow occasional short deviations outside the deadband (e.g., in the event of a forced outage). For example, compliance could be required for at least 95% of the time periods, consistent with the standard used for NERC's Control Performance Honor Roll. On the other hand, the benefits of tight limits on energy imbalance must offset the higher costs of metering, accounting, and billing. Any imbalance outside the deadband is handled with either standby service or unauthorized use, discussed below.

- Standby service would be contractually arranged beforehand between the customer and a supplier (not necessarily inside the local control area). Presumably, the provider of this service would impose both demand and energy charges for this service. See discussion later in this chapter.

- Unauthorized use would result in penalty charges imposed by the local control area if (1) the customer's load fell outside the deadband and (2) the customer had not arranged for standby service. Because unauthorized use is not a service, its charge would not be based on costs. Rather its price would be designed to encourage customers to obtain standby service and to discourage customers from leaning on the local control area.

Assuming that customers, on average, incur an energy imbalance outside the deadband equal to 1% of their loads, this service costs about 0.6 mills/kWh. These costs cover both the capital and operating costs of the generating units that provide the service.

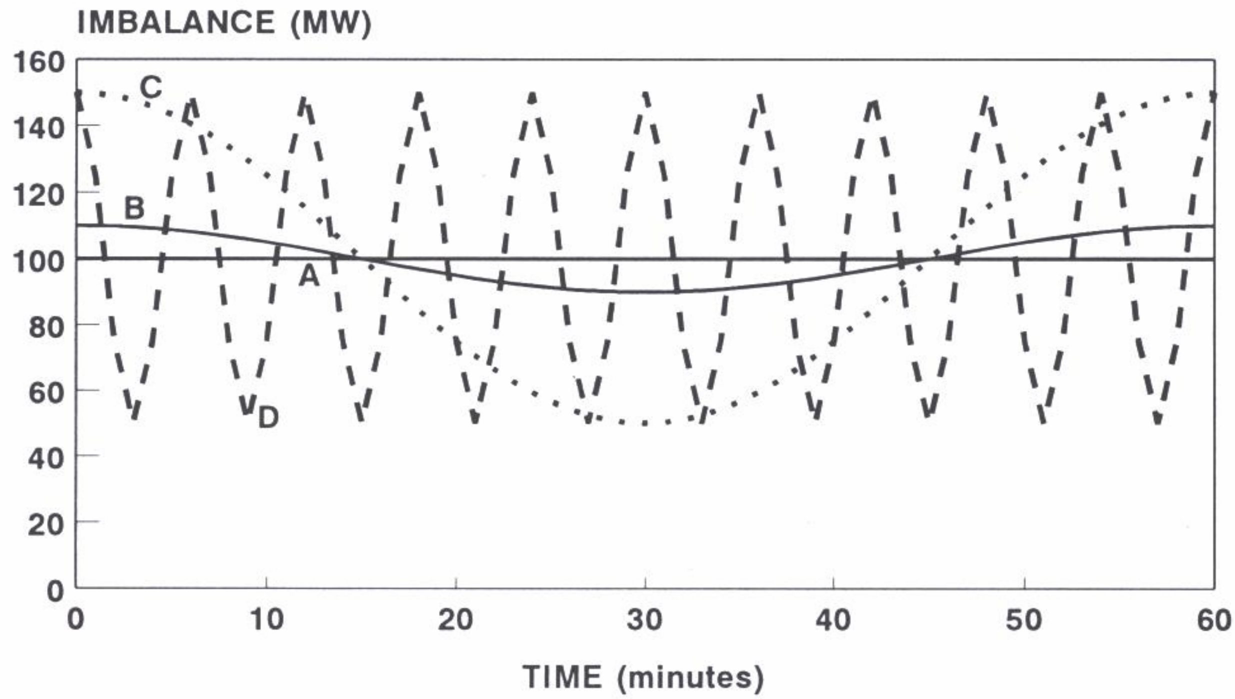

Fig. 6. Four hypothetical loads, each of which has zero energy imbalance for the hour. Load A meets its schedule exactly. Load B is above schedule at the beginning of the hour, below schedule during the middle, and above schedule at the end of the hour. Load C has the same frequency as load B but an amplitude five times greater. Load D has the same amplitude as load C but a frequency 10 times greater. In each case, the excesses and deficiency exactly balance, however.

REAL-POWER-LOSS REPLACEMENT

Real-power losses are the differences between generated real power and the real power delivered to customers. Moving power always results in losses because of the resistance of each element in the transmission system. The losses depend on network configuration, generator locations and outputs, and customer locations and demands.

Losses are composed of the excitation (no-load) and load losses of each element. Excitation losses are voltage dependent and essentially load independent. Load losses for most devices are a function of the square of the load. For a typical transmission system, losses average 2 to 3% of the system load. However, losses vary greatly as conditions on the network change. In particular, at times of system peak demands, losses are often much higher than under average loading conditions. The nonlinear nature and temporal variations in losses make it difficult to compute their costs and to assign them unambiguously to particular transactions.

Losses are generally not measured directly; they are calculated with load-flow computer models. In principle, these models can be used to calculate losses in near-real time. The results of these models can be used to allocate losses to individual transactions in different ways. Each customer could be charged for losses on the basis of the difference between the spot price at its location and the spot price at the generator's location. Alternatively, the models can calculate the incremental losses associated with each transaction (i.e., the losses that would occur if this transaction were the last one added to the system); these incremental losses could be scaled up or down to match total losses (Alvarado 1996).

Figure 7 shows how average and marginal losses vary with load for Northeast Utilities (1989). Marginal losses represent the losses associated with a small increment or decrement of load on the system. Marginal losses do not include no-load losses, which account for about 20% of total losses. Losses vary with load and also in response to changes in the operation of cogeneration facilities, nuclear plants, and pumped-storage facilities. The different locations of these generating units on the Northeast Utilities system affect losses.

