TRANSMISSION CAPACITY RESERVATIONS AND TRANSMISSION CONGESTION CONTRACTS

SCOTT M. HARVEY
Putnam, Hayes & Bartlett, Inc.
Cambridge, Massachusetts 02138

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

SUSAN L. POPE
Putnam, Hayes & Bartlett, Inc.
Cambridge, Massachusetts 02138

June 6, 1996
(Revised March 8, 1997)

CONTENTS

TRANSMISSION CAPACITY RESERVATIONS AND TRANSMISSION CONGESTION CONTRACTS

Scott M. Harvey, William W. Hogan and Susan L. Pope

EXECUTIVE SUMMARY

Open access to the transmission grid is a necessary support of a competitive market in electricity generation and supply. A key ingredient of open access is a system of capacity allocations for use of the transmission grid. A unique characteristic of electricity transmission is seen in the difficulty of defining transmission rights and matching these rights in a meaningful way to the actual use of the system. The old procedures from the era of vertically integrated utilities will not suffice for the new world of unbundling and competition. Strong network interactions coupled with user flexibility and choice require a new system of transmission capacity definition, reservation and use.

The Federal Energy Regulatory Commission (FERC) has proposed a system of point-to-point transmission capacity reservations that would govern use of the network. The emphasis is on specific performance and matching use to rights. An alternative contract network perspective would emphasize financial contracts for settlement relative to the actual use of the system. Under competitive market assumptions, the two approaches would be functionally and financially equivalent, although they appear different in their implementation. Exploring the connection between the alternative perspectives illuminates the challenges and provides another, easier path to the objective.

The contract network approach and its key elements -- access to essential facilities including the wires and pool dispatch, use of short-run marginal-cost pricing, and reliance on long-term contracts to provide economic hedges rather than specific performance -- are not far from actual operations or proposed reforms in other systems. The FERC proposed capacity reservation system moves very far in this direction, posing a point-to-point transmission reservation definition that does not depend on a decomposition and tracking of the actual flows. Decentralized trading of these capacity reservations would not be enough to support a competitive market, but this trading could be coordinated through the system operator. Opportunity cost pricing would define the trading criterion. Because of the strong and unavoidable network interactions, opportunity cost pricing for transmission and energy cannot be separated, but they would arise naturally as the result of bidding and economic dispatch offered by the system operator. Locational prices would define the opportunity costs for energy bid through the spot market, and the difference in locational prices would define the opportunity cost for transmission scheduled in addition to the spot market transactions. Then tradable point-to-point capacity reservations with opportunity cost pricing for unused or overused amounts would be functionally and financially equivalent to transmission congestion contracts. Conversely, transmission congestion contracts would be functionally and financially equivalent to the tradable point-to-point capacity reservations, would be easier to manage, and would fully support the competitive market while being fully consistent with the actual use of the transmission grid.

TRANSMISSION CAPACITY RESERVATIONS AND TRANSMISSION CONGESTION CONTRACTS

Scott M. Harvey, William W. Hogan and Susan L. Pope[1]

> "The proposed capacity reservation open access transmission tariff, if adopted, would replace the open access transmission tariff required by the Commission ..."[2]

INTRODUCTION

Open access to the transmission grid is a necessary support of a competitive market in electricity generation and supply. A key ingredient of open access is a system of capacity allocations for use of the transmission grid. A unique characteristic of electricity transmission is seen in the difficulty of defining transmission rights and matching these rights in a meaningful way to the actual use of the system. The old procedures from the era of vertically integrated utilities will not suffice for the new world of unbundling and competition. Strong network interactions coupled with user flexibility and choice require a new system of transmission capacity definition, reservation and use.

The present paper reviews the objectives and challenges of a workable capacity reservation system and develops the connection with the competitive electricity market. Starting with physical transmission rights and a trading system, the complications of network interactions lead naturally to a system that is financially equivalent to physical transmission rights but greatly simplified through the use of contracts consistent with the physical flows and financial

---

[1] Scott Harvey and Susan Pope are, respectively, Director and Principal of Putnam, Hayes & Bartlett, Inc., Cambridge MA. William W. Hogan is the Thornton Bradshaw Professor of Public Policy and Management, John F. Kennedy School of Government, Harvard University, and Director, 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, Judith Cardell, John Chandley, Doug Foy, Hamish Fraser, Geoff Gaebe, Don Garber, Stephen Henderson, Carrie Hitt, Jere Jacobi, Paul Joskow, Maria Ilic, Laurence Kirsch, Jim Kritikson, Dale Landgren, William Lindsay, Amory Lovins, Rana Mukerji, Richard O'Neill, Howard Pifer, 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 authors are or have been consultants on electric market reform and transmission issues for British National Grid Company, General Public Utilities Corporation, Duquesne Light Company, Electricity Corporation of New Zealand, National Independent Energy Producers, New York Power Pool, New York Utilities Collaborative, Niagara Mohawk Corporation, 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 the remaining errors are solely the responsibility of the authors.

[2] Federal Energy Regulatory Commission, "Capacity Reservation Open Access Transmission Tariffs," Notice of Proposed Rulemaking, RM96-11-000, Washington DC, April 24, 1996, p. 1.

transactions of the competitive equilibrium. The resulting transmission congestion contracts provide a well-defined foundation for measuring and allocating transmission capacity.[3]

TRANSMISSION RIGHTS

There is no controversy over the critical importance of transmission access as a necessary condition to support a competitive market in electricity.[4] Barring a policy that encouraged and anticipated construction of multiple, unconnected power grids covering the same region, open access would be necessary. Even in the exceptional case, system reliability would be improved by connecting the grids, making open access again necessary. Furthermore, open access with a competitive market for both buyers and sellers of power, would imply a set of rights and rules to govern transmission use. Motivated by the analogy of an isolated transmission line connecting two points, it has been widely accepted or assumed that there must be a system of transmission capacity reservations which could be defined, allocated, traded, and ultimately matched to the actual flows of power in the network. In the case of an interconnected grid, the challenge has been to move beyond an analogy applicable only to a single transmission line to a workable mechanism that would accommodate the realities of electricity network interactions.

There are many ways to summarize the basic requirements for a system of transmission capacity reservations and use that would support an efficient electricity market. One natural approach would be guided by the requirements as proposed by the Federal Energy Regulatory Commission (FERC).[5] The FERC principles go a long way towards describing the goal of a workable system for the competitive market, and means to reach this goal appear implicitly through an analysis of the logical requirements of the system. The full list of FERC principles appears in an appendix; however, a focus here on a few key elements can project the rest of the journey.

The FERC has recognized the limitations of the traditional contract path model of transmission capacity definition and use.[6] This old idea assumed that it would be an acceptable

---

[3] A transmission congestion contract is a financial instrument that entitles the holder to receive congestion payments for quantities of power at different locations. For instance, a balanced transmission congestion contract for 100 MW between A and B would entitle the holder to the difference in congestion costs at A and B for 100 MW for each period of application.

[4] Joskow, P, and R. Schmalensee, Markets for Power: An Analysis of Electric Utility Deregulation. Cambridge, MA: MIT Press, Chapter 12, 1983.

[5] Federal Energy Regulatory Commission, "Capacity Reservation Open Access Transmission Tariffs," Notice of Proposed Rulemaking, RM96-11-000, Washington DC, April 24, 1996.

[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,

approximation to describe transmission from one point to another in terms of the path that the power followed through the network. In reality, as examined below, the power flows on every parallel path, and the actual use of the system can deviate substantially from the fiction of the contract path. This is the loop flow complication, and its impact is both fundamental and substantial in a constrained network. In the world of vertically integrated monopoly, the inevitable problems of the contract-path fiction could be internalized by the monopoly. In a world of choice, with multiple users competing on the same integrated system, the old fiction must be replaced with a new reality.

One approach to providing firm transmission service is the point-to-point definition offered by the FERC as point 2 in its proposed principles:

2. Basic service concept

> All firm transmission service would be reserved, and all reserved service would be firm service. Reservations of transmission capacity should permit the customer to receive up to a specific amount of power into the grid at specified [Points Of Receipt], and to deliver up to a specific amount of power from the grid at specified [Points Of Delivery], on a firm basis. Individual PORs and PODs need not be "paired" with each other. The customer's capacity reservation would be the higher of either (1) the sum of the reservations at all PORs or (2) the sum of the reservations at all PODs. All nominations for a capacity reservation would be evaluated using the same standard; for example, the utility could apply a feasibility criterion that states that the grid must be able to accommodate the scheduled use of all capacity reservations simultaneously.[7]

This definition purposely sidesteps the complication of identifying where the power flows and suggests only the feasibility criterion that the collective reservations must be simultaneously feasible. This minimal feasibility test has the advantage that it poses a well-defined criterion. It is possible to determine if a given set of inputs and outputs to the grid would be simultaneously feasible. The routine calculations involve no more than computing a set of load flows for different contingencies and checking the status of the various constraints, without explicitly tracking or decomposing the individual transactions.

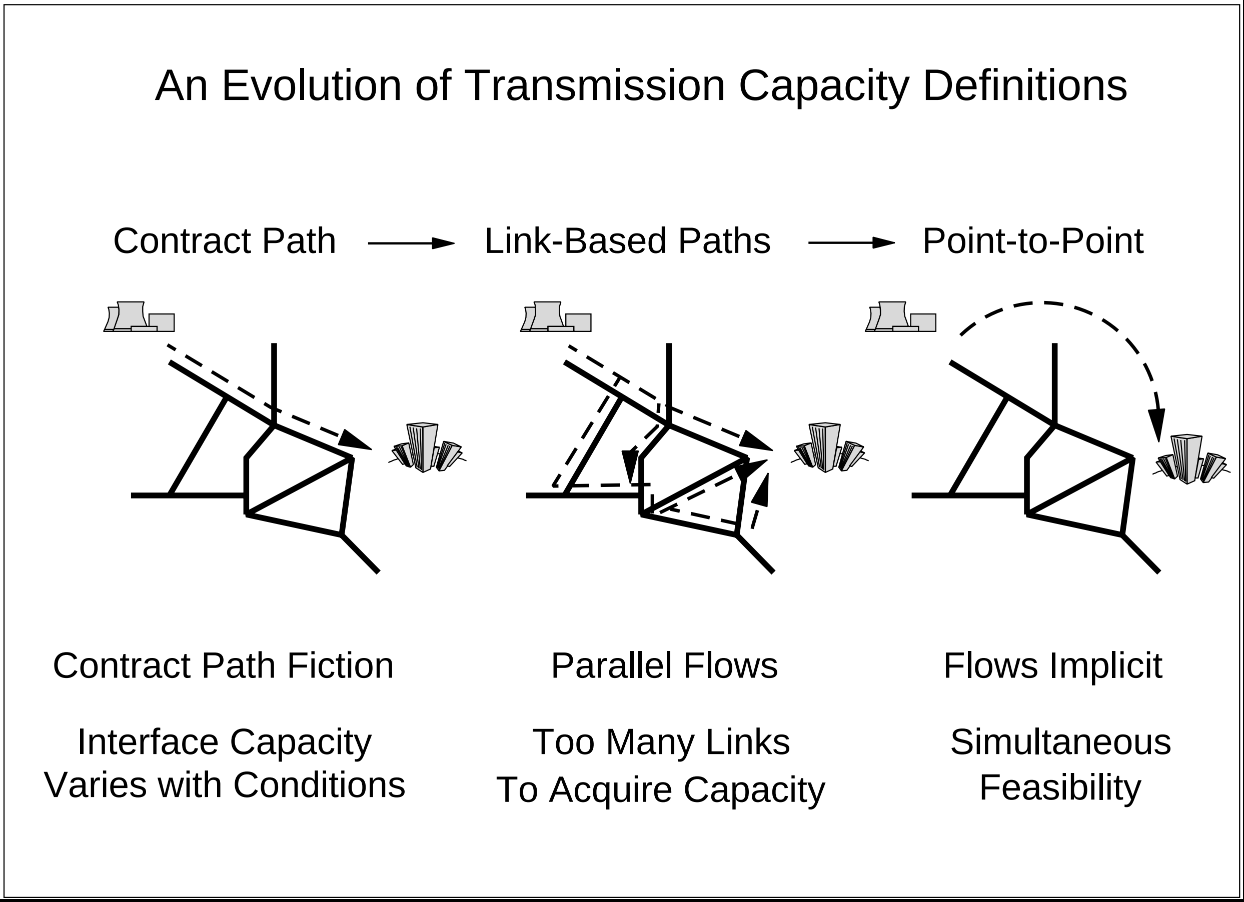
As indicated in Figure 1, this point-to-point approach to defining transmission service goes beyond another natural generalization of the contract path through a link-based method which would build on a measurement of the actual flows and voltages in the network to capture

---

April 24, 1996, pp.93-98. The contract paths are redefined as "posted paths" in Federal Energy Regulatory Commission, Open Access Same-Time Information System (formerly Real-Time Information Networks) and Standards of Conduct, Order No. 889, Final Rule, Washington, DC, April 24, 1996, p. 66.

