INDEPENDENT SYSTEM OPERATOR: PRICING AND FLEXIBILITY IN A COMPETITIVE ELECTRICITY MARKET

WILLIAM W. HOGAN

February 1998

Center for Business and Government
John F. Kennedy School of Government
Harvard University
Cambridge, Massachusetts 02138

CONTENTS

INDEPENDENT SYSTEM OPERATOR:
PRICING AND FLEXIBILITY IN A COMPETITIVE ELECTRICITY MARKET

William W. Hogan[1]

Parallel efforts in the United States to advance and implement the concept of an independent system operator provide a source of comparative insights on the principal choices and issues. A competitive electricity market requires supporting structures administered by the system operator. A key debate centers on the approach to transmission access and pricing. Efficient marginal-cost pricing in competitive electricity markets implies locational differences in the presence of transmission constraints. Aggregation of individual nodes into zones appeals as a putative simplification. In a sufficiently dense network, however, zonal aggregation provides less simplification than meets the eye. Illustration of the special features of electric networks and the implications for zonal aggregation suggest that nodal pricing is the simpler approach in a market with flexibility and choice.

INTRODUCTION

The development of Independent System Operators (ISOs) has proceeded steadily in the worldwide restructuring of electricity markets. There are significant advantages in this approach. There must be a system operator coordinating use of the transmission system. This control of the use of the transmission grid means control of the dispatch, at least at the margin, because adjusting the dispatch is the principal (or, in some cases, only) means of affecting the flow of power on the grid. That this system operator should also be independent of the existing electric utilities and other market participants is attractive in its simplicity in achieving equal treatment of all market participants. The ISO provides an essential service, but does not be compete in the energy market.

---

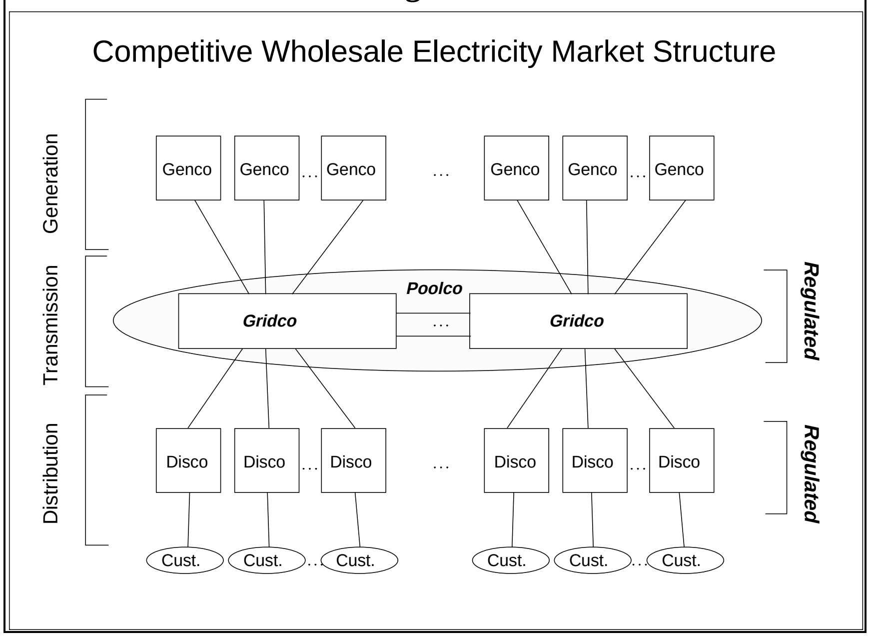
[1] Thornton Bradshaw Professor of Public Policy and Management, John F. Kennedy School of Government, Harvard University, and Senior Advisor, Putnam, Hayes & Bartlett, Inc. This paper draws on work for the Harvard Electricity Policy Group and the Harvard-Japan Project on Energy and the Environment. Many individuals have provided helpful comments, especially Robert Arnold, John Ballance, Jeff Bastian, Ashley Brown, Michael Cadwalader, Judith Cardell, John Chandley, Doug Foy, Hamish Fraser, Geoff Gaebe, Don Garber, Scott Harvey, Stephen Henderson, Carrie Hitt, Jere Jacobi, Paul Joskow, Susan Kaplan, Maria Ilic, Laurence Kirsch, Jim Kritikson, Dale Landgren, William Lindsay, Amory Lovins, Rana Mukerji, Richard O'Neill, Howard Pifer, Susan Pope, Grant Read, Bill Reed, Joseph R. Ribeiro, Brendan Ring, Larry Ruff, Michael Schnitzer, Hoff Stauffer, Irwin Stelzer, Jan Strack, Steve Stoft, Richard Tabors, Julie Voeck, Carter Wall and Assef Zobian.. The author is or has been a consultant on electric market reform and transmission issues for British National Grid Company, General Public Utilities Corporation (working with the "Supporting" PJM Companies), Duquesne Light Company, Electricity Corporation of New Zealand, National Independent Energy Producers, New York Power Pool, New York Utilities Collaborative, Niagara Mohawk Corporation, PJM Office of Interconnection, San Diego Gas & Electric Corporation, Trans Power of New Zealand, and Wisconsin Electric Power Company. The views presented here are not necessarily attributable to any of those mentioned, and any remaining errors are solely the responsibility of the author.

Establishment of objectives and criteria for an ISO has been followed by alternative approaches in implementation. The concept follows naturally from the principles established for open access to transmission systems.[2] However, there is variety in the details, especially in the institutions for short-term market operations. The experience in the United States provides a useful source of information on the continuing evolution of the idea. Here we summarize critical elements of this record and focus on a key set of issues that distinguish the alternative models in terms of pricing rules that allow for flexibility and choice. The issues addressed here revolve around the treatment of transmission access and transmission congestion, and have appeared in a debate over the use of locational marginal cost pricing versus various approaches to aggregation into zones.

ELECTRICITY MARKET

The defining characteristic of electricity restructuring to introduce competition in the market is the unbundling of the vertically integrated industry to separate the competitive elements (e.g., generation) from the essential facilities that will remain as monopolies (e.g., wires). However, the usual separation into generation, transmission, and distribution is insufficient. In an electricity market, the transmission wires and the pool dispatch are distinct essential facilities.

Figure 1



The conceptual distinction between the transmission wire connection function (Gridco) and the short-term dispatch function (Poolco) arises from the nature of power flows on an electric grid. In effect, there is no physical independence between transmission flows and power dispatch; one determines the other. These special conditions in the electricity system stand as barriers to an efficient, large-scale bilateral market in electricity. A pool based market model helps overcome these barriers by allowing both bilateral transactions and a coordinated spot market to

[2] W. Hogan, "An Independent System Operator: Criteria for Competitive Electricity Markets," Harvard-Japan Project on Energy and the Environment, Harvard University, February 1996.

deal with network interactions and limited transmission capacity.[3]

Is a system coordinator required in support of a competitive electricity market? The answer is yes, always. In England and Wales it is the "Pool." In New Zealand it is "Trans Power." In Norway and Sweden it is "Statnett Marked." Pool-based systems exist in Chile, Argentina, Alberta, Australia and so on. New names keep cropping up but the basic coordination functions will always be there, somewhere. The critical issues are access and pricing, especially transmission pricing. But a system coordinator or pool is required in support of a competitive market: "The importance of effective Pooling arrangements in a competitive [Electric Supply Industry] cannot be overstated."[4] In the United States, the ISO is a vehicle for providing this essential function.

THE ISO AND TRANSMISSION ACCESS

Simple independence from connection with individual market participants is not enough. The ISO and it's rules should support an efficient, competitive electricity market. Properly structured, an ISO can provide solutions to related problems that have appeared everywhere in the development of new electricity markets. These problems all have a common core, and they should not be treated separately. A review of the criteria for the ISO and their implementation presents an opportunity to unify seemingly independent conversations that are or should be about the same thing, namely how to deal with the special characteristics of electric networks in order to support a competitive market. Thus, while the concept of an ISO is hardly controversial, the specifics make a difference because of the problems of interaction in electric networks.

Network Interactions

Loop flow is the central network problem that is hard to face, often overlooked or dismissed, but critical in developing the rules for flexible competitive markets and electricity transmission access. The difficulty stems from the inherent and fundamental nature of system interactions in a transmission network. When power moves between one point and another on the transmission grid, it does not follow any single path. Rather the power moves in parallel along every path between source and destination; this is loop flow.

---

[3] Scott M. Harvey, William W. Hogan and Susan L. Pope, "Transmission Capacity Reservations and Transmission Congestion Contracts," Harvard-Japan Project on Energy and the Environment, Harvard University, February 1997.

[4] J. Moen, "Electric Utility Regulation, Structure and Competition. Experiences from the Norwegian Electric Supply Industry (ESI)," Norwegian Water Resources and Energy Administration, NVE, Oslo, April 1994, p. II-5.

The role of loop flow and its effects in the system needed to support a competitive market are important matters. The problems are fundamental in the presence of customer choice and competition. The principal implications of the ubiquitous effects of loop flow include:

No Property Rights. There is no workable system of property rights governing use of the transmission grid that would support a fully decentralized electricity market.

No Definition of "Available Transmission Capacity." It is not possible to define available transmission capacity (ATC) for a transmission interface without knowing everything about the use of the network at the time.

No Separation of Transmission Pricing and Spot Market. The opportunity cost of transmission depends critically on the marginal costs of power at different locations, and these costs are determined simultaneously with the dispatch and the spot market.

No Escape from the Network Externalities. There is a fundamental externality in transmission use, and decentralized markets do not deal well with externalities.

These are all facets of the same problem, and they strike at the very foundation of the decentralized, competitive electricity market. For example, we are familiar with the general economic problem of externalities. If there were no externalities, then competitive markets would be expected to find the efficient use of all our resources. However, as with environmental externalities, we know that even in a perfectly competitive market the market participants would not take into account the cost of externalities. Left to the market alone, therefore, resources like air and water would be misused. This is the reason we have environmental protection laws and agencies, not to supplant markets but to set the market rules to take account of externalities.

When someone transmits power in an electric grid with loops, parallel flows arise that can significantly affect the systems and dispatch of third parties not involved in the transaction. This is an externality. Sometimes it is a negative externality which increases the costs of the third parties, and sometimes a positive externality which lowers the costs. With existing networks and technology, there is no way to avoid this loop flow effect, and for reliability reasons this free flowing grid is an enormous asset that we should want to preserve.

Property rights in the transmission grid, in the usual sense, would allow the owners of the property rights to control the flow of power. If we could develop a workable system of physical property rights, we could internalize the externalities and achieve an efficient outcome through a decentralized competitive market. However, in the presence of loop flow and the free flowing grid, we can only control the use of the grid by controlling the dispatch, and there is no available system of decentralized property rights in terms of transmission alone. Ownership of individual lines in the grid would not create such property rights for use of the grid, and attempts to match such transmission line ownership with transmission use could exacerbate the problems of network interaction. Without such property rights, the externalities arise and efficiency suffers from the failures of markets.

Because of the loop flow effect, the short-run opportunity cost of transmission arises chiefly from the necessity to redispatch other generating units in the system in order to respect the many possible constraints in the transmission system. The redispatch can affect distant units in ways governed by the electrical distances for real and reactive power, not by geography. Hence, the opportunity cost of transmission use derives from the marginal costs of this redispatch, which would be determined simultaneously with the prices in the spot market. In getting the prices right, therefore, the simultaneity must be recognized and accommodated.

The combined effect of all this is that the traditional approach to unbundling a market to allow for competition will not work for the case of electricity. There must be a system operator, and the independent system operator can follow a set of protocols that addresses the special problems of electricity and support a competitive market.

Objectives

The ISO should advance the objectives outlined in many related proceedings considering the issues of transmission access. The United States is at the "beginning of the beginning" of its experience with the development and implementation of independent system operators. Motivated by the need to solve the problem of vertical market power, the ISO is the means for providing access to the essential facility. The relevant objectives include:

> **Reliability.** The ISO responsibilities should include coordinating short-term operations to ensure reliability while supporting the competitive spot market.
>
> **Independence.** The governance structure and incentives for the ISO should be designed to ensure that no one subset of the market participants is allowed to control the criteria or operating procedures.
>
> **Non-Discrimination.** Access to and pricing of services should be applied to all market participants without distinction as to customer identity or affiliation.
>
> **Unbundling.** Services should be unbundled when possible for acquisition from the competitive market and for utilization by the market participants.
>
> **Efficiency.** Operating procedures and pricing of services should support an efficient, competitive market for electricity. Attributable costs should be paid by the responsible parties. There should be no cost shifting. Joint costs should be allocated fairly with minimal impact on efficient incentives. Pricing and access rules should reinforce efforts to mitigate market power in generation.

The interesting part is in dealing with this "efficiency" objective. The importance of paying explicit attention to economic efficiency can be seen by reduction to the absurd case of minimizing use of the transmission grid, an operating policy which could meet the other objectives. The challenge is to describe efficient procedures for full use of the grid that would apply in the presence of transmission constraints; everything would be much easier with no transmission constraints, but we should not sweep the hard part under the rug. We can concentrate on two critical issues:

> **Economic Dispatch.** The short-term complexity of the interactions in the transmission grid requires the ISO to adjust the dispatch to meet transmission constraints and maintain balance in the system. The criterion for adjusting the dispatch should be to provide the most highly valued use of the grid based on the preferences of those in the market. In other words, users should provide bids, at least incremental and decremental bids around quantity schedules, and the ISO should use this information to determine the most economically rational use of the transmission system for the current dispatch.

> **Efficient Pricing.** The most significant attributable costs are the direct cost of power and the short-term cost of congestion in the transmission grid. The congestion cost arises when transmission constraints force some more expensive plants to operate. This cost of congestion would differ by location. Those causing the congestion at the margin should pay for it, and these prices should apply to everyone.

With efficient pricing, it is straightforward to achieve non-discrimination. Importantly, except in extreme circumstances, everyone in the market would be dispatched according to their stated preferences, and there would be no cost shifting due to congestion in the system. Use of the associated efficient prices would remove some of the most significant artificial incentives to game the system or exploit the externalities.

When the ISO performs these functions, the ISO is just another name for a Poolco. This approach is fully compatible with any kind of bilateral transactions that could be made without cost-shifting or without discrimination in favor of certain market participants. In fact, this approach expands the options of everyone in the market by making a virtue out of the necessity of central coordination.[5]

---

[5] Scott M. Harvey, William W. Hogan and Susan L. Pope, "Transmission Capacity Reservations and Transmission Congestion Contracts," Harvard-Japan Project on Energy and the Environment, Harvard University, February 1997.

Principles

The broad objectives have been elaborated into regulatory principles for ISOs. In its initial comprehensive treatment of the issue, the Federal Energy Regulatory Commission (FERC) outlined eleven principles for ISOs:[6]

1. The ISO's governance should be structured in a fair and non-discriminatory manner.

2. An ISO and its employees should have no financial interest in the economic performance of any power market participant. An ISO should adopt and enforce strict conflict of interest standards.

3. An ISO should provide open access to the transmission system and all services under its control at non-pancaked rates pursuant to a single, unbundled, grid-wide tariff that applies to all eligible users in a non-discriminatory manner.

4. An ISO should have the primary responsibility in ensuring short-term reliability of grid operations. Its role in this responsibility should be well-defined and comply with applicable standards set by NERC and the regional reliability council.

5. An ISO should have control over the operation of interconnected transmission facilities within its region.

6. An ISO should identify constraints on the system and be able to take operational actions to relieve those constraints within the trading rules established by the governing body. These rules should promote efficient trading.

7. The ISO should have appropriate incentives for efficient management and administration and should procure the services needed for such management and administration in an open and competitive market.

8. An ISO's transmission and ancillary services pricing policies should promote the efficient use and investment in generation, transmission, and consumption. An ISO or an RTG (regional transmission group) of which the ISO is a member should conduct such studies as may be necessary to identify operational problems or appropriate expansions.

9. An ISO should make transmission system information publicly available on a timely basis via an electronic information network consistent with the Commission's requirements.

---

[6] Federal Energy Regulatory Commission, Promoting Wholesale Competition Through Open Access Non-discriminatory Transmission Services by Public Utilities & Recovery of Stranded Costs by Public Utilities and Transmitting Utilities, Docket No. RM95-8-000 and Docket No. RM94-7-001, Order No. 888, Washington, DC, April 24, 1996, pp. 279-286.

10. An ISO should develop mechanisms to coordinate with neighboring control areas.

11. An ISO should establish an ADR (alternative dispute resolution) process to resolve disputes in the first instance.

These principles have served as the starting point for the development of ISOs in United States. However, given the innovation required, it is not surprising the some of the ISO developments have subsumed and then moved beyond these basic principles. Looking to the next level of detail, the practical implications emerge.

ISO DEVELOPMENTS IN THE UNITED STATES

A common discovery in the transitions to an ISO in the United States is the unanticipated level of difficulty of working out the details. The process reveals that there is more to the task than regulators and other institutional designers may have assumed. The black box of the vertically integrated industry is being opened, disassembled and rebuilt. There are many moving parts. And the gears have to do more then just turn--they also have to mesh to make the system as a whole work in the face of demands for greater flexibility and choice in the competitive market. At the beginning of 1998, therefore, there were many ISOs at various stages of development, and most were facing delays in implementation of the more ambitious efforts. In the case of California, the task was complicated by the need to set up an entirely new organization and control mechanism. But transition problems have also delayed implementation in the tight power pools that are facing the narrower but still significant task of reconstituting their rules and procedures.

ISO Summaries

The ISO proposals in the United States were in flux at the beginning of 1998. Some ISO plans had been approved by FERC and were in operation, but these were relatively simple in structure. Other more ambitious efforts were recently approved but not yet implemented, were awaiting regulatory approval, or were still under negotiation among the various parties. In all cases, there has been an intense process to select and specify the details. There are many common features, such as the use of access charges to cover the costs of transmission systems and the explicit treatment of bilateral transactions with scheduled transmission between locations. Governance is an issue, but not the central focus of this summary. The differences of greatest importance arise in the market design elements for coordinating the spot market and pricing transmission congestion. A brief snapshot summarizes key features related to market operations of the major ISO initiatives in the United States.

WEPEX. The Western Power Exchange (WEPEX) comprises the California ISO and Power Exchange (PX).[7] The FERC approved WEPEX's structure and governance mechanisms with independent boards and multiple advisory committees in October 1997, subject to extensive reporting requirements to monitor market performance and identify market power problems. Many details of access and pricing were unspecified or being changed up to the original January 1, 1998 start date, when system problems required delay of operations until at least March 31, 1998.