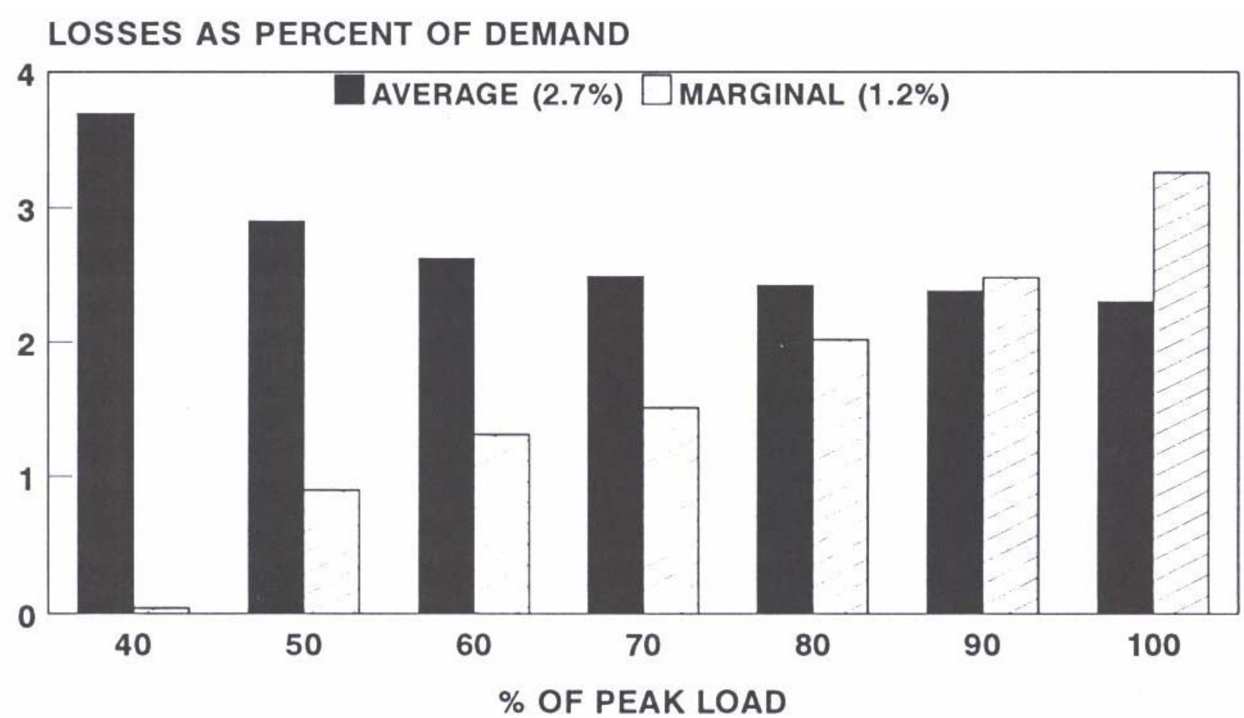

Fig. 7. Transmission-system losses for Northeast Utilities as a function of load.

Real-power losses must be made up by generators. The ISO could run the generators it controls to compensate for the losses, it could contract with another supplier to provide for the losses, or customers could contract with other suppliers to provide for the losses. Retail customers usually pay for losses on a system-wide basis. Point-to-point transaction customers (where a customer contracts with the system operator to move power from one point to another) can either pay the system operator for the losses or they can supply extra power to make up for system losses. The ISO must have online generation to compensate for real-time losses even if, on average, other suppliers make up these losses.

Typically, energy losses are paid for on a ¢/kWh basis and vary with time, based on the variable operating costs of generating units. Demand losses are paid for on a $/kW-month basis and reflect the costs of additional generating and transmission capacity. Only the system operator has sufficient information to know what the losses are at any time. On average, real losses cost about 1.3 mills/kWh.

VOLTAGE CONTROL

Reactive-power management and voltage control are the same service. System voltage control is used to maintain voltages within prescribed limits at various points in the transmission grid and to compensate for the reactive requirements of the grid. Injection and absorption of reactive power is also required to maintain system stability, in particular to protect against contingencies that could lead to voltage collapse. Enough reactive-power capacity must be available to meet expected demands plus a reserve margin for contingencies. Thus, voltage control is analogous to reliability spinning reserve. Local voltage regulation is a customer service intended to (1) meet customer reactive-power needs and (2) control each customer's impact on system voltage and system losses so power-factor problems at one customer site do not affect power quality elsewhere on the system.

We split the services into a local component and a system component because the customer has sufficient information at its location to control local reactive-power demand and the local voltage, while only the system operator has sufficient information to know what the voltage-regulation and reactive-power requirements are throughout the grid. Because local voltage control is a customer problem, not a grid problem, we do not consider it an ancillary service. Local voltage control can be managed by the customer, the ISO, or the local distribution company.

Reactive losses are much higher than real losses. Voltage drops are predominantly caused by the inductance of the lines and transformers (rather than the resistance), and can be compensated for by supplying reactive power. (Conversely, too much reactive compensation can produce excessively high voltages.) Because of the high inductance of lines and transformers, reactive power does not travel well through the transmission system, so reactive support must be provided much closer to reactive loads than real power needs to be provided to real loads.

Voltage regulation is aimed primarily at maintaining voltages within certain ranges, but is also concerned with minimizing temporal variations in voltage. Voltage regulation is generally more of a local problem than is provision of real power for the reasons given above. Finally, most devices (lines, transformers, breakers, etc.) are load limited by current rather than by real power. If they are carrying significant reactive power and reactive current, they have less capacity available to transport real power.

Voltage is controlled throughout the transmission system through the application and operation of ratio-changing devices (e.g., transformer taps and voltage regulators) and reactive-power-control devices (e.g., capacitors, reactors, static-var compensators, generators, and occasionally synchronous condensers). The system operator must monitor and control these voltages and supply the reactive-power requirements of the grid. At low-load times, VAR absorption is required to keep voltages from getting too high; at high-load times, VAR production is required to keep voltages from getting too low. At certain locations, it may be more economical for the utility to purchase reactive support from a customer or generator than to supply the reactive support directly. The equipment used to provide or absorb VARs can be categorized as dynamic (referring primarily to generating units) or static (referring primarily to transmission-system equipment) depending on its ability to change VAR output or absorption quickly. Per VAR, dynamic equipment is much more expensive than static equipment.

The cost of supplying reactive power is primarily the capital cost of the equipment (e.g., generators and capacitors) needed to meet expected requirements plus a reserve. In addition, the operating cost of over- or underexcitation of generating units should be assigned to reactive support. The primary cost of voltage support provided by generators is for the losses in the rotor, stator, exciter, and step-up transformer. In some cases, there may also be an opportunity cost associated with the reduction in real-power production capability caused by production or absorption of VARs. Transmission-related voltage-control devices have both capital and operating costs.

Because the cost of system voltage support cannot easily be assigned to individual customers, its cost should probably be included in the basic transmission tariff.[*] However, the system operator could purchase VAR support from generators as a separate service. Thus, voltage control is a service that, in our view, should be unbundled to suppliers but not to customers. Overall, voltage control costs about 0.4 mills/kWh.

OTHER SERVICES

Generating units and transmission systems provide other services beyond the basic services of energy supply, generating capacity, and power delivery plus the ancillary services discussed above. These additional services are discussed here either because we believe their costs should not be unbundled to customers, we are not sure how best to define and measure the service, the low cost of the service may not warrant the costs of metering and billing, or the

[*]The difficulties of assigning system voltage support to individual transactions are the same as those for assigning real losses to individual transactions; indeed, the same computer models are used to calculate both. The costs, however, differ because VAR costs are primarily capital while loss costs are primarily fuel.

service is highly site- and customer-specific. Also, the importance and pricing of these services will depend strongly on the future structure of the electricity industry, especially the role of the independent system operator.

Black-Start Capability

Under conditions of system collapse, it may not be possible to draw power from the grid. Generating units that can restart without taking power from the grid are called black-start units. After the unit starts itself, it can be used to start other units and to energize the transmission grid. Almost any generator can be used to start other units if it has black-start capability and is appropriately located on the grid.

Black start is a vital but inexpensive service. Its costs are primarily the capital cost of the equipment used to start the unit, the cost of the operators, the routine maintenance and testing of equipment, and the cost of fuel when the service is required. Because the system requires this service to be viable, all customers require this service. Therefore, we suggest that the cost of black-start capability be included in the basic transmission charge and not itemized, primarily because there is no way to determine different costs for different users. However, generators should be compensated explicitly for this service, perhaps through an auction conducted by the ISO every few years

Time Correction

Most electric clocks operate by counting the cycles in the power frequency. Although the frequency is maintained within tight limits (often within 0.01 Hz of the 60-Hz reference), the time displayed by power-system-supplied clocks would drift if not corrected. A 0.01-Hz average error would cumulate to 15 seconds a day. Generation must be adjusted periodically to correct for this error.