[7] Federal Energy Regulatory Commission; see Appendix I

Figure 1



the individual contribution of each transaction to each constraint. For example, under a link-based approach, the owner of each line would control the rights to send power over that line, subject to a thermal limit on the flow. Everyone who affected the flow on that line would have to obtain matching rights. This link-based approach would be correct in principle, at least for marginal impacts, but the obstacle to implementation of such a perspective is in the many thousands of constraints that would be affected, and thus the thousands of rights that would have to be acquired for even a simple transaction. By contrast, the point-to-point approach has intuitive appeal and lends itself to a practical implementation.

Under the FERC proposal, these point-to-point rights would be applied to match the use of the transmission grid:

3. Use of capacity reservations

> A customer with a capacity reservation could use the reservation to deliver or receive any type of power product (such as firm or non-firm power). That is, use of the capacity reservation should not be restricted to

> particular power products. Any such restriction would be inconsistent with unbundling. This would allow the capacity reservation holder to combine transmission and power products in any way that satisfies its needs.[8]

In practice, the patterns of load and generation, and the resulting demand for transmission service, will be changing, rapidly and significantly. The location of capacity reservations needed could differ from hour to hour, or even from minute to minute. Hence, in applying this point-to-point approach to operation of the transmission system, there must be a mechanism that allows for trading and reconfiguration of capacity reservations. Some such trading mechanisms are envisioned as critical elements of the FERC proposal. The easiest type of trading would allow for transfer of existing capacity reservations among market participants:

> 13. Reassigning reservations
>
> Customers would be allowed to reassign their reservations to other entities eligible to take service under the [Capacity Reservation Tariff] at no additional cost, subject to certain limitations, such as those in the Open Access Final Rule point-to-point tariff provisions.[9]

Presumably the intent is that the owner of an existing capacity reservation to inject some amount of power at one set of points and take from another set of points could transfer to another market participant the ownership or right to use this same capacity reservation. This form of exchange should be simple, and would support a more efficient electricity market. In fact, it would be difficult to prevent de facto exchanges of this simple type, which could take place in a decentralized market without any central coordination or approval.

Such simple reassignment would be necessary, but far from sufficient, to support the necessary trading in the electricity market. The desired pattern of flows would almost always be different than this simple exchange mechanism would allow. The nature of loop flow and the operation of the transmission system suggest the need for another type of trading that would be far more involved. The FERC recognizes that limiting transmission capacity trading to movements between a few pre-defined receipt and delivery points would substantially reduce the effective capacity of the transmission grid, and would in no way reflect the practice involved in traditional utility management of the flow of power. Hence, the FERC proposal calls for the ability not only to transfer ownership of existing reservations but also to reconfigure the reservations:

---

[8] Federal Energy Regulatory Commission; see Appendix I

[9] Federal Energy Regulatory Commission; see Appendix I

11. Service modifications

> Customers with a capacity reservation would be allowed to modify their capacity reservations at no additional charge if the modification can be accommodated without infringing upon any other firm capacity reservations. Modifications should not result in the customer's capacity reservation being exceeded. Modifications could include reallocation among the customer's already specified receipt and delivery points or reallocation from existing to new receipt and delivery points.[10]

It is the step of requiring a "reallocation from existing to new receipt and delivery points" that greatly expands the scope and importance of the proposed capacity reservation system. Here the proposal confronts one implication of the special characteristics of transmission grids. In particular, there is a complex relationship among the collection of feasible capacity reservations. It may be that a 1 MW reservation relative to points A and B could be exchanged for a 2.5 MW reservation between C and D "without infringing upon any other firm capacity reservations." The exchange ratio is not one-to-one. Furthermore, the exchange ratio is not fixed in advance. The exact exchange ratio would depend on many factors, including the current status of all other capacity reservations and all other exchanges that are requested.

Of course, these same problems would arise in any initial allocation of capacity reservations. Hence, whether in the initial allocation or in service modifications, there will be an essential and unavoidable role for the (independent) system operator in coordinating and approving exchanges and reconfigurations.[11] This coordination and adjudication process will require some criterion for selecting among and matching the many possible exchanges. A market based criterion is within reach and within the vision proposed by FERC, embedded in the explicit recognition of opportunity cost pricing for actual use of the capacity reservations:

14. Opportunity Cost Pricing

> Opportunity cost pricing would still be an option under a capacity reservation service. Under a CRT, a holder of a capacity reservation would not pay opportunity costs for use of its own capacity when the utility encounters a transmission constraint; instead, it would be eligible to receive opportunity cost payments if it did not use its full capacity reservation across the constrained interface. In contrast, a customer

---

[10] Federal Energy Regulatory Commission; see Appendix I

[11] That there must be a system operator is clear; the debate is over the functions and governance of the system operator. For FERC's principles for an independent system operator, see 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.

seeking a capacity reservation or using non-firm service might have to pay opportunity costs.[12]

This provision goes beyond the usual recognition that competitive market pricing will reflect opportunity costs, with equilibrium prices equal to marginal costs. This result would obtain naturally in a competitive market with simple exchanges of existing capacity reservations. If each customer is a price taker, then no market participant would sell capacity reservations for less than their opportunity costs or buy them for more than their opportunity costs. With an equilibrium market price, existing capacity reservations would trade at the marginal opportunity cost.

In the FERC proposal, however, this opportunity cost pricing principle is extended to reservations that are implicitly service modifications arranged through the system operator. These service modifications will result in some customers not using their original capacity reservation and who would, in turn, "receive opportunity cost payments" and others would request new capacity reservations or service and "might have to pay opportunity costs." As we shall see, this important opportunity cost pricing principle provides a critical part of the puzzle by establishing a decision criterion for the system operator in coordinating the complex trades that must occur in great volume and with great frequency if we wish to achieve an efficient outcome in the restructured electricity market.

Furthermore, this important provision allowing for actual use of the transmission system to deviate from the capacity reservations, with compensation, makes the reservation more like a financial instrument than a physical right intended to control use of the system. Assuming no artificial penalties, other than opportunity costs, for balancing the differences between reservations and use, this provision provides an enormous simplification for practical implementation of a system of transmission rights.

There are other details, but these few elements capture the essentials that could lead to a workable transmission capacity reservation system seen in outline here. Point-to-point reservations would be created and assigned. There would be a minimal test that the reservations should be simultaneously feasible. Those who actually use the system would schedule their flows with the system operator, presumably matching use to reservations. Reservations could be traded or reconfigured. Simple trading of ownership of existing reservations could take place with no more than notification requirements, but reconfiguration of capacity reservations would of necessity be coordinated through the system operator. And everything would come to an efficient competitive market equilibrium at opportunity cost prices, applicable over periods of an hour or less.

This is an ambitious agenda, but presumably an agenda that has been implicit all along in the developing vision of a restructured electricity market based on the presumed competitive

---

[12] Federal Energy Regulatory Commission; see Appendix I

potential in generation and supply. It is an agenda that can be realized, but only by exercising some care as to the details and the specific functions of the system operator in performing certain critical roles. The basic outline of the argument summarizing the problems points in the direction of the solution.

LOOP FLOW AND SERVICE MODIFICATION

The central difficulty facing the expansion of competition in the electricity market arises from the impacts of the strong network interactions collected under the general label of "loop flow." The difficulty arises from the inherent and fundamental nature of system interactions in a transmission network. When power moves between points 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. The focus here is to examine the significance of this unavoidable reality: why loop flow makes electricity different, why it is an especially important problem for a competitive market, why this problem should strongly influence the procedures, rules, and role of the system operator, and how it affects the implementation of transmission capacity reservations.

Loop Flow and Contract Paths

Under the contract path model, it would be possible to provide grid users information about the capacity of the various paths and the scheduled usage of the paths. The capacity would be defined in industry parlance as the "interface limit" for a set of lines, if not an individual line. Presumably, any user of the transmission system could look up the available capacity on an interface and make a decision and a commitment to use some of that capacity. The decision to use that interface could be made independently of any consideration of the limits on other interfaces elsewhere in the system. The new power flow could be identified and the change on that interface recorded. Similar decisions and commitments could be made by others for other interfaces in the network. Moving megawatts of electricity would be much like moving bushels of wheat.

How close is this stylized model to the real world? And how much would the differences matter? The answers are that the model is not at all close to the real world, and the differences matter a great deal. The interesting case is when the transmission system is congested. Electricity moves according to electrical laws, essentially following the path of least resistance at the margin. As a result of these physical laws, power moves across many parallel lines in often circuitous routes.[13] Hence, the "contract path" is a fiction. The actual flow of power may and often does diverge widely from the contract path. As a result, in the presence

---

[13] For an excellent discussion of the problems of loop flow, see Federal Energy Regulatory Commission, The Transmission Task Force's Report to the Commission. Electricity Transmission: Realities, Theory, and Policy Alternatives, Washington, DC, October 1989.

of system congestion, the supposed economics of the contract path may have little to do with the actual costs of the power transfer. Furthermore, these loop flows can affect third parties distant from the intended power flow, and under the current rules these third parties may and often do incur costs without compensation.

When loop flow was a small part of power economics, when informal swaps could balance out the effects over time, and when all the parties were members of the same transmission club, it was reasonable to employ the contract path fiction as a practical accommodation in crafting power contracts. These circumstances fit the past, and the contract path was a workable fiction. But all these conditions have changed. In FERC's terms, the contract path approach is "familiar," but this seems to be its only attraction.[14] The FERC capacity reservation proposal is an attempt to craft a new approach that will work in the future.

The problem of loop flow is ubiquitous and can invalidate some of the most important elements of transmission agreements. For example, what is the capacity of the network? The difficulty of defining the transfer capability of the power system is closely related to the economic problems of loop flow. Consider the comment of the New York Power Pool, one of many similar observations:

> ...actual transmission availability, or, more correctly, available transmission transfer capability, may be less than the thermal limits of the facilities, and the difference may change as conditions change. The Commission should make certain that all participants understand and accept these factors.[15]

For an illustration of the point, take the simplified example in Figure 2. This elementary network consists of three locations and three transmission lines. This is the smallest network that can illustrate the effects, which do not arise on radial connections between two points without loops. Unfortunately, the simple model of a single, isolated transmission line, which gives rise to the contract-path approach, is seriously misleading.

