

# Electricity Market Design: An Integrated Approach to Reliability Assurance

JOE H. CHOW, FELLOW, IEEE, ROBERT W. DE MELLO, MEMBER, IEEE, AND  
KWOK W. CHEUNG, SENIOR MEMBER, IEEE

## *Invited Paper*

*System reliability is an integral part of a properly designed deregulated electricity market, even though wholesale energy prices are its most visible piece. This paper discusses the resources, collectively known as the ancillary services, in a power system for maintaining reliability. It then discusses an integrated approach in systematically determining and securing the needed ancillary services and system configuration across multiple time spans, culminating in the cooptimization of energy and ancillary services for security-constrained generator unit commitment in the day-ahead market and security-constrained economic dispatch of the generators in real-time operation. The paper also provides a summary of the ancillary services market in several U.S. power markets and proposes several technical issues for further discussion and research.*

**Keywords**—Ancillary services, deregulated electricity market, multiple settlement system, reliability, security-constrained economic dispatch (SCED), security-constrained unit commitment (SCUC).

## I. INTRODUCTION

The responsibility of ensuring the reliability of a control area within an interconnected power system is shared by all the operating entities inside the power system. In a traditional regulated power system, the responsibility for assuring reliability is shared by the vertically integrated utilities. The reliability criteria are established by a committee consisting of all the interconnected utilities in the control area, and a system operating organization responsible for the overall power dispatch in the control area, commonly known as a

power pool, is in charge of enforcing the reliability rules. Most of the operating entities also belong to a larger regional reliability council, such as the Northeast Power Coordinating Council [1], so that neighboring control areas can operate with similar reliability criteria. The regional reliability councils form the North American Electric Reliability Council (NERC) [2], which establishes recommended standards on system reliability.

The power industry has been evolving toward a market-based approach in the United States and throughout the world for well over a decade [3], [4]. In addition to active long-term wholesale bilateral markets that exist in all regions of the United States, short-term wholesale markets with transparent prices and market structures that consistently produce just and reasonable prices of energy and ancillary services are crucial to the success of electricity markets. In a restructured or deregulated electricity market, a traditional utility is separated into three entities: generator companies (GENCOs), transmission owners (TOs), and load serving entities (LSEs). Along with other nonasset owners such as energy traders, they are collectively known as the market participants (MPs). The responsibility of ensuring the reliability of a control area is delegated to an independent system operator (ISO) or a regional transmission organization (RTO). In general, the MPs have responsibility for providing accurate data, certifying the performance of their equipment, and following the dispatch requested by the ISO/RTO. The ISO/RTO has the responsibility of ensuring that each MP meets its reliability rules and coordinating the dispatch of the electricity supply to meet the demand, such that the power system will also meet the operational reliability rules at the lowest possible cost. NERC has already included ISOs and MPs in its governance structure, with the traditional reliability criteria updated appropriately to meet the needs of restructured electricity markets.

A key premise of a successful competitive market is that the market works through the interaction of private, decentralized trading and investment decisions [5]. An

Manuscript received October 1, 2004; revised June 1, 2005. The work of J. Chow is supported in part by National Science Foundation under Grant 0300025.

J. H. Chow is with Rensselaer Polytechnic Institute, Troy, NY 12180 USA (e-mail: chowj@rpi.edu).

R. W. de Mello is with the New York Independent System Operator, Schenectady, NY 12303 USA (e-mail: rdemello@nyiso.com).

K. W. Cheung is with AREVA T&D, Bellevue, WA 98004 USA (e-mail: kwok.cheung@areva-td.com).

Digital Object Identifier 10.1109/JPROC.2005.857493

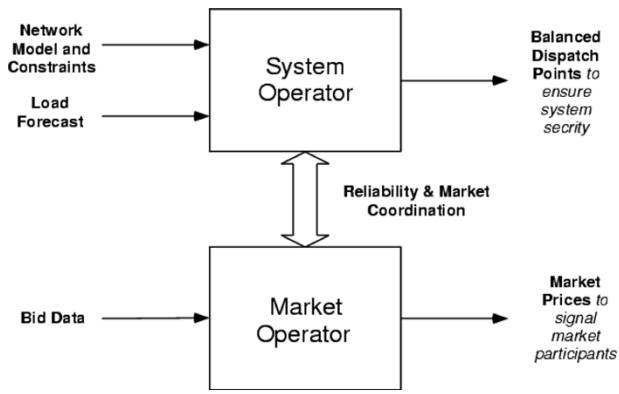


Fig. 1. System operator and market operator as separate entities.

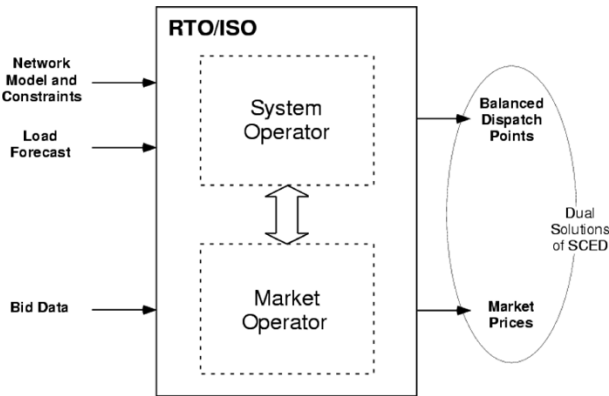


Fig. 2. Dual functions of RTO/ISO and dual solutions of SCED.

effective electricity market should allow substantial commercial freedom to market buyers, sellers, and various types of traders. Trading rules would then allow the market participants the freedom to fashion and implement various trading and risk management arrangements with each other, at prices to which they mutually agree, in pursuit of their respective commercial interests. Market design should, on the other hand, recognize the fact that the laws of physics dictate certain essential characteristics of system operations and the complexities of electricity networks require a degree of centralized coordination over system operations to ensure system reliability. In some power markets, the system operator, whole role is to maintain reliability, and the market operator, whole role is to settle supply and demand, are separate entities. This type of market structure requires a substantial amount of coordination between the two operators to be successful (Fig. 1). This is sometimes referred to as the unbundled approach [6], which was used in California (1998–2000). However, an electricity market in which the ISO or RTO functions both as the “system operator” for reliability coordination and the “market operator” for establishing market prices is more likely to succeed in allowing commercial freedom and centralized economic and reliability coordination to coexist harmoniously (Fig. 2). Within this framework there are a number of variations, such as the practice in Australia, and New Zealand, which do not procure reliability ahead of time and only run real-time markets, and other practices like in Texas which clear energy and ancillary services sequentially. This paper focuses on

the integrated approach, which has been adopted by most of the U.S. power markets, including NYISO [7], PJM [8], ISO-NE [9], and the new California market [10], in which reliability is addressed at many different levels and energy and ancillary services are optimized simultaneously [11], [12].<sup>1</sup>

For an ISO using an integrated approach, ensuring reliability of a power system is a multistage process. The part that has an immediate impact on a customer is the real-time dispatch (RTD). Less visible are the system planning and scheduling that prepare a power system for reliable operation. The envelope of power system reliability can encompass five major progressions starting from many years in the future, as shown in Fig. 3 [13].

Developing the electricity supply to meet future load growth needs to start at least five years ahead of time because of the long lead times required in the construction of generation and transmission equipment. Then, going forward three to six months, an ISO needs to secure the generating units and transmission facilities that are available to participate in the dispatch. Then, one to two weeks prior to the actual operation, an ISO will determine whether any unanticipated but deferrable maintenance on some equipment can be carried out. A day before operation, using the load bids submitted by the LSEs and the load forecast, an ISO will perform a day-ahead security-constrained unit commitment (SCUC) to determine the hourly setpoints of the generators based on the bid-in energy prices to serve the loads. In real time, an ISO needs to reconcile the differences between the committed load and the actual load on 5–15-min intervals with a 1–3-h look-ahead horizon. This requires a control room operator to execute a real-time security-constrained economic dispatch (SCED) to adjust the day-ahead generator dispatch schedules and perhaps solicit additional supply to ensure that the system is secure.