American Electric Power, Southern California Edison, and the North Texas Security Center monitor real and apparent time for the eastern interconnection, the western interconnection, and Texas, respectively. When the difference between real and apparent time reaches two to eight seconds, the time monitor notifies each control area in the interconnection. All control areas then adjust their frequency setpoint of the AGC by 0.02 Hz from 60 Hz (up if apparent time is behind real time or down if apparent time is ahead of real time). The frequency setpoint is returned to 60 Hz when real and apparent time are once more synchronized. This is a small, inexpensive, mandatory, and important service. Because its costs are so small, it should not be unbundled to either generators or customers.

Standby Service

Standby service refers to generating capacity that is reserved ahead of time to provide energy and capacity when contingencies occur. This capacity could be used when a customer’s generation supply is interrupted (e.g., a forced outage), when that generation is offline for planned maintenance, or when the customer's demand exceeds the contracted amount (Toulson and Mauldin 1995).

Standby service includes both capacity and energy components. The capacity component refers to the generating capacity that is available to serve customers. Customers pay for this capacity (in $/kW-month) regardless of whether they actually use it (a form of insurance). When they do call on this capacity, they would pay an additional charge (in ¢/kWh) for the energy they consume. The usage characteristics for standby and supplemental power can be quite different, with standby involving large demands infrequently (only when a generation outage occurs) and supplemental involving much smaller demands much more often (whenever customer loads change). Therefore, these two services could be priced differently.

Standby service is important and expensive. We do not consider it an ancillary service because customers can decide for themselves whether to acquire the insurance that this service offers, risk facing large energy-imbalance charges, rely on the spot market, or curtail operations.

Planning Reserve

Most discussions of wholesale power markets focus on time-varying spot prices, with prices set on the basis of generator bids for each hour. Recent proposals in New York and the mid-Atlantic states add a planning-reserve component to these energy-only spot prices.

For example, the PJM Interconnection Association (1995) proposed a reserve-sharing agreement based on the pool's installed-capacity requirement. Each year, the ISO (the successor to PJM, in this case) would develop or collect load forecasts to determine expected system demands for the coming year. Based on the pool's reliability requirement (e.g., loss-of-load probability of no more than one day in 10 years), the ISO would determine the amount of generating capacity required beyond that needed to meet the system's expected peak demand for the coming year. Based on the quantity and quality (e.g., planned and forced outage rates and unit size) of the generating units used to meet customer demands, each entity with a responsibility to meet customer loads would be assigned a capacity requirement for the coming year.

Putnam, Hayes & Bartlett (1995) suggested a method to calculate the price of this planning reserve. Their proposal would pay generators a fixed amount ($/kW-year) equal to the cost of the units needed to maintain the pool's reserve-margin requirement.

We do not list planning reserve as an ancillary service because we are not sure whether it will serve any purpose in competitive electricity markets. Many economists argue that changing spot prices will be enough to balance supply and demand in real time. When supplies are limited relative to demand, the spot price will increase far above the marginal cost of the last unit online. That higher price will both equilibrate supply and demand and provide enough compensation to generation owners that some “extra” generating capacity or interruptible load will be made available. Many engineers, on the other hand, are skeptical that customers can respond quickly enough to price changes to maintain reliability. They would prefer to charge for extra generating capacity up front to be sure that the lights do not go out. Only time will tell which model works better in balancing economics with reliability.

Redispatch

The least-cost dispatch may not be possible because of transmission constraints (i.e., voltage, thermal, or stability limits), sometimes referred to as congestion. The lowest-cost production may come from a remote plant whose energy must be imported into the load center over long-distance transmission lines. If the transmission system is composed of a few high-capacity lines, loss of one of these lines may limit the import capacity to the point that service reliability may be unacceptable. To resolve this problem, a utility may reduce output from low-cost but remote generation units (said to be constrained off) and, instead, operate more expensive units near the load center (said to be constrained on) to provide backup for the transmission system. [Several generating units in New England sometimes operate out of merit order to maintain reliability in northeastern Massachusetts (New England Electric System 1995). Compared to expansion of the transmission system, it is cheaper to operate these units even though their variable costs exceed the market price of power.]

Pacific Gas and Electric (1989) identified several constraints that force operation of generation away from least cost because of reliability, safety, and environmental considerations:

- PG&E limits operation of certain plants at certain times to protect striped bass. These limits reduce cooling-water intake to prevent problems of entrainment or impingement in the plants’ once-through cooling systems.

- PG&E operates certain plants within San Francisco out of merit order to protect against transmission-system contingencies within the city. At least 50% of the local load is met with local generation.

- The Humboldt Bay area is remote, served by two 115-kV transmission lines that pass through rugged, mountainous terrain subject to severe weather conditions (wind, snow, and lightning). Therefore, PG&E limits imports into the area to 70 MW, which requires local generation operated out of merit order.

- Regional conditions require some generation in the East Bay region because of its large industrial loads. These local generating requirements are used for intrahour load following, to provide local voltage support, and to protect against overloading regional transmission lines.

Any properly located generator can provide this local-area security. If the transmission limitation is thermal rather than stability, customers might help provide the service by allowing loads to be automatically interrupted, but only the ISO has sufficient information to know when control actions are required. Compensation for redispatch might be based on the difference in cost to operate the constrained-on units relative to the cost to operate the constrained-off units. Redispatch is not considered an ancillary service because it is very site- and case-specific.

Transmission Services

In our earlier report (Kirby, Hirst, and VanCoevering 1995), we identified several transmission-related ancillary services that we now think should be bundled within the basic transmission tariff. These services include transmission-system monitoring and control; transmission reserves; repair and maintenance of the transmission network; and metering, billing, and communications. We no longer consider them ancillary services because they cannot readily be unbundled and priced separately to different customers. That is, these functions cannot be assigned differentially to different customers. For example, there is no way to identify customer-specific costs for outage restoration. If all customers face the same costs for these services, there is no reason to unbundle them.

Power Quality

Power quality refers to the provision of an uninterrupted power supply with a pure sinusoidal waveform to customers. Unlike the transmission services discussed above, this is a service that can be provided uniquely and separately to individual customers. However, we dropped this service from our list because we no longer view it as an *ancillary* service and because power-quality requirements and costs are highly customer-specific.

Planning, Engineering, and Accounting Services

Houston Lighting & Power (1995) and the New York Power Pool (1995) include several planning, engineering, and accounting functions in their lists of ancillary services. The planning services include load forecasting, scheduling and coordination of the maintenance of generating units, scheduling of pumped-storage and hydroelectric generating units, and coordination of transmission-facility maintenance and outages. Engineering services include black-start studies, load-flow analyses, and planning for expansion of the bulk-power system. Finally, accounting services include scheduling, billing, contract administration, and reporting to various regulatory bodies (state regulatory commission, FERC, and NERC). We do not consider these ancillary services because they are long-term in nature and often integral to the provision of other services.