The WEPEX structure involves novel mechanisms for coordination among the ISO, the PX and parallel scheduling coordinators, with several markets for energy and services. A system design objective is to separate the "market" from "transmission" by assigning some of the dispatch functions to the PX and by restricting the operations of the ISO. Market participants trade through the PX or the competing scheduling coordinators. The PX is a special, default scheduling coordinator, and all the traditional utilities are required to use the PX for the immediate future. The PX and scheduling coordinators deal in turn with the ISO to secure transmission access and ancillary services. Coordination though the ISO is intended to be limited to the minimum requirements for reliability with an assumption that the market alone will achieve economic efficiency. The principal restrictions prevent the ISO from arranging trades across scheduling coordinators in the forward market. This has been justified as necessary for a number of reasons, including an intent to allow the traders to price discriminate and, in effect, exercise market power.[8] Although many other electricity systems include semantic distinctions between a "power exchange" and an ISO, all others have the short-term dispatch and spot market coordinated by a single entity, which is the functional system operator. The WEPEX system is the only electricity system in the world which purports to divide and separate these functions.[9]

The separation of the PX and ISO markets has been accompanied by a design feature to allow for extensive iteration of market clearing mechanisms, as a means of price discovery and convergence to a balanced dispatch. This innovation arises in part to accommodate relatively simple structures for market bids by generators and other participants.[10] Apparently, software and other problems of implementation will prohibit use of the iteration protocols until sometime after initial operation of the WEPEX market. The iteration feature does not depend on the market

---

7 Western Power Exchange Internet page — www.caiso.com (go to “About the ISO” or “ISO News”).

8 Paul R. Gribik, George A. Angelidis and Ross R. Kovacs, "Transmission Access and Pricing with Multiple Separate Energy Forward Markets," IEEE Meeting Paper, Tampa, Florida, February 3, 1998, pp 1-2.

9 W. Hogan, "A Wholesale Pool Spot Market Must Be Administered by the Independent System Operator: Avoiding the Separation Fallacy," The Electricity Journal, December 1995, pp. 26-37.

10 The WEPEX system uses only "one-part" bids, meaning energy price and quantity bids to constitute a supply curve. This is as opposed to "multi-part" bids which in addition to simple energy bids include start-up costs, minimum loads, and ramping constraints. The market in England and Wales employs multi-part bids. Victoria in Australia uses what are essentially one-part bids. No other market has introduced formal iteration with repeated bid revision.

separation, and if implemented could provide information about the any tradeoffs in efficiency and transaction costs compared to other systems which use more complicated bid structures but less formal iteration.

The WEPEX system includes an extensive system of "must run" contracts that are call contracts for specific generating units that might otherwise enjoy local market power. There is no separate system for securing generating capacity by ensuring long-run reserve margin requirements, with this function left to the market.

Transmission congestion pricing is based on a zonal system, with congestion giving rise to different prices for different zones. The transmission usage charge for movement between zones is the difference in the zonal prices. Congestion within zones is handled through a system of paying generators not to run in constrained areas and paying other more expensive generators elsewhere to balance the dispatch.[11] The congestion payments are recovered through an average charge to loads in the zone. The WEPEX designers initially planned a system of accompanying transmission congestion contracts for inter-zonal congestion, but this has been deferred and there are no transmission rights or equivalent forms of price protection administered by the ISO.

PJM. The Pennsylvania-New Jersey-Maryland (PJM) Interconnection is an ISO with no separate power exchange.[12] The PJM has a two-tiered governance structure with an independent board and a member committee to advise the board. The market structure approach is more conventional in the use of the ISO to coordinate short-term operations through bid-based economic dispatch with multi-part bids. A major innovation of the PJM plan includes locational marginal cost pricing for energy transactions through the spot market and fixed transmission rights (FTR) to deal with transmission access and price certainty.[13] The FERC approved the PJM plan in November 1997 for January 1998 implementation. The PJM ISO requested a delay until April 1, 1998 to accommodate training and other transition issues.

The FERC asked for some modifications to improve price protection for market participants, and PJM plans to adopt features such as a day-ahead forward market, introduction of market hubs, auctions of FTRs, and other modifications that had been discussed with market participants. The PJM system includes a capacity reservation system that is required of load-serving entities as a way of ensuring sufficient generating reserve margins. An initial allocation of FTRs accompanies the generating capacity reservations, with other FTRs to be allocated through firm transmission rights or an auction.

---

[11] This intra-zonal system is essentially the approach used in England and Wales. A key feature is paying generators not to run.

[12] The PJM ISO Internet page— www.pjm.com.

[13] FTRs are identical to transmission congestion contracts. An FTR provides the for payment of the congestion cost difference between two locations.

The existing PJM utilities have adopted a market power mitigation plan that limits the bid price for their generating units. In effect, this is similar to the must-run contracts in WEPEX as a means for addressing the possible ability of units in constrained areas to influence market prices.

The PJM system constitutes a major step forward in the development of ISOs and the details are examined further below.

NYPP. In January 1997, the New York Power Pool (NYPP) filed with FERC to establish an ISO, Reliability Council, and Power Exchange.[14] Like PJM, but unlike WEPEX, the ISO would administer the dispatch and spot market for the Power Exchange. The Power Exchange serves principally as a default aggregator, and market participants can use the Power Exchange or another load aggregator. Like PJM, the New York plan includes locational marginal cost pricing and transmission congestion contracts to deal with transmission access. A day-ahead forward market with multi-part bids, a two settlement system, and a transmission congestion contract auction provide added flexibility and price protection for market participants. There is a separate generation capacity market to deal with long-term reserve margins. The initial allocation of TCCs is based in part on existing transmission contracts or ownership of transmission facilities. Remaining TCCs will be sold in an auction, with previously allocated TCCS available through a coordinated secondary market.

The NYPP submitted a revised filing to FERC at the end of 1997, providing for an ISO board of directors unaffiliated with any market participant; the ISO management committee, handling day-to-day operations decisions, will be made up of market players. A New York State Reliability Council would be established to oversee system reliability. Industry stakeholders have nevertheless told FERC that they are not satisfied that NYPP's governance structure is independent of transmission owner control or representative of the industry. A FERC decision on the proposal had not been reached by the end of February 1998.

NEPOOL. New England Power Pool.[15] In June 1997, FERC granted conditional approval for creation of "ISO New England." The FERC approved interim operation for July 1997, with the electricity market scheduled to open to competition on April 1, 1998. The governance structure includes an independent board of directors. The ISO will administer a bid-based dispatching system with no separate power exchange. The bidding system will include seven separate markets for energy services. Unlike PJM and New York, or WEPEX for inter-zonal transmission, the transmission congestion system would spread the costs across all users and appears similar to an earlier failed experiment in PJM. In effect, the entire system would be treated as a single zone. Difficulties in software development and other systems resulted in an postponement of initial operation until later in 1998, although no new deadline had been set by February. Apparently the market delay will not interfere with implementation of

---

[14] New York Power Pool Internet page—www.nypowerpool.com (go to “ISO Filing”).

[15] ISO New England Internet page—www.iso-ne.com (go to “Document Library”).

retail competition this year.

ERCOT. The Electric Reliability Council of Texas (ERCOT) is an ISO that has been in operation in 1997 for the greater part of the state of Texas.[16] This version is an example of a working minimalist ISO. The ISO schedules transmission usage and administers a cost-sharing scheme to deal both with a small amount of current congestion and planned transmission expansion. Transmission congestion is reported to be small and localized. In 1997, the ISO received and approved 23,171 requests for unplanned transactions. ERCOT's Technical Advisory Committee approved ancillary service registration and verification procedures developed by the ISO. The ERCOT model has been under pressure because of the cost-sharing schemes and the disconnect between the benefits for transmission investment and the sharing of the costs.

INDEGO. The Independent Grid Operator.[17] A group of 21 utilities in the Pacific Northwest and the West signed a memorandum of understanding to create an ISO called the Independent Grid Operator (IndeGO). The proposed system includes use of transmission zones and transmission capacity reservation trading in the form of financial contracts to manage access and congestion between zones.[18] In January 1998, the 21 signatories hosted public comment meetings to present an overview of the current IndeGO proposal and to take public comments. In order to address implementation problems, IndeGO delayed its planned November 1997 FERC filing to early 1998. In February, however, eight members withdrew, and those remaining faced a deadline to commit to the group so that it determine whether it still has enough support to go ahead with the FERC filing. Some of the remaining members advocate an independent grid scheduler, a simpler and cheaper entity that would provide only scheduling and security. Those who left cited concerns about high costs and a large geographic spread. Given the importance of this large region, there will be some reconstitution of this initiative. Since the approach falls somewhere in between the WEPEX and PJM systems, the developments in INDEGO could provide valuable information and deserve close study.

MISO. Midwest ISO covering at one time as many as 26 utilities in several states.[19] After a long process of working towards a December 1997 filing at FERC, the Midwest ISO (MISO) appeared to collapse at the last minute, with only four of the original 26 utilities pledging to support implementation. However, it came back to life in January, with backers filing at FERC for approval. As proposed then, the MISO has the backing of seven investor-owned

---

16 Electric Reliability Council of Texas Internet page—www.ercot.com (go to "Independent System Operator").

17 Independent Grid Operator Internet page—www.idahopower.com/ipindego1.htm.

18 The developing proposals for transmission capacity reservations in INDEGO are financial rather than physical rights. These financial contracts are similar to transmission congestion contracts, but have a different definition based on distribution factors and multiple transmission paths. Unlike TCCs, the INDEGO financial contracts do not provide a perfect congestion hedge between locations.

19 Midwest ISO Internet page—www.midwestiso.org.

utilities and two cooperatives. Its miles of transmission lines, megawatts of generation, customers, and territory would make it the largest ISO proposed so far. The group pressed FERC to find that a larger group, including the eleven that pulled out in December (the GAPP group, discussed below), would make MISO more efficient, competitive, and reliable. MISO's request raises issues about FERC's authority to define an appropriate ISO and require it to form. Provisions of the MISO proposal include non-pancaked zonal rates for the first six years of operation. As with PJM and ISO New England, members would elect a disinterested board to oversee operations. The details of market operations have not been specified.

GAPP. Eleven of the original MISO group. The eleven that pulled out of MISO announced in December 1997 that they would pursue an alternative ISO based on the longstanding investigations of the General Agreements on Parallel Paths (GAPP). This group of utilities covers the eastern and northern borders of the MISO group. No details have been specified.

DesertSTAR. Desert Southwest Transmission and Reliability Operator.[20] Participating utilities in DesertSTAR concluded in late October 1997 that an ISO for the areas of Arizona, New Mexico, southern Nevada and West Texas would be feasible. In December 1997, parties signed a development agreement. Thirty-one companies have signed the agreement, and others have expressed interest. Four working groups, formed in the feasibility phase, are continuing with their work, with pricing issues currently the main focus. The group hopes to have the basic design set by July, with a FERC filing scheduled for the end of 1998.

Wisconsin. The Wisconsin Public Service Commission (WPSC) urged creation of an ISO in 1996. In September 1997, the WPSC told three Wisconsin utilities that if the state is to approve their proposed merger, they must form an ISO by 2000. In December 1997, the merger partners announced their own proposal for a regional ISO for Wisconsin and the Midwest. In January, the WPSC staff concluded that the proposed ISO does not meet WPSC standards concerning control area operation, governance, and system planning. If the Commission agrees with the staff, the partners will have to submit to FERC the ISO proposal which was earlier drawn up by Wisconsin Public Power Inc. and found to meet the WPSC's standards.[21] General principles have been announced, but the market operations details have not been specified.

Virginia. A bill introduced by a Virginia state legislator in January would create an ISO and regional power exchange for that state starting in January, 2001. Full competition would be implemented among generators, and utilities, co-ops and other generators in Virginia and elsewhere would offer power to the pool on a daily basis. While the bill does not call for retail choice, that is seen as following later. The market operations details have not been specified.

---

[20] Desert Southwest Transmission and Reliability Operator Internet page—www.swrta.org.

[21] Wisconsin Public Power Inc. Internet page—www.wppisys.org.

Flexible Markets and Consistent Pricing

The PJM plan approved by FERC and its accompanying order constitute a major breakthrough in the development of ISOs in the United States. Although the detail is difficult to extract from the formal tariffs written in the language of a different era, the FERC order on the PJM restructuring includes important elements for transmission access and pricing that support a competitive electricity market.

> Under the PJM Transmission Tariff, PJM-OI (the ISO) will offer pool-wide open access transmission service throughout the PJM Pool via the facilities of the eight PJM Companies. All transmission services will be subject to a single, non-pancaked rate based on the costs of the individual utility's transmission system where the point of delivery is located. Supporting Companies propose the locational marginal pricing approach for calculating and recovering the costs of transmission congestion. In general, under locational marginal pricing, transmission congestion costs are calculated based on differences in the marginal price of generation at each location on the transmission grid. [22]
>
> ... We are directing PJM-OI to implement Supporting Companies' proposal prospectively, effective January 1, 1998, subject to further modification in accordance with our findings herein.[23]

The developing PJM plan includes a number of features that support both open access and competition in the electricity generation market.

- Open access with non-discriminatory pricing. Any qualified participant in the market may trade in the bid-based spot-market or schedule transmission for bilateral transactions. The usage charges for transmission, including congestion costs, apply to everyone.

- Transmission fixed cost recovered primarily through system-wide (but not necessarily uniform) network service charges. In effect, these are intended to be transmission access charges. These charges cover the cost of the transmission system. They are postage-stamp rates in the sense that each load customer pays one charge for use of the entire

---

[22] Federal Energy Regulatory Commission, "Pennsylvania-New Jersey-Maryland Interconnection: Order Conditionally Accepting Open Access Transmission Tariff and Power Pool Agreements, Conditionally Authorizing Establishment of an Independent System Operator and Disposition of Control over Jurisdictional Facilities, and Denying Rehearings," Docket No. OA97-261-000, et al., Washington, D.C., November 25, 1997, p. 6.

[23] FERC, PJM Order, pp. 13-14.

transmission system. The charge does not vary depending on the source of any particular transmission schedule within the PJM system. The charges are non-uniform in the sense that they reflect the historical investment of each utility in the transmission system. Customers in each existing franchise area pay a network charge based on these historical embedded costs. There is a possibility to move to a uniform charge over a number of years.

- The ISO administers both a spot market and bilateral schedules, while maintaining reliability under principles of bid-based, economic, security-constrained dispatch. Those participating through bilateral transactions schedule transmission through the system operator. Those using the coordinated dispatch can provide multi-part bids for economic dispatch. Participants can combine bilateral schedules and spot-market bids.

- Transmission congestion charges are determined by locational prices from the bid-based spot market. Purchases and sales through the spot market occur at the locational prices. Transmission schedules are charged at the difference in locational prices for the congestion component. The locations will include defined hubs as fixed weight aggregations of individual locations. Trading and pricing can "take place" at the hubs which are treated as additional locations.

- Fixed transmission rights (i.e., transmission congestion contracts) are available for congestion costs between locations. The ISO administers a system of payments, collecting congestion usage charges from the actual users of the system and disbursing the payments to the holders of FTRs. An FTR is a financial contract which is a hedge against the changing congestion costs between two locations.

- Future developments to include multi-settlement systems for forward markets, FTR auctions and coordinated FTR trading.

Essentially the same innovations can be found in the proposed system for New York. The PJM and New York reforms are important, in part, because of the lessons learned in the process of adoption and in their demonstration of the ability to include features that should support a competitive market.

A major objective is to support both a bilateral market and a bid-based economic dispatch with the purpose of providing market participants with the greatest possible degree of flexibility. A challenge for transmission pricing and access is to balance the goals of reliability, commercial practicality, and flexibility in customer choice.

In the vertically integrated electricity system with bundled products, the details of energy, transmission, and distribution pricing did not matter greatly. The final price was important, to be sure, and regulatory attention focussed on the rates final customers would pay. But the decompostion of the final price into its components did not matter nearly as much. For most purposes, a slightly higher generation price and a lower transmission price would have no effect. Customers could not respond to the different component prices by changing the mix and location of generation. And the system operators typically did not use these prices in making their dispatch decisions. They used costs and efficiency criteria, with the pricing being determined after the fact according to complicated allocation rules such as "split-savings." Few understood these pricing rules, but this ignorance had little impact since the prices affected no major operating decisions.

This comfortable ignorance and cursory treatment of the connection between prices and operating decisions should be among the major casualties of the electricity restructuring process. If customers have flexibility in the choice of generation, spot purchases, bilateral transactions, and so on -- then prices matter and prices should reflect marginal cost impacts. In large part, control of operating decisions is moving from engineers motivated by principles of efficiency, to market participants motivated by profit. This is a major purpose of electricity restructuring -- to change the locus of these key decisions. However, if we want the market to be guided by prices, and we expect and intend for people to take these prices seriously, it becomes important to follow the usual advice of economists to "get the prices right."

If prices reflect marginal costs, then participants will see their individual profit motives more closely aligned with the objective of overall efficiency. They will respond to the prices and we will reap the benefits of competition. If prices do not reflect marginal cost impacts, customers will still respond according to their own profit incentives. However, in this case the result would be a less efficient and costlier system, with a gradual accretion of administrative fixes that lead to reduced choice and the de facto return of greater regulation.

The latter threat is serious and should concern all who are interested in electricity restructuring. Seemingly small inefficiencies in pricing may appear innocuous on the surface, but they can produce large changes in the response of market participants. Although the immediate economic effect may seem modest, the impact on system operations could be substantial. To the extent that the changes in generation and transmission use affect reliability, system operators have only two alternatives: (i) modify the pricing rules or (ii) impose administrative limits to foreclose the response to the prices distortions. It is one or the other. The easy or only path for the ISO may be the use of administrative rules, and this leads to constriction of the market.

Seeing the Future in PJM

The debate over transmission pricing and the use of a zonal approach in the PJM system provides a stark illustration of the difficulty and the challenge. In March of 1997, the FERC approved an interim transmission access and pricing system for the PJM ISO. Faced with opposition to the full locational pricing and congestion charging mechanism, the FERC endorsed the locational approach in principle but adopted temporarily an alternative model proposed by Philadelphia Electric Company (PECO) and others. The PECO approach minimized the importance of transmission congestion and rejected the locational pricing model as too complicated and unnecessary. Instead, the PECO model would treat the entire PJM system as a single zone.

In essence, the PECO model priced all transactions through the spot-market at the "unconstrained" price, based on a hypothetical dispatch. To the extent that the actual dispatch encountered transmission constraints, the PECO model would pay the more expensive generators to run and average these congestion costs over all users.