The stages in Fig. 3 can be applied to both regulated and deregulated electricity markets. The analysis techniques and software are similar, but the inputs and the decision processes are very different. For example, in long-term planning, a regulated utility would include the rate of return set by the public utility commission when considering generation and transmission expansion, which are considered as capital investment projects. In a deregulated market, however, an MP needs to request an ISO to perform studies to determine whether a new generating or transmission facility expansion plan is technically feasible. If the project is technically feasible, then the MP will evaluate the investment opportunity, based on historical and predicted energy prices in peak and off-peak hours, and arrange for project financing. There are typically many ongoing potential projects under study by the ISO, with only a small fraction of the projects eventually coming to fruition. Although the process of planning and the return on investment differ between regulated and deregulated markets, the analytical tools used for the studies belong to the so-called production planning software [14], [15].

<sup>1</sup>We must emphasize that although energy and ancillary services are optimized simultaneously in these markets, there are still regional differences in the implementations.

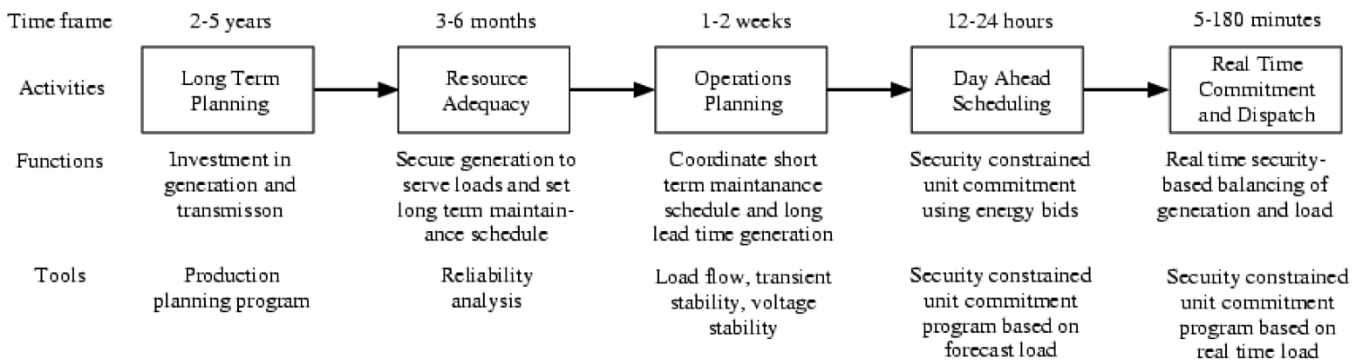


Fig. 3. Security functions in different time frames and power system reliability functions.

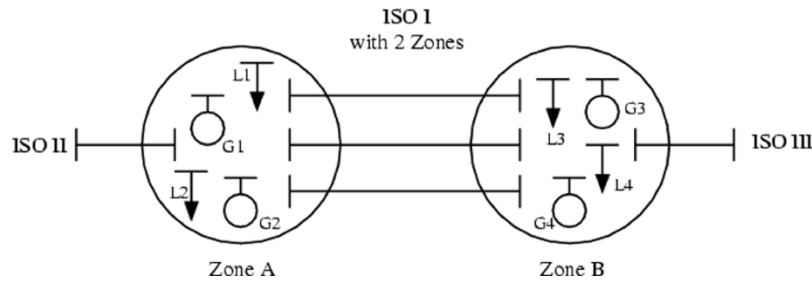


Fig. 4. ISO with two zones and showing connections to other ISOs.

This paper focuses on four ancillary services to meet the needs of the reliability criteria in four time horizons. Ancillary services include regulation, reserve (both spinning and nonsynchronous), voltage support, and black start capability. The time horizons cover those of resource adequacy, operations planning, day-ahead scheduling, and RTD. In some discussions, we will use an example of an ISO with two zones (Fig. 4) for illustration. The zones, each represented by two generators and two loads, are connected by several parallel transmission paths. During peak load hours, these transmission paths are congested, and are constrained to certain precomputed limits. Note that in a practical power system, depending on the load patterns, different sets of transmission paths can be congested, resulting in the zones having different boundaries. In Fig. 4, we also show the connections of the ISO (called ISO I) to other ISOs, namely, ISO II and ISO III.

The paper is organized as follows. In Section II, some of the NERC reliability criteria relevant to this paper are summarized. The other sections are organized according to the time frames in Fig. 3. The availability of resources for maintaining reliability is addressed in Section III, including a summary of the current practices of several ISOs. Generator operating reserves and outage coordination are discussed in Section IV. Section V discusses a multiple settlement system for the day-ahead market (DAM) and the RTD. Several potential investigations for improving reliability are presented in Section VI. A mathematical formulation of a SCED problem is given in the Appendix.

## II. RELIABILITY CRITERIA

Reliability criteria, as detailed in NERC documents [2] and commonly practiced in power systems, address both

technical and organizational aspects of reliability in planning and operation. The technical aspects include control area performance, transmission systems, system coordination between control areas, and operations planning. The organizational aspects include procedures for coordinating reliability within a control area and between control areas, communication of data, and operation staff training and certification. In this paper, we will focus mostly on the technical aspects of the reliability criteria.

### A. Control Area Generation Control and Performance

Each control area needs to be able to regulate generation to meet the daily variations of load demand. This generation regulation has several aspects.

- 1) *Frequency response*: All generating units above a certain size must be equipped with a governor capable of responding to system frequency deviation due to load ramps and generator trips. Governors typically provide a 5% droop and should be responsive to frequency deviations outside of a prespecified band.
- 2) *Area control error (ACE)*: ACE [15] represents the shift in the generation in the control area required to restore frequency and the net interchange to its desired value, and is given by

$$ACE = -\Delta P_{\text{net int}} - 10B\Delta f \quad (1)$$

where  $\Delta P_{\text{net int}}$  is the deviation in megawatts of the interchange from the desired value,  $\Delta f$  is the frequency deviation in hertz, and  $B$  (in megawatts/0.1 Hz) is the bias set as close as possible to the control area's frequency response.

- 3) *Automatic generator control (AGC)*: AGC is typically installed at a central location, such as the power grid operator, as a means of coordinating the generation available in a control area to restore ACE to zero. The control logic aims at driving the overall ACE as well as individual unit generation deviation to zero. Generation set-points from AGC calculation are typically issued every 2–6 s.
- 4) *SCED*: Like AGC, SCED is installed in a central location and coordinates generation. SCED's control logic, however, aims at determining the generation schedules that will securely and most economically meet load. SCED typically has a cycle time of 5 min.
- 5) *Generation reserves*: A control area needs to maintain sufficient amount of generating capacity on spinning and standby, allowing the AGC to reduce the ACE to an acceptable value in case of single or double contingencies.

#### B. Transmission System Performance

Each control area should have a reliable transmission system that can recover from single or multiple contingencies without losing stability, uncontrolled separation, and cascading outages.

- 1) *Single and multiple outages*: The power system shall survive through the most severe single contingency and credible multiple contingencies.
- 2) *Voltage and reactive control*: Each control area shall implement voltage control strategies to maintain acceptable voltage levels under normal operating conditions and contingency conditions. It needs to have adequate reactive resources to provide the required voltage support. The reactive resources need to be location specific because the effect of reactive support tends to be local. The system operator shall take corrective actions, including load shedding, to prevent voltage collapse when reactive resources are insufficient.
- 3) *Monitoring*: The system flows and voltages should be monitored by the grid operator and the TOs to ensure that no operating limits, such as stability and thermal limits, are violated.
- 4) *Protection*: Protective systems, including relays, should be of minimum complexity to achieve the desired objectives and should not mis-operate on unrelated events. Redundancy should be built into the protective systems to provide additional system security.
- 5) *Maintenance*: Maintenance of all equipment must be performed at regular intervals. The maintenance needs to be scheduled so that system reliability will not be compromised.

#### C. Operating Planning

The operations planning function in a power control center needs to have a set of current plans to evaluate all its options and set procedures for reliable operation through a reasonable future time period. These operational plans can be separated into four regimes [16]–[18].

- 1) *Normal operations*: Bulk power system studies need to be performed to coordinate the power system operation for the current day, the next day, and seasonal

variations. The studies should include planned and unplanned changes in system configuration, generation dispatch, interchange scheduling, and demand changes. The unplanned changes should include the most severe single contingencies and all credible multiple contingency cases. The results will determine the capacity and energy reserve requirements, voltage and reactive support limits, and total transfer capability.