DISCUSSION

UNRESOLVED ISSUES

Although considerable progress has been made in identifying, defining, and quantifying the costs of ancillary services, much remains to be done. Uncertainties remain concerning which services can be competitively acquired, which ones (and how much of each) are required to maintain electric-system reliability, which ones must be under the control of the system operator, and which ones depend on industry structure.

How the Current System Operates

A useful starting point would be a clear explanation of the current NERC reliability requirements. As examples, what are the bases for the A1 and A2 criteria used to limit area-control error? What is the rationale for the L_d criterion, and how does the amount of load-following service depend on this factor? How do utilities today (and how might ISOs in the future) derive their requirements for load-following service, reliability reserve, and supplemental reserve from these criteria? Why must frequency be maintained so close to 60 Hz at all times (Fig. 8)? A 1000-MW generation loss leads to roughly a 0.03-Hz decline in the eastern interconnection and a 0.14-Hz decline in the western interconnection. Frequency stabilizes after such a disturbance within about 7 seconds in the east and 14 seconds in the west. Are these responses fast enough? Faster than necessary? Should the A1 and A2 criteria be replaced or supplemented with a statistically based measure of control-area performance, such as the product of ACE and frequency deviations from 60 Hz (Hoffman and Illian 1996)? We suggest that NERC and the regional reliability councils take the lead in explaining clearly what the current “rules of the road” are and in defining the engineering and economic bases for these procedures, guidelines, and criteria.

Tradeoffs Between Reliability and Costs

We also need to understand better the tradeoffs between cost and reliability associated with these rules. How much money would electricity consumers save and by how much would reliability deteriorate if the A1 criterion were relaxed to require a zero-ACE crossing only once every 20 minutes instead of once every 10 minutes? In a more competitive electricity industry, in which customers have more influence over the products and services they purchase, the tradeoffs between cost and reliability might be different from those that apply today.

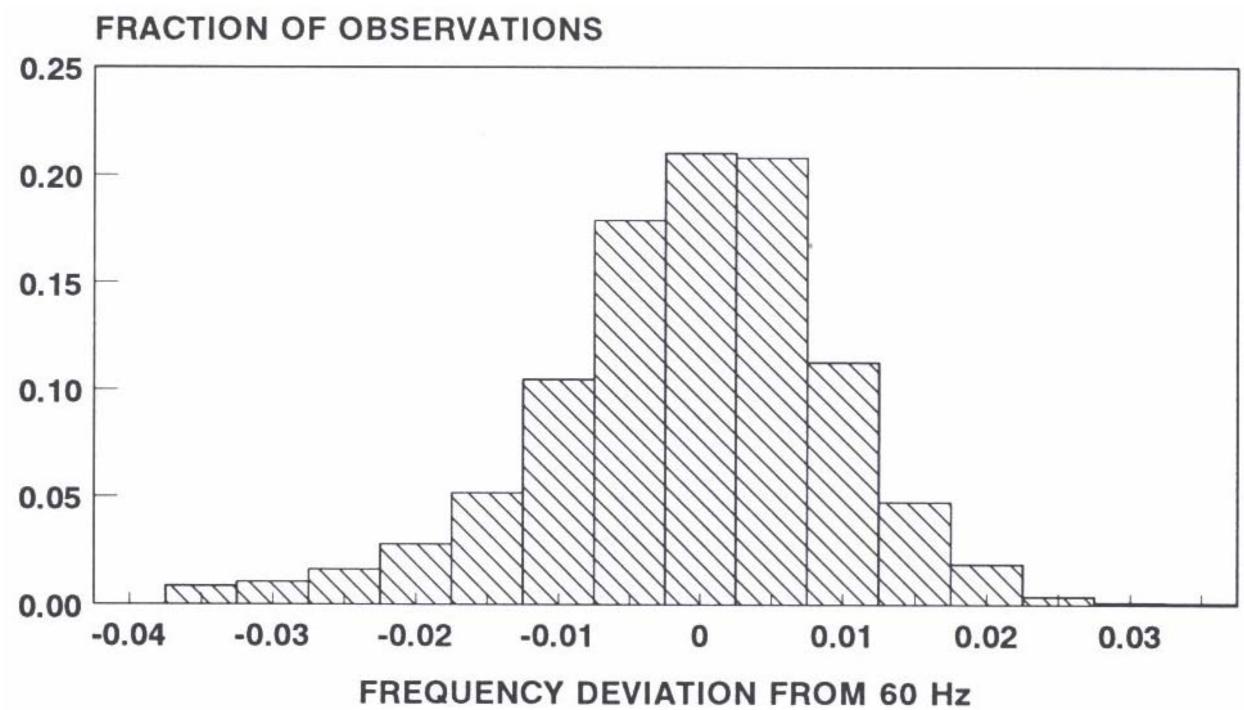

Fig. 8. Frequency deviations from 60 Hz (measured at 10-second intervals) for the eastern interconnection on December 24, 1995. Frequency was within ±0.02 Hz 93% of the time.

Operating Requirements in a Future Industry Structure

Today's control-area operating procedures are based on a particular industry structure. That structure is dominated by a small number of vertically integrated utilities, each of which has an "obligation to serve" its customers. For example, the ~100 control areas in the eastern interconnection are the key actors in establishing the engineering standards for participation in electricity markets and for maintaining system reliability. In the future, ownership and operation of generating units, system control, transmission networks, and customer services may be split among various entities. These structural changes will increase greatly the number of parties that interact with system operators. And the obligation to serve may disappear, replaced by an obligation only to connect. These changes may reduce the ability of any one party to discipline and resolve disputes-to ensure that generation is sufficient to match loads and to maintain system reliability.

Can today's procedures work well in the future when the electricity industry will exhibit much greater diversity in size, resources, responsibility, and interests among industry participants? For example, today's practice of returning inadvertent interchange "in kind" may fail FERC's comparability standard and be replaced with a cash-based system. What procedures should be used to maintain interconnection frequency (and how tightly should frequency be controlled) and to match actual and scheduled transactions? More broadly, what entities (e.g., NERC, regional reliability councils, regional transmission groups, or the ISO) will have the responsibility, ability, authority, and accountability to maintain system reliability and to establish operating requirements?

Ancillary Services in a Future Industry Structure

The relationships between industry structure and the definitions, provision, time period, and pricing of ancillary services warrants exploration. For example, assume a future structure that includes an active power exchange with hourly spot prices and an ISO that provides ancillary services. With such a structure, many of the ancillary services might (1) disappear and be rolled into an uplift charge that is the same for all customers or (2) be charged to customers on the basis of their usage and the current spot price. And these prices would likely have no capacity component but rather would reflect energy only (¢/kWh). The ISO might acquire load-following, reliability, and supplemental reserves on an hourly basis, with the prices of each service tied to the current spot price. Energy imbalance and losses could disappear in such a structure, with the ISO charging customers the current spot price for all deliveries not covered by bilateral contracts. This example illustrates why developing unambiguous definitions for ancillary services is so difficult. The definitions-indeed, the existence-of these services depend strongly on the structure and operation of the electricity industry. The industry structure, as well as the cost and capability of metering and telecommunication equipment, will determine the appropriate time periods (e.g., 10 minutes or an hour) with which to measure and cost these services.