The PECO model included two other notable features. First, the generators that were constrained not to run, even though they had bids below the "unconstrained price," would be paid nothing. There was objection to adopting any system that depended on paying generators not to run. Second, market participants had the option to schedule bilateral transactions separate from the bid-based dispatch, with a separate payment for their share of the total congestion cost. This flexibility to use bilateral transactions or to participate in the coordinated spot market was a major design objective not to be abandoned.

This PECO pricing system is representative of a zonal approach, and has much in common with zonal systems adopted elsewhere in the world. However, should the system become constrained, the two exceptional features noted above would create a powerful and perverse incentive. If there were no transmission constraints, there would be no transmission congestion and everything would work as with the locational pricing system. But when congestion appeared, everything would be different. The PECO supporters argued that the total cost of congestion would be small, summed over the year, and therefore any inefficiencies could be safely ignored.

Ignoring a difference between prices and marginal costs is a safe practice in a regulated world without flexibility and choice. The incentives don’t matter and the small costs get lost in the larger system. But the cost of ignoring a gap between prices and marginal costs in the world of choice can be large indeed. Witness the events when the PJM system became constrained, starting in June 1997.

The data for a representative constrained dispatch found the marginal cost in the east at about $89 per MWh, when at the same time the marginal cost in the west was $12 per MWh. At the same time, the "unconstrained" price was approximately $29 per MWh. The incentives were clear. A customer could buy from the spot-market dispatch at $29, or it could arrange a bilateral transaction with a constrained-off generator in the west at a price closer to \$12.[24] The small average congestion cost would be the same either way, and would not affect the choice.

Faced with these incentives, constrained-off generators quickly arranged such bilateral transactions and scheduled their power for delivery. This, in turn, would require the ISO to constrain the output from some other generator, who would then follow the same direct path to a bilateral schedule rather than sit idle and collect nothing. Soon the ISO had no more controllable generating units with which to manage the transmission constraints. Unable to fix the perverse pricing incentives, the ISO resorted to administrative mechanisms to prohibit bilateral transactions or declare a "minimum" generation emergency during the peak generation period. In effect, the result of restructuring to facilitate a market was to create administrative rules to prohibit the market from responding to the price incentives.

The point was made in a dramatic way. The important issue is not the total cost of congestion, which may be small on average. The point is the incentives at the margin when the system is constrained. In designing the rules for transmission pricing and access for a competitive market, it matters little what the rules are for periods when the system is unconstrained. The important question is how the rules deal with the market when the system is constrained. Even though the total cost of congestion might be modest over the year, the gap between \$29 and \$12, or \$89 and \$12, is more than sufficient to get the attention of market participants. Given the margins in this business, they will change their behavior for \$1. And the changes in behavior can substantially affect system operations; in fact, the whole point of electricity restructuring is that changes in behavior can affect system operations.

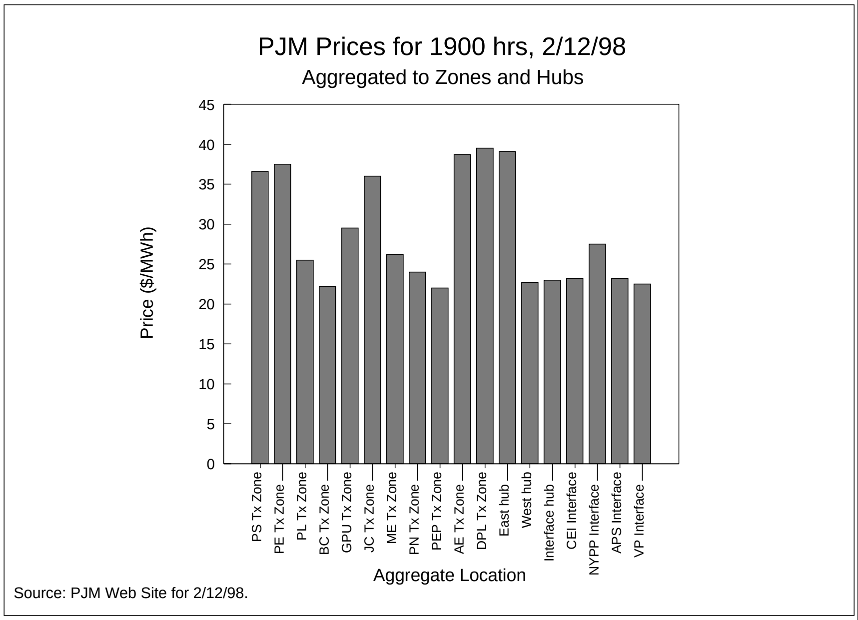
In the locational pricing system, the perverse incentives would not arise. Given the same facts as above, the locational prices would equal the marginal costs. Those customers purchasing power from the spot market in the east would have seen \$89 as the price. True, they could have arranged a bilateral transaction with a generator in the west, paying \$12 for the energy. But they would then face a transmission charge of (\$89-\$12), making them indifferent at the margin, just as intended. Likewise, customers in the west would pay \$12 and have no incentive to change. Every generator would be producing at its short-run profit maximizing output, given the prices. The market equilibrium would support the necessary dispatch in the presence of the transmission constraints. Spot-market transactions and bilateral schdeules would be compatible. Flexibility would be allowed and reliability maintained consistent with the choices of the market particpants.

The PJM ISO was fully aware of these perverse incentives and the problems they created, but without the authority to change the pricing rules it had no alternative but to restrict the market. The restrictions should now be lifted, once the locational pricing system goes into operation.

---

24 Power Markets Week, September 1, 1997, p. 13.

Figure 2



Source: PJM Web Site for 2/12/98.

In anticipation of this reform, in February 1998 the PJM ISO began posting the locational prices for 1277 locations plus a number of aggregations into utility service areas, hubs and interface connections. In most hours during this unseasonably warm month, the system was unconstrained and the locational prices (ex losses) were identical everywhere, even though they changed rapidly from hour to hour. For a few hours in February 1998, relatively modest constraints arose, and locational price differences appeared. The chart in Figure 2 reports on one such hour for the aggregations reported by PJM. The details for the full set of locations are available from PJM.[25] These aggregate zones are to be used for billing and settlement purposes, but the prices are actually determined for every location on the high voltage grid system.

These readily available locational prices in PJM should put and end to the myth that determining locational prices is complicated. The PJM ISO has long maintained that locational pricing required only a modest and fully auditable extension of calculations already done as a part of system operations, demanding very little data storage or computation other than that which is already necessary for system monitoring and control. Once the dispatch is determined, it is a relatively simple matter to determine the locational prices.[26]

Furthermore, as is evident from the data, transmission constraints on open interfaces can lead to different prices at every location. The typical network system does not lend itself to the intuition of other transportation systems. This experience and evidence from the PJM experiment highlight the issues that arise in the debate over the use of aggregate transmission zones versus determination of prices at every location or node.

If the connection between two regions is as shown in Figure 3, then our usual intuition

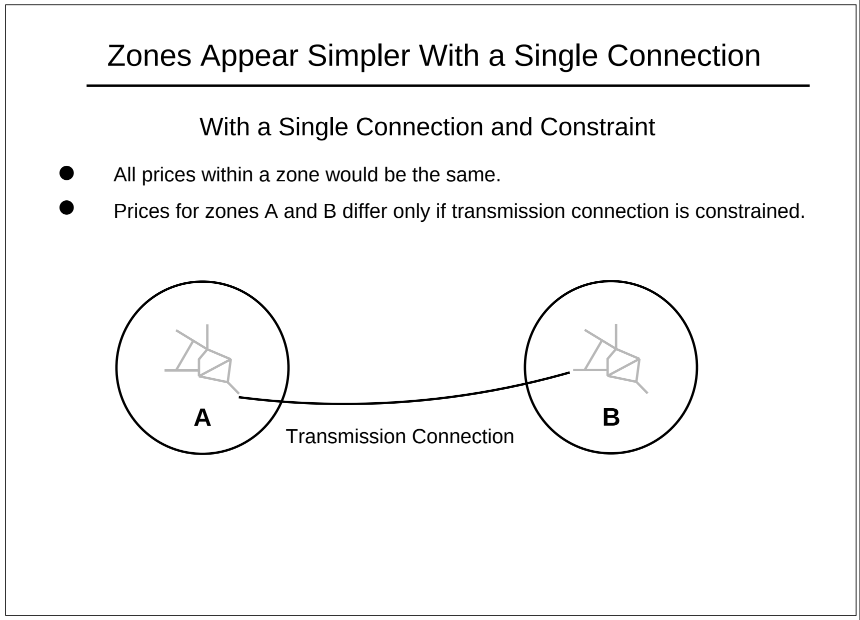
---

[25] The Inernet address is "www.pjm.com/ferc/filings/19971231/lmp_info.html".

[26] William W. Hogan E. Grant Read, and Brendan J. Ring, "Using Mathematical Programming for Electricity Spot Pricing," International Transactions in Operational Research, Vol. 3, No. 3/4, 1996, pp. 209-221.

applies. We have a radial link with a single line. The two grids in each region might be complicated, but they are assumed to be unconstrained. The only possible constraint is on the single line between the regions. In this special case, we could sensibly aggregate into two zones, A and B. The prices at all locations within a zone would be the same. The prices between the zones would differ only when the single line between them became congested. Everything would be easy. The flaw in the PECO model might be seen as not in the zonal approach itself, but in having only one zone. Decomposition into more zones had been suggested.

Figure 3



In the real world, however, the obvious decomposition into zones may not be available. The real meshed networks in the United States do not consist of simple radial connections. For reliability purposes, it is much better to design free-flowing transmission grids so that they include multiple parallel connections, both locally and across regions. With parallel connections, failure of a single line need not lead to interruption of service due to failure of the system. Power flows would redistribute automatically across the parallel lines while emergency adjustments were being made. The interconnected system would be robust.

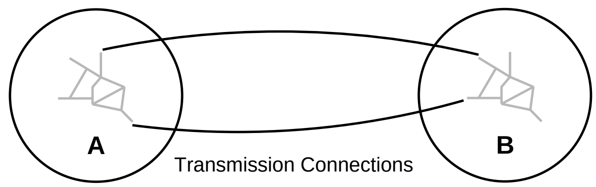
The natural evolution of transmission systems, therefore, is away from simple radial connections and towards a more interconnected design with multiple parallel paths. This reality is evident in many developments in the electricity market. For instance, the effect of parallel paths and the need for recognition of the effects of those network interactions is at the root of the extended debate about reliability and the "line loading" relief protocols being developed for managing inter-regional trade and reliability. The earlier debates over the meaning, or lack of meaning, of the concept of "available transmission capacity" sprang directly from the same root. This feature of electricity systems is unavoidable. It is not new. System operators in the vertically integrated utilities have always been aware of the network interactions and the need to modify the dispatch to respect the effects of loop flow. What is new is the exposure of this reality in the unbundled market and the recognition of its importance in the design of market institutions and transmission pricing.

Figure 4

Zones Are Not As Simple With Parallel Connections

With a Single Connection and Constraint

- All prices within a zone would be the same.
- Prices for zones A and B differ only if transmission connection is constrained.



With a Parallel Connection and Constraints

- *None of the above.*
-

These positive reliability benefits of interconnected systems come with a burden for our intuition when analyzing market pricing for transmissions systems. In the event that there is a parallel connection between zones, as in Figure 4, then our intuition misleads. None of the previous simplifications applies. A single constraint can lead to multiple prices; constraints within zones can change prices between zones; and so on. The world can be so different that the plausible conjecture that zonal aggregation would simplify would be turned on its head. In fact, for an interconnected system, zonal aggregation actual makes the problem more complicated. The simplicity of zones is deceptive; in the end, locational pricing at every node, as in PJM and as proposed for New York, is really simpler in the context of competitive markets and customer choice.

Having seen the perverse and unintended effects of the PECO model of zonal pricing, the FERC followed its early judgment about the appeal of full locational pricing and approved the locational model for PJM. The FERC order for PJM has set us on the right path with locational prices at nodes -- "We have seen the future and it is PJM."[27] The New York proposal, awaiting FERC approval at the end of February 1998, would apply the same locational pricing model and additional features to facilitate a competitive market. Other ISOs could adopt the same system. The PJM experience and comparison with other ISO proposals to date offers the opportunity and the motivation to look closely again at locational pricing and the tradeoff between zones and nodes.

---

[27] James J. Hoecker, Chairman, Federal Energy Regulatory Commission, November 25, 1997.

SHORT-RUN TRANSMISSION PRICING

Examples of pricing in networks illustrate the issues that accompany transmission congestion in a competitive electricity market.[28] In theory, pricing in a competitive electricity market with price-taking participants is at marginal cost. The competitive model is equivalent to a market with a central coordinator operating a pool. The many potential suppliers compete to meet demand, bidding energy supplies into the pool. The dispatcher chooses the welfare-maximizing combination of generation and demand to balance the system.[29] This optimal dispatch determines the market clearing prices. Consumers pay this price into the pool for energy taken from the spot market and generators in turn are paid this price for the energy supplied.

Inherently, energy pricing and transmission congestion pricing are intimately connected. A series of examples of pricing in the competitive electricity market model illustrates the determination of prices under economic dispatch in a network and relates transmission constraints to congestion rentals that lead to different prices at different locations. Use of the real nodes in the network appears to be a requirement of locational pricing that captures the marginal costs of congestion in a competitive market. The asserted complexity of using the real nodes leads frequently to proposals to aggregate the individual nodes into zones that would appear to be simpler for commercial purposes.[30] The examples here explore this issue to question the reality of the "simplification." In a market with choice, it is important to get the prices right. To the extent that prices differ from true marginal costs, there will be profit incentives to exploit the inconsistency. These incentives then lead to rules to constrain the most perverse behavior. The rules then add a new form of complexity and restrict the market. In the end, to the extent that zonal prices differ from locational marginal costs, the zonal system would not be a simplification, and locational pricing at the actual locations would be simpler and allow for greater market flexibility.

---

[28] These examples illustrate the elements of locational marginal cost pricing. They are adapted from William W. Hogan, "Contract Networks for Electric Power Transmission," Journal of Regulatory Economics, Vol. 4, No. 3, 1992; William W. Hogan, "A Wholesale Pool Spot Market Must Be Administered by the Independent System Operator: Avoiding the Separation Fallacy," The Electricity Journal, December 1995, pp. 26-37; William W. Hogan "Transmission Pricing and Access Policy for Electricity Competition," The Harvard-Japan Project on Energy and the Environment, February 1996; Scott M. Harvey, William W. Hogan and Susan L. Pope, "Transmission Capacity Reservations and Transmission Congestion Contracts," Harvard-Japan Project on Energy and the Environment, Harvard University, February 1997.

[29] The welfare maximizing formulation is the natural extension of traditional least-cost dispatch to include flexible demand. See F. C. Schweppe, M. C. Caramanis, R. D. Tabors, and R.E. Bohn, Spot Pricing of Electricity, Kluwer Academic Publishers, Norwell, MA, 1988. On the same point, but with examples to illuminate the critical importance of the phenomenon of loop flow in interconnected electrical grids, see W. Hogan, "Contract Networks for Electric Power Transmission," Journal of Regulatory Economics, Vol. 4, No. 3, 1992; Hung Po Chao, and Stephen Peck, "A Market Mechanism for Electric Power Transmission," Journal of Regulatory Economics, Vol. 10, No. 1, 1996, pp. 25-59.

[30] S. Walton and R. Tabors, "Zonal Transmission Pricing: Methodology and Preliminary Results form the WSCC," Electricity Journal, November 1996, pp. 34-41.

An ISO can implement a pricing regime to support the competitive market. This pricing and access regime can accommodate both a pool-based spot market and more traditional "physical" bilateral contracts. The key is in how the ISO provides balancing services, adjusts for transmission constraints and charges for transmission usage. The ISO would match buyers and sellers in the short-term market. The ISO would receive "schedules" that could include both quantity and bidding information. For the participants in the pool, these schedule-bids would be for loads or generation with maximum or minimum acceptable prices. For the self-nominations of bilateral transactions, the schedule-bids would be for transmission quantities with increment and decrement bids for both ends of the transaction. These incremental and decremental bids would apply only for the short-term dispatch and need not be the same as the confidential bilateral contract prices.

The responsibility of the ISO would be to integrate the schedules and the associated bids for deviations from the schedules to find the economic combination for all market participants. This range of schedule-bids would be more varied and flexible, giving everyone more choices.

Basic Transmission Pricing Examples

A set of examples can illuminate the treatment of spot-market transactions and bilateral transactions, under the ISO’s responsibility to achieve an economic dispatch. These examples are simple, but they capture the essential points in terms of the alternatives available for bilateral transactions. The test of no conflict of interest and non-discrimination is that, other things being equal, there should be no incentive in the dispatch or pricing mechanism to favor either the spot market or the bilateral transaction.