- 2) *Emergency operations*: The system operator needs to have plans to deal with: 1) insufficient generating capacity; 2) transmission system overloaded beyond thermal and/or stability (transient/voltage) limits; 3) the proper use of load shedding to prevent frequency decay; and 4) system restoration after a power disruption has occurred.
- 3) *Load shedding*: Each operating center shall develop, in coordination with neighboring control areas, plans to perform automatic load shedding, to minimize the risk of uncontrolled separation and system shutdown. The load shedding is typically performed based on frequency and voltage levels, with activation occurring in multiple steps. An operator also should have the means to manually activate load shedding in any appropriate time frame.
- 4) *System restoration*: Each system operator, in coordination with the transmission and distribution operators, needs to be able to reestablish the electric power system after a system disruption. In case of tie-lines between control areas, the coordination extends to the neighboring control areas. An integral part of the restoration plan is a quick and reliable source of generators, typically hydro-turbine and diesel units, to provide start-up power.

#### D. Reliability Coordination

Each system operating entity should have a reliability coordinator with broad responsibility to oversee the compliance of the entity with NERC reliability criteria. The reliability coordinator is also responsible for working with reliability coordinators from other control areas to ensure the reliable operation of the interconnected system.

#### E. Standard Market Design (SMD)

The FERC's SMD [19] envisions an independent operator (ISO or RTO) that acts as reliability coordinator, system operator, and market operator. Other than an independent operator, some key components of SMD related to reliability include the following.

- A voluntary short-term spot market for energy based on competitive bids and offers with transparent prices. The spot market would include both a day-ahead function to coordinate generator starts and stops (unit commitment) and a real-time balancing function.
- Locational marginal pricing (LMP) of energy that assigns a value to transmission constraints.
- Markets for certain ancillary services.
- Resource adequacy requirements.
- Regional transmission planning.

### III. RESOURCE ADEQUACY

A distinction can be made between contingency reserves and planning reserves [20]. The former, which includes spinning reserves, nonsynchronous reserves, etc., must typically be available with a lead time of 30 min or less. Contingency reserves relate to short-term security and are typically the concern of day-ahead planning and real-time operation. Contingency reserves provide for reliable operation right now, or tomorrow. Planning reserves, on the other hand, are designed to avoid difficulties that might result from generator outages, whether planned or forced, or from load growth. Planning reserves take a longer term view of system reliability, to provide for reliable operation over the next month, months, year, or years. Day-ahead and real-time electricity markets deal with contingency reserve, while resource adequacy requirements and outage scheduling procedures target planning reserves. The analytic tools associated with the operation of the electric power system and contingency reserve (unit commitment, economic dispatch, security analysis, etc.) become secondary to economic and policy considerations as the period of time in question grows from days to months or years. In short, resource adequacy generally considers only the ability of resources to provide energy. The ability of resources to provide ancillary services is generally outside the scope of resource adequacy consideration.<sup>2</sup>

#### A. Generation Resource Adequacy

Both resource adequacy requirements and coordinated outage scheduling are used to ensure adequate generation capacity in the near future. Many variations exist, and some feel resource adequacy requirements are unnecessary. The resource adequacy rules described below are characteristic of the electricity markets of the northeastern United States [21]–[23]. The Australian electricity market [24] lacks a resource adequacy requirement and is the basis for the discussion of energy-only markets.

Resource adequacy requirements attempt to establish the economic conditions that encourage investment by suppliers, where “suppliers” encompass both generating resources and demand reduction resources. Day-ahead and real-time (spot) energy and ancillary service prices tend to reflect only suppliers’ short run marginal costs and, at times, scarcity rents. These may not represent suppliers’ total costs. Hence, reliance solely on spot markets for compensation may not provide suppliers with sufficient incentive for investment. The markets that are concerned with resource adequacy requirements are intended to better reflect suppliers’ long-term marginal costs. The energy-only market design, a design without a capacity market, relies solely on scarcity rents to provide sufficient incentive for new investment. While all energy markets incorporate some sort of upper limit (cap) on the offer price of energy, the energy-only markets typically have a higher cap than other designs to allow for higher scarcity rents. Nevertheless, even the energy-only market may place a cap on the amount of scarcity rent that can be

<sup>2</sup>ISO-NE has defined a requirement for capacity that can be started in 10 min or less.

collected. The key aspects of programs that implement resource adequacy requirements are: 1) capacity requirements; 2) bilateral market; 3) market-based auctions to augment the bilateral markets; 4) a market-based balancing auction; 5) supplier obligations to be available; and 6) market administration and oversight.

Capacity requirements may be determined by an independent agent, such as an ISO or reliability council. Requirements are generally based on a predetermined methodology such as the one day in ten years loss of load probability. The capacity requirement recognizes that forecasts of load growth are uncertain and that electrical generators have nonzero outage rates. A typical capacity requirement might be 118% of predicted peak load. LSEs must acquire capacity to cover peak load plus the previously determined margin. The duration of the capacity requirement varies; durations as long as five years have been proposed, but a one-month duration is more typical. Capacity requirements may also be locational.

The need of a locational requirement can be illustrated with the example of Fig. 4. Suppose that peak load in each of Zone A and Zone B is 1000 MW with only 500 MW of transfer capability between the two zones. Suppose further that Zone A has a glut of inexpensive generation and could, if the transmission capacity were available, supply the peak load of both Zone A and Zone B with a comfortable margin left over. The locational requirement that LSEs in Zone B acquire at least 50% of their installed capacity locally would guarantee that the zone’s peak load can be served. The locational requirement recognizes the inherent limitation in transfer capability between the two zones. A similar locational requirement could be imposed on Zone A. The requirement for LSEs in Zone A is unnecessary because we have supposed that the zone contains a glut of inexpensive generation, and, it would make little sense for an LSE in Zone A to acquire the relatively more expensive capacity of Zone B.

Each LSE must acquire capacity to meet its obligation. Generally, capacity can be acquired bilaterally or through an auction. A typical auction matches bids to purchase capacity with offers to sell capacity. The auction determines the amount of each bid that is accepted and the amount of each offer that is accepted. Clearing prices are also determined. The selection is done to maximize the social welfare, that is, the quantity (benefit minus cost), where:

- benefit is the sum of accepted bids to purchase times the bid price;
- cost is the sum of accepted offers to sell times the offer price.

When posed as a constrained optimization problem the following constraints must be observed.

- The accepted amount of each offer to sell must be between zero and maximum amount in the offer.
- The accepted amount of each bid to purchase must be between zero and maximum amount in the bid.
- There must be a balance between buyers and sellers.  
— Sum of accepted amounts of all buyers’ bids must equal the sum of accepted amounts of all sellers’ offers.

— Where there are locational requirements, the sum of accepted amounts of buyers' bids must be less than or equal to the sum of accepted amounts of sellers' offers in the location. That is, a capacity resource in a restricted location can (theoretically) satisfy a capacity requirement outside that location.

- Import restrictions must be observed.
  - For each external control area, total award to suppliers in that control area must be between zero and a predetermined maximum amount.
  - For all external control areas in aggregate, total award to suppliers in those control areas must be between zero and a maximum predetermined amount.

Deficiencies are settled in a balancing auction. LSEs that may have failed to acquire adequate capacity and suppliers that may have failed to perform are required to purchase additional capacity in a balancing auction. Deficient parties are treated as price takers in the balancing auction, while suppliers may offer their capacity at a price, as in the normal auction. With no elasticity on the demand side, the balancing auction's results (clearing prices) can be quite volatile. To mitigate this phenomenon, administratively determined demand-side elasticity can be introduced into the balancing auction. With elasticity on the demand side, the balancing auction's clearing prices tend to stabilize. Prices do not get too high because demand is reduced in response to high prices; and prices cannot get too low because demand is increased in response to low prices.

Suppliers of capacity generally have an obligation to participate in the day-ahead and/or real-time energy and ancillary services markets, that is, to make their generating units available for commitment. Their long-term marginal costs having (presumably) been met in the capacity market, they participate in the day-ahead and real-time markets concerned only with short-run marginal costs.

Market administration and oversight goes far beyond the simple administration of auctions, accounting of capacity holdings, and verification of bilateral contracts. Since no market can remain efficient when market power is abused, market administration also requires that market power be monitored and mitigated when needed. Mitigation generally takes two forms: 1) a requirement that suppliers offer capacity, that is, physical withholding is not allowed, and 2) caps on offer prices, that is, economic withholding is not allowed. In some instances suppliers with market power have been prohibited from entering into bilateral contracts and forced into capacity auctions.

### B. Black Start Resources

Black start-capable generators must start-up and self-sustain continued operation without an outside energy source. In the event of widespread loss of electric service (blackout), black start resources must supply the power to other generators so they can be started, or restarted, during system restoration. Resource adequacy generally does not have a requirement for black start resources. Rather, compensation of black start resources is cost-based and generally considers only the cost of facilities within the generators that provide the black start capability and the cost of restoration training for operators.