Competition vs Engineering Standards

There is general agreement that some services can be provided by competitive markets and that, where possible, competition rather than regulation should set prices. There is less agreement on which services can be provided competitively and which ones can only be provided by one or a few entities. Because reactive losses are 10 to 20 times greater than real power losses, reactive power must be supplied and absorbed locally. This geographical restriction on voltage control limits the number of suppliers to the transmission entity and those with suitably located generating units. Thus, the pricing of voltage control will likely remain regulated, with prices often based on embedded costs. On the other hand, makeup for real-power losses can be provided by any generating unit connected to the grid over a much broader area. Thus, the pricing of electricity to make up for losses can be set by markets (based either on long-term contracts or on spot prices) because there are many potential suppliers for this service.

The tradeoffs between engineering standards and economic incentives in the procurement and provision of ancillary services is an important issue (Hirst and Kirby 1995b). To illustrate, the system operator could specify a minimum power-factor capability (e.g., 90% or lower) for all generators that connect to the grid, or the system operator could conduct periodic auctions and accept bids from generators and transmission-system owners to supply reactive power.

System operators must decide how to obtain generation reserves. These services, provided by unloaded generating capacity plus governors on the generators and AGC systems, allow the control area to respond to changes in generation and load, both within the control area and in other parts of the interconnection. The rapid ramp rate (MW/minute) of some generators is a valuable feature that the system operator needs to meet the ACE requirements. (Generators vary enormously in their ramp rates. Hydro units can respond at 50 to 100% per minute, combustion turbines at 10 to 20% per minute, and coal units at 1 to 3% per minute; nuclear units are generally not used for ramping.) And the inertia of some generators (i.e., the ability of the unit to provide additional power almost immediately by slowing down slightly) aids system stability. How should these various services be obtained and paid for?

In a competitive market, utilities could charge for load following at a uniform rate for all customers regardless of their load variations. Alternatively, utilities could charge each customer for the specific load-following burden it placed on the control area. The former approach is simple, consistent with the current system, and inequitable. An electric-arc furnace (Fig. 5), with a load that varies by 100 MW from second to second, would pay the same for load following as a paper mill, with a near-uniform load. The second approach is more complicated, consistent with the notions of a competitive market, and equitable. Each customer would pay for the load-following burden that it placed on the control area. Unfortunately, metering and calculating that burden could be complicated because it is not the *individual* load variation that imposes costs on the control area, but the *aggregate* load variation. Thus, the system operator would have to calculate the covariance between a customer's time-varying load and that for the total control-area load. Given such information, the customer could choose to pay the system operator for its load-following costs. Or it could reduce its burden on the system by more carefully managing its load to reduce the fluctuations seen by the control area or by installing energy storage or generation equipment to counteract its load fluctuations.

A similar situation could apply to reliability reserve. In a competitive market, utilities could charge suppliers (generators or power marketers) a uniform rate for this service, or they could impose customized charges based on each generating unit's characteristics (e.g., its reliability as measured by its forced outage rate and its size), or they could impose a fixed reserve requirement on all suppliers. In addition, customers could sell "reserves" if they are willing to allow the system operator to cut load at that customer's facilities in exchange for lower electricity costs. The former approach is simple, consistent with the current system, and inequitable (because it imposes the same costs on all generators, regardless of their reliability levels). The second approach would provide a clear incentive to generators to maintain their units at economically efficient levels of reliability. That is, generation owners would pay for enough maintenance so that the marginal cost of reducing outages would equal the marginal cost of reliability reserves.

Metering and Models

Finally, the industry should identify the types and amounts of metering that will be required to unbundle generation and transmission. Additional meters and computational procedures might be needed to measure the amounts and timing of the provision and consumption of ancillary services. That is, today's metering, which focuses on the control area as the unit of observation, may need to be supplemented with real-time data on the outputs of individual generators and the time-varying consumption by individual loads. The feasibility, costs, and benefits of such additional metering should be addressed. The extent of unbundling ancillary services will likely depend in part on the costs and availability of metering equipment.

CONCLUSIONS

We developed a revised set of services, based on the services suggested by FERC, Houston Lighting & Power, the New York Power Pool, the Michigan Public Service Commission, NERC, and our earlier ORNL analysis. Our seven services include:

- Scheduling and dispatch, the control-area operator functions that schedule generating units and transactions before the fact and control these units in real time to maintain reliability.

- Load following, the use of online generating equipment that is equipped with governors and automatic generation control to track the moment-to-moment fluctuations and the hourly trends in customer loads. In so doing, load-following service (along with reliability reserve) helps to maintain interconnnection frequency and generation/load balance within the control area.

- Reliability reserve, the use of spinning and fast-start generating equipment that can be fully available within 10 minutes to correct for generation/load imbalances caused by generation and transmission outages.

- Supplemental-operating reserve, the use of generating equipment and interruptible load that can be fully available within 30 minutes to back up reliability reserves.

- Energy imbalance, the use of generating equipment and fuel to match any differences between actual and scheduled transactions between suppliers and their customers.

- Real-power-loss replacement, the use of generating equipment and fuel to compensate for the transmission-system losses associated with power flows from generators to customers.

- Voltage control, the use of generating and transmission-system equipment to inject or absorb reactive power to maintain voltages on the transmission system within required ranges.

We also discussed several services that, after considerable consideration, we decided not to include in our list of key ancillary services. These functions include: black-start capability; time correction; standby service; planning reserve; redispatch; transmission services; power quality; and planning, engineering, and accounting services. We excluded these services either because they are very inexpensive, they are not required to maintain reliability or to complete a transaction, or they are highly site- and customer-specific.

ACKNOWLEDGMENTS

We thank Steven Naumann, Antonio Sammut, and Jerry Smith for their frequent and insightful advice during the course of this study. We thank Fernando Alvarado, Gregory Basheda, Paul Carrier, Kurt Conger, Patrick Connors, Raymond Coxe, Murray Davis, Dyke Farrow, Richard Kafka, Charles King, Laurence Kirsch, David LaPlante, Derek Mauzy, Alfred Mistr, Steven Naumann, Paul Rodzianko, Antonio Sammut, Jerry Smith, Carl Sunser, and Lawrence Worobec for their helpful reviews of a draft of this report. We thank Fred O'Hara for editing this report and Ethel Schorn for managing its clearance, printing, and distribution.

REFERENCES

F. L. Alvarado 1996, *Methods for the Quantification of Ancillary Services in Electric Power Systems*, draft, University of Wisconsin, Madison, WI, February.

Edison Electric Institute 1995, *Ancillary Services, FERC Open Access NOPR Technical Conference Report*, Washington, DC, November.

E. Hirst and B. Kirby 1995a, *Ancillary Services*, Comments to the Federal Energy Regulatory Commission in Docket Nos. RM 95-8-000 and RM 94-7-001, Oak Ridge National Laboratory, Oak Ridge, TN, June.