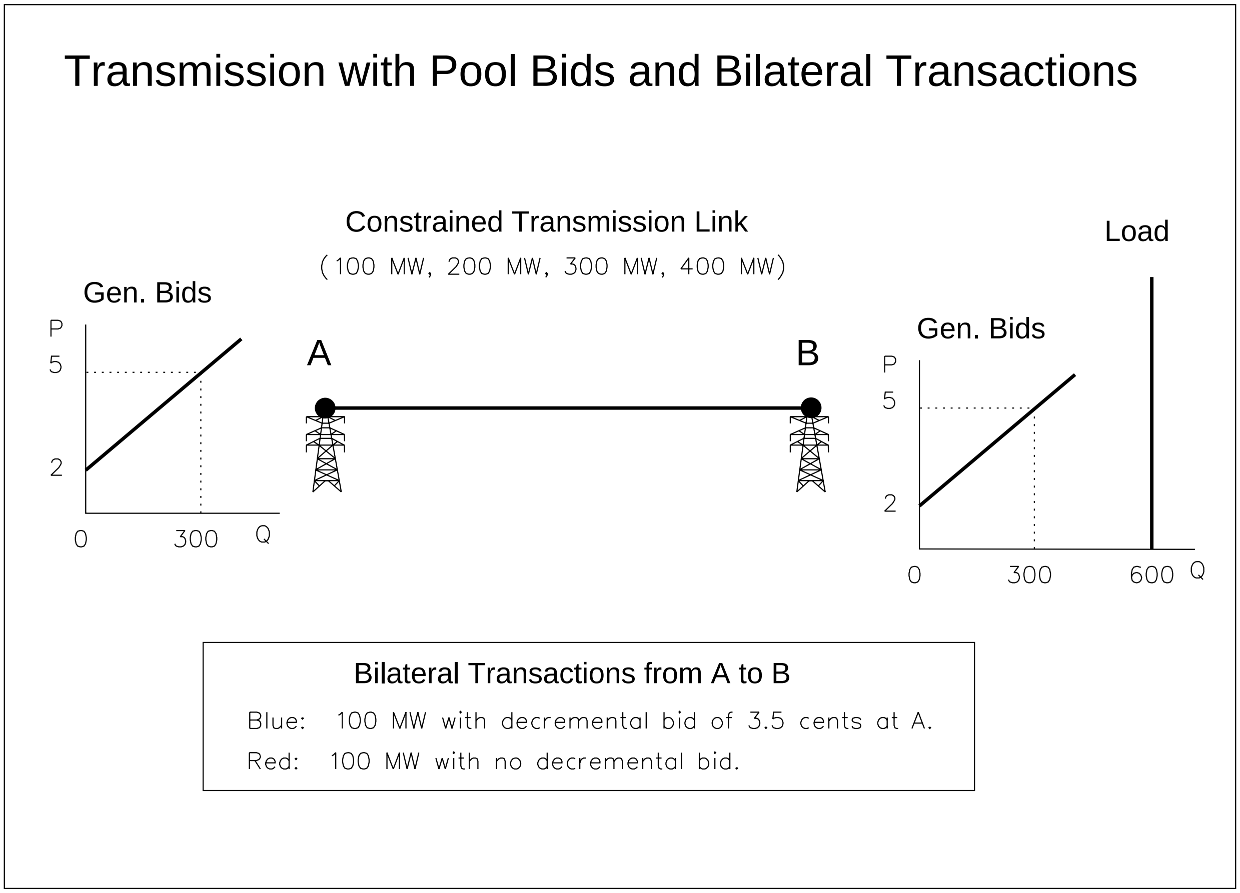
For simplicity, we ignore here any complications of market power or long-run issues, such as the creation of transmission congestion contracts, and focus solely on the short-run dispatch and pricing issues. A market with a single transmission line, as shown in the accompanying Figure 5, allows an illustration of the basic principles. What is less obvious, however, is that these same principles in no way depend on the special case of a single transmission line. Unlike many other approaches, such as ownership and physical control of the line, or the contract-path fiction, as expanded below in further examples for a grid, these pricing principles extend to a framework to support open access in a complicated network that includes the effects of loop flow.

The assumptions include:

- Two locations, A and B.

- Total load is for 600 MW at location B. For simplicity, the load is fixed, with no demand bidding.

Figure 5



- A transmission line between A and B with capacity that will be varied to construct alternative cases.

- Pool bid generation at both A and B. To simplify, each location has the same bid curve, starting at 2 cents/kWh and increasing by 1 cent/kWh for each 100 MW. Hence, a market price of 5 cents at A would yield 300 MW of pool-based generation at that location. Likewise for location B.

- Two bilateral transaction schedules, Blue and Red, each for 100 MW from A to B. Each bilateral transaction includes a separate contract price between the generator and the customer; the ISO does not know this contract price.

  Blue provides a (completely discretionary) decremental bid at A of 3.5 cents. In other words, if the price at A falls to 3.5 cents, Blue prefers to reduce generation and, in effect, purchase power from the pool. Blue may do this, for example, if the running cost of its plant is 3.5 cents, and it would be cheaper to buy than to generate.

Red provides no such decremental bid, and requests to be treated as a must run plant.

The ISO accepts the bids of those participating in the spot market at A and B and the bilateral schedules. The load is fixed at 600 MW. The bilateral transactions cover 200 MW, or the person responsible for the bilateral transaction must purchase power at B to meet any deficiency. The remaining 400 MW of load must be met from the spot market to include production at A or B, and use of the transmission line.

In determining the economic dispatch, the system operator treats the pool generation bids in the usual way. The Blue bilateral transaction is treated as a fixed obligation, with the 3.5 cent decrement bid as an alternative source of balancing adjustment at A. The Red bilateral transaction is treated as a fixed obligation, with no such balancing adjustment.

Power Flows and Locational Prices					
	Alternative Cases				
Link Capacity A to B	MW	400	300	200	100
Total Load at B	MW	600	600	600	600
Price at A	cents/kWh	4	3.5	3	2
Price at B	cents/kWh	4	5	6	7
Transmission Price	cents/kWh	0	1.5	3	5
Pool Generation at A	MW	200	150	100	0
Pool Generation at B	MW	200	300	400	500
Blue Bilateral Input at A	MW	100	50	0	0
Red Bilateral Input at A	MW	100	100	100	100

Assuming that the net of the fixed obligations with no balancing adjustments is feasible, which is the interesting case, we can vary the capacity on the link to see the results of the economic dispatch and the payments by the participants. The examples cover four cases, starting at 400 MW of transmission capacity, and reducing in increments of 100 MW. The details are in the accompanying table.

400 MW. In the case of 400 MW of transmission capacity, the economic dispatch solution is just balanced with no congestion. Everyone sees the same price of 4 cents. The payments for each party include:

- Pool Generation at A: Paid 4 cents for 200 MW.

- Pool Generation at B: Paid 4 cents for 200 MW.

- Pool Load at B: Pays 4 cents for 400 MW.

- Blue Bilateral: Pays zero cents for transmission of 100 MW.

- Red Bilateral: Pays zero cents for transmission of 100 MW.

Everybody is happy.

300 MW. In the case of 300 MW of transmission capacity, the economic dispatch solution encounters transmission congestion, and the prices differ by location. The price at A drops to 3.5 cents, and the price at B rises to 5 cents. The opportunity cost of transmission is 1.5 cents. The payments for each party include:

- Pool Generation at A: Paid 3.5 cents for 150 MW.

- Pool Generation at B: Paid 5 cents for 300 MW.

- Pool Load at B: Pays 5 cents for 400 MW.

- Blue Bilateral: Pays 1.5 cents for transmission of 50 MW. Blue makes up the remaining 50 MW obligation at B at a price of 5 cents.

- Red Bilateral: Pays 1.5 cents for transmission of 100 MW.

Everybody would prefer less congestion, but everyone is paying the opportunity cost of the transmission congestion. Note that at these prices, Blue is indifferent to bidding in its generation at 3.5 cents in the pool at A, or continuing as a bilateral transaction. Further, note that the ISO reduced both pool and Blue transactions. There is no artificial bias induced by the ISO fulfilling the directives of the economic dispatch.

200 MW. In the case of 200 MW of transmission capacity, the economic dispatch solution encounters more transmission congestion, and the prices differ more by location. The price at A drops to 3 cents, and the price at B rises to 6 cents. The opportunity cost of transmission is 3 cents. The payments for each party include:

- Pool Generation at A: Paid 3 cents for 100 MW.

- Pool Generation at B: Paid 6 cents for 400 MW.

- Pool Load at B: Pays 6 cents for 400 MW.

- Blue Bilateral: Prefers not to generate and has no transmission. Blue makes up the 100 MW obligation at B at a price of 6 cents.

- Red Bilateral: Pays 3 cents for transmission of 100 MW.

Everybody would prefer less congestion, but everyone is paying the opportunity cost of the transmission congestion. Note that at these prices, Blue is better off than if it had actually generated. Of course, Blue would still be indifferent to bidding in its generation at 3.5 cents in the pool at A, or continuing as a bilateral transaction. Further, note that the ISO reduced both pool and Blue transactions. There is no artificial bias induced by the ISO fulfilling the directives of the economic dispatch.

100 MW. In the case of 100 MW of transmission capacity, the economic dispatch solution encounters transmission congestion to the point of eliminating everything other than the must run plant, and the prices differ more by location. The price at A drops to 2 cents, and the price at B rises to 7 cents. The opportunity cost of transmission is 5 cents. The payments for each party include:

- Pool Generation at A: No generation.

- Pool Generation at B: Paid 7 cents for 500 MW.

- Pool Load at B: Pays 7 cents for 400 MW.

- Blue Bilateral: Prefers not to generate and has no transmission. Blue makes up the 100 MW obligation at B at a price of 7 cents.

- Red Bilateral: Pays 5 cents for transmission of 100 MW.

Everybody would prefer less congestion, but everyone is paying the opportunity cost of the transmission congestion. Note that at these prices, Blue is better off than if it had actually generated. Of course, Blue would still be indifferent to bidding in its generation at 3.5 cents in the pool at A, or continuing as a bilateral transaction. Further, note that the ISO reduced both pool and Blue transactions. There is no artificial bias induced by the ISO fulfilling the directives of the economic dispatch.

The net spot-market payments that are made to and from the ISO are summarized in the accompanying table. Note that the cases of transmission congestion include net payments to the ISO. These net payments are equal to the value of the constrained transmission capacity. These are the congestion payments which would be redistributed through a system of transmission congestion contracts, as illustrated below in further examples.

Power Flows and Locational Prices					
	Alternative Cases				
Link Capacity A to B	MW	400	300	200	100
Price at A	cents/kWh	4	3.5	3	2
Price at B	cents/kWh	4	5	6	7
Transmission Price	cents/kWh	0	1.5	3	5
Payments to Independent System Operator					
Pool Load at B (400 MW)	cents (x1000)	1,600	2,000	2,400	2,800
Contract Load at B (200 MW)	cents (x1000)	0	0	0	0
Generation at A	cents (x1000)	(800)	(525)	(300)	0
Generation at B	cents (x1000)	(800)	(1,500)	(2,400)	(3,500)
Blue Transmission	cents (x1000)	0	75	0	0
Blue Imbalance at B	cents (x1000)	0	250	600	700
Red Transmission	cents (x1000)	0	150	300	500
Red Imbalance at B	cents (x1000)	0	0	0	0
Net to ISO	cents (x1000)	0	450	600	500

Implications

These examples for a single, isolated line are simple, but they capture the essential features. These features generalize to a more complicated network under the economic dispatch model in the sense that participants can provide bids at their discretion. Some of the bids can be "must run." The locational prices are easily determined from the economic dispatch considering all the bids and schedules, not just those included in the power exchange. And although everyone would prefer a less congested system, all users would pay the short-run opportunity costs of their contribution to the congestion. Other things being equal, there would be no bias between spot market and bilateral transactions.

Note that if Blue and Red did not pay the opportunity cost of transmission, there would be a substantial bias in favor of the bilateral transactions. Furthermore, the locational prices are consistent with the efficient competitive outcome, as is best illustrated by Blue's willingness to adjust a bilateral transaction.

Contrary to a common argument -- that the ISO would have a bias in favor of spot market transactions -- the treatment of the Red bilateral transaction might lead to an accusation that there is a reverse bias in favor of the bilateral transaction. However, there are two important features of the pricing and access rules that run counter to this assertion.

First, the spot market participants could achieve the same result by bidding in generation at A at a zero reservation price, or lower. In fact, in performing the economic dispatch, the ISO treats the Red transaction as just this type of bid. Under these circumstances, the price at A could drop to zero, or lower, with a corresponding increase in the opportunity cost of transmission.

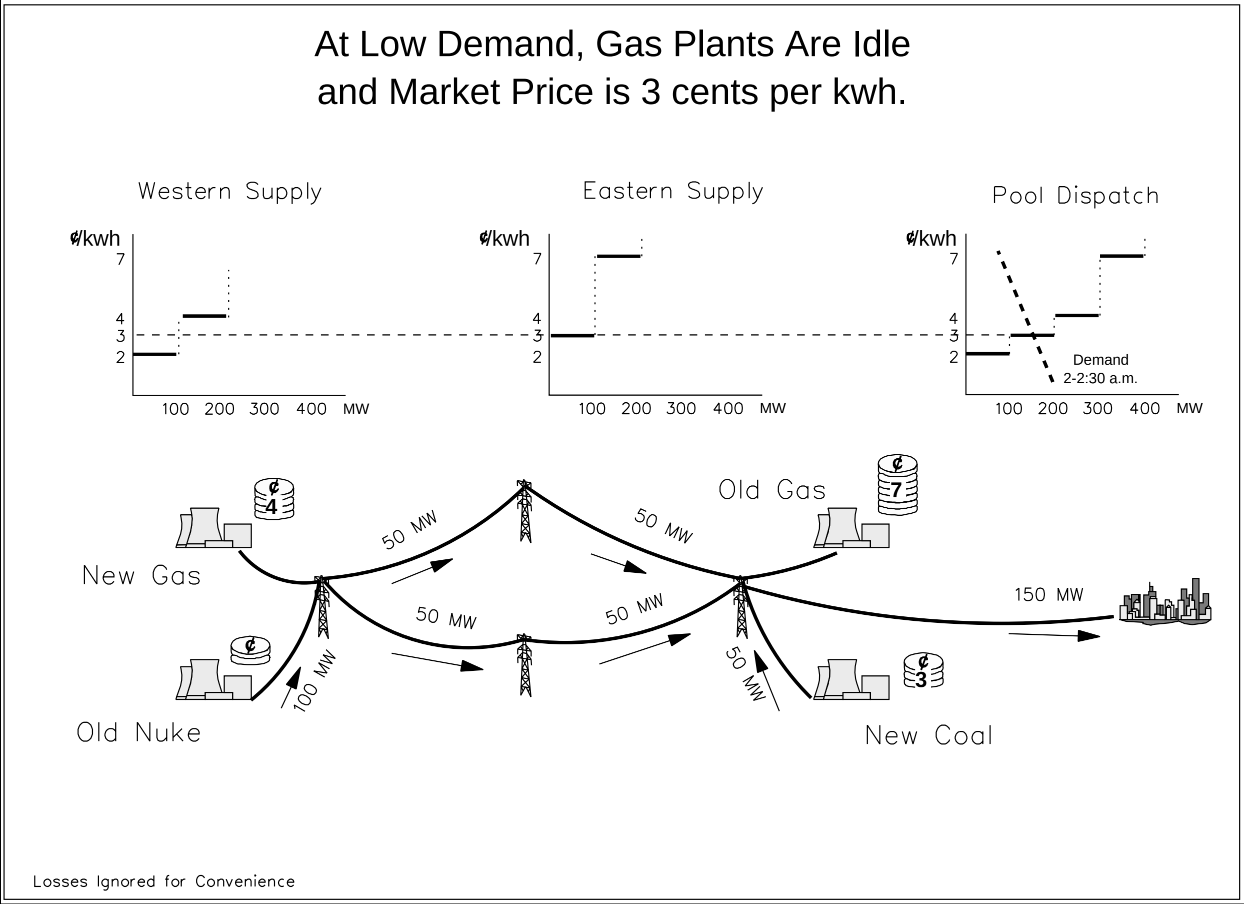
Furthermore, suppose that Red's true short-term generation cost is 3 cents, but it refused to make a decremental bid to the ISO. Then in the 100 MW case above, Red would have acted irrationally and would be worse off than if it offered such a decremental bid. It can also be shown that the cost thus imposed on Red is at least as large as the total cost imposed on everyone else in the market. Thus Red would pay for its own mistakes; the effect would be a net gain for the other generators and load (although there could be winners and losers, in aggregate everyone else would win).

Hence, the single line examples illustrate the use of locational prices for the various types of transactions that might take place in the short-term market. Locational pricing provides the opportunity costs price signals and transmission pricing is at the difference between the locational prices at source and sink. This ideas generalize from the single, radial to a more general network that includes free-flowing loops and the strong network interactions that are characteristic of electric grids. However, most of the intuition about the determination of prices and the relationships of prices across locations do not extend to the real grid. In the presence of loop flow, the interactions are complicated and important. This reality needs to be understood to appreciate the arguments for and against zonal aggregation of spot prices.

ECONOMIC DISPATCH ON A GRID

The pricing results for network can be quite different than those found for a single transmission line or a radial connection. The key difference is in the existence of loops that give rise to network interactions and create the phenomenon of "loop flow." Analogies built on the case of a single line can be misleading. The determination of market clearing prices at equilibrium, equal to the marginal costs that would arise from an economic dispatch, follows from the same principles. But the application and interpretation of these principles requires an extension of our intuition.

Figure 6



Consider the simple market model in Figure 6, which will serve as the starting point for a set of a succeeding examples for a grid that move from the analogy of a single line to a grid with multiple loops. In this market there is one load center, a city in the east, supplied by generators located far away in the west, connected by transmission lines, and by local generators who are in the same region as the city customers. The plants in the west consist of an "Old Nuke" which can produce energy for a marginal cost of 2 ¢/kWh and a "New Gas" plant that has an operating cost of 4 ¢/kWh. These two plants each have a capacity of 100 MW, and are connected to the transmission grid which can take their power to the market in the east.

The competing suppliers in the east are a "New Coal" plant with operating costs of 3 ¢/kWh and an "Old Gas" plant that is expensive to use with a marginal cost of 7 ¢/kWh. Again these eastern plants are assumed to have a capacity of 100 MW. The two plants in the west define the "Western Supply" curve, and the two plants in the east define the corresponding "Eastern Supply" curve. These supply curves could represent either engineering estimates of the operating costs or bids from the many owners of the plants who offer to generate power in the competitive market. For simplicity, we ignore transmission losses and assume that the same supply curves apply at all hours of the day.