### C. Reactive Power and Voltage Control Resource Adequacy

In a three-phase power system, reactive power is crucial for the transfer of active power over the transmission system [25]. Equipments in a power system that can affect reactive power and hence perform voltage control include [26]:

- generators;
- synchronous condensers;
- switched capacitors and reactors;
- static Var compensators (SVCs) [27];
- voltage-sourced converter-based flexible ac transmission system (FACTS) controllers such as static compensators (STATCOMs) and unified power flow controllers (UPFCs) [27].

The scheduling of reactive power such that the system voltage for a particular power transfer level in a control area is within prescribed values is typically dispatched by an ISO. A proper voltage profile is important for minimizing active power losses and for the system to survive contingencies. For reliability, an ISO or LSE is required to secure the necessary reactive power resources. The cost of using these resources will be paid by the loads.

The reactive power resources, which include generators, synchronous condensers, SVCs, and FACTS controllers, and switched capacitors and reactors, have different capabilities and operational characteristics. Generators, synchronous condensers, SVCs, and FACTS controllers are relatively fast and thus are more valuable. Mechanically switched capacitors and reactors are slower and their reactive power output may not be adjusted rapidly. Nevertheless, generally only generators are paid for the ability to supply reactive power. In the case of FACTS controllers such as a STATCOM, in lieu of a reactive power payment, it can be paid with transfer rights [for example, transmission congestion contracts (TCC)] if their application results in higher power transfer levels [28]. In peak load periods, such transfer rights can be more valuable than reactive power payments. Generators usually are contracted to operate at certain power factors, and are paid with energy prices for active power generation and separately for reactive power generation. In the case when the ISO needs a particular generator to supply more reactive power than it is contracted to do so, and as a result, the generator has to reduce its active power generation, the generator will be compensated with an energy payment equivalent to its active power reduction. The payment is known as the lost opportunity cost and is socialized among the loads.

### D. Current Industry Practice

Table 1 summarizes treatment of capacity and ancillary services in several ISOs/RTOs.<sup>3</sup> While some ISOs/RTOs operate a capacity market and others do not, all provide cost-based payments to generators that supply reactive power and to providers of black start capability. Regulation and reserves will be discussed in the next section.

<sup>3</sup>Table 1 is accurate at the time of writing this paper. The payment methods may change as the market restructuring evolves in these ISOs/RTOs.

**Table 1**  
Capacity and Ancillary Service Payments in Several ISOs/RTOs

Operating entity	Capacity availability	Ancillaries			
		Regulation	Reserves	Voltage support	Black start
NYISO	market-based	market-based	market-based	cost-based	cost-based
ISO-NE	market-based	market-based	market-based <sup>3</sup>	cost-based	cost-based
CAISO	Cost-based	market-based	market-based	cost-based	cost-based
PJM	market-based	market-based	market-based	cost-based	cost-based
ERCOT	n/a <sup>4</sup>	market-based	market-based	cost-based	cost-based

<sup>3</sup>ISO-NE has a capacity requirement for quick start generation.

<sup>4</sup>Energy transactions are all bilateral contracts.

#### IV. OPERATIONS PLANNING

For reliable real-time operation, a power grid operator must ensure the availability of additional system resources that can be dispatched within an appropriate time frame following contingencies. These resources, because they do not receive energy payment, are grouped into the ancillary services category, to allow their readiness to serve to be compensated.

##### A. Generation Operating Reserves

One of the ancillary services resources is generator operating reserves [29], the total amount of which is equal to the largest amount of power lost in the most severe credible contingency case. If a single contingency is used, then the reserve would be equal to the power output of the largest generating unit. Generation operating reserves can be classified into several subcategories.

1. *Regulating reserves to maintain system frequency*— This service is in general provided by generators such as hydro turbines that can respond rapidly to changing power setpoints, which are dispatched by the ISO every 2–6 s.
2. *Fast reserves that can be available within 10 min*— Frequently these are already dispatched steam turbines that have upward generating capacity (spinning) or fast-start gas turbine units that can be brought on-line very quickly.
3. *Slower reserves that can be available within 30 min*— These are typically off-line gas turbines that require a longer time to ramp up.

The value of generator capacity reserves decreases as the time to its availability increases. Thus regulating reserve bids tend to be higher than those of the slower reserves. If a lower quality reserve is exhausted, the requirement can be substituted by a higher quality reserve (known as down substitutability), with the anticipation that the clearing price of that lower quality reserve be at the same level as the higher quality reserve. All generators committed to providing reserves will be paid based on market values when they are committed in the DAM (see Table 1). If in addition, they are requested to provide energy during real-time system operation, they will also be paid for the energy that they provide to the grid.

##### B. Locational Aspect of Resources

Although the total amount of operating reserves required is defined by the contingency analysis, for a control area that experiences congestions between its zones, such as the ones

shown in Fig. 4, the generation reserves must be located in the appropriate zones so that bringing up the reserves on-line after a contingency would not cause the congested lines between the zones to be overloaded. As a result, one would expect reserves in congested zones to be more valuable. The proportioning of the locational operating reserves depends on the system operating condition, and thus should be determined by a contingency evaluation program. For expediency the system operator may wish to set the locational reserves at fixed amounts in each zone, so as to reduce the complexity in computing the system dispatch. Such a practice is acceptable if the overall dispatch cost would not be very different and generation investments in the congested zones are not adversely affected.

##### C. Outage Coordination

The coordination of transmission and generation outages in a competitive environment seems to be in its infancy, because these outages affect both reliability and congestion rents. Currently the ISO or RTO of each deregulated control area confirms that scheduled outages will not cause a reliability problem. That is, the system must be able to withstand contingencies when selected equipment is out of service. The ISO/RTO has veto power over outages requested by transmission and generation asset holders if reliability would be compromised. Contingency analysis software is capable of making such assessments. In a deregulated environment, the challenge will be to schedule outages to minimize the impact on the electricity markets with the constraint that reliability is not compromised.

#### V. DAY-AHEAD AND REAL-TIME OPERATIONS

To be successful, a power market must operate with a reliable supply of energy and stable competitive energy prices [30]. To achieve this objective, the grid operator must approach the dispatch of power to serve loads in multiple steps. As discussed earlier, the grid operator must ensure that the resources are adequate to meet the real-time demand. Then, because of the start-ups and ramp rates of generators, supply schedules need to be provided to the generators with a sufficient lead time. In determining the supply schedule, not only the application of the principle of marginal costs (based on the bids provided by the generators) but also system reliability needs to be considered. However, because the real-time loads cannot be predicted with certainty, the grid operator needs to adjust in real time the generator schedule to accommodate for either excess or inadequate supply. Ensuring

consistency between energy prices and reliability-based dispatch is important for market transparency and will feed-back the appropriate price signals to the MP to achieve a stable, self-correcting market. Efficient price signals consistent with reliability also provide a stage upon which commercial freedom can be exercised freely via bilateral contracts or the arrangements of self-scheduling and self-commitment. The intent is to place bilateral and spot trading on an equal footing, leaving the mix between long- and short-term options solely as decisions of the market participants. For systems using the self-commitment approach without the DAM like in New Zealand or Australia, only the RTD mode is needed. In this section, we will discuss several aspects of bids and settlements with reliability consideration.

#### A. SCUC and Dispatch

The operation of a deregulated electricity market requires a multistage security-constrained optimization program based on LMP [5] to settle energy prices and provide ancillary services [7]–[9], [12], [31], [32]. This program, either in the unit commitment mode or the RTD mode, simultaneously balances the system generation, loads, and losses at the least as-bid cost, honors all transmission constraints, and provides market price signals over an appropriate operating horizon. In this framework, the ISO is using bids/offers from market participants to arrange a dispatch to achieve reliable operations while the market participants are using bid-based dispatch from the ISO to buy and sell energy at market prices to support their commercial objectives. The same algorithm applies to short-term forward markets (e.g., DAM) and real-time markets to ensure market convergence.