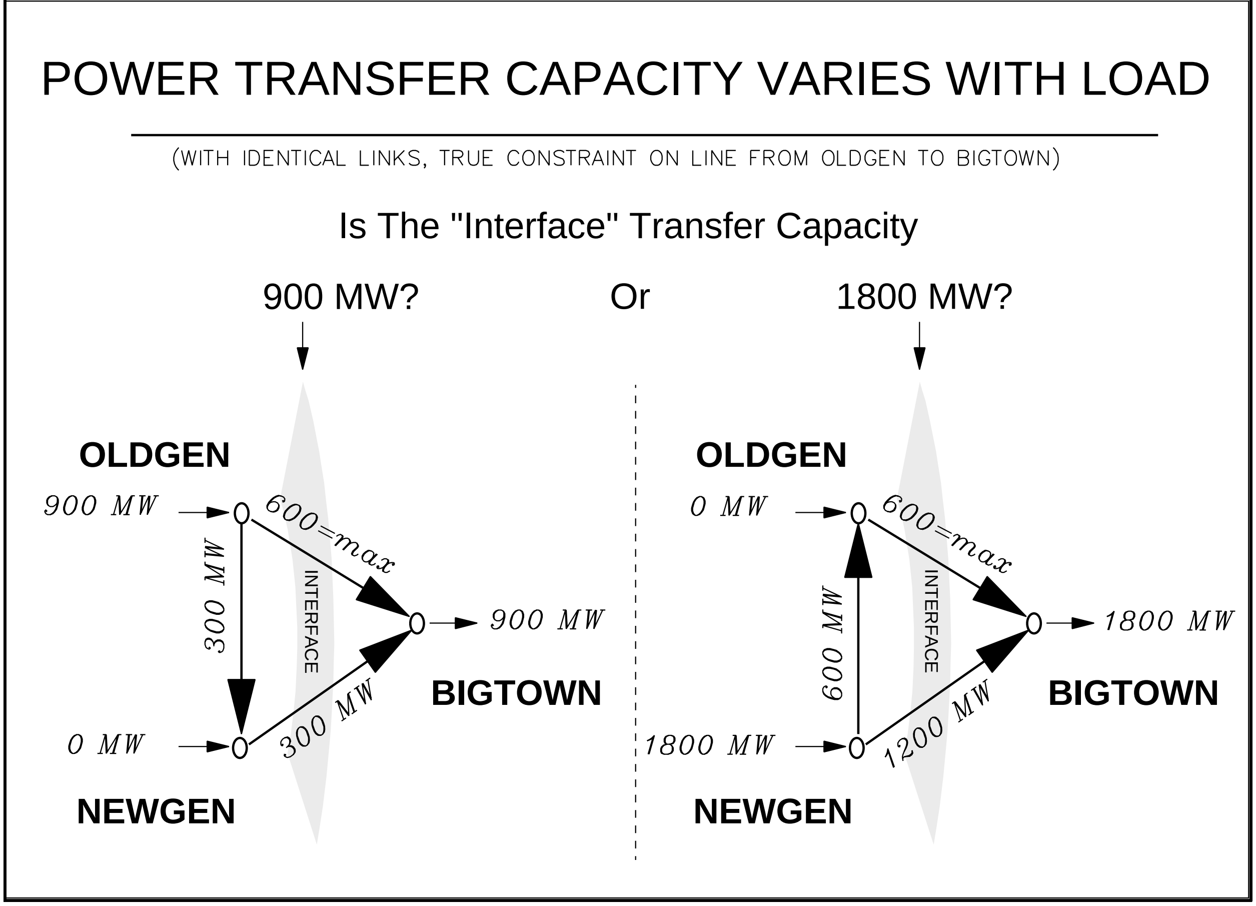
Think of the two generating locations, "OLDGEN and "NEWGEN," as being in one region, with many customers located in "BIGTOWN." As is customary, we could speak of the transmission "interface", as indicated by the shaded boundary in Figure 2, that separates the producing region from the consuming region. For the sake of illustration, assume that the transmission lines are identical except that a thermal constraint on the line connecting OLDGEN

---

[14] 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, p. 96.

[15] Comments of Member Systems of the New York Power Pool, Request for Comments Regarding Real-Time Docket Information Networks, No. RM95-9-000, Federal Energy Regulatory Commission, July 5, 1995, p. 9-10.

Figure 2



and BIGTOWN limits the flow on that line to a maximum of 600 MegaWatts (MW). This is a stylized but not unrepresentative situation. And it would be natural to assume that there is an easy answer to the question: What is the transfer capability across the transmission interface?

The two panels in Figure 2 depict two different load patterns that exhaust the capacity of the constrained line. As seen by a comparison of the two panels, the estimate of the interface transfer capability depends on the configuration of the generation. In the left panel, the total demand at BIGTOWN is for 900 MW, and these 900 MW are provided by the low cost generator at OLDGEN. The flow of power follows the physical laws. Since the path OLDGEN→NEWGEN→BIGTOWN is twice as long as the path OLDGEN→BIGTOWN, it has twice the electric resistance. Hence 600 MW move along the path OLDGEN→BIGTOWN, while 300 MW move along the parallel path OLDGEN→NEWGEN→BIGTOWN. This is loop flow.[16] There is no power generated at NEWGEN and none can be added there without violating the 600

[16] Here we are ignoring losses and use the conventional "DC-load" approximation for purpose of the illustration. Since the lines are identical, path OLDGEN→NEWGEN→BIGTOWN has twice the resistance of path OLDGEN→BIGTOWN, which makes it easy to verify the power flows.

MW constraint on the line between OLDGEN and BIGTOWN. Because of the constraint, as long as we choose to generate 900 MW at OLDGEN, we cannot satisfy any more demand at BIGTOWN. In a real sense, therefore, the power transfer capability of the interface might be viewed as 900 MW.

If demand increases at BIGTOWN, there is no choice but to generate power at NEWGEN *and* reduce the generation at OLDGEN; otherwise the power flow along OLDGEN→BIGTOWN would exceed the maximum thermal limit. In the extreme, as shown in the right panel of Figure 2, if demand rises to 1800 MW, the *only* solution is to generate all the power at NEWGEN and none at OLDGEN. Hence, the interface power transfer capability might be viewed as 1800 MW.

One solution to this problem might be to award 900 MW of capacity reservations and let the generators buy and sell these reservations to match the schedules from either OLDGEN or NEWGEN. This "solution" to the problem, however, would forgo large cost savings, artificially raise the price of power at BIGTOWN, and increase the market value of transmission across the interface by creating artificial scarcity.

In a real network, the conditions in the left panel of Figure 2 might represent the economics and availability of generators at one time and the right panel would apply at another. And the announced transfer capability might be somewhere in between. But in the words of the North American Electric Reliability Council (NERC) -- "The actual transfer capability available at any particular time may differ from that calculated in simulation studies because in the simulation studies only a limited set of operating conditions can be evaluated, whereas in real time, widely different conditions may exist."[17] -- even when there has been no change in the transmission system. A contract for 600 MW between OLDGEN and BIGTOWN may have relied on a contract path along the direct connection, but the higher demand case would have precluded this use of the system and the contract could not be honored. Evidently changing load patterns could play havoc with specific performance--actually delivery of power from specific plants for specific customers--for transmission rights expected to extend over many years.

Loop Flow and Network Interactions

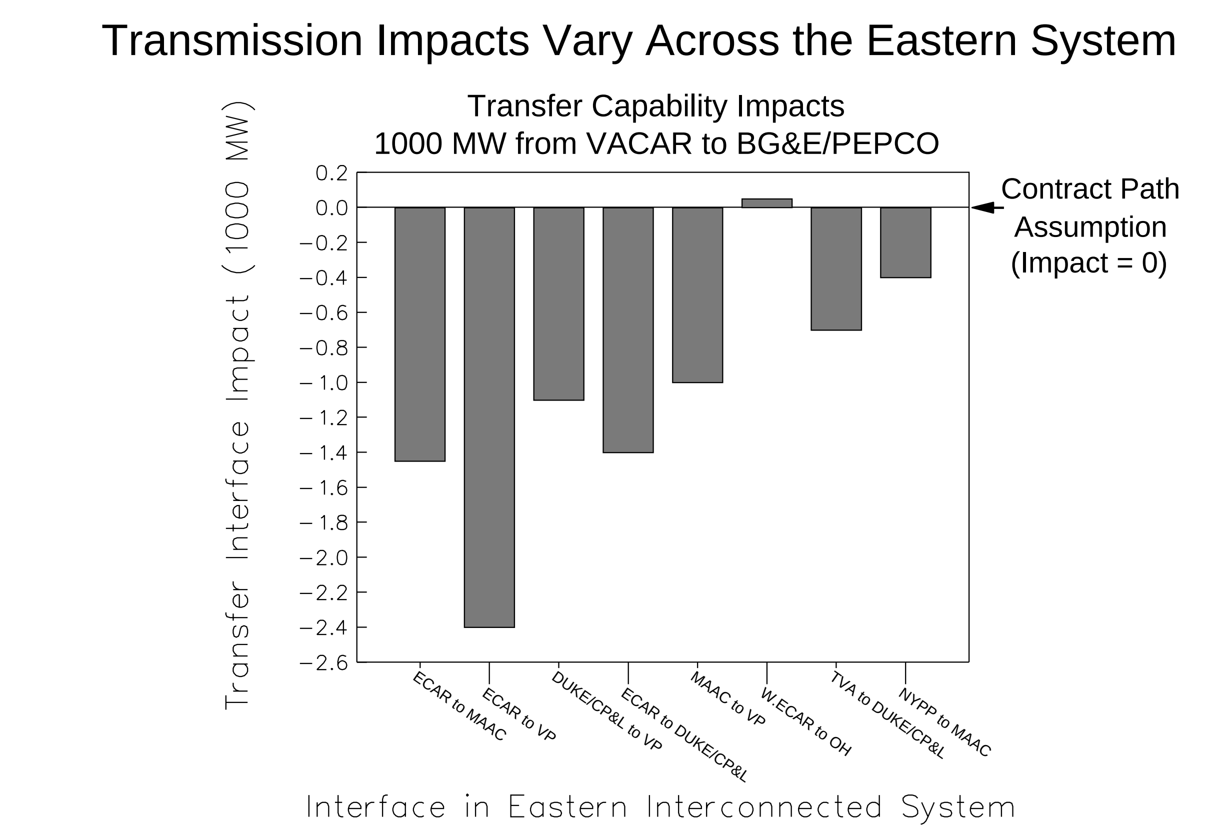
The small system in Figure 2 makes the network interactions apparent. Of course, it would be possible to remove the line between OLDGEN and NEWGEN, or put in a switch, but this would compromise the reliability benefits of the network that come from Nature's ability to instantaneously redirect the flow. And such *ad hoc* fixes, designed to dismiss the illustrative example, are mere debating tactics which provide little comfort for large real networks where it can be difficult to assign the loop flows or untangle the effects of other network interactions.

---

[17] Comments of the North American Electric Reliability Council, "Promoting Wholesale Competition Through Open Access, Non-discriminatory Services by Public Utilities," Transmission Docket No. RM95-8-000, Federal Energy Regulatory Commission, Washington, DC, August 7, 1995, pp. 10-11.

When the system is used only a little, anything can be done and the contract path fiction can be accommodated. However, when the system is constrained, loop flow may cause other generators not on the contract path to redispatch at higher cost. Symmetrically, the rights on the contract path could be preempted without compensation by others who are using the system. There is a dramatic result totally at odds with the contract-path model and successful operation of a competitive market.

Figure 3



Source: VEM, Winter Operating Study, December 1993.

With the small example as the explanation, we can review analyses of real systems to demonstrate that the problems are ubiquitous and can be of great significance. The system operators in the eastern interconnected grid regularly conduct joint studies of the transmission transfer capabilities of various interfaces. One of these types of exercises was conducted by the VEM Study Committee which examined the impact of various power transfers under peak load operating conditions.[18] A central task undertaken in the VEM study was an evaluation of the impacts of a power transfer across one interface on the transfer capabilities across other

---

[18] Virginia-Carolinas (VACAR) (Subregion Electric Reliability Council), East Central Area Reliability Coordination (ECAR), Mid Atlantic Area Council (MAAC), Winter Operating Study, December 1993.

interfaces. For example, what would be the impact of a 1000 MW transfer from the Virginia-Carolinas (VACAR) region to Baltimore Gas & Electric (BG&E) and Potomac Electric Power (PEPCO)? The assumption of the contract-path model is that there would be no impact on the transfer capability of other interfaces. Under the contract-path fiction, users could use the capacity on one interface without worrying about the limits on other interfaces. As Figure 3 summarizes, however, the actual effects elsewhere would be far from zero, and certainly not negligible. The impacts would range from a gain of 50 MW to a loss of 2400 MW, depending on the locations of the other interfaces. Clearly parties quite distant from the transaction would experience major effects, sometimes larger than the originating transaction, and have a keen interest in the decision to move 1000 MW from VACAR to BG&E/PEPCO. This problem cannot be avoided, and must be faced in the design of the institutions for an efficient competitive market.