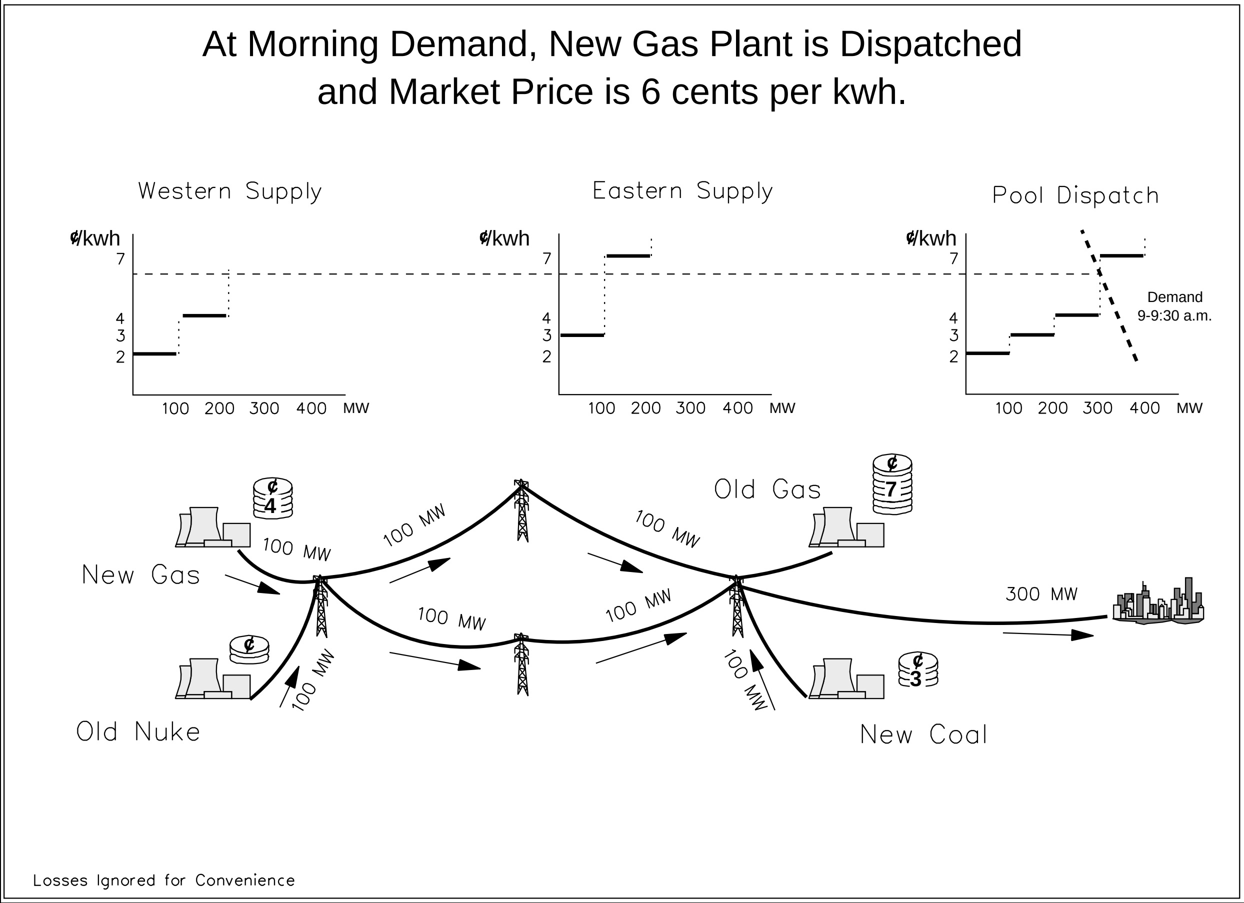
Under low demand conditions, as shown in Figure 6 for the early hours of the morning, the supply curves from the two regions define an aggregate market supply curve that the pool-based dispatchers can balance with the customer demands. The aggregate market supply curve stacks up the various generating plants from cheapest to most expensive. The pool-based dispatchers choose the optimal combination of plants to run to meet the demand at this hour. In Figure 6, the result is to provide 150 MW. The inexpensive Old Nuke plant generates its full 100 MW of capacity, and the New Coal plant provides another 50 MW. The New Coal plant is the marginal plant in this case, and sets the market price at 3 ¢/kWh for this hour. Hence the customers in the city pay 3 ¢/kWh for all 150 MW. The New Coal plant receives 3 ¢/kWh for its output, and this price just covers its running cost. The Old Nuke also receives 3 ¢/kWh for all its 100 MW of output. After deducting the 2 ¢/kWh running cost, this leaves a 1 ¢/kWh contribution towards capital costs and profits for Old Nuke owners.

In this low demand case, and ignoring losses, there is no additional opportunity cost for transmission. The 100 MW flows over the parallel paths of the transmission grid. But there is no constraint on transmission and, therefore, no opportunity cost. Hence the price of power is the same in the east and in the west. In the short run, there is no charge for use of the transmission system.

If demand increases, say at the start of the business day, the system operator must move higher up on the dispatch curve. For example, consider the conditions defined in Figure 7. This hour presents the same supply conditions, but a higher demand. Now the pool-based dispatchers must look to more expensive generation to meet the load. The Old Nuke continues to run at capacity, the New Coal plant moves up to its full capacity, and the New Gas plant in the west also comes on at full capacity. The New Gas plant in the west is the most expensive plant running, with a marginal cost of 4 ¢/kWh. However, this operating cost cannot define the market price because at this price demand would exceed the available supply, and the system operator must protect the system by maintaining a balance of supply and demand.

In this case, the result is to turn to those customers who have set a limit on how much they are willing to pay for electric energy at that hour. This short-run demand bidding defines the demand curve which allows the system operator to raise the price and reduce consumption until supply and demand are in balance. In Figure 7 this new balance occurs at the point where

Figure 7



the market price of electricity is set at 6 ¢/kWh. Once again, the customers who actually use the electricity pay this 6 ¢/kWh for the full 300 MW of load at that hour. All the generators who sell power receive the same 6 ¢/kWh, which leads to operating margins of 2 ¢/kWh for New Gas, 3 ¢/kWh for New Coal, and 4 ¢/kWh for Old Nuke.

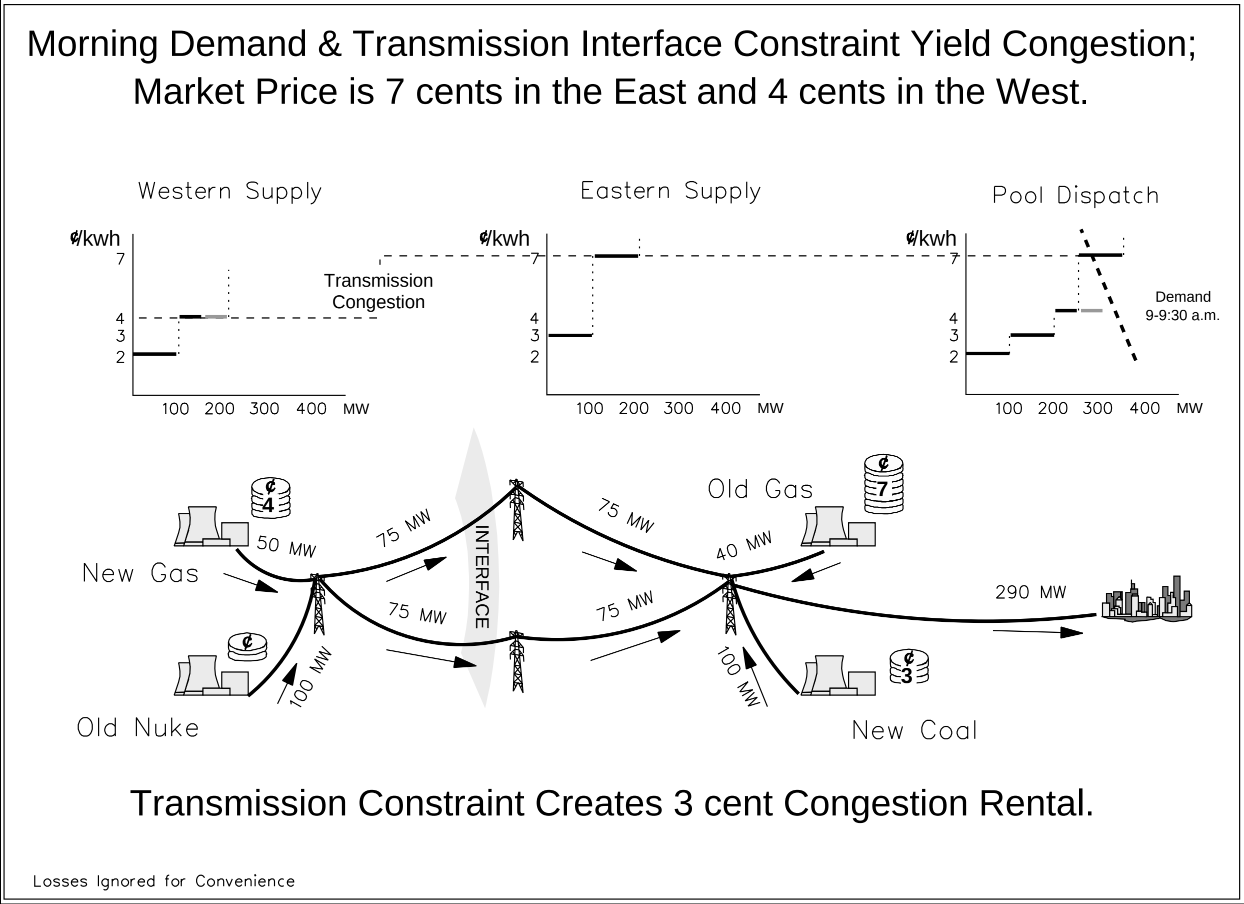
Once again, the pool-based dispatch in Figure 7 depends on excess capacity in the transmission system. The plants in the western region are running at full capacity, and the full 200 MW of power moves along the parallel paths over the grid to join with New Coal to meet the demand in the east. There is a single market price of 6 ¢/kWh, and there is no charge for transmission other than for losses, which are ignored here for convenience in the example.

Transmission Constraints

With the plants running at full capacity, there might be a transmission constraint. To illustrate the impact of a possible transmission limit, suppose for sake of discussion that there is an "interface" constraint between west and east. According to this constraint, no more than

150 MW of power can flow over the interface.

Figure 8

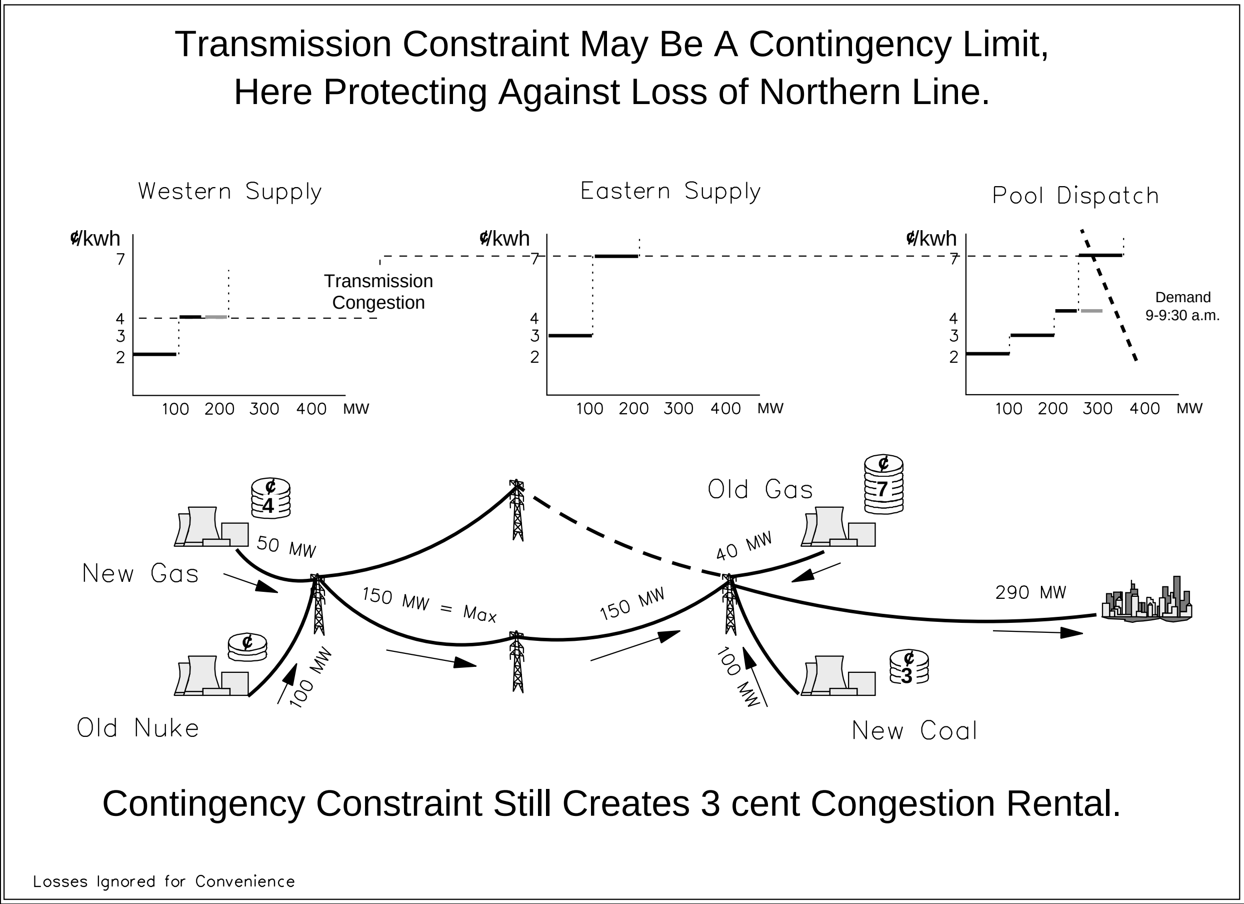


As shown in Figure 8, this transmission constraint has a significant impact on both the dispatch and market prices based on short-run marginal costs. In Figure 8 the level of demand from the city in the east is assumed to be the same as in the case of Figure 7. However, now the pool-based dispatcher faces a different aggregate market supply curve. In effect, only half of the New Gas output can be moved to the east. To meet the demand, it will be necessary to simultaneously turn off part of the New Gas output and substitute the more expensive Old Gas generation which is available in the East. This new dispatch increases the market price in the east to 7 ¢/kWh and necessarily induces a further reduction in demand, say to a total of 290 MW. The New Coal and Old Gas plants receive this full price of 7 ¢/kWh for their 140 MW, which provides a 4 ¢/kWh operating margin or short-run profit for New Coal and allows Old Gas to cover its operating costs.

In the western region, however, a different situation prevails. The transmission interface constraint has idled part of the output of the New Gas plant. Clearly the market price in the west can be no more than the operating cost of the plant. Likewise, since the plant is running at partial output, the market price can be no less than the operating cost of 4 ¢/kWh. This is the price paid to New Gas and Old Nuke, which covers New Gas operating costs and provides Old Nuke an operating margin of 2 ¢/kWh.

The 3 ¢/kWh difference between the market price in the east and the market price in the west is the opportunity cost of the transmission congestion. In effect, ignoring losses, the marginal cost of transmission between west and east is 3 ¢/kWh, and this is the price paid implicitly through the transactions with the system operator. Electricity worth 4 ¢/kWh in the western region becomes worth 7 ¢/kWh when it reaches the eastern region.

Figure 9



The transmission "interface" constraint is a convenient shorthand for a more complicated situation handled by the pool-based dispatchers. The interface limit depends on a number of conditions, and can change with changing loads. Typically it is not the case that there is a 75 MW limit on one or both of the parallel lines through which power is flowing in the grid. In normal operation, it may well be that the transmission lines could individually handle much more flow, say 150 MW each or twice the actual use. At most normal times, the lines may be far from any physical limit. However, the pool-based dispatchers must protect against contingencies--rare events that may disrupt operation of the grid. In the event of these contingencies, there will not be time enough to start up new generators or completely reconfigure the dispatch of the system. The power flow through the grid will reconfigure immediately according to the underlying physical laws. Hence, generation and load in normal times must be configured, and priced, so that in the event of the contingency the system will remain secure.

For instance, suppose that the thermal capacity of the transmission lines is 150 MW, but the pool-based dispatchers must protect against the loss of a northern transmission line. In this circumstance, the actual power flows may follow Figure 8, with 75 MW on each line, but the pool-based dispatchers must dispatch in anticipation of the conditions in Figure 9. Here the northern line is out, and in this event the flow on the southern line would hit the assumed 150 MW thermal limit. This contingency event may never occur, but in anticipation of the event, and to protect the system, the system operator must dispatch according to Figure 9 even though the flows are as in Figure 8. In either case, the transmission constraint restricts the dispatch and changes the market prices. The price is 4 ¢/kWh in the west and 7 ¢/kWh in the east, with the 3 ¢/kWh differential being the congestion-induced opportunity cost of transmission. This "congestion rental" defines the competitive market price of transmission.

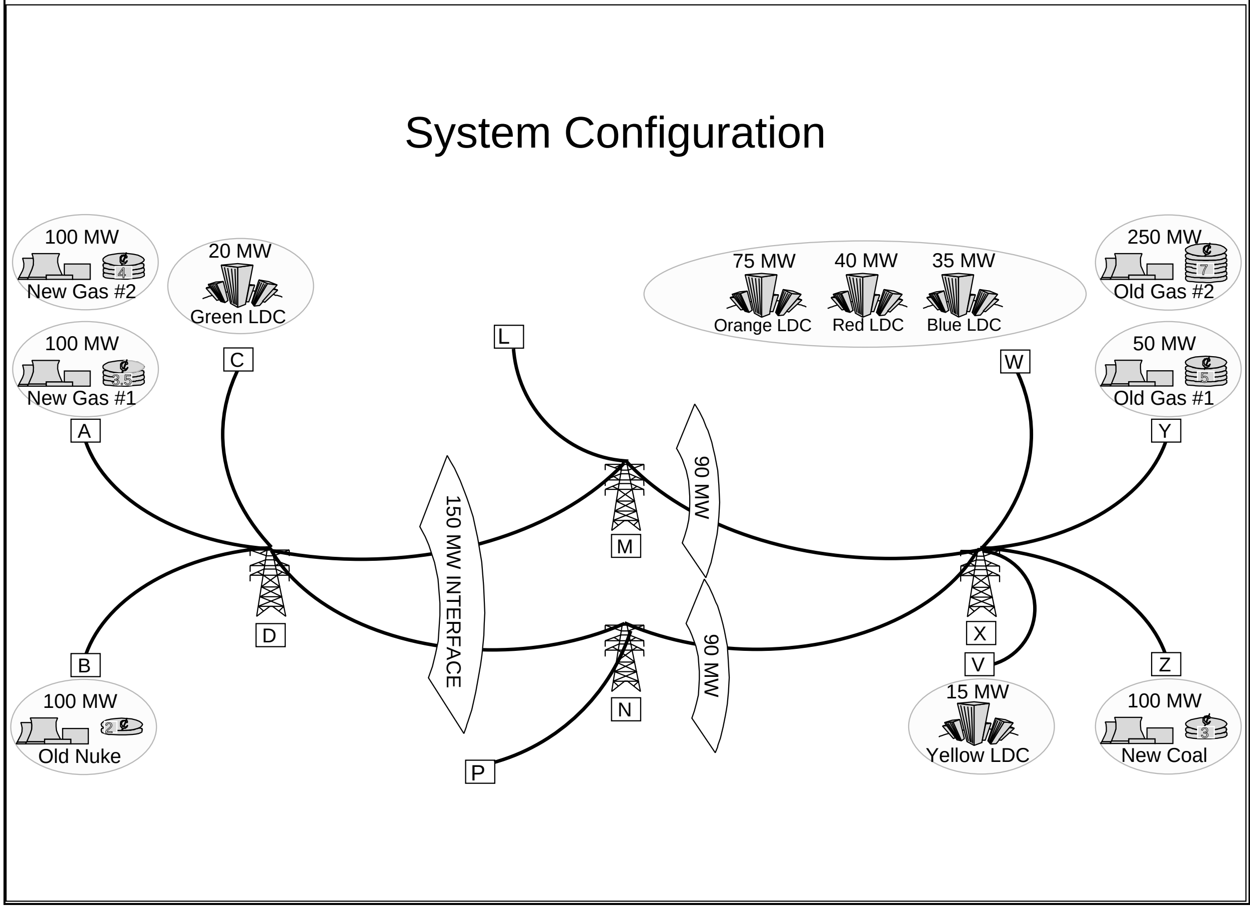
Buying and selling power at the competitive market prices, or charging for transmission at the equivalent price differential provides incentives for using the grid efficiently. If some user wanted to move power from east to west, the transmission price would be negative, and such "transmission" would in effect relieve the constraint. The transmission price is "distance- and location-sensitive," with distance measured in electrical rather than geographical units. And the competitive market prices arise naturally as a by-product of the optimal dispatch managed by the system operator.