The energy and ancillary service commitment and dispatch should ideally be formulated as a nonlinear programming problem with the objective to maximize the market benefits, which is equivalent to minimizing the sum of costs to energy and ancillary service offers subject to the real-time power network operating constraints and equipment physical characteristics. The network operating constraints include power transfer limits across congested interfaces, which are predetermined based on off-line contingency evaluations, both for steady-state and dynamic stability conditions. Equipment physical characteristics include generator start-up times and ramp rates, and maximum line current loading from thermal considerations. Under a cooptimized framework of energy and reserve markets, system-wide and locational reserve constraints are enforced for the market operator to procure enough reserves to cover preselected and credible contingency events. In most ISOs, external transactions are considered for energy supply, but the ancillary services are provided solely by resources within the control area.

With recent, significant advances in optimization methods such as interior-point methods [33], it is becoming more practical to treat the dispatch problem as a nonlinear problem. To meet stringent market clearing deadlines of large-scale systems, however, the dispatch production code

typically still uses a linear programming formulation with cost functions and constraints that are piecewise linear so that a dispatch solution can be obtained reliably. Power transfer limits are typically fixed at worst case values, instead of depending on loading conditions. The power network is generally represented as a dc network or with generation shift factors. In some production code, an ac network is used, with a provision to default to the dc solution when convergence problems are encountered. As an illustration, a simplified linear programming formulation of a security-constrained energy and reserve dispatch optimization is given in the Appendix.

#### B. Multiple-Settlement Systems

Electricity is a commodity that cannot be effectively stored and the energy-supplying generators have limits on how quickly they can be started and ramped up or down. As a result, both the supply and demand become more inelastic and the electricity market becomes more volatile and vulnerable as it gets closer to real time [34]. To achieve a stable margin as well as to maintain the system reliability, a forward market is needed to provide buyers and sellers the opportunity to lock in energy prices and quantities and the ISO to secure adequate resources to meet predicted energy demand well in advance of real time. Thus architecturally, many ISOs (e.g. PJM, ISO New England, New York ISO) take a multisettlement approach for market design. Fig. 5 shows a typical architecture of a multiple-settlement system.

The two main energy markets, each producing a financial settlement, in a multisettlement system, are the following.

- 1) *DAM*: schedules resources and determines the LMPs for the 24 h of the following day based on offers to sell and bids to purchase energy from the market participants.
- 2) *Real-time market*: optimizes the clearing of bids for energy so that the real-time system load matching and reliability requirements are satisfied based on actual system operations. LMPs are computed for settlement at shorter intervals, such as 5–10 min.

Fig. 6 shows the timeline of the multiple-settlement systems used in NYISO, PJM, and ISO-NE, which are typical of those used in practice. Supply and demand bids are submitted for the DAM, typically 12–24 h ahead of the real-time operation. Then the day-ahead energy prices are computed and posted, 6–12 h ahead of real-time operation. In some systems such as PJM and ISO-NE, from the end of the DAM settlement to the beginning of the real-time market, a resource adequacy assessment (RAA) will be performed to revise, if needed, the generator schedule to meet the forecasted load and operating reserve. RAA will be discussed further in the next subsection.

1) *DAM*: The DAM typically consists of supply and demand bids on an hourly basis, usually from midnight to the following midnight. The supply bids include generation supply offers with start-up and no-load costs, incremental and decremental bids, and external transactions schedules. The demand bids are submitted by loads individually or collectively through load-serving entities. In scheduling the supply to meet the demand, all the operating constraints



such as transmission network constraints, reserve requirements, and external transmission limits must not be violated. This process is commonly referred to as an SCUC problem, which is to determine hourly commitment schedules with the objective of minimizing the total cost of energy, start-up, and spinning at no-load while observing transmission constraints and physical resources' minimum runtime, minimum downtime, equipment ramp rates, and energy limits of energy-constrained resources. Based on the commitment schedules for physical resources, SCUC is used to clear energy supply offers, demand bids, and transaction schedules, and to determine LMPs and their components at all defined price nodes including the hubs, zones, and aggregated price nodes for the DAM settlement. The SCUC problem is usually optimized using a Lagrangian relaxation (LR) or a mixed-integer programming (MIP) solver.

In some deregulated electricity markets, in order to improve the market liquidity, nonasset holding energy traders are permitted to bid in the DAM virtual supply and virtual load. When a virtual supply offer is competitive, it will displace an equivalent amount of physical generation. For reliability, in SCUC, the total amount of committed physical generation must not be less than the total amount of physical load. In case when this is not true, the deficit will be made up by securing physical generators spinning at no-load. This will not affect energy prices, but an uplift charge must be levied on the virtual supply parties causing the physical generation imbalance to pay for the additional generators spinning unloaded. The calculation and distribution of the uplift can be embedded in the SCUC program [35].

A critical part of the DAM is the bid-in loads, which is a day-ahead forecast of the real-time load. The load estimate depends on the season, day type (weekday, weekend, holiday), and hour of the day. Most ISOs have sophisticated load forecasting programs, some with neural network components [36], [37], to predict the day-ahead load to within 3%–5% accuracy and the load forecasts are posted. LSEs with fully hedged loads through long-term bilateral contracts tend to bid in the amount corresponding to the ISO predicted loads. Some other LSEs may bid in loads that are different from those posted by the ISO. In such cases, if the LSE bid load exceeds the ISO load, the LSE bid load is taken as the load to be dispatched. Otherwise, the ISO load will supersede the LSE bid load and the SCUC will commit generators to supply the ISO forecasted load in a reliability stage. Then the generation levels of the committed generators will be allocated to supply LSE bid loads. Committing extra generators outside the DAM will be treated as uplifts and be paid by the LSEs.

For some ISOs that do not supersede the lower LSE bid load with the ISO forecasted load, the SCUC will be cleared without committing enough generators to supply the forecasted load. For these ISOs, it would be necessary to perform an RAA to commit additional generators using SCUC to meet the ISO forecasted load. In addition, a security-constrained reliability analysis (SCRA) analyzes the next day's transmission system security based on the ISO load forecast and the RAA unit commitment schedules. The SCRA func-

tions can be performed on an ongoing basis after the DAM is cleared and continuing up to the RTD hour of the next operating day.

As presented above, without any reliability considerations, day-ahead least bid-cost unit commitment can be a single-pass optimization. With reliability considerations, SCUC is typically a multiple-pass commitment process designed to minimize the least added cost to the energy demand to satisfy reliability constraints. In addition to the reconciliation of bid and forecasted loads and virtual supply bids, reliability constraints of transfer capability can be imposed on both the transmission and subtransmission levels. To simplify the unit commitment process, SCUC is first solved without maintaining the subtransmission constraints. In the next pass, the SCUC solution is revised to satisfy the subtransmission constraints. There is no prescribed fixed approach to the multiple-pass SCUC solution for DAM—the sequencing of the passes is dependent on the market rules and the characteristics of the power control area.

Nevertheless, the sequencing should aim for least increases in unit commitment cost for reliability and be able to consistently clear the market.

2) *Real-Time Market*: In real-time operation, a system operator has to revise the day-ahead generator schedule so that the power supply will match precisely the load demand, while keeping enough reserves and not overloading the transmission system. The real-time balancing of generation and load is carried out in the real-time market, preferably using tools having similar capabilities as the SCUC program and similar enforcement of the reliability constraints, so that the price consistency can be carried from the DAM to the real-time market. In contrast to the DAM scheduling, which is optimized on an hourly interval over 24 h, the real-time balancing uses a shorter time interval, such as 5–10 min, over a shorter look-ahead horizon, such as 1–3 h.