E. Hirst and B. Kirby 1995b, "The Devil *Is* in the Details," *The Electricity Journal* **8**(10), 12-18, December.

S. P. Hoffman and H. Illian 1996, "Control Performance Criteria," Commonwealth Edison, Lombard, IL, February.

W. W. Hogan 1995a, personal communication, Kennedy School of Government, Harvard University, Cambridge, MA, April.

W. W. Hogan 1995b, "A Wholesale Pool Spot Market Must Be Administered by the Independent System Operator: Avoiding the Separation Fallacy," *The Electricity Journal* **8**(10), 26-37, December.

Houston Lighting & Power 1995, Comments to the Federal Energy Regulatory Commission in Docket Nos. RM 95-8-000 and RM 94-7-001, Houston, TX, August.

IRP Report 1995, "Texas Plans to Step into Deregulation-But How Far?" **3**(10), 3-5, October.

B. Kirby, E. Hirst, and J. VanCoevering 1995, *Identification and Definition of Unbundled Electric Generation and Transmission Services*, ORNL/CON-415, Oak Ridge National Laboratory, Oak Ridge, TN, March.

B. Kirby and E. Hirst 1996, *Ancillary-Service Costs for Ten U.S. Electric Utilities*, ORNL/CON-427, draft, Oak Ridge National Laboratory, Oak Ridge, TN, January.

Michigan Public Service Commission 1995, *Opinion and Order After Remand*, Case Nos. U-10143 and U-10176, Lansing, MI, June 19.

Mid-American Interconnected Network 1995, *MAIN Guide No. 5A, Operating Reserve*, Lombard, IL, May.

New England Electric System 1995, *Choice: New England, The NEES Companies' Proposal for Restructuring the Electric-Utility Industry*, Westborough, MA, October.

New York Power Pool 1995, *New York Power Pool Presentation on Ancillary Services*, Prepared for FERC Dockets RM95-8-000 and RM94-7-001, Schenectady, NY, November.

North American Electric Reliability Council 1992, *Control Area Concepts and Obligations*, Princeton, NJ, July.

North American Electric Reliability Council 1995, *Interconnected Operations Services Reference Document*, Version 1.0, Princeton, NJ, November.

Northeast Utilities 1989, *Determination of Loss Factors for the Northeast Utilities System*, Hartford, CT, October.

Pacific Gas and Electric Company 1989, *Forecast of Electric Operations, Appendix B: Electric System Operating Constraints Report*, Before the California Public Utilities Commission, Docket No. 89-04-001, April 3.

Pennsylvania-New Jersey-Maryland Interconnection Association 1995, *Supplemental Comments*, Submitted to the Federal Energy Regulatory Commission, Docket Nos. RM95-8-000, and RM94-7-001, Norristown, PA, November.

Putnam, Hayes & Bartlett, Inc. 1995, *Above-Market Cost Estimation in the Competitive Opportunities Proceeding*, Energy Association of New York, Albany, NY, October.

Texas Public Utility Commission 1995, "Chapter 23 Substantive Rules," *Texas Register* **20**, 9520-9527, Austin, TX, November 17.

D. Toulson and M. Mauldin 1995, *Winning Retail Strategies: Beyond Innovative Rate Design*, EPRI TR-105226, Electric Power Research Institute, Palo Alto, CA, June.

U.S. Federal Energy Regulatory Commission 1995, *Promoting Wholesale Competition Through Open Access Non-Discriminatory Transmission Services by Public Utilities, Recovery of Stranded Costs by Public Utilities and Transmitting Utilities*, Docket Nos. RM95-8-000 and RM94-7-001, Washington, DC, March 29.

R. A. Wakefield, H. D. Limmer, D. J. LeKang, and J. S. Graves 1995, *Transmission Services Costing Framework, Volume 1: Technical and Economic Fundamentals* and *Volume 2: Framework Description and Application*, EPRI TR-105121, Electric Power Research Institute, Palo Alto, CA, April.

K. N. Zadeh and R. C. Meyer 1995, *Proceedings: Forum on Ancillary Services*, EPRI TR-105686, Electric Power Research Institute, Palo Alto, CA, October.

ALTERNATIVE SETS OF ANCILLARY SERVICES

Here we summarize the definitions and classifications of ancillary services developed by FERC, ORNL, Houston Lighting & Power, the New York Power Pool, the Michigan Public Service Commission, and the North American Electric Reliability Council.

FEDERAL ENERGY REGULATORY COMMISSION

FERC's March 1995 NOPR included six ancillary services. FERC divided these services into three categories: "actions taken to effect the transaction (such as scheduling and dispatching services) ... , services that are necessary to maintain the integrity [reliability] of the transmission system (such as load following, reactive power support, and system protection services) ... , [and] services needed to correct for the effects associated with undertaking a transaction (such as loss compensation and energy imbalance services)."

FERC combined scheduling and dispatching into one service. It defined scheduling as the "pre-determined (before-the-fact) use of generation and transmission facilities to meet anticipated load ... ." FERC's definition of dispatching is "the control room operation of all generation and transmission resources and transmission facilities on a real-time basis to meet load within the transmission provider's designated service area ... ."

Scheduling could be expanded to include unit commitment, the *a priori* decisions on which generating units to run during certain time periods. On a weekly basis, the utility forecasts expected load hour-by-hour and then schedules generation to supply that expected load. This procedure is necessary because some units require several hours to be brought online. The utility selects from among the generation resources available to obtain the lowest-cost mix of generation resources. Unit commitment is constrained by unit startup costs, startup times, minimum operating loads, minimum operating times, etc. It can also be constrained by transmission limitations, reliability requirements, emissions limits, and restrictions on the operation of hydro units. Decisions are made as to which units should be used throughout the week, which units should be kept warm for possible use, and which units can be turned completely off.

The weekly schedule is usually reexamined daily to determine which units to bring on- and offline to serve the expected load for the following day. This process primarily accommodates an updated forecast of the following day's load (based largely on a better weather forecast). It also accommodates changes in unit availability (caused by forced outages or early returns to service). Changes in the external power market can also influence the unit commitment if someone unexpectedly has power for sale at a low price or may be willing to pay a high price for additional power.

Dispatching could be expanded to include economic dispatch, the continuous real-time decision-making function in which the system operator, given the actual mix of generating units and power purchase/sell opportunities, attempts to meet current customer demands at the lowest variable cost while meeting NERC reliability requirements, emissions restrictions, and the terms of customer and interutility contracts.

FERC defined load following, its second service, as the “continuous balancing of resources with load under the control of the transmission provider ... accomplished by increasing or decreasing the output of on-line generation ... to match moment-to-moment load changes.” FERC suggests that the transmission provider (i.e., control-area operator) “may be uniquely positioned to provide load following reserve.”

FERC’s third service was system protection, defined as “operating reserves or other system protection facilities available in order to maintain the integrity of its transmission facilities [caused by] unscheduled [transmission or generation] outages.” This service, which many people call operating reserves, is in some ways the supply side analog of load following. Both load following and system protection use online generating resources to respond to discrepancies between loads and generation. Load following is used to respond to the many small changes in load, while system protection is used to respond to the fewer but much larger changes in supply.

FERC’s fourth service was loss compensation, defined as replacement of the “capacity and energy losses [that] occur when a transmission provider delivers electricity across its transmission facilities for a transmission customer.” Although these losses are caused by the transmission system, they are replaced by generation. Losses vary with time and location as a function of the transmission network’s configuration, the location and output of generators, and the location and demand of customers. The cost to replace losses also varies with time.