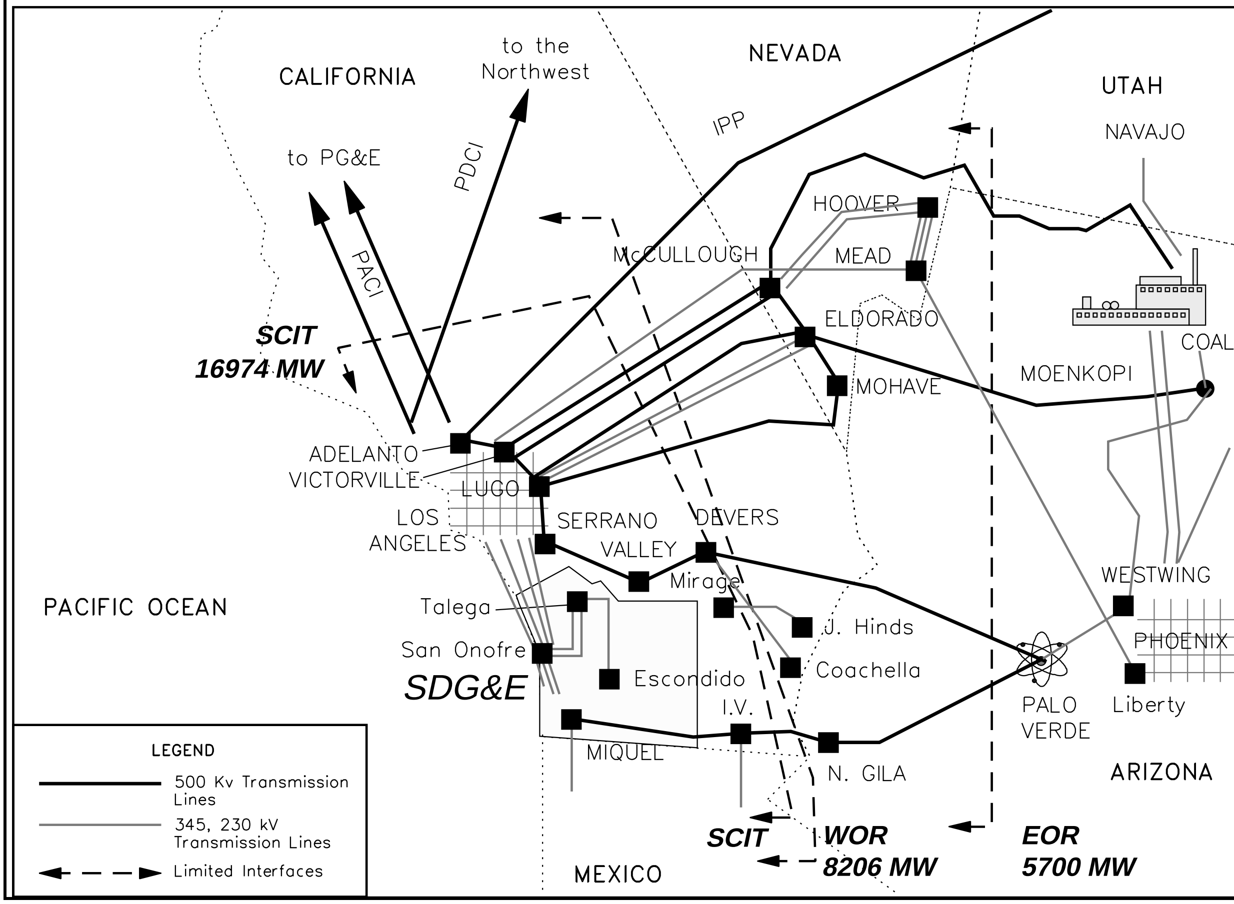
The complex network interaction or "loop flow" effect is caused by the nature of the highly interconnected grid and the current state of technology governing power flows. There are many, interacting, nonlinear constraints that limit operations in power systems. The reduction to "interfaces" is a simplification that is used for network management in a highly coordinated system. The interface metaphor and contract-path fiction are not suitable for a decentralized market.

Furthermore, the problems arise in any interconnected grid, not as sometimes argued just in the highly networked system in the eastern part of the United States. Consider, for example, the simplified map of southern California as shown in Figure 4. The map provided by San Diego Gas & Electric (SDG&E) illustrates the location of major power plants, loads, and transmission lines. The schematic includes three interfaces with associated maximum transfer limits: The East of River (EOR) with a maximum of 5700 MW, West of River (WOR) with a maximum of 8206 MW, and the Southern California Import Transmission (SCIT) with a maximum of 16974 MW.

Under the contract path model, presumably it would be possible to post these interface capacities and allow individual utilities or users to make decisions on how much capacity to use on each interface. In principle, the participants might assume that they could use both the 5700 MW on EOR and the 16974 MW on SCIT, simultaneously. Unfortunately, the indicated capacities are not all achievable simultaneously. In actual use of the system, there are further limits that are summarized in the "nomogram" of Figure 5. This figure reports on the net effect of limits on simultaneous flows on the EOR and SCIT interfaces. Because of the interaction of load patterns with a number of physical limits such as stability and voltage control, the allowable flow on one interface cannot be determined independently of knowing the flow on another. Furthermore, the limits on the flows depend on other factors such as the "inertia" of the available power plants operating in southern California and the status of the nuclear units at Palo Verde.

The interactions are complicated and large. In order to achieve the full SCIT limit, for instance, the EOR capability must be reduced from 5700 MW to 700 MW. Or in order to use the full EOR limit, the SCIT flows must be cut in half. And when we note that the flows over

Figure 4



the EOR would be counted again in the SCIT flows, the reduction of the non-EOR imports across the SCIT could be by as much as a factor of seven! The model of the contract path and the assumed independence of interfaces is seriously misleading.

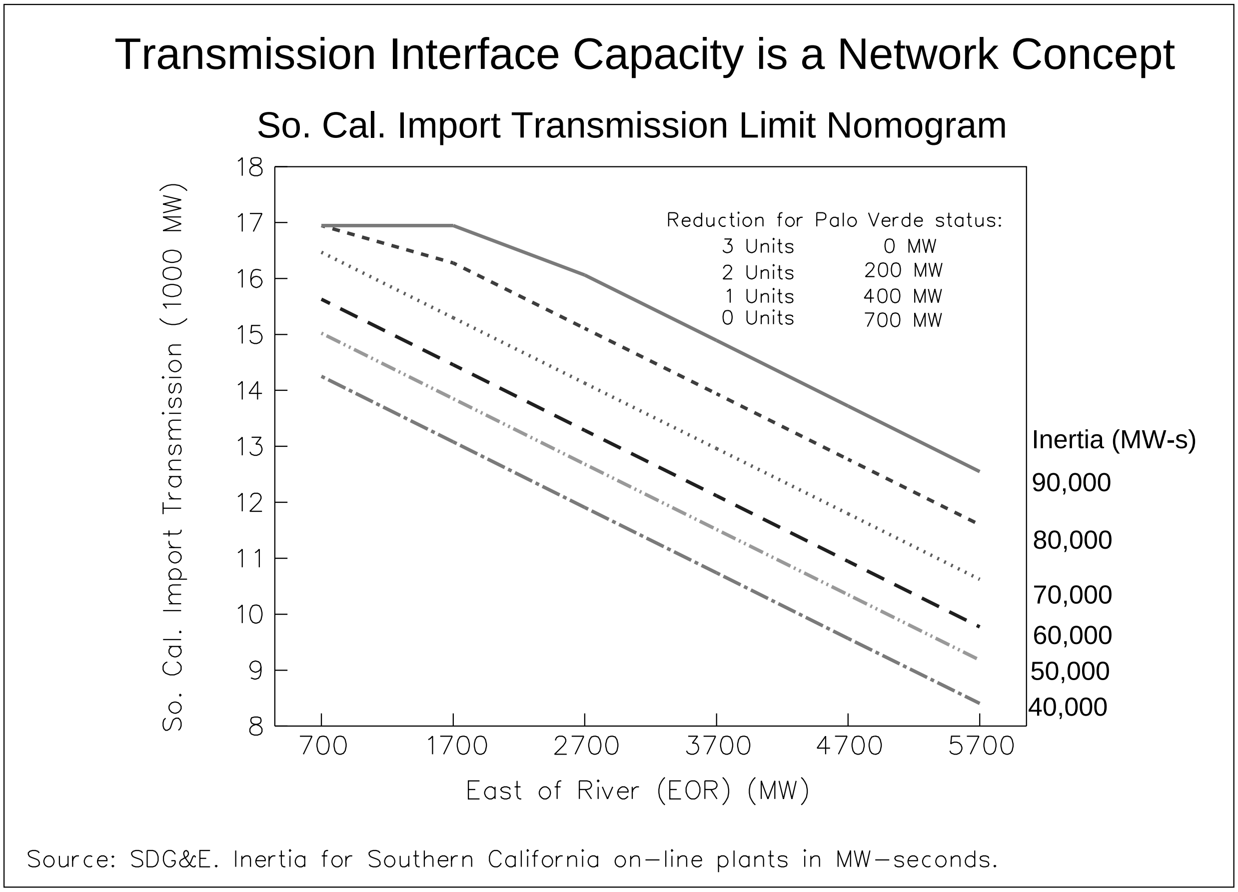
Hence, the examples illustrate the problems of defining and using transmission rights in the conventional way. The implications are far reaching.

Loop Flow and Externalities

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 and important 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.

Figure 5



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. We can approach this problem starting with any of these issues. For example, we are all 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 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 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.[19]

The same loop flow effects explain why there has been so much difficulty in defining ATC in the various attempts to provide access to transmission networks. For example, in the FERC's open access decision, the definition of ATC remains a troubling and unsolved problem. The open access decision refers to the companion information system order for the definition of ATC.[20] The information system order,[21] in turn, directs the filing utilities to provide a definition of ATC, leaving the reader of the lengthy documents with no guidance on how to do what cannot be done. This passing of responsibility is understandable, given the impossible nature of the task, and must provide a large part of the motivation for the capacity reservation tariff notice of

---

[19] Joskow, P. "Comments at FERC Technical Conference Concerning Independent System Operators and Reform of Power Pools under the Federal Power Act, in re FERC RM95-8-000 and RM94-7-001." Washington, DC, January 24, 1996.

[20] 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, p. 97.

[21] Federal Energy Regulatory Commission, Open Access Same-Time Information System (formerly Real-Time Information Networks) and Standards of Conduct, Order No. 889, Final Rule, Washington, DC, April 24, 1996, p. 58.

proposed rulemaking.[22] Fortunately, the FERC and many commenters in the transmission access discussion have been forthright about the conundrum, avoiding the strong temptation to avert their eyes because the problems are too hard. The problems are unavoidable. The capacity on any particular transmission interface depends importantly on the flows on all interconnected interfaces. There is no way to say what the actual capacity will be at any time in the future without specifying all the flows on the system. But without a stand-alone measure of the grid capacity, it is not possible to partition that capacity and assign property rights to its use. There is no escaping this fundamental physical fact, and the effort to define and allocate physical transmission capacity along a contract path is a conceptual dead end.

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, this simultaneity must be recognized and accommodated.

The combined effect of all this is that the traditional approach to unbundling a market to allow for fully decentralized competition will not work for the case of electricity. This does not mean that unbundling and competition cannot be achieved; quite the contrary. However, it does mean that a different approach is required that employs a degree of central coordination organized by the system operator. The focal point is in the definition of the rules for the system operator.

This reality has been recognized in many evaluations of electricity restructuring, most recently in the ongoing debate in California, where the California Public Utilities Commission (CPUC) committed to just such a set of rules for the system operator, rules which the CPUC recognized must be approved by FERC.[23] The result is just the type of innovative proposal that FERC has been calling for, a good model that shows the way, at least for the independent system operator. Further details have been provided in the filings on the Western Power Exchange (WEPEX).[24] Similar details are found in an earlier FERC filing by San Diego Gas & Electric, the recent filing of the PJM power pool, in announcements from New York, and so on. Working systems with many of these elements can be found in Chile, Argentina, Norway and Sweden, and under active investigation in other countries such as Australia and New Zealand. The essence

---

[22] Federal Energy Regulatory Commission, "Capacity Reservation Open Access Transmission Tariffs," Notice of Proposed Rulemaking, RM96-11-000, Washington DC, April 24, 1996, p. 1.

[23] California Public Utilities Commission, December 20, 1995.

[24] Joint Application of Pacific Gas and Electric Company, San Diego Gas & Electric Company, and Southern California Edison Company for Authorization to Convey Operational Control of Designated Jurisdictional Facilities to an Independent System Operator, before the Federal Energy Regulatory Commission, Docket EC96-__-000, April 29, 1996.

of the approach is to use the system operator to internalize the network externalities, get the prices right, and employ workable financial contracts to stand in place of the unavailable, strictly physical property rights.

Link-Based Rights Have Too Many Paths

If a single contract path is not good enough, perhaps many paths would be better. Since power flows along many parallel paths, there is a natural inclination to develop a new approach to transmission services that would identify the links over which the power may actually flow, and to define transmission rights according to the capacities along these links. This is a tempting idea with analogies in markets for other commodities and echoes in the many efforts in the electricity industry for MW-mile proposals, the General Agreement on Parallel Paths (GAPP), and related efforts that could go under the heading of transmission services built on link-based rights.

For any given total set of power injections and withdrawals, it is possible to compute the total flows across each line in the transmission network. Under certain simplifying assumptions, it would be possible further to decompose the flows on the lines and allocate an appropriate share of the flows to individual transactions that make up the total loads. If we also knew the capacity on each line, then presumably it would be possible to match the flows against the capacities and define transmission services. Transmission users would be expected to obtain rights to use the individual lines, perhaps from the transmission line owner.