In the real-time market, a system operator typically performs two functions: RTD and real-time commitment (RTC). The RTD provides market-clearing functions to balance generation and load, and to meet reserve requirements based on actual real-time system operating conditions using the SCED algorithm and fast-start generator commitment approach. SCED dispatches generators and bilateral transactions based on their respective bids, while observing various system security constraints. It clears energy and various ancillary service markets in an optimal fashion using a linear programming (LP)-based optimization engine. The most recent state estimator solution may be used as the RTD base case. Other input data for RTD include interchange, reserve requirements, load forecast, generator regulation, self-scheduled generation, and actual/potential transmission constraints. Using the man-machine interface of RTD, the system operator has the capability to manage real-time resource operational plans such as changing unit status, overriding resource operating limits, regulation assignments, and limited energy resource parameters, etc., to adapt to system condition changes. RTD determines desired dispatch points (DDP) and reserve designations for generators and

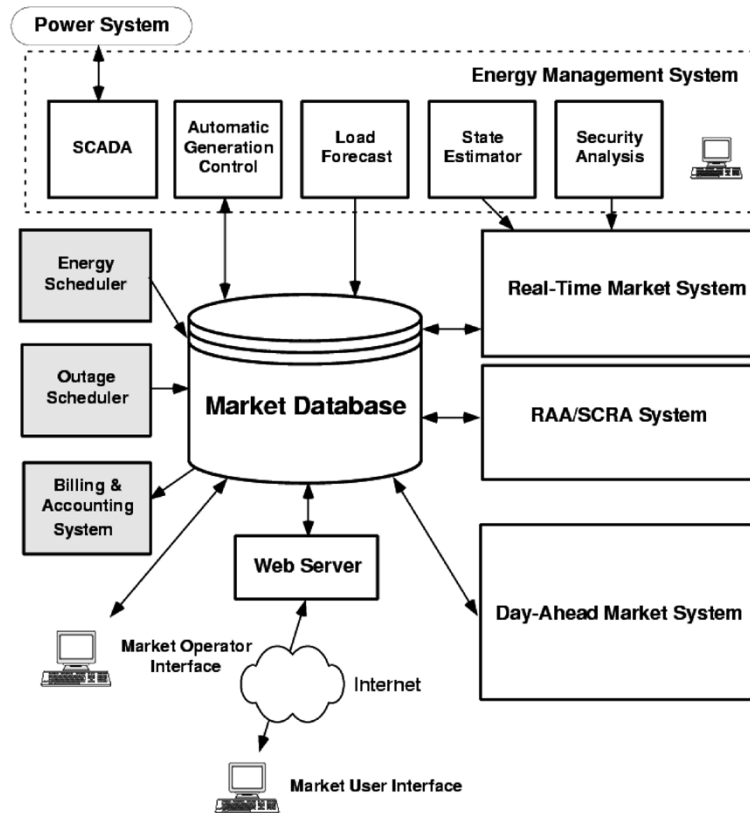


Fig. 5. System overview of a typical multisettlement system.

dispatchable loads. RTD also computes *ex ante* LMPs and zonal reserve prices for a specific time period into the future (e.g. 5 min, 15 min etc.). DDPs and *ex ante* LMPs of the approved RTD solution case will be sent to regional control centers or directly to the power plants.

Concurrently, a system operator will run the RTC task, which is similar to RTD, except for a longer time horizon, typically 1–3 h. It aims to commit resources that need a longer transition time.

## VI. LOOKING AHEAD

Before deregulation, most regulated electricity markets operated their power systems with a high level of reliability, backed by a significant amount of reserves and a well-maintained transmission and distribution power network. In the transition period from a regulated market to a deregulated one, this reliability level is most likely to be maintained, as long as the existing resources are being properly compensated for participating in ancillary services markets. In the longer term as load increases, we need to examine whether the existing incentives will be sufficient to encourage new investments in equipment capable of providing ancillary services, in addition to energy. For example, reactive power support from generators is paid at a fixed amount and reactive power support from switched shunt capacitors is most likely not being compensated at all. It would be of interest to study whether an energy-dependent payment of reactive

power support for generators and some compensation for new shunt capacitor installation would be effective in encouraging new reactive power investments.

There are also questions on how the reserves should be procured. For example, most ISOs secure generation operating reserves from resources located exclusively in their own control areas and the amount secured is independent of how much reserves exist in the neighboring ISOs. Should an ISO secure reserves from neighboring ISOs and should the neighboring ISOs share the available reserves, as long as the external interfaces can support the flows? Furthermore, should the amount of reserves to be secured be a function of prices? For example, if the last 5 MW of reserves is several times more expensive than the first 995 MW, should an ISO just purchase 995 MW of reserves instead of the fixed 1000 MW? The notion of a reserve demand function is worth investigating.

As pointed out earlier, coordinating equipment outages with the forecasted load will impact on the amount of resources available for ancillary services as well as the locational aspect of reserve requirements. Most ISOs would allow a piece of equipment to be taken off-line provided there are enough resources available to clear the market and operate the system, without directly optimizing the cost. It would be beneficial if an ISO has a maintenance scheduling tool that would optimize the staggering of equipment maintenance on a monthly or yearly basis, using historical load patterns.

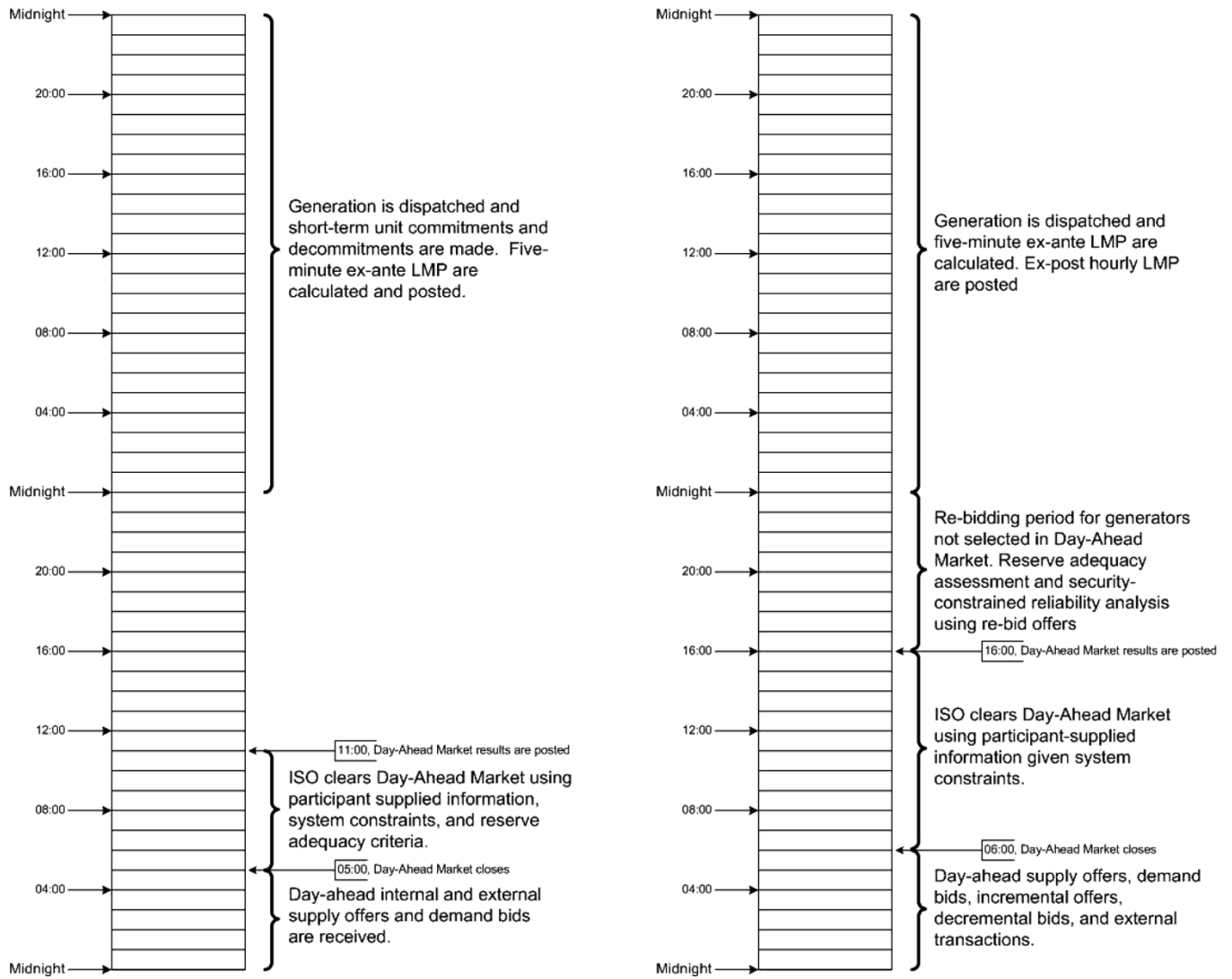


Fig. 6. Market timelines of two multiple-settlement energy markets: the left column is for NYISO, and the right column is for PJM and ISO-NE.

## VII. CONCLUSION

In this paper, we have provided an overview of reliability criteria and how these criteria are being met in deregulated electricity markets. It is important to recognize that a successful deregulated market needs to treat power system reliability as a multiple time horizon process where building new generation and transmission facilities will take years and operating the market has a time span of hours and minutes. In addition, the power market needs to be segmented to allow for a forward market and a real-time market, so that the cost of securing energy and reliability can be optimized. We have not discussed the details of the computational issues associated with optimizing these markets subject to reliability constraints. We encourage the reader to consult the references mentioned in this paper and also the operation manuals of the ISOs/RTOs, which can be found on their Web sites.