The simplified networks in Figure 6 through Figure 9 illustrate the economics of least-cost dispatch and locational prices. However, these networks by design avoid the complications of loop flow that can be so important in determining prices and creating the difficulties with physical transmission rights. These examples differ from the single line case only in the explicit representation of the parallel flows on the lines, but as yet this has no effect on the prices. The extension of these examples and the basic pricing properties to more complicated networks includes the possibility of inputs and load around loops in the system. Here assume a transmission system as before but with the basic available generations and loads as shown in Figure 10. Our attention will focus on the prices at L-M and N-P, where the introduction of generators and load will reveal the impacts of the loops. The generators in Figure 10 define a basic supply configuration with quantities and prices, coupled with the associated loads, and all with the following characteristics:

- Generation available at four locations in the East (Y, Z) and West (A, B).

- Load in the East, consisting of the Yellow LDC at V and the Orange, Red and Blue LDCs at W.

Figure 10

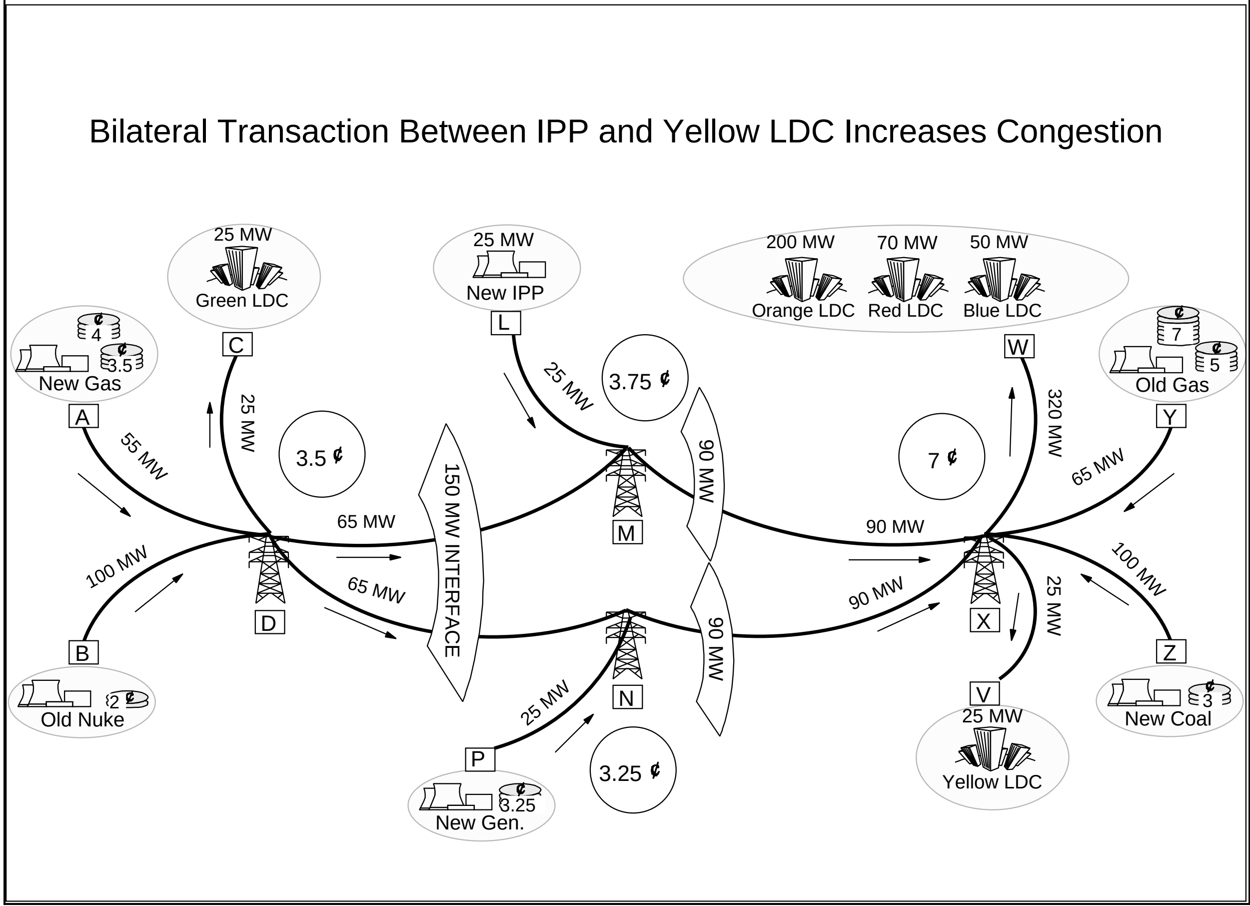


- Load in the West, consisting of a Green LDC at C.
- Interface constraint of 150 MW between bus D and buses M and N.
- Thermal constraints of 90 MW between M and X and between N and X.
- The New Gas and Old Gas generating facilities each consist of two generating units whose marginal costs of production differ.

Loads in Figure 10 are illustrative and will vary systematically in each example. For convenience, losses are ignored in all examples.

The first example to introduce the effect of loop flow involves a new sources of supply at a location on the loop. Here a low cost, large capacity generator becomes available in Figure 11 at bus "P." An IPP at bus "L" has bid in a must run plant at 25 MW, having arranged a corresponding sale to the Yellow distribution company at bus "V". Were it not for the IPP sale, more power could be taken from the inexpensive generators at bus "P" and at bus "A".

Figure 11

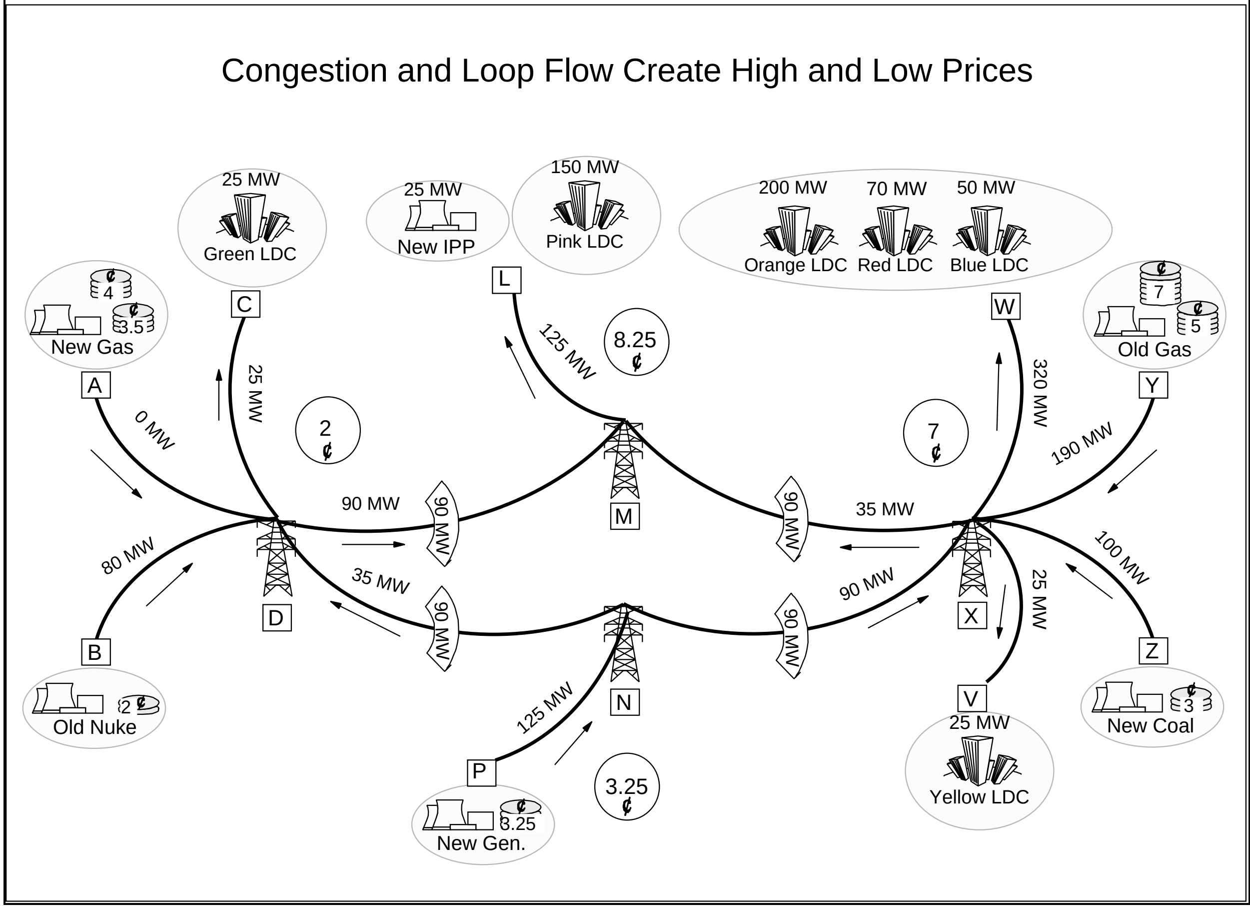


However, because of the effects of loop flow, these plants are constrained in output, and there are different prices applicable at buses "D", "M", "N", and "X".

In a further example the constraints are modified to replace the interface limit with limits on the flows on individual lines. Here every line in the main loop is constrained by a thermal limit of 90 MW, replacing the interface limit. With these constraints in Figure 12, an added load of 150 MW at bus "L" alters the flows for the market equilibrium. In this case, the combined effect of the increased load and the constraints leads to a price of 8.25¢ per kWh at bus "L". This illustrates that it is possible to have market clearing prices at some locations that are higher than the 7¢ marginal running cost of the old gas plant at bus "Y", the most expensive plant in the system. The interaction of the network constraints is such that with a reduction of load at bus "L" it would be possible to reduce out put of the most expensive plant by even more, and make up the difference with cheaper sources of supply, causing the high price for load at bus "L".

Changing the network further adds new loops and even more examples of the effect on prices and dispatch caused by the network interactions. In this case, a new line has been

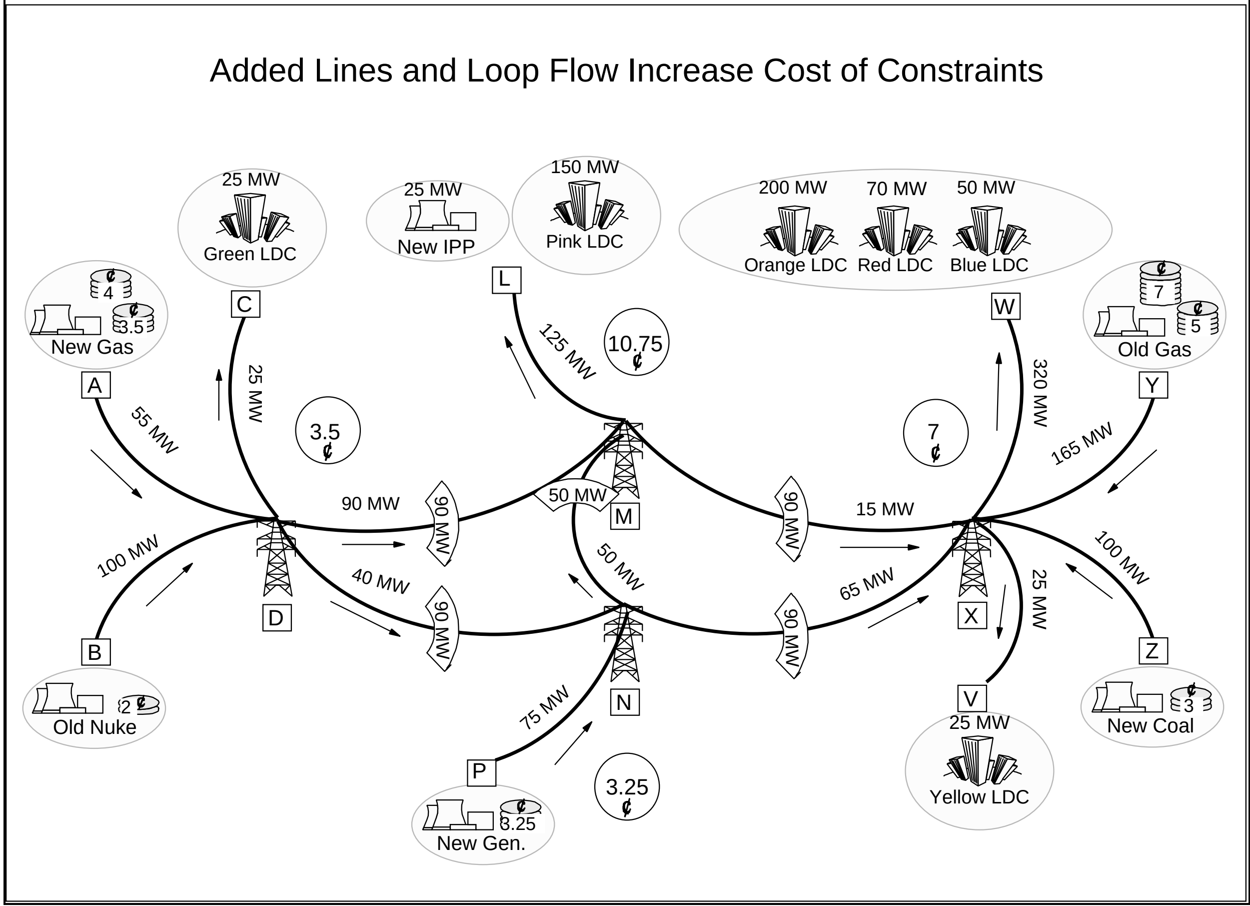
Figure 12



added to the network in Figure 13, connecting bus "N" to bus "M". This line is assumed to have a thermal limit of 50 MW. The new line adds to the capability of the network in that the new pattern of generation lowers the overall cost of satisfying the same load. The total cost reduces from $20,962.50 in Figure 12 to $19,912.50 in Figure 13. Although the average cost of power generation fell, the marginal cost of power increased at bus "L", where the price is now 10.75¢ per kWh. The new loop provides more options, but it also interacts with other constraints in the system. This set of interactions is the cause of the high price as it appears at bus "L".

As a final example that confirms the sometimes counterintuitive nature of least-cost dispatch and market equilibrium prices, add a new bus "O" between bus "M" and bus "N" in Figure 14, and lower the limit to 30 MW between bus "O" and bus "M". Bus "O" has a small load of 15 MW. The increased load of 15 MW at bus "O" actually lowers the total cost of the dispatch, as reflected in the negative price. Each additional MW of load at bus "O" changes the flows to allow a dispatch that lowers the overall cost of meeting the total load. The optimal solution would be to pay customers at "O" to accept dump power, thereby relieving congestion elsewhere and providing benefits to the overall system.

Figure 13

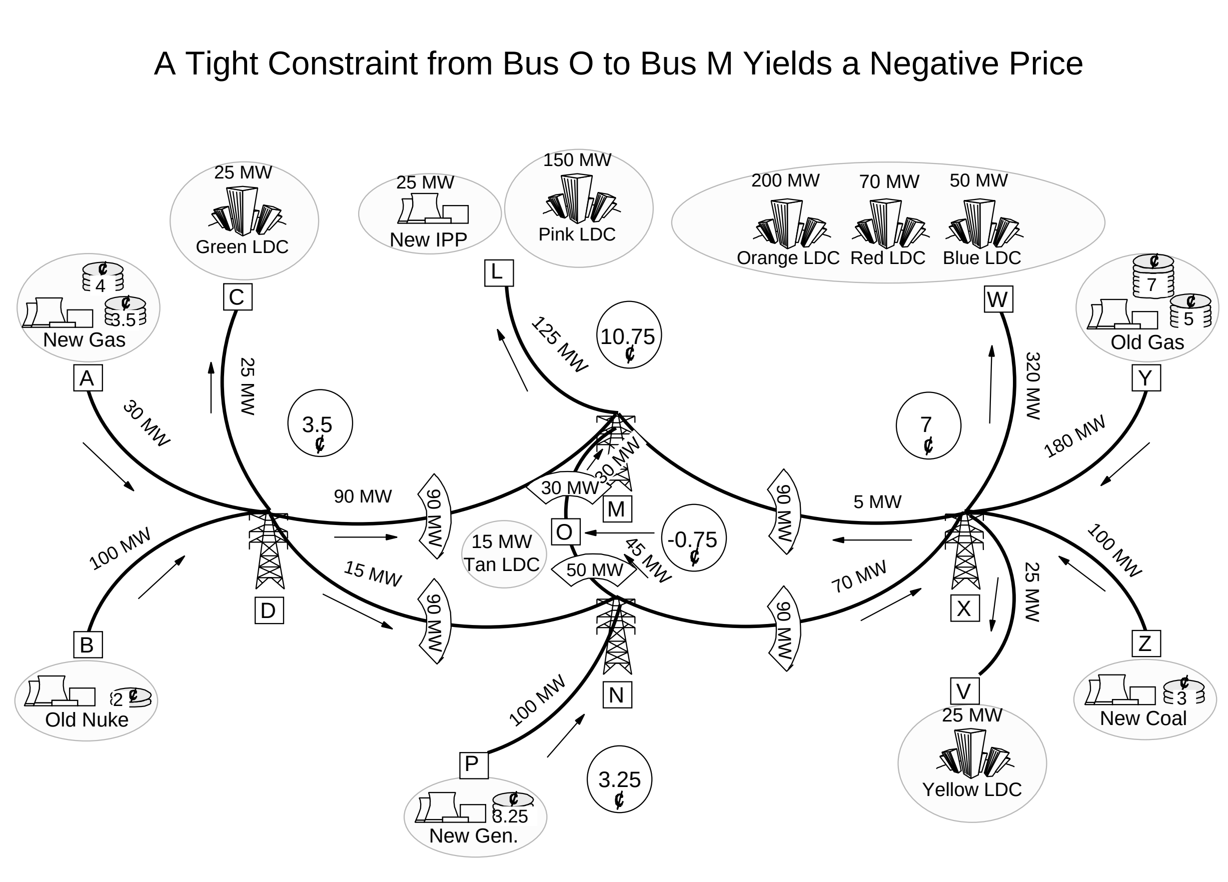


This final example, therefore, illustrates and summarizes the types of interactions that can develop in a network with loop flow. Power can flow from high price nodes to low price nodes. The competitive market clearing price, equivalent to the marginal costs for the least-cost dispatch, can include simultaneously at different locations prices higher than the cost of the most expensive generation and lower than the cost of the cheapest generation source. Application of the principle of locational pricing implies that transmission congestion would lead to many prices. Even with only a single constraint, there could be a different price at each location.