In principle, these rights on each line might be seen as supporting a decentralized market.[25] Associated with each link would be a set of capacity allocations to (many) capacity right holders who trade with the (many) users of the system who must match their allocated flows with corresponding physical capacity rights. Within this framework there are at least two interesting objectives. First, that the trading rules should lead to an efficient market equilibrium for a short period; and second, that the allocated transmission capacity rights would be useful for supporting the competitive market for geographically dispersed buyers and sellers of power.

As a matter of principle, it is likely that the first objective could be met. There should be some system of tradable property rights that would be sought by users of the system, and in so doing would lead to an efficient short-run dispatch of the system. This would seem to be nothing more than an application of the principles of competitive markets with well-defined property rights and low transactions costs. There is a general belief that this short-run efficiency would be available in principle: "Efficient short-run prices are consistent with economic dispatch, and, in principle, short-run equilibrium in a competitive market would reproduce both these prices

---

[25] The theory of one such an approach has been presented in 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.

and the associated power flows."[26] The problem has always been with the natural definitions of the "physical" rights: these are cumbersome to trade and enforce. The property rights are hard to define, and the transaction costs of trading would not be low.

The second objective is perhaps more important. Presumably the allocated transmission capacity rights would extend over many short-run periods, for example, even only a few days, weeks or months of hourly dispatch periods.[27] Presumably one natural characteristic that would be expected of these link-based rights would be that a seller of power with a known cost of power production could enter into an agreement with a distant buyer to deliver a known quantity of power at a fixed price, including the out-of-pocket cost for transmission using the transmission right. Many other contracts could be envisioned, but this minimal possibility would seem to be essential; and it is broadly taken for granted that this capability will exist in the future open-access transmission regime. However, any approach that defines tradable physical capacity rights based on flows on individual links faces an obstacle that appears to make it impossible to meet this minimal test.

There are many variants of such link-based transmission rights that one can imagine, and the industry has been struggling with these ideas for years. Despite the appeal of a move closer to the actual underlying reality of the transmission network, however, these generic methods built on link-based rights encounter a hidden trap. The amount of the rights that must be acquired over each line under a link-based system would generally not be simply the amount of power that flows in the actual dispatch. The binding constraints on transmission generally are on the level of flows or voltage post-contingency and flows in the actual dispatch are limited to ensure that the system could sustain a contingency. Operation of a link-based system would, therefore, require a trader to acquire the rights on each link sufficient to cover its flows on that line in each post-contingency situation. This might nevertheless be feasible if the flows over each line in each post-contingency situation could be calculated once for each potential receipt and delivery, but this is not the case. Instead, the flows over the line and voltages at the buses will depend on all the other receipts and deliveries on the grid. Thus, the flow over a particular line that can be attributed to a particular transaction will be changing all the time, so it will be difficult to know how much of a right is required or how much would be used.

Another way to view this trap is as an example of the curse of dimensionality. Under current practice, the system operators typically adhere to "(n-1) contingency" constraints on power flows through the grid. This means that the allowed power loads at every location in the transmission system must be such that in the event one of series of possible contingencies occurs, the instantaneous redistribution of the power flows that results will still meet minimum standards for thermal limits on lines and will still avoid voltage collapse throughout the system. We can

---

[26] W. Hogan, Contract Networks for Electric Power Transmission," Journal of Regulatory Economics, Vol. 4, 1992, p. 214.

[27] This is apart from the problems encountered with changes of the grid capacity or configuration. Link-based rights have other substantial problems for dealing with system expansion.

think of the terminology as coming from the notion that one of the "n" lines in the system may drop out of service, and the system must still work with the (n-1) lines remaining. The actual contingencies monitored can be more diverse, but this interpretation conveys the basic idea of an (n-1) contingency-constrained power flow.

Depending on conditions, any one of many possible contingencies could determine the current limits on the transmission system. During any given hour, therefore, the actual flow may be, and often is, limited by the impacts that would occur in the event that the contingency came to pass. Hence, the contingencies don't just limit the system when they occur; they are anticipated and can limit the system all the time. In other words, analysis of the power flows during contingencies is not just an exception to the rule; it is the rule.[28]

Hence, a single line may have a normal limit of 100 MW and an emergency limit of 115 MW.[29] The actual flow on the line at a particular moment might be only 90 MW, and the corresponding dispatch might appear to be unconstrained. However, this dispatch may actually be constrained because of the need to protect against a contingency. For example, the binding contingency might be the loss of some other line. In the event of the contingency, the flows for the current pattern of generation and load would redistribute instantly to cause 115 MW to flow on the line in question, hitting the emergency limit. No more power could be dispatched than for the 90 MW flow without potentially violating this emergency limit. The 90 MW flow, therefore, is constrained by the dispatch rules in anticipation of the contingency.

There are many possible contingencies. The Achilles' heel of link-based rights appears when we recognize that the flows on the individual lines can be different for every contingency, even for an unchanging transaction. Hence, the logic of matching flows on lines against capacities on lines means that users would have to determine the flows separately for each possible contingency and obtain the corresponding capacity right. On real systems with thousands of individual, potentially constrained, links, and hundreds of seriously considered contingencies, the possible combinations could run into the millions. Even a judicious selection of the requirements for a simple long-term firm transaction between two locations might involve assembling many hundreds or thousands of link-based rights covering the possible conditions that might exist over the course of a contract. Furthermore, the flows and associated requirements for rights would change with every change in the configuration of the grid through later investment intended to avoid congestion. Hence, transmission expansion would always entail a contentious reallocation of existing rights in ways that have been difficult to address in simpler times, and may be impossible to accommodate in a more competitive market.

---

[28] Comments of Member Systems of the New York Power Pool, Request for Comments Regarding Real-Time Information Networks, Docket No. RM95-9-000, Federal Energy Regulatory Commission, July 5, 1995, pp. 5-6.

[29] Expressing the limits in terms of MW and real power is shorthand for ease of explanation. Thermal limits are actually in terms of MVA for real and reactive power.

Point-to-Point Capacity Reservations

The escape from the contract path trap, therefore, is not likely to be found in a system of link-based rights. These systems of link-based rights present a trap of their own. This does not mean that there is no available set of "physical" rights that can be assigned and traded. It simply means that the "rights" would have to depend on something other than the flows on the lines. For example, a "physical right" could be defined as the right or obligation to put power in at one location and take power out at another. By definition, this type of physical right would meet the minimal long-term contracting test. Once obtained, the holder could use this right independent of all the other changing flows in the system, and secure the long-term delivered cost of power.

This perspective is embraced in the FERC principles and its definition of point-to-point rights that is intentionally silent on the paths the power follows. The paths would be important in evaluating the simultaneous feasibility of the rights, but this is a technical calculation and importantly not part of the definition. The point-to-point approach would exploit a critical advantage over traditional approaches to property rights. Although it is impossible to identify the available capacity for any interface, without knowing everything that is happening in the network, it is possible to evaluate the simultaneous feasibility of a combined set of point-to-point capacity reservations. There would be no unique set of point-to-point reservations, but any feasible set of reservations could be described in a meaningful way as being within the capacity of the network. And once the reservations were established, they could be traded to adjust to the changing patterns of loads.

This point-to-point form of a firm transmission right would be far easier to use in the long run, but may not be much easier to trade in the short run. The remaining difficulties go beyond the substantial problems of dimensionality and extend to the way we seek to define firm transmission service as somehow related to way the power actually flows, and matching rights to use. A more flexible approach may be required.

The alternative approach builds on an important set of properties of the competitive equilibrium. Because of the strong network interactions, it will not be possible to assign transmission rights and have a fully decentralized system. There must be coordination through the system operator. However, the coordination rules and associated pricing can be made consistent with the equivalent outcome that would arise in the competitive market with active trading of transmission capacity reservations.

COMPLICATIONS AND SOLUTIONS

The complications arise because of the unavoidable need for the system operator to recognize and accommodate the interactions in the transmission network. We could begin with an assignment of point-to-point capacity reservations to the various market participants, and then move to a particular hour for actual use of the system. If life were simpler, the various participants would pursue decentralized trade of these reservations to arrive at a market equilibrium and notify the system operator of the resulting power schedules. In real life, however, if we want to use all of the capacity of the system, this is not possible:[30]

Decentralized Capacity Reservation Trading Will Not Suffice

Many readers of the FERC's capacity reservation proposal initially interpreted the principles as requiring a strict, physically based system for trading capacity reservations with a companion requirement for a close or exact equation of each capacity reservation and the actual transactions and associated power flows. In an extreme, the vision was that the multitude of individual traders would operate through the marketplace to exchange and reconfigure their current transmission capacity reservations to *equal* their intended power inputs and outputs for the next period. In other words, in a constrained situation, the actual movement of power would have to align with the capacity reservations. Each such process of trading through exchange and reconfiguration would require many iterations in the search for equilibrium prices and the associated efficient allocation of the capacity reservations. The vision was of a decentralized process of traders organizing complicated reconfigurations that would actually control the use of the system, depending at most on minimal interaction with the system operator.

With the periods of adjustment over which power transactions are defined and managed as short as a few minutes, the implied volume and frequency of such trading staggers the imagination of traditional system operators and seems impossible on its face. This strictly physical interpretation of transmission capacity rights seems out of touch with reality, and the polite language of discourse with FERC cannot hide the concern that the impossible problem of defining ATC has been replaced with the impossible solution of providing fully decentralized operation of the electric system.

Although it is possible to read the FERC principles and come to this view as to their intent, this cannot be the ultimate solution nor is it the only model that is consistent with the FERC proposal. The fact that it cannot be the ultimate solution is easy to explain. The interpretation that it is not the FERC's intent can be supported by considering the combined impacts of the individual elements of the FERC proposal.

Because of loop flow effects and other network interactions, except for trades that involve no reconfiguration of reservations, only the system operator could know which trades would be feasible. Furthermore, the feasibility test would depend on knowing all the trades that were to be made. Hence, comparability and efficiency would require that the trades for a particular hour be considered simultaneously. Only the system operator, or the functional equivalent, could coordinate these trades. It follows, therefore, that fully decentralized trading is not possible. There must be coordination through the system operator.

---

[30] With a sufficiently conservative allocation of capacity reservations, leaving a good deal of unassigned capacity to serve as a buffer, more decentralized trading would be possible.

The form of this coordination through the system operator cannot be solely an evaluation of a simple feasibility test for each individual trade. The multitude of traders cannot simply queue up at the system operator's (virtual) door and ask if a particular trade would be acceptable. The evaluation of the feasibility of one trade of capacity rights would depend on the others that would be allowed. The problems of loop flow and network interactions appear here in the interdependence of all the trades among all the market participants. To evaluate the trades and the possible exchanges, nondiscrimination would require simultaneous consideration of the nominated trades and some criterion for choosing among the proposals to find the set that would be feasible, consistent with the preferences of the participants, and consistent with a competitive market equilibrium.

Opportunity Cost Pricing Defines the Trading Criterion

When confronted with the potential capacity reservation reconfiguration requests, comparability and non-discrimination provide only partial guidance in choosing among many feasible but mutually exclusive reconfigurations of capacity reservations. Opportunity cost pricing and the goal of economic efficiency point to a workable criterion that should govern the trading of transmission capacity reservations to achieve the competitive equilibrium outcome. And the general principle of opportunity cost pricing is supported explicitly in the principles enunciated as part of the FERC proposal.

In theory, we could imagine opportunity cost pricing operating in the following way. Those who hold transmission capacity rights offer them for sale for the next period under consideration. Those who plan to use the transmission system make offers to purchase the needed transmission rights for the same period. The system operator coordinates this process by considering the offers for sale and purchase to find the acceptable trades that are both simultaneous feasible and consistent with market equilibrium.

Suspending disbelief for the moment, the form of this auction might be through an iterative process. Imagine the auctioneer, the system operator, announcing market prices for each point-to-point right, and then checking to see if the accepted sales and purchases would be feasible. The system operator would somehow search among the tentative price offers until the feasibility test was met, and a market equilibrium obtained in each period; say every hour, or every five minutes!

An alternative and more workable approach would be for the system operator to accept bids for transmission rights in the form of confidential reservation prices, and then to solve directly for the market equilibrium which would be the same as the allocation having the highest value as expressed by the bids of the participants. The bids might be of the form of "1 cent per kw for 100 MW between locations A and B." Here the system operator would have responsibility for centrally determining all the uses of the system, based on the bids submitted by the participants. The resulting allocation would yield the set of market clearing prices for the sales and purchases of the various transmission capacity reservations.