FERC defined energy imbalance, its fifth service, as the “difference [that] occurs between the hourly scheduled amount and the hourly metered (actual delivered) amount associated with a transaction. Within a deadband of ±1.5% of the scheduled transaction, FERC proposed that the energy could be returned in kind within 30 days. Imbalances outside the deadband would be subject to financial charges. FERC was unclear about possible differences in charges for undergeneration (in which the customer takes energy and power from the transmission provider) and compensation for overgeneration (in which the customer delivers energy and power to the transmission provider).

The sixth and final service that FERC identified was reactive-power/voltage control, defined as the “reactive power support necessary to maintain transmission voltages within limits that are generally accepted in the region and consistently adhered to by the transmission provider." FERC recognized that reactive power can be produced and absorbed by both generators and transmission-system equipment.

OAK RIDGE NATIONAL LABORATORY

Our March 1995 report (Kirby, Hirst, and VanCoevering 1995) identified 19 ancillary services. The 11 generation-related services include unit commitment, economic dispatch, load-following service, reliability-spinning reserve, supplemental-operating reserve, unscheduled energy, time correction, stability enhancement, local-area security, nonoperating reserve, and black-start capability. We identified eight transmission-related services, including transmission monitoring and control; real-power loss replacement; system reactive power management and voltage regulation; local reactive power management and voltage regulation; transmission reserves; repair and maintenance of the transmission network; metering, billing, and communications; and power quality.

FERC's six services correspond closely to 12 of the 19 services we identified (Table 2); see Hirst and Kirby (1995a).

FERC's reactive-power/voltage control service appears to encompass both system and local functions. We split this service into two pieces in the belief that only the system operator can have sufficient information about voltages throughout the transmission network to determine whether, where, and when reactive power must be injected or extracted from the system to maintain system voltages within their required ranges. On the other hand, customers should have sufficient information to maintain local voltage control and to meet their own reactive-power needs. Thus, we consider it possible to split this service into two separate pieces, only the first of which must be under the control of the system operator. Also, our definition implicitly accepts the possibility that a customer or third party could supply transmission- or distribution-related reactive power as well as generation-related reactive-power/voltage control support.

The FERC and ORNL definitions of real losses are the same.

FERC's definition of scheduling and dispatching services are equivalent to our definitions of unit commitment and economic dispatch.

The FERC and ORNL definitions of load-following services are similar, although, as discussed in Chapter 2, our view of load following is more expansive than FERC's.

FERC's system protection service includes system-operator control of any resources (both generation and transmission) needed to maintain the stability and integrity of the interconnected electric network. This single service includes several of the ORNL services, including reliability-spinning reserve (fully available within 10 minutes), supplemental operating reserve (fully available within 30 minutes), stability-enhancement reserve (equipment at generating units and in transmission systems used to maintain dynamic stability of the electric system), local-area security reserve (operation of generating units out of merit order to respond to transmission system constraints), transmission reserve, and reliability spinning reserve. All these reserves are used to maintain system stability and integrity.

Table 2. Comparison of FERC and ORNL classifications of ancillary services

FERC	ORNL
Reactive power/voltage control	System reactive-power management and voltage control Local reactive-power management and voltage control
Loss compensation	Real-power-loss replacement
Scheduling and dispatching	Unit commitment Economic dispatch
Load following	Load-following service
System protection	Reliability spinning reserve Supplemental operating reserve Stability-enhancement reserve Local-area security Transmission reserves
Energy imbalance	Unscheduled energy
Services not identified by FERC	Time correction; nonoperating reserve; black start; transmission monitoring and control; repair and maintenance of network; metering, billing, and communications; power quality

FERC's energy imbalance is equivalent to our unscheduled energy service.

We identified several ancillary services not included in FERC's list. Time correction, control actions to ensure that the electric-system frequency averages 60 Hz, is a very small, but important, service. In a competitive electricity industry, nonoperating reserves may not be considered a service at all. Rather, these generating units may be the "inventory" that a generating company uses to meet future contingencies. Black start, the capability of generating units to restart themselves after shutdown without taking power from the grid, is an important service that the system operator should purchase from some of the generating units in its control area.

We identified transmission monitoring and control as well as repair and maintenance of the network as two separate services. We surmise that FERC considered these to be part of the basic transmission service, the costs of which are included in the tariffs for network-integration service or point-to-point service. Provision of appropriate levels of power quality to individual customers is clearly a service, which we surmise FERC wants to treat outside the scope of its transmission considerations. Presumably, maintenance of sine-wave-quality electricity on the transmission grid is considered part of the basic transmission service to be achieved through engineering standards.

HOUSTON LIGHTING & POWER

Houston Lighting & Power (1995), as part of a case before the Texas Public Utility Commission, developed a set of 60 ancillary services, which it subsequently reduced to 20 in number. The company categorized these services on the basis of the number of providers for each. The first category includes services that can only be provided by an entity in a specific location (e.g., voltage support, tie-line telemetry, and generation redispatch). The second category includes services that can only be provided by a control area (but not necessarily the local control area); through appropriate telemetry and control equipment, any control area could provide services such as energy imbalance and generator scheduling. The third category includes services that can be provided by any party willing and able to do so (i.e., those connected to the grid with the ability to affect the supply of or demand for power). Table 3 lists the HL&P services, shows how the company classified them, and indicates whether the service is optional or required.

HL&P also developed definitions for each of its services. As was true for the ORNL services, there is considerable overlap between the HL&P services and the six defined by FERC (Table 4); HL&P also identified several services not included in FERC's six.

Based in part on the HL&P set of services, the Texas Public Utility Commission (1995) proposed a rule that included seven ancillary services, the six that FERC proposed plus "reporting and accounting." Subsequently, the commission expanded its list to 11 services: responsive reserve, spinning reserve, generation energy imbalance, load energy imbalance, scheduled backup, automatic backup, emergency energy, load following, load regulation, static scheduling, and dynamic scheduling.

NEW YORK POWER POOL

The New York Power Pool (1995) defined ancillary services as "the physical equipment and human resources that are essential to support transmission service across a bulk power system, while maintaining the system's reliability as required by local and regional criteria." NYPP identified five sources of ancillary services: generation equipment, control centers, people, transmission facilities, and distribution facilities.

NYPP's system of classification identified five types of services provided by generating equipment:

Table 3. Houston Lighting & Power classification of ancillary services

Service	Who can provide? Control area	Other party[a]	Optional or required?
Generation reserves			
1. Responsive reserve	Can	Can	R
2. Spinning reserve	Can	Can	O
3. Planning reserve	Can	Can	R
Energy/capacity			
4. Losses	Can	Can	R
5. Generation energy imbalance	Must[b]		R
6. Load-energy imbalance	Must[c]		R
7. Inadvertent energy payback	Can	Can	R
8. Scheduled backup	Can	Can	O
9. Automatic backup	Can	Can	R
10. Emergency energy	Can	Can	O
11. Voltage/VAR support[d]	Must[c]		R
Regulation/control			
12. Load following	Can	Can	O
13. Load regulation	Can	Can	O
Generation scheduling			
14. Static scheduling	Must		Either one required
15. Dynamic scheduling	Must		
16. Short-notice scheduling	Must		O
17. Redispatch[d]	Can	Can	O
Transportation			
18. Wheeling	Must		R
Metering/accounting			
19. Tie-line telemetry[d]	Must		O
20. ERCOT reporting	Can	Can	R

Source: Houston Lighting & Power (1995).