Zonal Versus Nodal Pricing

The use of locational prices has been described as being too complex, with the implication that an alternative approach would produce a simpler system. A common response to this assertion has been to recommend a "zonal" approach that would aggregate many locations into a smaller number of zones. The assumption has been that this would tend to reduce complexity. However, in the presence of real constraints in the actual network, the zonal

Figure 14



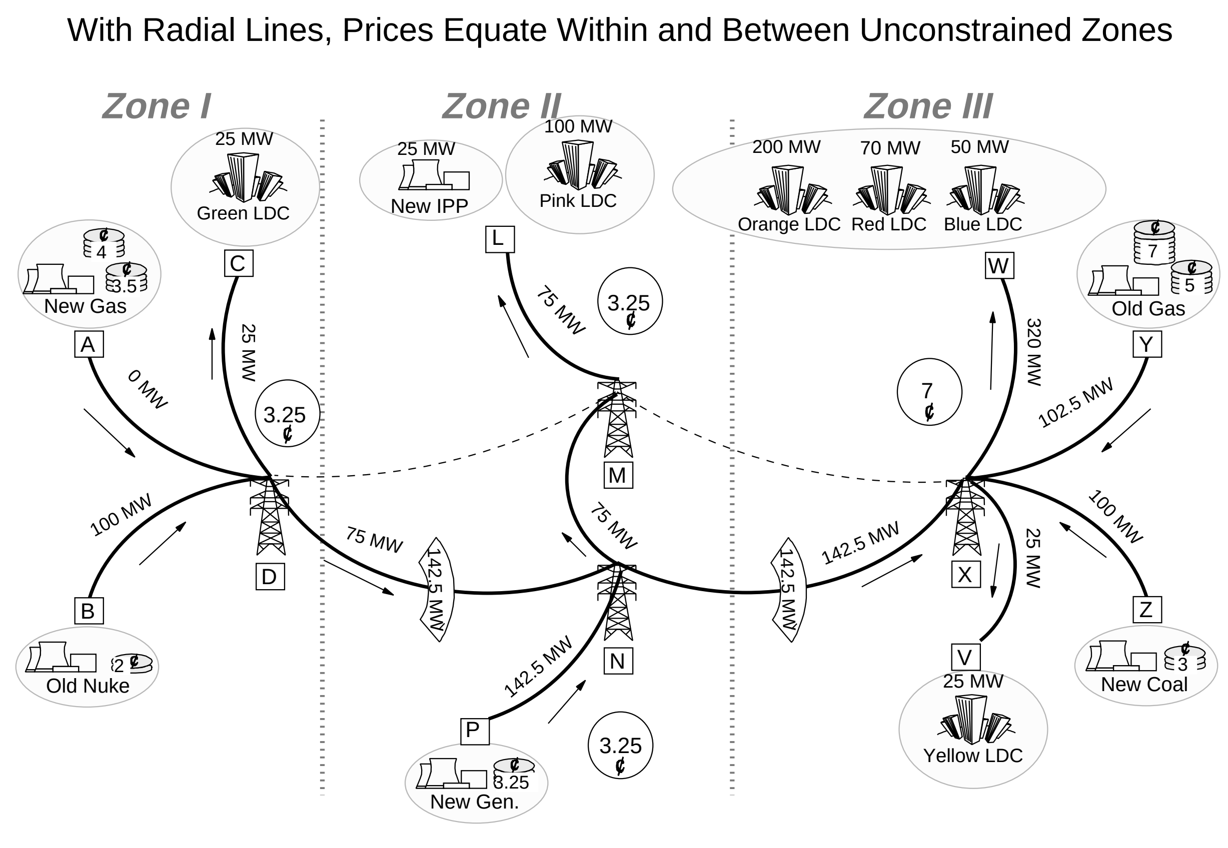
approach may not be as simple as it might appear without closer examination.[31]

The difficulties would arise in the context of a competitive market where participants have choices. If the actual operation of the network system does not conform to the pricing and zonal assumptions, there will be incentives created to deviate from the efficient, competitive solution. In the presence of a vertical monopoly that can ignore the formal pricing incentives, this has not been a problem. But under the conditions of a market, where participants will respond to incentives, the complications created by a zonal approach may be greater than any complications that would exist with a straight locational approach to pricing and transmission charging.

Consider the simplified example in Figure 15. The network has been constructed so that there are only radial connections. With strictly radial connections, locations within and

[31] For a similar analysis with similar conclusions, see Steven Stoft, "Analysis of the California WEPEX Applications to FERC," Program on Workable Energy Regulation, University of California, PWP-042A, October 15, 1996.

Figure 15

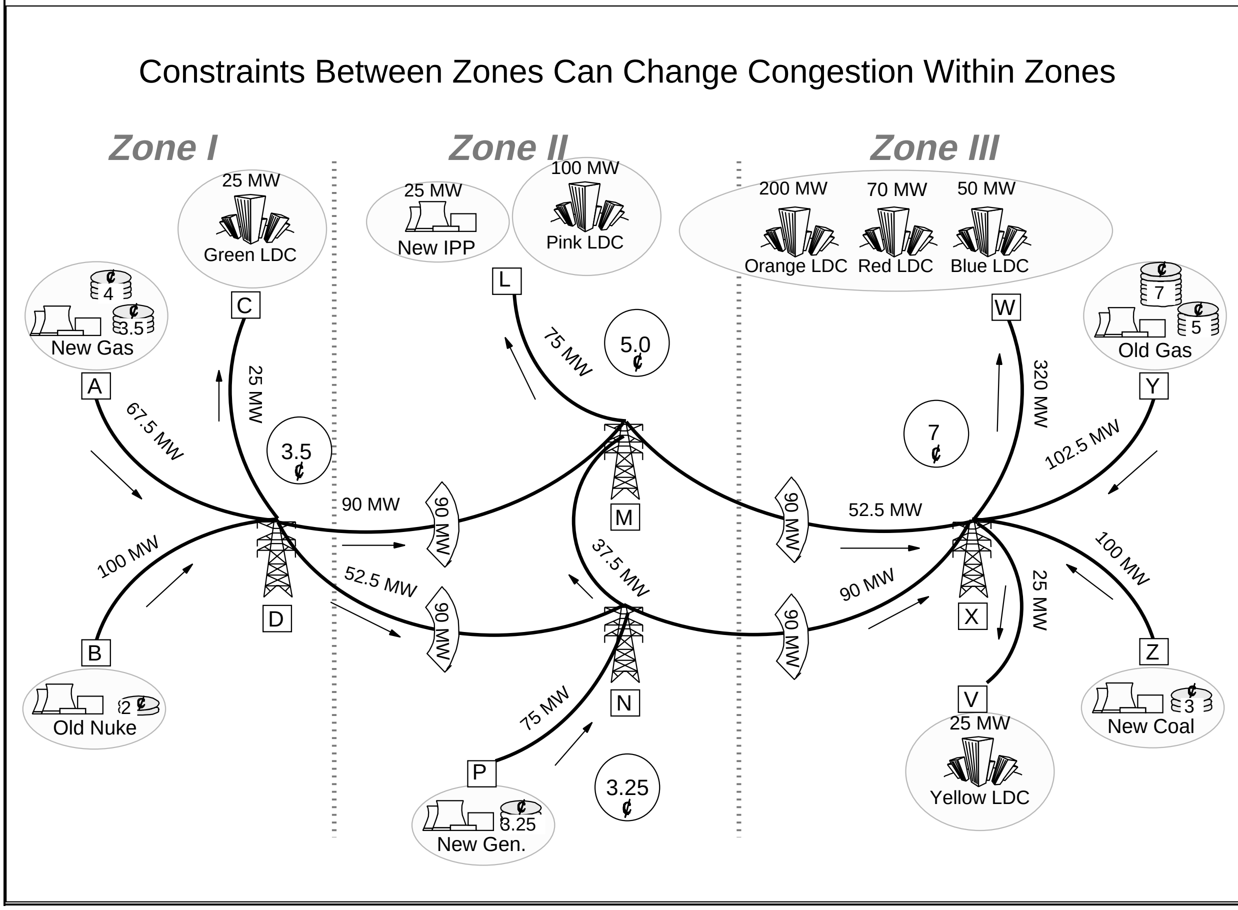


between unconstrained zones would have a common price. Hence, aggregation of locations offers an apparent simplification by reducing to a few distinct zones. This motivation from a typical radial examples leads to the assumption that in general there could be areas in a real network that would have the same prices and, therefore, these locations could be aggregated into zones that would be simpler for participants in market operations.

There are two problems with this line of argument. First, if the multiple locations truly do have the same prices, then there is no need to aggregate into zones. The point of the aggregation was to reduce the number of prices, and in the case where the assumption of common prices holds, aggregation would be unnecessary.

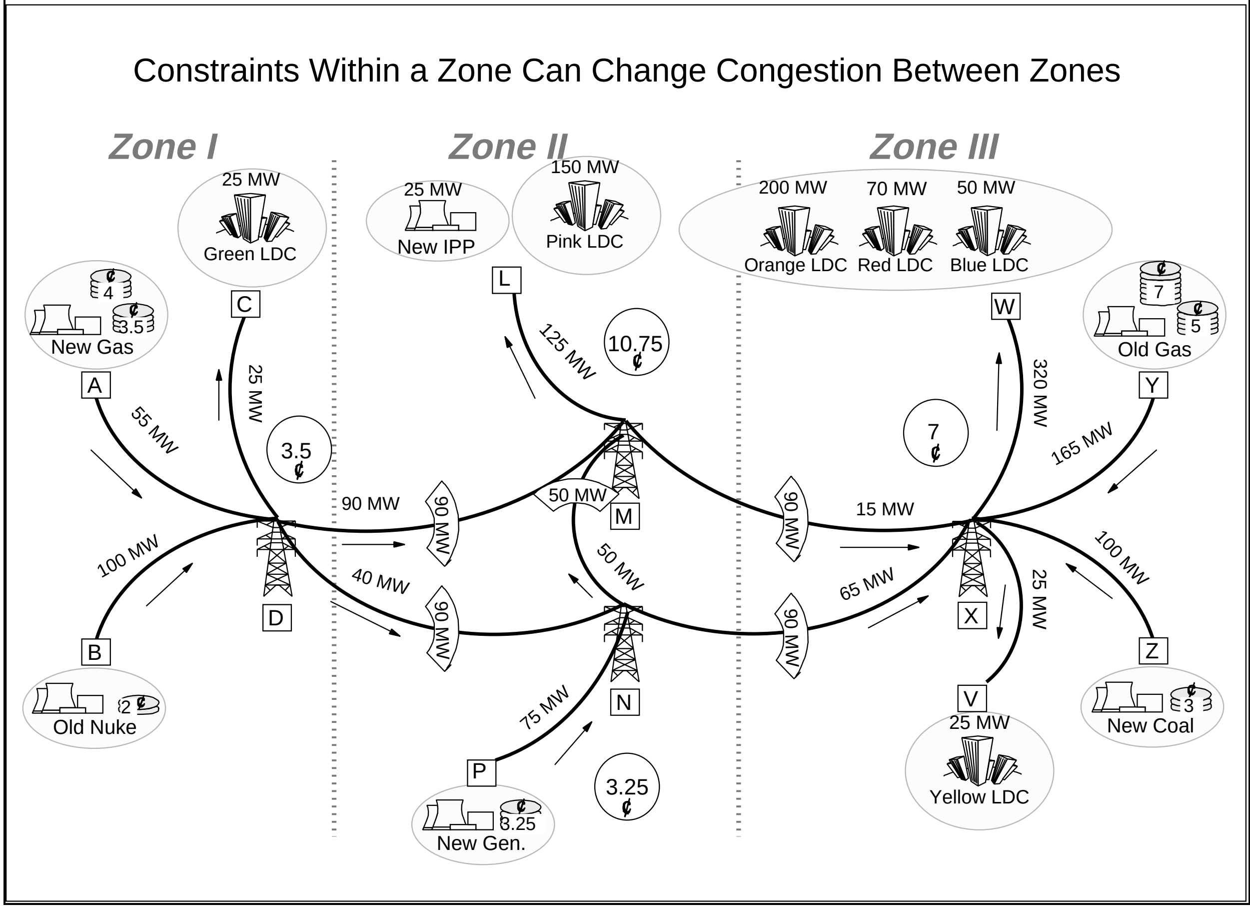
Second, the definition of a zone, which appears easy in the case of a radial network, becomes more problematic in the case of a more realistic network with loop flows. The radial examples can be a poor guide to thinking about interactions in networks. For example, it is often argued, or assumed, that congestion or differences in prices between zones would be caused only by transmission constraints that could be defined for lines that connect the zones. Furthermore, it is often assumed that differences in prices within zones can only be caused by congestion on lines within the zone. Under these simplifying assumptions, therefore, it is assumed that zones can be well defined and that what happens within a zone can be treated independently of what happens between zones, or independently of what happens in other zones. When we move beyond the radial examples, however, these assumptions and the associated conclusions can be false.

Figure 16



With the more typical case of loops in a network, prices could differ within and between "unconstrained" zones due to the indirect effects of "distant" constraints. Consider the slightly modified example in Figure 16. In this case, the zones developed from the radial analogy produce a very different outcome from the assumptions derived from the radial case in Figure 15. In this example, the prices within "Zone II" differ, but there is no binding constraint in the zone. The lines within the zone are operating below their thermal limits. The difference in prices between buses M and N arises not due to constraints within the zone but because of the loop flow effects interacting with the binding constraints between the zones. Apparently the determination of prices within a zone can not be made independent of the effects on constraints outside the zone.

Figure 17



A symmetric result appears in Figure 17 with a different pattern of loads and flows. In this case, there is no constraint binding between Zones II and III, but the price in Zone III differs from the prices in Zone II. Again this effect cannot be seen in radial networks, but it is easy to create in real networks with loop flow. The price in Zone III differs from all the other prices in part because of the interaction with the constraints in Zone II. In a sufficiently interconnected network, these examples suggest that a wide variety of pricing patterns would be possible. In fact, with loop flow, it is possible for a single binding constraint to result in different prices at every location in the system, reflecting the fact that every location has a different impact on the constraint.

Aggregation into zones may add to complexity and distort price incentives. The assertion that conversion to zones will simplify the pricing problem is not supported by analysis of the conditions that can exist in a looped network. Furthermore, aggregating networks presents a number of related technical problems that follow from the fact that exact aggregation requires first knowing the disaggregated flows. In other words, the first step in calculating consistent aggregate flows and prices is to calculate the disaggregated flows and prices. Hence aggregation produces no savings in computation, and no additional simplicity. If no price dispersion exists, no aggregation is necessary. And if price dispersion does exist, aggregation only sends confused price signals. In the end, the simplest solution may be to calculate and use the locational prices at the nodes, without further aggregation.

NODES AND ZONES FOR SHORT-RUN PRICING

With this somewhat tedious review of locational pricing principles available, we can turn to the issue of the "simplification" offered by the proposal to implement short-run pricing by aggregating into nodes into zones. The approach is to postulate and then examine a series of canonical arguments or questions.

If Zones are Defined by Nodes with Common Prices, Why Bother?

The definition of a zone is sometimes offered as a group of locations that would have the same locational price. As we have seen from the above examples, for this to be literally true the connections must be radial connections. In a sufficiently interconnected system, with parallel flows, there will be different prices across any collection of locations, even with constraints that appear to be external to the putative zone. Hence, the case of true equality of the locational prices would be a limited and special set of conditions.

The more general argument would be to aggregate locations where the locational prices do not differ by very much. There may be collections of such locations, but with this definition the natural question arises as to what is the need for the zones? If the locational prices differ by only trivial amounts, then locational pricing and zonal aggregation produce the same end result. There would be no need to aggregate.

Hence, it appears that the real application of zonal aggregation must be in situations where the underlying rationale is compromised. In other words, zonal aggregation would be interesting only in those situations where the aggregation results violated the premise of the creation of the zone. In particular, zonal aggregation produces a material difference from locational pricing only in those cases where the locational prices being aggregated are materially different, in contradiction of the original justification for the definition of the zone.

How Would We Define the Zonal Prices?

If the real application of zones is important only when there is a material difference in the locational prices, then there must be some rule specified for determining the price in the zone. The answer is not obvious. Here the intuition derived from the analysis of the single radial transmission line connection can mislead. Contrary to the case of the single line, the price in the zone is not determined by the highest cost generator operating in the zone. Rather, as illustrated above, it is entirely possible for the true price to be more than the operating cost of the most expensive generator in the zone, driven by the effects of loop flow and the interaction with transmission constraints.

The usual proposals for price aggregation involve some form of averaging. A typical approach, as in England and Wales, is to determine a hypothetical unconstrained price in the zone, which is then charged to all customers and paid to all generators. For those generators that are needed but operate a locations with truly higher prices, they would be paid their opportunity costs. The difference above the "unconstrained" price is collected in an uplift that is averaged over all customers.

Any such rule immediately raises the question of what happens to the constrained off generators, those that would have run at the unconstrained price but whose true locational value is less than the unconstrained price. Clearly they would want to operate, but they cannot be allowed to run because of the transmission constraint.

The solution to this problem in England and Wales has been to pay the generators the profit they would have made if they had run and received the unconstrained price, with the costs added to the uplift and collected again from all customers. This adds to the total cost as seen by all customers and has the perverse effect of providing and incentive to build or maintain generation in locations where there is excess capacity.