Although this latter market clearing auction may seem like a natural way to organize trading of transmission capacity rights, there is a fundamental deficiency in thinking about transmission in this way and organizing an auction based on transmission rights alone. In addition, there must be a recognition of the close interaction with the energy market.

Transmission Trading Cannot Be Separated from the Energy Market

The opportunity cost of transmission over the hour is directly connected to the opportunity cost of energy at different locations over that same period. The competitive market equilibrium in one segment would be dictated by the competitive market equilibrium in the other. In other words, it would not be possible in general to submit an efficient reservation price bid for bid for transmission alone without knowing the equilibrium solution in the energy market. From the perspective of transmission alone, a bid of the form of "1 cent per kw for 100 MW between locations A and B" would be consistent with energy prices of 2 cents at A and 3 cents at B, or just as well, with energy prices of 4 cents at A and 5 cents at B. However, the actually willingness to pay for transmission would be quite different in the two cases.

For example, suppose that the generator's variable cost were 3 cents. Then in the lower price case it would be better not to generate, avoid the cost of transmission, and purchase the power at B at 3 cents in the spot market. But in the higher energy price case, it would be appropriate to use the transmission system and pay the 1 cent for moving the power.

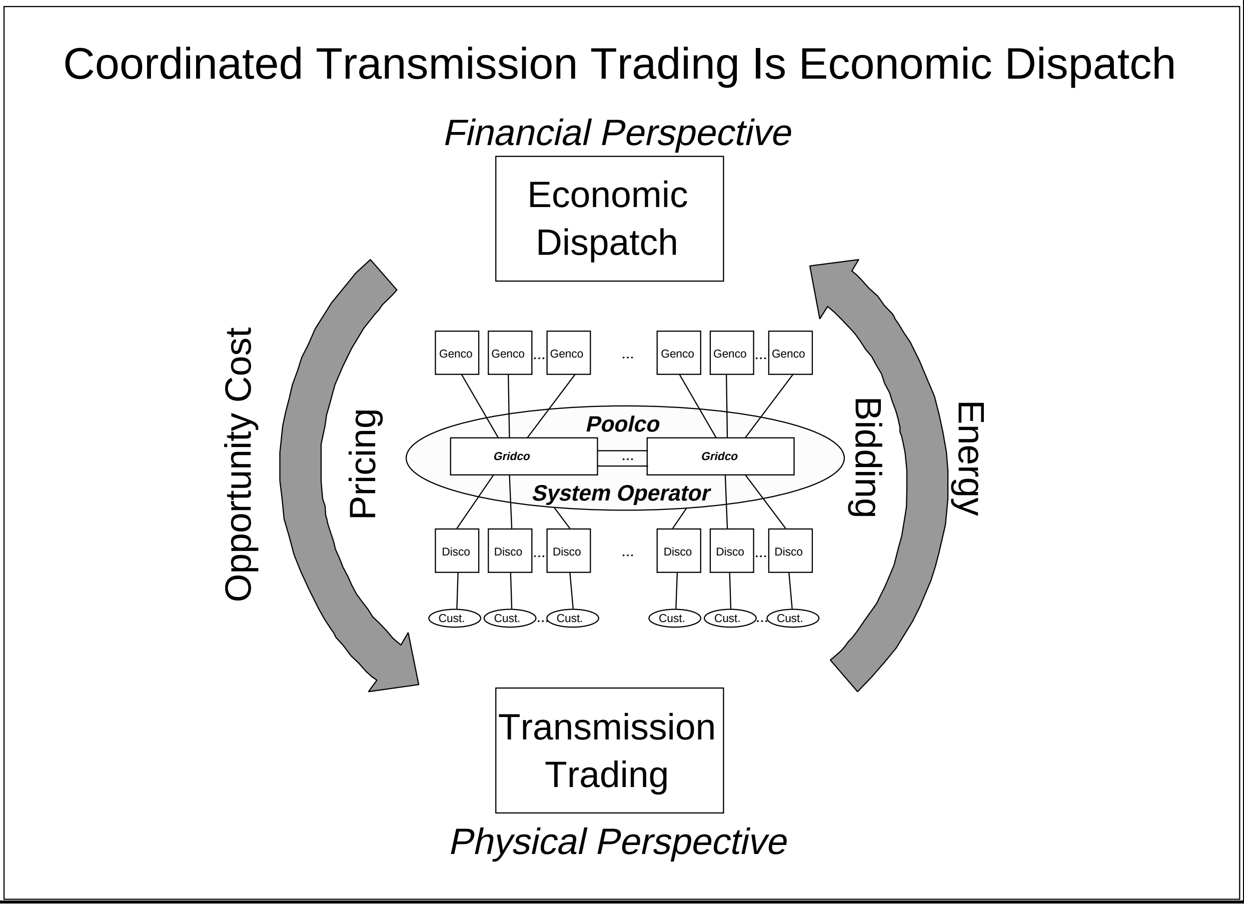
In theory, again, it would be possible to allow decentralized trading in the energy market and iterative auctions in the transmission market, with the tentative coordinated exchange of transmission rights constantly revised to consider the developing opportunities in the energy market. In practice, however, it would be simpler to restate the form of the transmission reservation bids in terms of increments and decrements for both the PORs and the PODs, consistent with the form of the capacity reservations in the FERC proposal. In this way, the "transmission right" bids could be made compatible with the corresponding reservation prices in the energy market. In other words, the bids would be for purchase and sales of capacity rights at PORs and separately for PODs, with only the requirement that the overall collection of all PORs and PODs be simultaneously feasible, a determination that would be guaranteed by the system operator's solution of the combined auction. Hence, in trading transmission reservations, the "transmission" bid might be for "sell for 3 cents at POR A and buy for 6 cents at POD B." In the higher price case, the two rights would be priced at 4 cents and 5 cents, and the total transmission opportunity cost payment would be one cent. In the lower energy price case, however, only the POD would be purchased, and the bidder would receive power from the grid at B at the price of 3 cents. Hence, from this bidder's perspective, there would be no transmission in the lower energy price case.

The reader familiar with the theory of bid-based economic dispatch would recognize that this form of bidding and auction for the transmission capacity right trades would both allow the desired coordination with the energy market and be equivalent in both form and function to a pool-based, spot-market organized through economic dispatch with energy price reservation bids for generation and load. The observation is correct and points to the resolution of the complications that brings together competitive equilibrium in the energy market, economic dispatch and transmission congestion contracts.

Competitive Market Equilibrium Corresponds to an Economic Dispatch.

The competitive market equilibrium trading of transmission and energy would support equilibrium opportunity cost pricing of transmission capacity reservations. Although decentralized trading is not possible, the system operator could coordinate the transmission and energy trading market according to the well known principles of economic dispatch. As summarized in Figure 6, coordinated transmission trading through the system operator is equivalent to bid-based economic dispatch.

Figure 6



The end result of this outline implies that the system operator should accept both transmission schedules for receipts and deliveries and energy bids from market participants.[31] The energy bids are the functional equivalent of the bids for PORs and PODs. The bids contain all the information needed to coordinate any desired reconfiguration of transmission capacity reservations. The final use of the system will achieve the competitive market equilibrium and the efficient outcome of the economic dispatch. "[A] holder of a capacity reservation would not pay opportunity costs for use of its own capacity." Capacity reservations not actually used by transmission schedules would in effect be purchased at the equilibrium opportunity cost prices, reconfigured by the system operator to match actual use, and sold to others at the corresponding equilibrium opportunity cost prices.

This last step moves the final distance in the transition from the strictly physical interpretation of transmission capacity reservations to the financial interpretation of transmission congestion contracts. The actual use of the transmission grid does not have to equate to the allocation of transmission capacity reservations achieved through decentralized trading. Such trading would occur, but we would not depend on this trading to manage the use of the system. The actual use of the system would be based on the principles of security constrained economic dispatch. To the extent that the actual use deviated from the allocation of capacity reservations, the settlement process would compensate the holders of capacity reservations at opportunity cost prices for anything they did not use themselves, and charge others who actual used the system at the same consistent set of opportunity cost, locational prices. The complicated, impossible to implement, iterative trading process assumed under the strictly physical interpretation of transmission capacity reservations would be replaced by the necessary reservation price bidding process for energy, at least at the margin, and the determination of locational marginal cost prices. The capacity reservations are then seen as the same thing as financial transmission congestion contracts. By the definition of transmission congestion contracts, the holder of the contract receives the opportunity cost price of congestion between locations.[32] If the holder of a transmission congestion contract actually uses the system, the net payments cancel and "a holder of a capacity reservation would not pay opportunity costs for use of its own capacity." If the holder of the transmission congestion contract does not use the full allocation, there is automatic compensation for the difference at the opportunity cost price of congestion.

The actual mechanics of the system would be simple. Implementing this system in existing power pools would require no more than replacing cost-based dispatch systems with bid-based dispatch systems, allowing "physical" transmission schedules, and applying a new pricing mechanism consistent with a competitive electricity market. In its essential details, this is the

---

[31] To anticipate some of the common misperceptions, this may be the place to remind the reader that the transmission schedules can include bilateral "physical" transactions; energy bidding is voluntary; the coordination is only for the short term period of the dispatch; and so on. All this is compatible with a truly competitive electricity market.

[32] For a detailed discussion of transmission congestion contracts, see S. Harvey, W. Hogan and S. Pope, "Transmission Capacity Reservations and Transmission Congestion Contracts," Harvard University, June 6, 1996, (revised); or earlier, D. Garber, W. Hogan, L. Ruff, "An Efficient Electricity Market: Using a Pool to Support *Real* Competition," The Electricity Journal, Vol. 7, No. 7, September 1994, pp. 48-60.

"contract network" approach.[33] The system would be workable, practical, and in its major parts, familiar. Most importantly, the resulting pricing would be consistent with the actual use of the system, the conditions of competitive equilibrium, and comparable open access to the transmission grid.

The contract network approach and its key elements -- access to essential facilities including the wires and pool dispatch, use of short-run marginal-cost pricing, and reliance on long-term contracts to provide economic hedges rather than specific performance -- are not far from actual operations or proposed reforms in other systems. The FERC proposed capacity reservation system moves very far in this direction, posing a point-to-point transmission reservation definition that does not depend on a decomposition and tracking of the actual flows. Decentralized trading of these capacity reservations would not be enough to support a competitive market, but this trading could be coordinated through the system operator. Opportunity cost pricing would define the trading criterion. Because of the strong and unavoidable network interactions, opportunity cost pricing for transmission and energy cannot be separated, but they would arise naturally as the result of bidding and economic dispatch offered by the system operator. Locational prices would define the opportunity costs for energy bid through the spot market, and the difference in locational prices would define the opportunity cost for transmission scheduled in addition to the spot market transactions. Then tradable point-to-point capacity reservations with opportunity cost pricing for unused or overused amounts would be functionally and financially equivalent to transmission congestion contracts. Conversely, transmission congestion contracts would be functionally and financially equivalent to the tradable point-to-point capacity reservations, would be easier to manage, and would fully support the competitive market while being fully consistent with the actual use of the transmission grid.

This approach, therefore, is consistent with the FERC transmission capacity reservation proposal; it must be what FERC intended, and will embrace.

MARKET EQUILIBRIUM AND ECONOMIC DISPATCH

There is a close connection between the conditions of competitive market equilibrium and the achievement of an economic dispatch. Often this connection is implicit, but it is so important in the discussion of decentralized trading and coordinated dispatch, that there is an advantage in making the connection explicit. In particular, when decentralized trading is not sufficient or practical, coordinated economic dispatch provides an alternative for describing and achieving a competitive market equilibrium outcome.

---

[33] W. Hogan, Contract Networks for Electric Power Transmission," Journal of Regulatory Economics, Vol. 4, 1992. See also W. Hogan, "Coordination for Competition in an Electricity Market," Response to an Inquiry Concerning Alternative Power Pooling Institutions Under the Federal Power Act, Docket No. RM94-20-000, March 2, 1995.

Competitive market equilibrium results in a set of generation, loads, transmission use, and equilibrium prices where there are no further profitable opportunities for trade. Economic dispatch produces a set of generation, loads, transmission use, and equilibrium prices that maximize the total net benefits of all customers. Under certain regularity conditions, the results are the same. However, the process of arriving at the result can be quite different. In many cases, with transmission congestion and strong network interactions, it may be practically impossible to converge to a market equilibrium through decentralized trading. Under the same conditions, however, a pool-based economic dispatch using customer bids would be both practical and no more than a minor modification of the procedures long-used in existing power pools.