[a]Other parties can include the transmission owner, a generator, the load customer, or another party.

[b]Only the supply control area can provide these services.

[c]Only the load control area can provide these services.

[d]These services cannot be provided competitively.

Table 4. Comparison of FERC and HL&P classifications of ancillary services

FERC	HL&P
Reactive power/voltage control	Voltage/VAR support
Loss compensation	Losses
Scheduling and dispatching	Static scheduling, dynamic scheduling, short-notice scheduling, redispatch
Load following	Load following, load regulation
System protection	Responsive reserve, spinning reserve, automatic backup, emergency energy
Energy imbalance	Generation energy imbalance, load-energy imbalance, inadvertent energy payback
Services not identified by FERC	Planning reserve, scheduled backup, wheeling, tie-line telemetry, ERCOT reporting

- voltage services (voltage support and voltage control);
- regulation (frequency control and tie-line regulation, load following, energy imbalance, and deadband protection);
- reserves (10-minute operating, 10- to 30-minute operating, installed, standby, and black start);
- transmission (support, flow balancing, and losses); and
- power quality (dynamic reserve and dynamic backup demand).

Altogether, NYPP listed 16 generator ancillary services.

NYPP listed three classes of services provided by control centers:

- real-time system-security management (tie-line and frequency regulation, inadvertent energy management, time-error management, interchange scheduling implementation, system energy management, administration of intercontrol-area energy transactions, load shedding, transmission-system operation, online generation dispatch, communications, and system restoration);

- capacity management (installed capacity criteria, online capacity management, operating-reserve management, operating-reserve scheduling, generator-outage scheduling, transmission-outage scheduling, generator-outage coordination, transmission-outage coordination); and

- administration [engineering services, operations planning (including load forecasting, unit commitment, pumped-storage scheduling, and contract scheduling), training, administrative charges for transmission agreements, and administrative functions].

Altogether, NYPP listed 23 control-center ancillary services.

In addition to defining each of these 39 services, NYPP provided draft answers to the following questions for each service: Is the service required, optional, or location-dependent? Who can provide the service under the present system and with open access, the local utility, the power pool, or a third party? Can/should the service be unbundled? Should the price for the service, in the short run and in the long run, be based on embedded costs, on location-specific marginal costs, or on the market?

MICHIGAN PUBLIC SERVICE COMMISSION

As part of its retail-wheeling experiment, the Michigan Public Service Commission (1995) adopted tariffs for ancillary services for Consumers Power and Detroit Edison. Its decision was based on extensive testimony from both utilities, the Commission staff, industrial companies, and other intervenors.

The Commission considered operating reserves, voltage control, and reactive support to be mandatory services that all retail-wheeling customers would have to purchase from the local utility. Optional services included line-loss replacement, deadband service, standby service, and an unauthorized-use charge. The nomenclature and tariff specifics differed between the two utilities.

Operating reserves cover regulating margin, fast-start reserves, and supplemental operating reserves. The Commission estimates that, in aggregate, these generating reserves will total about 6% of a utility's peak demand. It approved charges for reactive power ($0.11/kW-month for Consumers Power and $0.05/kVAR-month for Detroit Edison) but none for voltage support and control. Although not listed as either mandatory or optional, the Commission allowed a charge for system control and load dispatch in its discussion of mandatory services.

The Commission allowed both demand and energy charges for transmission-system losses. The charges vary with the voltage level at which power is taken. For example, the assigned losses are 1.9% at the 345- or 138-kV level and 3.7% at the subtransmission level for Consumers Power.

Deadband protection, standby service, and unauthorized use are three services that deal with situations in which a customer's actual transaction does not match its scheduled transactions because of differences in supply, consumption, or both. Deadband protection is intended to allow for temporary fluctuations in customer load above and below the scheduled level. The Commission approved a 10-minute ±3% deadband period (compared with FERC's 60-minute, ±1.5% deadband for energy imbalance), beyond which any imbalance would incur a load-regulation charge.

Standby service refers to a prior arrangement to provide generating capacity in case a customer's supplier experiences an outage. Failure to purchase standby service from either the local utility or another provider could invoke the unauthorized-use charge. The Commission decided to use existing standby-service tariffs for this retail-wheeling experiment. It also decided that "unauthorized use charges should serve as a weighty deterrent to using utility power outside of the parameters of the experiment." In other words, failure to acquire insurance (i.e., standby power) could result in substantial charges, $50/kW-month demand plus the greater of 5¢/kWh or incremental energy costs plus 1¢/kWh.

NORTH AMERICAN ELECTRIC RELIABILITY COUNCIL

NERC (1995) divides its interconnected operations services into three groups: interchange scheduling, generation, and transmission. Under scheduling, NERC includes control and dispatch, dynamic scheduling, and accounting. Under generation, NERC includes eight services: regulation, imbalance, frequency response, backup, reserves, reactive, losses, and restoration. Finally, under transmission, NERC includes reactive power and facilities.

Interchange scheduling services can be provided only by the local control area. System control and dispatch includes "interchange scheduling confirmation and implementation with other control areas ... ." Dynamic scheduling includes "the metering, telemetering, computer software, hardware, communications, engineering, and administration required to electronically move a [transaction's] generation or load out of the control area to which it is physically connected and into a different control area." Accounting services are just what the name implies.

The generation services can be supplied either by a control area or by another generation supplier. Regulation is the "real-time generation control that follows ... real-time variations in load ... . It is the generating capacity that is used to follow the greatest change in this load (the difference between the maximum and minimum) during the specified period." Energy imbalance, according to NERC, involves provision of energy, but not capacity, "to correct for the mismatch between ... supplied energy and the load ... ." The distinction between capacity to follow load variations and energy to balance may reflect a NERC view that because energy imbalance is an after-the-fact accounting, it does not include any capacity reservation.

Frequency response is the real-time use of generating resources to respond to frequency deviations in the interconnection. Changes in either loads (compensated for with regulation) or generation (compensated for with operating reserves) can affect frequency. Together, frequency response, regulation, and operating reserves are probably the same as FERC's load-following and system-protection services.

Backup supply provides resources to replace the loss of a customer's generation supply or to cover that portion of a customer's load that exceeds its generation supply. This is a long-term service (akin to insurance) that is arranged beforehand.

Operating reserves, the same as FERC's system protection, include spinning plus supplemental reserves used to respond to generation or transmission outages and to meet reliability-council requirements.

Real-power-loss service is the same as FERC's loss compensation.

Reactive supply and voltage control services are the generation part of FERC's reactive-power and voltage-control service. The remainder of FERC's service is covered by NERC's transmission provision of these services. Thus, NERC explicitly recognizes that voltages are maintained within limits throughout the transmission network with a combination of generating and transmission equipment.

Restoration is equivalent to what we call black-start capability, which refers to the ability of a generating unit to restart without the use of electricity from the grid.

The two transmission services that NERC identifies are facilities use (the basic use of the transmission network to transport power from a supplier to a customer) and reactive power.