In addition, the English pool prohibits bilateral transactions to avoid the problem of the constrained-off generators going around the price averaging system. This obviously limits the flexibility of the market and has been a principal source of complaint by market participants in that system. However, the importance of such rules was dramatically illustrated by the events in the Pennsylvania-New Jersey-Maryland Interconnection (PJM) system starting in June 1997.

During June, the PJM system was operating under an interim tariff provision that followed English zonal model with two critical exceptions. The PJM market was treated as one zone, with a hypothetical unconstrained dispatch setting the so-called market clearing price. In the event of transmission congestion, some generators bidding into the pool were constrained on while others were constrained off. The cost of the more expensive generation was rolled into an average congestion charge that applied to all loads. So far this was the same as the English system. The difference is that constrained-off generators were not compensated for their lost profits, and non-firm bilateral transactions were allowed at only the price of the average congestion cost uplift.

The result when the system because constrained was predictable and predicted. Some low cost generators bidding into the pool were constrained off. The corresponding loads would then be charged the "unconstrained" price of 2.9 cents plus the uplift. However, the marginal cost of the constrained off generators was as low as 1.5 cents. If the same loads arranged a bilateral transaction with these constrained off generators, who would now withdraw from the pool dispatch, the loads could have the energy for this 1.5 cents price plus the same uplift. Given this incentive, this is precisely what happened. The loads and constrained-off generators arranged bilateral transactions that brought the generators back on line. This, in turn, forced the ISO to back off other generators, who then faced the same incentive to leave the pool and schedule themselves.[32]

The end state of this downward spiral was that the ISO was left with virtually no controllable generation to redispatch in order to respect the transmission constraints. Under the existing tariff provisions, there was nothing the ISO could do to correct the perverse pricing incentives. Hence, the ISO invoked a "temporary fix" by immediately modifying the rules to prohibit such bilateral transactions, which it did with subsequent FERC approval.[33]

The net effect was to subsidize consumers in the high price areas, by charging more to consumers in low price areas, and to remove the flexibility for bilateral transactions that was an objective of the market restructuring and would be a natural result of locational pricing. Apparently, the zonal simplification contained a hidden complexity.

Would Locational Prices Be Hard to Calculate and Come from a Black Box?

The locational prices would be determined by the actual dispatch, which makes the problem simple. The computations are easy, and have been available for years in power pools; they just haven't been used for pricing purposes. Calculating locational marginal costs for the actual dispatch is easier than the familiar and widely used split-savings methodology. Furthermore, since locational pricing is already done (almost in full) in Argentina, Chile, New Zealand, and Norway, there is a demonstration that the technical computation is straightforward.

Once the method is explained, system operators always say the prices could be computed easily. Part of the misunderstanding on this point is the distinction between determining an economic dispatch (difficult) and determining the prices given the dispatch (easy). The hard part in dispatching is both unavoidable and already done. The easy part of calculating the prices is a detail. At a FERC Technical Conference, the ISO for PJM explained how to calculate the prices and described the software which is operating in parallel to determine the prices.[34] An independent auditor verified that the system was understood and auditable.[35]

---

[32] PJM Supporting Companies, "Motion to Intervene and Comments of the PJM Supporting Companies in Support of PJM interconnection, Inc. Filing." FERC Docket No. ER97-3463-000, p.5.

[33] PJM Interconnection, Inc., FERC Section 201 filing, June 27, 1997.

[34] Pennsylvania-New Jersey-Maryland Interconnection, "Responses to FERC's March 28 Questions on Implementation of Locational Marginal Pricing," FERC Docket Nos. OA97-261-000 and ER97-1082-000, April 14, 1997.

[35] Price Waterhouse comments, FERC Technical Conference, Docket Nos. OA97-261-000 and ER97-1082-000, May 9, 1997.

This brings us to the issue of the perception and comprehension of the market participants. At the moment the majority of market participants would claim that the idea of using locational prices is too complicated. However, the view of the moment should not be all that concerns us. So far, every simple alternative proposed has turned out to be pretty complicated, once the implications of the full package unfolded to include the extensive regulatory rules needed to negate the incentives of incorrect prices.

Would It Be an Easy Matter to Set and Later Change the Zonal Boundaries?

The rationale for zones rests in part on the assumption that the zones would be easy to define and would be stable for long periods. However, when conditions changed, the zones would be redefined to come back into compliance with the original definition that there would be no difference in locational prices within the zones.

Each of these points raises a number of complications that must be recognized. First, it is not so obvious where the zonal boundaries should be set. For example, recent PJM "[O]perating data show that, during the past 14 months, 70 percent of the out-of-merit costs for transmission control in PJM resulted from thermal contingencies."[36] These thermal limits are exactly the type of constraints that create the looped interactions as illustrated in the previous sections by the numerous examples. They are not typically radial lines, and the impact of the constraints give rise to different prices throughout the system.

If the zones are not stable, then there would be again little or no distinction between the prices reported by a locational pricing system and the zonal prices. In Norway, for example, the system is described as a zonal system, but the system operator can and does change the definition of the zones daily or hourly. Hence, the Norwegian systems is more like a locational pricing system.

If the zones are intended to last for extended periods, but change when there is a material and sustained difference in locational prices within the zone, then a number of other complications arise. For instance, it would still be necessary to calculate the locational prices on a regular basis just to evaluate the suitability of the zonal definition. This means that there will be regular information available that some people are being subsidized and other people are paying the subsidy required by the zonal configuration. The reconfiguration step, by definition, amounts to rearranging the pattern of these subsidies just when the threshold criteria indicates that the reconfiguration really matters.

Since establishment and reconfiguration of the pattern of subsidies will depend on extensive analysis of prospective conditions, there will be many assumptions and points of debate about what the appropriate boundaries should be next week or next year. Although the

---

[36] PJM Supporting Companies, "Motion to Intervene and Motion to reject CCEM Filing Submitted by the PJM Supporting Companies," FERC Docket Nos. OA97-261-000 and ER97-1082-000, July 23, 1997 p. 30.

computational challenge of determining locational prices for the actual dispatch is trivial, the process of forecasting these prices is another matter entirely, one that promises to be controversial. At the risk of understatement, there is little in past regulatory experience which gives confidence that this creation and rearrangement of subsidies will be either swift or simple. Policy makers who think that zonal aggregation and cross subsidies will simplify the process should look again.

Is Transmission Congestion a Small Problem?

To argue that transmission congestion is and will be minor is to argue that there should be no interest in gaining transmission rights. Given the keen interest in tradable transmission capacity rights, the behavior of the market participants already contradicts the assertion that this is a minor issue. Furthermore, if it is a minor issue, then the locational prices will not differ most of the time, except for losses, and nothing could be simpler than this outcome. Even if congestion costs were small in the past under the regime of vertically integrated utilities, the incentives will be different in the competitive market where customers have choices. As seen elsewhere, small differences in costs could be a large part of the profit on a transaction, and would lead to substantial differences in behavior. If we give market participants choices, such as between pool and bilateral transactions, it will be important to get the prices right.

Furthermore, the improving understanding of the importance of this matter indicates that when the constraints do apply, the price differences can be surprisingly large. In August for example, the reports were that PJM single zone again was operating with "dispatch rates," which would be similar to the locational prices if they were being charged, that were $89 per MWh in the constrained-on regions and $12 per MWh in the constrained-off regions of the zone. The "widely differing dispatch rates were repeated for several days last week."[37] When constraints bind, therefore, the incentives created can be much larger than most people imagine. If participants were given the choice and flexibility that we think of as appropriate for the competitive market, these incentives would overwhelm the system as long as the prices charged diverged from the underlying locational marginal prices.

Would Zonal Pricing Mitigate Market Power?

To the extent that there is a high concentration of control of generation or load, there will continue to be a potential for an exercise of market power. This potential creates demand for continued regulatory oversight. The analysis of market power in the face of significant

[37] Power Markets Week, September 1, 1997, p. 13.

transmission constraints is a broader subject.[38] However, an advantage of the market model with opportunity cost pricing at locations is the ability to expand the range of options available to address potential problems of market power without compromising other goals in the development of a competitive electric market.

This argument appears counterintuitive at first glance, and there would appear to be advantages to aggregation into zones. As the argument goes, the use of locational pricing would imply small local markets. By contrast, it seems logical that aggregation into zones would expand the geographic scope of the market and bring more actors into competition, thereby mitigating market power.

If the separation into local markets and locational marginal cost differences were simply an artificial institutional constraint, there might be something to this story, especially if the local generators were not competing with other generators in the network. In the present case, however, the facts are different. The constraints are real, and aggregation into zones would not remove the transmission constraints. Aggregation into zones would be likely to hide the market power and remove some of the most important limitations on market power; namely, the demand side response and the ability of new entrants to challenge the dominant generator.

Under locational pricing, the ISO provides open access to the grid at opportunity cost prices. This unbundles the system and eliminates vertical market power. Horizontal market power arises from concentration of ownership of generation plants. The auction mechanism in the bid and dispatch system does not create market power; a dominant firm would not need the auction to manipulate market prices. Furthermore, compared to charging locational marginal cost prices, all the alternatives involve some form of price averaging, which would both enhance and hide horizontal market power.[39] With locational pricing, customers at the location would face the higher price and this would create two beneficial incentives. First, customers would have an incentive to reduce their demand and thereby weaken the power and profits of the dominant firm. Second, customers would have an incentive to sign long-term contracts with new entrants that would support entry and mitigate market power. With zonal aggregation, however, these incentives would be removed or substantially attenuated. The generator with market power would still be paid a high price. Customers would not see the high price, they would see only an average price spread across those at other locations. In fact, the generator with market power would benefit from this disguise. By confronting a less responsive demand curve, the generator would see its market power enhanced within the zone. With zonal averaging, new entrants would

---

[38] Judith B. Cardell, Carrie Cullen Hitt and William W. Hogan, "Market Power and Strategic Interaction in Electricity Networks," Resource and Energy Economics., Vol. 19, 1997, pp. 109-137; William W. Hogan, "A Market Power Model with Strategic Interaction in Electricity Networks," The Energy Journal, Vol. 18, No. 4, 1997, pp. 107-141.

[39] If locational prices differ in a zone, the rule might be to charge all customers and pay all generators the highest price in the zone, allowing the scope of the generators market power to expand. However, most zonal proposals are based on some form of averaging to soften the impact of higher prices.

face the problem of entering a market that was subject to manipulation by the dominant generator but have no customers prepared to sign a long-term contract, because no individual customers would see the higher price. Hence, zonal aggregation would increase the need for regulation.

Zonal aggregation would not expand the real geographic scope of competition unless the aggregation rule implied setting all prices at the price of the dominant firm, which would create another set of problems. Hence, locational marginal cost pricing would reduce market power relative to the common zonal alternatives, and locational pricing would make the exercise of market power more transparent.

Can the Market Operate With a Simpler System?

Locational marginal cost pricing lends itself to a natural decomposition. For example, even with loops in a network, market information could be transformed easily into a hub-and-spoke framework with locational price differences on a spoke defining the cost of moving to and from the local hub, and then between hubs. This would simplify without distorting the locational prices. As shown in Figure 18, a contract network could develop that would be different from the real network without affecting the meaning or interpretation of the locational prices.[40]

With the market hubs, the participants would see the simplification of having a few hubs that capture most of the price differences of long-distance transmission. Contracts could develop relative to the hubs.[41] The rest of the sometimes important difference in locational prices would appear in the cost of moving power to and from the local hub. Commercial connections in the network could follow a configuration convenient for contracting and trading. The separation of physical and financial flows would allow this flexibility.

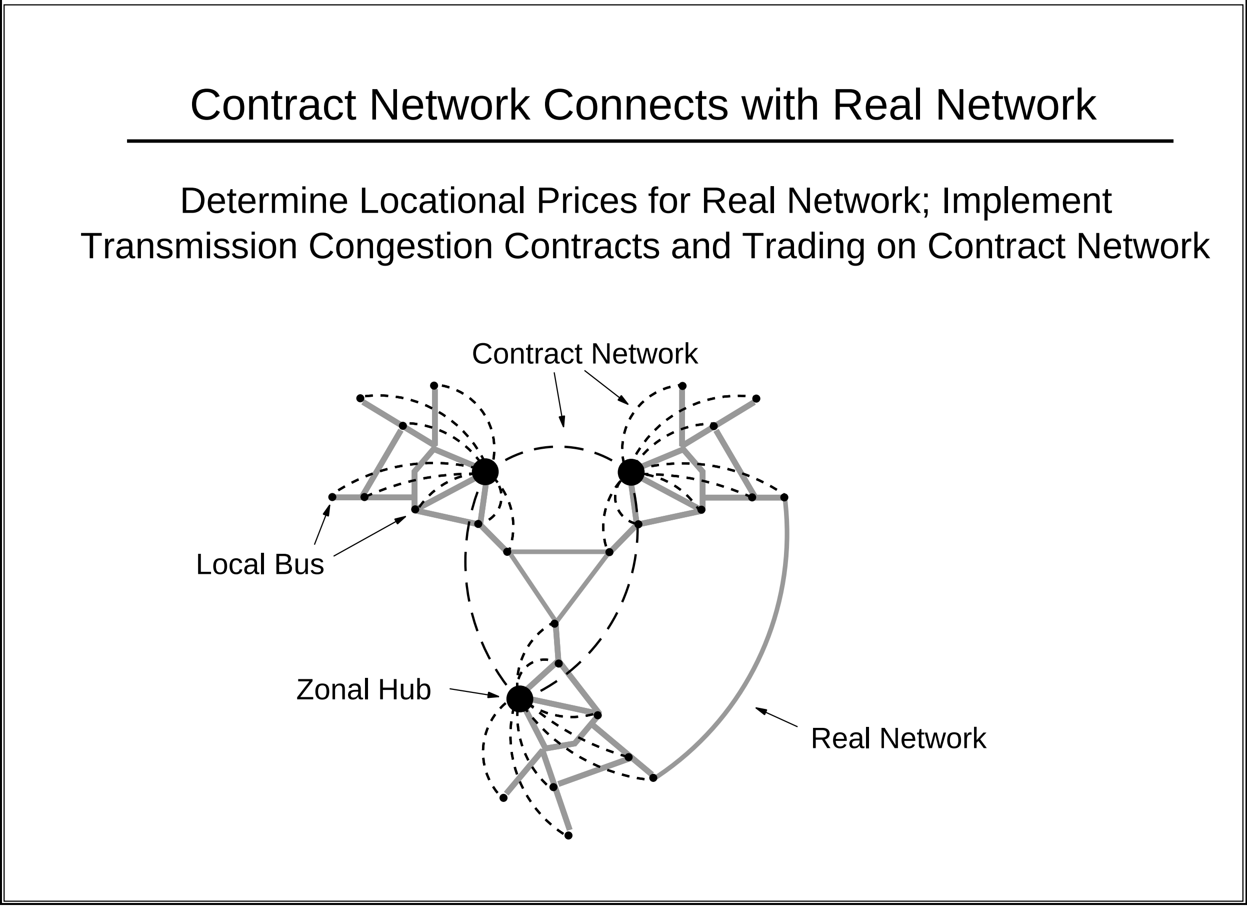
The creation or elimination of hubs would require no intervention by regulators or the ISO. New hubs could arise as the market requires, or disappear when not important. A hub is simply a special node within a zone. The ISO still would work with the locational prices, but the market would decide on the degree of simplification needed. However, everyone would still be responsible for the opportunity cost of moving power to and from the local hub. There would be locational prices and this would avoid the substantial incentive problems of averaging prices. The hub-and-spoke approach appears to give most of the benefits attributed to zones without the costs, and it implies that the ISO works within a locational pricing framework.

---

[40] For further details on long-run transmission congestion contracts, see Scott M. Harvey, William W. Hogan and Susan L. Pope, "Transmission Capacity Reservations and Transmission Congestion Contracts," Harvard-Japan Project on Energy and the Environment, Harvard University, February 1997.

[41] For a similar argument, see S. Walton and R. Tabors, "Zonal Transmission Pricing: Methodology and Preliminary Results form the WSCC," Electricity Journal, November 1996, pp. 40.

Figure 18



A version of this hub-and-spoke system appeared in the discussions among market participants in the PJM system in January 1998. Although the market participants could create hubs without the participation of the ISO, there was a popular request that the ISO identify and post hub prices. In order to reduce the price volatility that might be present in selecting a single location as a hub, the ISO responded to the request to create hubs consisting of a fixed-weight average of a number of underlying locations. Hence, purchases and sales at the hub are equivalent to a portfolio of purchases and sales at the underlying locations, with the portfolio composition in terms of the fixed weights. Likewise, transmission between any location and the hub is the equivalent of transmission between the given location and the portfolio of locations that make up the hub. The PJM ISO began posting prices for such hubs in the preliminary price reporting system that began operation in February 1998.[42] If the market participants find these hubs convenient, then contracts would be written using the composite hub prices as references, and futures contract could be constructed relative to the same spot-market hub prices. However, nothing limits the markets to the hubs selected by the ISO, and anyone is free to utilize another

---

[42] The Internet address is "www.pjm.com/ferc/filings/19971231/lmp_info.html".

hub and organize trades relative to this hub price.

CONCLUSION

The development of ISOs in the United States in early 1998 was at the beginning of the beginning, with major new proposals either approved or under consideration. Implementation was about to occur for some of the most well-known examples. The comparison across regions of the United States highlighted a number of differences in approach. A key distinction was found in the use of locational pricing versus the aggregation of regions into zones. Efficient marginal-cost pricing in competitive electricity markets implies sometimes substantial locational differences in the presence of transmission constraints. Aggregation of individual nodes into zones for short run pricing appeals as a putative simplification. However, in a sufficiently dense network, zonal aggregation provides less simplification than meets the eye. Zonal pricing itself would be controversial and complicated, more complicated in practice than straight locational pricing. Furthermore, locational pricing would avoid perverse incentives that lead to restrictive rules. The path to greater flexibility in commercial transactions, therefore, is through locational pricing. And the hub-and-spoke approach that can evolve as the market requires, can obtain the simplification hoped for with zones while avoiding the restrictions and regulation that zonal approaches would require. It would be better if the world were less complex, but the reality of the electric network cannot be avoided if we want to have a competitive market with a maximum degree of flexibility. To support choice, as always it is important to get the prices right. And in electrical networks with transmission constraints, there is an essential and significant locational element that can be accommodated, but not easily suppressed.

- end -