Any efficient system for organizing the electricity market should include economic dispatch as a centerpiece. To be sure, the economic dispatch concentrates only on the short-run and the greater part of the value of a competitive system is to be found in the long-run decisions that will control contracting and investment. However, economic dispatch based on participant bids is the ideal short-run outcome that would appear in a competitive market if it were possible for all the many participants to define the appropriate property rights and conduct all the complex trades in the network. Because of the complexity of these trades or the lack of workable definitions of key physical property rights, the common judgment is that a system operator is needed to coordinate the dispatch, at least for some fraction of the flexible plants. Since the operator must function to provide coordination services, economic dispatch provides the natural framework that replicates as close as possible the ideal outcome of the short-term competitive market. And working from this starting point, the other features needed for the market can be derived within a consistent framework. Energy bidding and economic dispatch provide naturally the level playing field for all market participants, both large and small. There would be no special advantage to size in benefitting from dispatch diversity and acquiring backup supplies. These services would be available to all on the same basis. The separation of ownership from control would guarantee open access to the dispatch and related services to facilitate entry and the pursuit of the forces of competition.

For the purpose of this discussion, assume that the generation and load markets are competitive. In the presence of market power, pricing and access rules might not rely fully on the competitive model, but this competitive case provides the foundation for a broader discussion. In addition, this is the implicit assumption which conditions most debates about transmission pricing and access rules.

Network Representation

The focus here is on the dispatch of power plants and use of the transmission system for a given short-run period. The configuration of the grid is presumed to be known, and the value of demand and supply can be determined by the market participants. If the demand for (real and reactive) power at each location or bus is "d" and the generation is "g", let y = d-g be the vector of net real and reactive loads at each bus.[34]

If we know the net load at each bus, then it is possible to perform a standard load flow calculation to determine the power flows on the lines and the voltages at the buses. We assume that the power system must operate to meet certain constraints on these variables. The typical practice is to limit the net loads so that the constraints are met under a variety of contingencies, such as the loss of a line. Each contingency, in principle, will change the flows on the grid and the impact on the different constraints.

Let ω index the possible contingencies, then for each contingency there is a set of constraints,

$$K^{\omega}(y) \leq b^{\omega} \quad ,$$

where $K^{\omega}(y)$ is the set of line flows, voltages and other results as determined by the electrical laws governing power systems, and that must meet the various physical limitations of the system.[35] In principle, the are a many contingencies and a large number of constraints for each contingency.

The balance requirement for real and reactive losses, $L(y)$, is $L(y) + e^t y = 0$, where matrix "e" is the elementary matrix for summing the respective net real and reactive loads. In other words, the total generation must equal to the total load plus losses for both real and reactive power. To simplify the notation, we can represent this total power balance requirement as two inequalities:

$$L(y) + e^t y \leq 0$$
$$-L(y) - e^t y \leq 0 \quad .$$

Let these loss balancing requirements be included as the first constraints on net loads. Then with the loss balance and the set of all contingency constraints, we can represent the combined vector

---

[34] The sign convention reverses the approach in Schweppe et al., but simplifies the interpretation of prices. F. C. Schweppe, M. C. Caramanis, R. D. Tabors, and R.E. Bohn, Spot Pricing of Electricity, Kluwer Academic Publishers, Norwell, MA, 1988.

[35] Notably, the power flow follows Kirchoff's laws, hence the notation "K." The aggregate power flows and voltages, and therefore the elements of K, are defined implicitly as the solution to a set of nonlinear equations. For a review of basic circuit theory, see D. A. Bell, Fundamentals of Electric Circuits, 4th ed., Prentice Hall, Englewood Cliffs, New Jersey, 1988. The basic results for "per unit" systems and transmission lines are developed in A. J. Wood and B. F. Wollenberg, Power Generation, Control, and Operation, John Wiley and Sons, New York, 1984, p. 75; or in more detail in A. R. Bergen, Power Systems Analysis, Prentice Hall, Englewood Cliffs, New Jersey, 1986, chapters 5 and 6; or in O.I. Elgerd, Electric Energy Systems and Theory, McGraw-Hill Book Company, 2nd. ed., New York, 1982. See also P. M. Anderson and A. A. Fouad, Power System Control and Stability, Iowa State University Press, Ames, Iowa, 1977, for an earlier development and further discussion of stability issues.

of constraints as

$$K = \{L(y) + e^t y, -L(y) - e^t y, K^1, K^2, \ldots, K^\omega, \ldots, K^m\}, \quad \text{and } b = \{0, 0, b^1, b^2, \ldots, b^\omega, \ldots, b^m\},$$

with

$$K(y) \le b.$$

The corresponding rows of K and b identify constraints indexed by i. Typically there will be a very large number of potential constraints. In principle, if every line and every bus is potentially constrained, in every contingency, then the number of constraints would be of the order of the sum of the number of lines and buses times the number of contingencies. In the eastern interconnect, typically modeled with 10^4 buses and of order 10^3 lines, the potential number of constraints could be of order 10^8. If, as in practice, attention is restricted to only a limited number of the most significant contingencies, with a large number of possible constraints for each contingency, the number of monitored constraints might still be very large.

The net loads "y" define the inputs and outputs of the grid. Implicit in the set of inputs and outputs is the implicit point-to-point transmission service. For example, if each of the users schedules receipts and deliveries, these individual inputs and outputs amount to a vector of net loads y_k for the kth market participant, summarizing that user's net inputs and outputs at each location. The simultaneous feasibility condition amounts to $K(\sum y_k) \le b$.

Competitive Market Equilibrium

Assume that each market participant has an associated benefit function for electricity defined as $B_k(y_k)$, which is concave and continuously differentiable.[36] In the FERC terminology, the market participants are the transmission service customers. The customers' benefit functions can arise from a mixture of load or demand benefits and generation or supply costs. In this framework, the producing sector is the electricity transmission provider, with customers putting power into the grid at some points and taking power out of the grid at other points. The system operator receives and delivers power, providing transmission service across locations.

The competitive market equilibrium applied here is based on the conventional partial equilibrium framework that stands behind the typical supply and demand curve analysis.[37] The market consists of the supply and demand of electric energy and transmission service plus an aggregate or numeraire "good" that represents the rest of the economy. Each customer is assumed to have an initial endowment w_k of the numeraire good. In addition, each customer has an ownership share s_k in the profits "π" of the electricity transmission provider, with $\sum s_k=1$.

Under the capacity reservation system, each customer has a capacity reservation t_k which is the combination of net loads at each bus reserved under the point-to-point rights. If the customer uses more or less of its capacity reservation, say y_k, it pays for or sells the difference of $y_k - t_k$ at the market price, or $p^t(y_k - t_k)$.

An assumption of the competitive model is that all customers are price takers. Hence, given market prices p, the customers choose the level of consumption of the aggregate good, c_k, and electric energy including the use of the transmission system according to the individual

---

[36] A sufficient condition for these to obtain would be that the demand and supply functions at each node are continuous, additively separable and aggregate into a downward sloping net demand curve. The benefit function would be the area under the demand curves minus the area under the supply curves in the usual consumer plus producer surplus interpretation at equilibrium. To avoid notational complexity, the assumption here is that each participant has a continuously differentiable concave benefit function defined across the net loads at every location. Concavity is important for the analysis below of the equivalence of economic dispatch and market equilibrium, if there is a market equilibrium. This would eliminate from this competitive market analysis the related unit commitment problem which includes nonconcave start-up conditions. As is well known, in the presence of nonconcave benefit functions there may be no competitive market equilibrium. Differentiability can be relaxed, with no more than the possibility of multiple equilibrium prices. Restricting the benefit function to definition at a subset of the locations would be more realistic, but different only in the need to account for the corresponding variable definitions. It would not affect the results presented here. In practice, as is often assumed, the benefits functions may be separable across locations.

[37] The partial equilibrium assumptions are that electricity is a small part of the overall economy with consequent small wealth effects, and prices of other goods and services are approximately unaffected by changes in the electricity market. See Mas-Colell, A., M.D. Whinston, and J.R. Green, Microeconomic Theory, Oxford University Press, 1995, pp. 311-343. Importantly, we adopt here a relaxed set of assumptions that do not include convexity of the set of feasible net loads.

optimization problem maximizing benefits subject to an income constraint:[38]

$$(CB_k) \qquad \max_{y_k,\, c_k} \; B_k(y_k) + c_k$$

$$\text{s.t. } c_k + p^t(y_k - t_k) \leq \; w_k + s_k\pi \; .$$

In this simple partial equilibrium model of the economy, there is only one producing entity, which is the system operator providing transmission service. Under the competitive market assumption, the producer is constrained to operate as a price taker who chooses inputs and outputs that are feasible and that maximize profits. The inputs are the capacity reservations that can be traded or reconfigured, t_k, and the outputs are the net loads at the buses, y_k. The profits amount to $\pi = p^t\Sigma(y_k - t_k)$. Hence, the transmission system operator's problem is:

$$(TO) \qquad \max_{y_1, \ldots, y_n} \; p^t\Sigma(y_k - t_k)$$

$$\text{s.t. } K(\Sigma y_k) \leq b.$$

Of course, the transmission service provider is a monopoly and would not be expected to follow the competitive assumption in the absence of regulatory oversight. However, the conventional competitive market definition provides the standard for the service that should be required of the system operator.[39]

Given the capacity reservations t_k, the initial endowment of goods w_k, and the ownership shares s_k, the competitive market equilibrium is defined as a vector of prices, p, and a set of net loads $y^e_1, \ldots, y^e_n$, which simultaneously solve (CB_k) and (TO).

The competitive equilibrium will have a number of important properties that we can exploit. First, note that $\Sigma c_k = \Sigma w_k$, which is implied and necessary for feasibility. Furthermore, every customer's income constraint is binding and the derivative of each benefit function will equal the common market prices, $\nabla B_k^t = p$. Hence, the equilibrium price at each location is equal to the market clearing marginal benefit of net load and the marginal cost of generation and redispatch to meet incremental load.

By assumption, the allocation of the capacity reservations is simultaneously feasible,

---

[38] The variables under the "Max" identify the decision variables. The abbreviation "s.t." stands for "subject to" and identifies the additional constraints that limit the choices for the decision variables.

[39] It is the standard formulation to include (CB_k) and (TO) as part of the definition of competitive market equilibrium. Failure to follow this well established convention leads to confusion when the term "market equilibrium" is applied excluding the producing sector in (TO), as in Wu, F., P. Varaiya, P. Spiller, and S. Orren, "Folk Theorems on Transmission Access," Journal of Regulatory Economics.Vol. 10, No. 1, 1996, pp. 5-24.

$K(\sum t_k) \leq b$. Hence, the market equilibrium must satisfy $p^t\sum(y^e_k - t_k) \geq 0$. In other words, the amount the system operator collects from the users of the system at opportunity cost prices, $p^t\sum y^e_k$, is at least as large as the obligation to pay holders of capacity reservations at opportunity cost prices, $p^t\sum t_k$. Hence, even though the actual use of the system, y^e_k, may be very different than the capacity reservations, t_k, the payments for the actual use of the system always equal or exceed the obligations to the holders of the capacity reservations. All the network interactions are internalized in the dispatch and pricing. Load is always met, so there is physical delivery of energy, either through long distance transmission or displacement and local generation. Unlike the problems of guaranteeing physical delivery of both energy and transmission, it is easy to guarantee physical delivery of energy and financial delivery of transmission. In short, the system operator always sees the trading through the transmission system as revenue adequate.[40]

Economic Dispatch

The problems in (CB_k) are the customers' individual benefit maximization problems with an income constraint. Problem (TO) has three interpretations. First we have the conventional partial equilibrium model of the profit maximizing transmission service provider, taking prices as given.[41] Second, in the context of a system of capacity reservations, we could interpret (TO) as describing a system operator coordinating trades of transmission rights at the equilibrium price. Under this second interpretation, at equilibrium we have the no arbitrage condition: the transmission coordinator reports that at the current prices there are no further transmission capacity reservation trades or possible reconfigurations which would produce a positive profit. All the existing t_k have been traded or reconfigured into y^e_k, and there is no other configuration of system use that would be both feasible and profitable at prices p.