## APPENDIX FORMULATION OF AN ENERGY AND RESERVE DISPATCH OPTIMIZATION PROBLEM

A security-constrained energy and reserve dispatch optimization problem consists of minimizing the costs of serving the load demand and securing adequate reserves subject to the power network and flow constraints. The cost is given in the form of an objective function

$$C = \sum_{i \in G} \sum_q (c_{i,q} P_{i,q}) - \sum_j \sum_r (b_{j,r} D_{j,r}^d) + \sum_z \sum_w (c_{z,w}^L S_{z,w}^L) + \sum_t \sum_a \sum_v (c_{t,a,v}^S S_{t,a,v}^S) \quad (A1)$$

where

- $i$  the index for all the generators;
- $j$  the index for all the price-sensitive demands;

- $t$  the index for all the reserve categories, for example, 10-min spinning, 10-min nonspinning, and 30-min operating;
- $z$  the index for all the zonal reserve curves;
- $q$  the index for all the energy blocks of generators;
- $r$  the index for all the energy blocks of price-sensitive demands;
- $w$  the index for all the blocks of zonal reserve curves;
- $v$  the index for all the blocks of area reserve curves;
- $P$  the generation output;
- $D^d$  the price-sensitive demand;
- $D^f$  the fixed demand;
- $S^L$  the local reserve demand curve;
- $S^S$  the area reserve demand curve;
- $c$  the offer price for generator or reserve;
- $b$  the bid price for price-sensitive demand;
- $a$  the index for all categories of control area reserve curves;
- $G$  the set of the all on-line generators.

The constraints include the following.

*Energy Balance Constraint:*

$$(\lambda) : \sum_{i \in G} P_i - P_{\text{loss}} = \sum_j D_j^d + \sum_n D_n^f \quad (\text{A2})$$

where  $P_i = \sum_q P_{i,q}$ ,  $D_j^d = \sum_r D_{j,r}^d$ , and  $P_{\text{loss}}$  is the system losses.

*Transmission Constraints:*

$$(\mu_k) : BF_k \leq BF_k^{\text{max}} \quad (\text{A3})$$

where

- $BF$  the power flow over a transmission line;
- $\text{max}$  the superscript that indicates the maximum value;
- $k$  the index for all the transmission constraints.

*Locational Reserve Constraints:*

Typically a locational reserve constraint is used to cover the second contingency event caused by the loss of generator in a local area or the loss of second line of an interface of the local area. Therefore, the unused tie-line capacity can be used to cover the zonal contingency within 30 min.

$$(\alpha_z) : \sum_i \sum_t R_{t,i} + \sum_j \sum_t R_{t,j} + (\text{IF}_z^{\text{max}} - \text{IF}_z) \geq \sum_w S_{z,w}^L \quad (\text{A4})$$

where  $R$  is the reserve quantity, and  $\text{IF}$  is the real power flow over an interface.

*Control Area Reserve Constraints:*

$$(\beta_{t,a}) : \sum_i \sum_{t \in \Lambda_t} R_{t,i} + \sum_i \sum_{t \in \Lambda_t} R_{t,i} \geq \sum_v S_{t,a,v}^S \quad (\text{A5})$$

where  $\Lambda_t$  is the set of categories of reserves have quality equal to or higher than  $t$ .

*Capacity Constraints for On-Line Units:*

$$(\gamma_i^{\text{max}}) : P_i + \sum_t R_{t,i} \leq P_i^{\text{max}} \quad \text{for any } i \in G \quad (\text{A6})$$

$$(\gamma_i^{\text{min}}) : P_i \geq P_i^{\text{min}} \quad \text{for any } i \in G \quad (\text{A7})$$

where  $\text{min}$  is the superscript that indicates the minimum value.

*Reserve Capacity Constraints for Dispatchable Loads:*

$$(\gamma_j^{\text{min}}) : \sum_t R_{t,j} \leq D_j^d - D_j^{d,\text{min}} \quad \text{for any } j \quad (\text{A8})$$

$$(\gamma_j^{\text{max}}) : D_j^d \leq D_j^{d,\text{max}} \quad \text{for any } j. \quad (\text{A9})$$

*Ramp Constraints for On-Line Units:*

$$(\eta_i^{\text{up}}) : P_i - P_i^o \leq \text{RC}_i \quad \text{for any } i \in G \quad (\text{A10})$$

$$(\eta_i^{\text{dn}}) : P_i^o - P_i \leq \text{RC}_i \quad \text{for any } i \in G \quad (\text{A11})$$

where

- $\text{RC}$  the ramp capability for a generator or load;
- $\text{up}$  the superscript that indicates the upper bound;
- $\text{dn}$  the superscript that indicates the lower bound.

*10-Min Ramp Capability Constraints:*

$$(\eta_i^{10}) : \sum_{t \in \Omega_{10}} R_{t,i} \leq \text{RC}_i^{10} \quad \text{for any } i \quad (\text{A12})$$

$$(\eta_j^{10}) : \sum_{t \in \Omega_{10}} R_{t,j} \leq \text{RC}_j^{10} \quad \text{for any } j \quad (\text{A13})$$

where 10 is the superscript or subscript that indicates a value that can be achieved in 10 min, and  $\Omega_{10}$  is the set of reserve categories for 10 min.

*30-Min Ramp Capability Constraints:*

$$(\eta_i^{30}) : \sum_{t \in \Omega_{30}} R_{t,i} \leq \text{RC}_i^{30} \quad \text{for any } i \quad (\text{A14})$$

$$(\eta_j^{30}) : \sum_{t \in \Omega_{30}} R_{t,j} \leq \text{RC}_j^{30} \quad \text{for any } j \quad (\text{A15})$$

where 30 is the superscript or subscript that indicates a value that can be achieved in 30 min, and  $\Omega_{30}$  is the set of reserve categories for 30 min.

The parameters  $\lambda$ ,  $\mu$ ,  $\alpha$ ,  $\beta$ ,  $\gamma$ , and  $\eta$  in parenthesis in the above equations are the shadow prices for their corresponding constraints, respectively.

The LMP for each bus  $n$  and the reserve-clearing price (RCP) for each reserve can be defined as

$$\text{LMP}_n = \lambda \left( 1 + \frac{\partial P_{\text{loss}}}{\partial D_n^f} \right) - \sum_k \mu_k \frac{\partial B F_k}{\partial D_n^f} + \sum_z \alpha_z \frac{\partial I F_z}{\partial D_n^f} \quad (\text{A16})$$

$$\text{RCP}_{t,a} = \sum_{t \in \Gamma_t} \beta_{t,a} \quad (\text{A17})$$

$$\text{RCP}_{t,z} = \sum_{t \in \Gamma_t} \beta_{t,a} \Big|_{z \subseteq a} + \alpha_z \quad (\text{A18})$$

where  $\Gamma_t$  is the set of reserve categories that have quality equal to or lower than  $t$ .

The objective function  $C$  (A1) is suitable for the dispatch in one time period. For the DAM with hourly dispatch, the objective function will be a summation of  $C$  over 24 periods. Additional constraints to represent the transition between successive periods, such as the power output level of a generator at the end of one period is also its level at the start of the next period, need to be included in this multistage optimization problem. A similar optimization problem is solved for the RTC and RTD, except the time interval is smaller, the time horizon is shorter, and the initial load and generation are actual real-time values or obtained from the state estimator.

#### ACKNOWLEDGMENT

The authors would like to thank their colleagues at NYISO and AREVA for discussions on system reliability issues. The views expressed in this paper are solely those of the authors, and do not necessarily represent those of NYISO and AREVA. The authors would also like to thank the reviewers for providing useful comments in revising the paper.

#### REFERENCES

- [1] Northeast Power Coordinating Council [Online]. Available: <http://www.npcc.org>
- [2] "Operating policies and planning standards" North American Electric Reliability Council (NERC), 2004 [Online]. Available: <http://www.nerc.com>.
- [3] F. C. Schweppe, M. C. Caramanis, R. D. Tabors, and R. E. Bohn, *Spot Pricing of Electricity*. Boston, MA: Kluwer, 1998.
- [4] M. Ilic, F. Galiana, and L. Fink, *Power System Restructuring: Engineering and Economics*. Boston, MA: Kluwer, 1998.
- [5] W. Hogan, "Competitive electricity market design: A wholesale primer," J. F. Kennedy School of Government, Harvard Univ., 1998 [Online]. Available: <http://ksghome.harvard.edu/~whogan/>
- [6] T. Wu, M. Rothleder, Z. Alaywan, and A. D. Papalexopoulos, "Pricing energy and ancillary services in integrated market systems by an optimal power flow," *IEEE Trans. Power Syst.*, vol. 19, no. 1, pp. 339–347, Feb. 2004.
- [7] "Day ahead scheduling operations manual" New York Independent System Operator (NYISO) [Online]. Available: <http://www.nyiso.com/>
- [8] "Scheduling operations manual" Pennsylvania–New Jersey–Maryland (PJM) Interconnection [Online]. Available: <http://www.pjm.com/>
- [9] "Market operations," Independent System Operator (ISO) New England [Online]. Available: <http://www.iso-ne.com/>
- [10] "CAISO 28-Jun MD02 tariff filing of clean tariff sheets and errata," [Online]. Available: <http://www.caiso.com/>

- [11] A. I. Cohen, V. Brandwajn, and S. K. Chang, "Security constrained unit commitment for open markets," in *Proc. 21st Int. Conf. Power Industry Computer Applications* 1999, pp. 39–44.
- [12] K. Cheung, P. Shamsollahi, D. Sun, J. Milligan, and M. Potishnak, "Energy and ancillary service dispatch for the interim ISO New England electricity market," *IEEE Trans. Power Syst.*, vol. 15, no. 3, pp. 968–974, Aug. 2000.
- [13] "Course notes, New York market orientation course," NYISO [Online]. Available: <http://www.nyiso.com/>
- [14] L. L. Garver, P. R. Van Horne, and K. A. Wirgau, "Load supplying capability of generation-transmission networks," *IEEE Trans. Power App. Syst.*, vol. PAS-98, no. 3, pp. 957–962, May/June 1979.
- [15] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation, and Control*, 2nd ed. New York: Wiley, 1996.
- [16] T. E. DyLiaco, "The adaptive reliability control system," *IEEE Trans. Power App. Syst.*, vol. PAS-86, pp. 517–528, 1967.
- [17] L. H. Fink and K. Carlsen, "Operating under stress and strain," *IEEE Spectr.*, vol. 15, no. 5, pp. 48–53, May 1978.
- [18] R. P. Schulz and W. W. Price, "Classification and identification of power system emergencies," *IEEE Trans. Power App. Syst.*, vol. PAS-103, no. 12, pp. 3471–3479, Dec. 1984.
- [19] "Report to Congress: Impacts of the federal energy regulatory commission's proposal for standard market design," U.S. Department of Energy, Apr. 30, 2003 [Online]. Available: [http://www.energy.gov/engine/doi/files/dynamic/205200310489\\_DOES0138SMDfinal.pdf](http://www.energy.gov/engine/doi/files/dynamic/205200310489_DOES0138SMDfinal.pdf)
- [20] E. Hirst, "Long-term resource adequacy: The role of demand resources," Jan. 2003 [Online]. Available: <http://www.ehirst.com>
- [21] "Installed capacity manual" New York Independent System Operator (NYISO) [Online]. Available: <http://www.nyiso.com>
- [22] "PJM reserve requirements" PJM Regional Transmission Operator [Online]. Available: <http://www.pjm.com>, Manual 20
- [23] "Installed capacity" ISO-NE Regional Transmission Operator [Online]. Available: <http://www.iso-ne.com>, Manual 20
- [24] "An introduction to Australia's national electricity market" National Electricity Market Management Co. Ltd. [Online]. Available: <http://www.nemmco.com.au>
- [25] A. R. Bergen and V. Vittal, *Power System Analysis*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 1999.
- [26] B. Kirby and E. Hirst, "Ancillary service details: Voltage control," Oak Ridge National Laboratory, Rep. ONRL/CON-453, 1997.
- [27] N. G. Higorani and L. Gyugyi, *Understanding FACTS: Concepts and Technology of Flexible AC Transmission Systems*. New York: IEEE Press, 2000.
- [28] B. Fardanesh, M. Henderson, B. Shperling, S. Zelingher, L. Gyugyi, B. Lam, R. Adapa, C. Schauder, J. Mountford, and A. Edris, "Convertible static compensator: Application to the New York transmission system," presented at the CIGRE Paris Session, Paris, France, 1998, Paper 14-103.
- [29] E. Hirst and B. Kirby, "Ancillary service details: Operating reserves," Oak Ridge National Laboratory, Rep. ONRL/CON-452, 1997.
- [30] R. B. Wilson, "Architecture of power markets," *Econometrica*, vol. 70, pp. 1299–1340, 2002.
- [31] X. Ma, D. I. Sun, and K. W. Cheung, "Evolution toward standardized market design," *IEEE Trans. Power Syst.*, vol. 18, no. 2, pp. 460–469, May 2003, invited paper.
- [32] N. S. Rao, "Optimal dispatch of a system based on offers and bids—a mixed integer LP formulation," *IEEE Trans. Power Syst.*, vol. 14, no. 1, pp. 274–279, Feb. 1999.
- [33] E. D. Castronuovo, J. M. Campagnolo, and R. Salgado, "New versions of interior point methods applied to the optimal power flow problem," in *Proc. IEEE Power Engineering Soc. Transmission and Distribution Conf. Exposition* 2002 [Online]. Available: [http://www.optimization-online.org/DB\\_HTML/2001/11/405.html](http://www.optimization-online.org/DB_HTML/2001/11/405.html)
- [34] P. Cramton, "Electricity market design: The good, the bad, and the ugly," in *Proc. Hawaii Int. Conf. System Sciences* 2003 [Online]. Available: <http://www.hicss.hawaii.edu/HICSS36/HICSSpapers/CSMAR08.pdf>
- [35] "Allocation of uplift costs to load and other entities associated with virtual trading," New York ISO, Tech. Bulletin 82, 2001.
- [36] A. D. Papalexopoulos, S. Hao, and T. M. Peng, "An implementation of a neural network based load forecasting model for the EMS," *IEEE Trans. Power Syst.*, vol. 7, no. 4, pp. 250–257, Nov. 1992.
- [37] P. Shamsollahi, K. W. Cheung, Q. Chen, and E. H. Germain, "A neural network based very short term load forecaster for the interim ISO New England electricity market system," in *Proc. 2001 Power Industry Computer Applications Conf.* pp. 217–222.



**Joe H. Chow** (Fellow, IEEE) received the M.S. and Ph.D. degrees from the University of Illinois, Urbana-Champaign, working on singular perturbation techniques.

He joined the General Electric Power System, Schenectady, NY, in 1978. He joined Rensselaer Polytechnic Institute in 1987 and is a Professor of Electrical, Computer, and Systems Engineering. He is also a consultant to NYISO. His research interests include multivariable control and power system modeling, dynamics, and control. Most

recently he has been working on voltage-sourced converter-based FACTS Controllers.



**Robert W. de Mello** (Member, IEEE) received the B.S. degree in electrical engineering from the Johns Hopkins University, Baltimore, MD, in 1971, and the M.S. degree in control system science and engineering from Washington University, St. Louis, MO, in 1973.

From 1973 to 1975, he worked at Systems Controls Inc., where he developed a reduced-order, coherency-based modeling technique that increased the speed of transient stability simulation. He joined Power Technologies Inc.

in 1975, where he developed software for the simulation and evaluation of generating units and created novel software components for the analysis of unbalanced electrical power systems. He joined the New York Independent System Operator, Schenectady, NY, in 2000, where he defines and documents market design issues and solutions.

Mr. de Mello is a Member of Tau Beta Pi and Eta Kappa Nu.



**Kwok W. Cheung** (Senior Member, IEEE) received the B.S. degree in electrical engineering from National Cheng Kung University, Taiwan, in 1986, the M.S. degree in electrical engineering from the University of Texas, Arlington, in 1988, and the Ph.D. degree in electrical engineering from Rensselaer Polytechnic Institute, Troy, NY in 1991.

He has been with AREVA T&D Inc. (formerly ALSTOM ESCA Corporation), Bellevue, WA, since 1991. For the past decade, he has been developing wholesale electricity market applications, consulting in market design and leading deregulation projects for ISOs and RTOs including ISO New England, PJM, and ERCOT. He is currently Manager, Core Market Applications for AREVA's Market Management Systems.

Dr. Cheung is a Registered Professional Engineer in the State of Washington since 1994.