This interpretation as coordinated trading of transmission capacity rights includes a provision for a reconfiguration of the existing rights and the creation of new rights that are feasible but have not been allocated. Opportunity cost payments for the rights traded consist of $p^t t_k$. Payments for the newly created rights constitute the objective function, $p^t\sum(y_k - t_k)$. In general, $p^t\sum(y_k - t_k)$ is greater than zero, and these payments would be distributed according to the shares s_k. For example, s_k might be determined according to the payment for the fixed charges of the network, and the payments for unassigned rights would be distributed in the form of reduced access charges for the system. In the special case of $\sum t_k=0$, none of the rights have been assigned, all the acquisition of new rights comes from the unused capacity in the network, we would have $\pi = \sum p^t y_k$, and the transmission capacity right sale at opportunity cost prices would

---

[40] For further details, see W. Hogan, "Contract Networks for Electric Power Transmission," Journal of Regulatory Economics, Vol. 4, No. 3, September 1992, pp. 211-242. W. Hogan, "Markets in Real Electric Networks Require Reactive Prices," Energy Journal, Vol. 14, No. 3, 1993, pp. 171-200. W. Hogan, "Electric Transmission: A New Model for Old Principles," The Electricity Journal, March 1993, pp. 18-29.

[41] Ginsburgh, V.A. and J.L. Waelbroeck, Activity Analysis and General Equilibrium Modeling, North-Holland, 1981.

yield revenue $\sum p^t y_k$ that would be netted against the network fixed charges.

The third interpretation is as an economic dispatch. If there is a market equilibrium p and $y^e_1, \ldots, y^e_n$, and if each B_k is concave and continuously differentiable, then we also have a solution to the problem:[42]

$$\max_{y_1, \ldots, y_n} \quad \sum B_k(y_k)$$

$$\text{s.t.} \quad K(\sum y_k) \le b.$$

Given the benefit functions, $B_k(y_k)$, this problem is simply the economic dispatch problem.[43] If the customer demand curves, $y_k(p)$, are all invertible, we could simplify the statement of the economic dispatch problem by using the notation $B(y) = \sum B_k(y_k)$ and $y = \sum y_k$. With the requirement that $\nabla B_k^t = p$ for all k, the individual allocations would follow immediately. Hence, we can refer subsequently to the notationally simpler economic dispatch problem as:

$$\text{(ED)} \qquad \max_{y} \quad B(y)$$

$$\text{s.t.} \quad K(y) \le b.$$

Without knowing the benefit functions, searching for an equilibrium price vector, p, while repeatedly solving the decentralized problems (CB_k) as well as (TO), would amount to an iterative real-time pricing approach as described by Schweppe et al.,[44] and would likely be beyond the capability of the existing system. However, given B(y), workable methods for approximately solving the economic dispatch problem in (ED) are familiar and well developed.[45]

Under reasonable assumptions, an acceptable approximation of the benefit functions

---

[42] To verify this, assume it is false. Then there must be another feasible $y'_1, \ldots, y'_n$, such that $\sum B_k(y'_k) > \sum B_k(y_k)$. Therefore, by concavity of B_k, $p^t\sum(y'_k - y^e_k) = \sum \nabla B_k(y'_k - y^e_k) \ge \sum B_k(y'_k) - \sum B_k(y^e_k) > 0$. Hence, $p^t \sum y'_k > p^t \sum y^e_k$, a contradiction of the optimality of $y^e_1, \ldots, y^e_n$ in (TO).

[43] This is also called the optimal power flow problem. For fixed demands, this reduces to the "least cost" dispatch. The terms are often used interchangeably, with "least cost" not intended to foreclose demand responses.

[44] Schweppe, F.C., M. C. Caramanis, R. D. Tabors, and R.E. Bohn, Spot Pricing of Electricity, Kluwer Academic Publishers, Norwell, MA, 1988.

[45] System operators have a great deal of experience in approximately solving operational economic dispatch problems, often with the support of computer models. See W. F. Tinney and D. I. Sun, Optimal Power Flow: Research and Code Development, Electric Power Research Institute, EL-4894, February 1987. For an interesting application to the case of reactive power pricing, see Dandachi, N.H., M.J. Rawlins, O. Alsac, M. Prais, B. Stott, "OPF for Reactive Pricing Studies on the NGC System," IEEE PES PICA Conference, May 1995.

can be obtained through bidding mechanisms such as that employed in the England and Wales system and elsewhere, and being proposed in various parts of the United States.[46] Under the competitive assumption that no individual customer can significantly affect the price, the policy of accepting supply and demand bids at various locations, constructing the benefit function from these bids, solving for an economic dispatch, and then paying the market price, provides the right bidding incentives, and the best strategy for each customer would be to bid its true opportunity cost.[47]

With enough participant bids available, the system operator can solve the economic dispatch problem, recognizing and accommodating the complex network interactions. The network constraints are many and highly nonlinear. In general, they should be continuously differentiable and easily satisfy one of the strong constraint qualifications.[48] With these mild regularity requirements, the solution of the economic dispatch problem would yield a set of prices that meet a necessary condition for both the market equilibrium condition and the optimal dispatch; namely that

$$p^o = \nabla B_k^t = \nabla K^t \mu, \text{ with } \mu \geq 0.$$

Note that the Lagrange multipliers μ are non-zero only for the binding constraints. Typically, although there would be many potential constraints, the number of binding constraints would be small. Therefore, the marginal cost prices for all the locations could be obtained knowing the prices, or bounds on the prices, at a few locations and the identity of the binding constraints in K. Hence there will be a set of prices associated with the economic dispatch problem. Furthermore, although the economic dispatch problem may be difficult to solve, it would be an easy matter to determine a consistent set of prices for a given economic dispatch.[49]

---

46 H. Rudnick, R. Varela and W. Hogan, "Evaluations of Alternatives for Power System Coordination and Pooling in a Competitive Environment," IEEE Transactions on Power Systems, forthcoming.

47 Strictly speaking, with discontinuous supply and demand curves. for the bidder at the margin the best strategy would be to bid the next highest bidder's opportunity cost, and for everyone else the best strategy would be to bid their true opportunity cost. If a bidder owns large or multiple plants, and can affect the market prices, the usual market power arguments apply and the bidding may not be incentive compatible.

48 The requirement would be no more than an independence of the gradients of the real and reactive loss functions and the existence of some otherwise unconstrained solution in the neighborhood of the optimal solution. See Fiacco, A.V., and G. P. McCormick, Nonlinear Programming: Sequential Unconstrained Minimization Techniques, John Wiley and Sons, New York, 1968, p. 21.

49 Solving for the prices for an optimal dispatch reduces to the solution of a small set of linear equations. If the benefit function is concave but not differentiable everywhere, solving for an equilibrium set of prices reduces to relatively small linear program. For a further discussion, see Hogan, W. , E.G. Read and B. J. Ring, "Using Mathematical Programming for Electricity Spot Pricing," Presented at "Energy Models for Policy and Planning", London Business School/IFORS First Joint International Symposium, London, July 18-20, 1995, forthcoming International Transactions in Operational Research.

Ignoring losses and in the absence constraints, there would be a single market clearing price in the system. In the presence of constraints, transmission congestion and the natural differences in the costs of generation or the value of load interact in the network to create different prices at different locations. It is common to speak of the effect of congestion as separate from the marginal losses. However, treating the total system balance requirement as just another set of constraints emphasizes the similarity. Losses could be viewed as a continuous form of congestion, a constraint that is always binding. Congestion in the grid creates the differences in locational marginal costs. And in a sufficiently interconnected system, which is often the case, a single constraint could, and often would, give rise to a different price at every location.

In general, if y^o is the solution of the economic dispatch problem, something more would be required to guarantee that (p^o,y^o) corresponds to a market equilibrium. Convexity of the set $\{y \mid K(y) \leq b\}$, at least over the normal operating range, would suffice. But this is not true in general, not even for the component functions other than the loss components. It is easy to construct examples where some elements of K might be strictly concave rather than convex. Hence, we cannot establish in general that (p^o,y^o), a solution for (ED), would also be a solution for (TO). In other words, there may be no price that produces a market equilibrium and satisfies the no arbitrage condition. For the same reason, revenue adequacy might not be guaranteed.

However, the conditions on K for which there would be no market equilibrium would seem to be unusual, caused by low load conditions or other special cases that would not arise in practice, at least not simultaneously with a constrained transmission system. In the event such non-convexities did arise, it would seem better to use (p^o, y^o) from the economic dispatch as though it were also a complete equilibrium solution. It would always be true that this would be a feasible solution, and would be consistent with the individual customer preferences in (CB_k). A principal anomaly might be that π could be negative, and there might be a requirement to pay into the system operator enough money to guarantee payments for the existing capacity reservations t_k. In addition, the economic dispatch problem might be particularly difficult to solve. In practice, the methods for solving the economic dispatch problem produce a local optimum, but the existence of significant non-convexity in the constraints would complicate the problem of ensuring a global optimum. It is a conjecture that this circumstance would be rare, or non-existent in practice. However, it cannot be ruled out in theory.[50]

As a practical matter, therefore, we describe the competitive market equilibrium as a no arbitrage condition that is equivalent to solving the economic dispatch problem based on

---

[50] Voltage magnitude can be a concave function of net loads. A suitable upper bound on magnitude would produce a nonconvex set of feasible loads. With an appropriately chosen benefit function, we could have a global optimal solution of (ED) that was not a solution of (TO). Note that this type of nonconvex problem is distinct from the nonconvexity in B that would arise in the unit commitment problem. For the full unit commitment problem, market pricing alone may not be sufficient to cover costs, and a separate "uplift" is typically required as, for example, in the market in England and Wales. This unit commitment problem is beyond the scope of the analysis of a market clearing price and competitive equilibrium.

customer bids and transmission schedules. To the extent that the equivalence is not complete, the economic dispatch formulation is preferred.

This structure and its key elements -- access to essential facilities including the wires and pool dispatch, use of short-run marginal-cost pricing and reliance on long-term contracts to provide economic hedges rather than specific performance -- are not far from actual operations or proposed reforms in other systems. This competitive market is the essence of the design of the current U. K. system, with the notable difference of the lack of locational short-run prices.[51] Locational prices are applied in Chile and New Zealand with an explicit treatment of losses and implicit use of congestion costs. Norway applies both losses and congestion costs. Transmission congestion contracts to hedge against locational cost differentials appear in several pooling proposals.

TRANSMISSION CONGESTION CONTRACTS

The capacity reservations t_k have two natural interpretations. The "physical" interpretation of these point-to-point rights would emphasize the inputs and outputs and match use against reservation, and "[a] customer with a capacity reservation could use the reservation to deliver or receive any type of power product." Customers could trade their capacity rights to allow different receipts and deliveries, and the net trades would be equal to $y_k - t_k$, with a total payment of $p^t(y_k - t_k)$. The individual customers would make their choices subject to the income constraint in (CB_k):

$$(\text{Ip}) \qquad c_k + p^t(y_k - t_k) \leq w_k + s_k\pi .$$

The "financial" interpretation of these point-to-point rights would emphasize the payments, with the holder of transmission reservations receiving $p^t t_k$ in payment for those reservations and the users of the system paying for actual net loads at the same equilibrium prices, or $p^t y_k$. The individual customers would make their choices subject to the income constraint:

$$(\text{If}) \qquad c_k + p^t y_k \leq p^t t_k + w_k + s_k\pi .$$

Under the market theory, the customers take prices, endowments, ownership shares and profits as given and have their capacity reservations, choosing only the total consumption of goods, c_k, and the use of the transmission system, y_k. It follows, therefore, that the problem in (CB_k) is unchanged if we use the physical interpretation in (Ip) or the financial interpretation in (If).

---

[51] S. Littlechild, "Competition, Monopoly and Regulation in the Electricity Industry," U.K. Office of Electricity Regulation, June 1993. U.K. Office of Electricity Regulation, "Pool Price Statement," Birmingham, England, July 1993. The United Kingdom design is plagued by incentive problems created in unusual and unnecessary features of the calculation of market prices coupled with too few generators to ensure competition.

The difference between the two interpretations arises not in terms of the nature of the market or the equilibrium outcome. The difference would arise most likely in the institutions designed for implementation. In the extreme version of the "physical" rights, there would have to be continuous trading, in real time, with the system operator involved in a complicated effort to match or reconfigure many individual capacity reservations. The information requirements and the transaction costs would be daunting. With no more assistance from the system operator than a determination that an isolated trade would be possible, it remains an open question whether this system would ever converge to an equilibrium, and the queue of trading requests might never empty.

The "physical" trading process might be simplified, and made more comparable, if the system operator coordinated an open auction for trading and reconfiguring transmission reservations. This approach would recognize that the system operator must be constantly involved to ensure the feasibility of all but the simplest (and most uninteresting) trades of capacity reservations. Taking the inevitable involvement of the system operator to the point of accepting bids to buy and sell capacity reservations, and then finding the market clearing price for all such trades, would make the system workable.

The next step, from a coordinated auction of capacity reservations to the treatment of the point-to-point reservations as "financial" contracts, is primarily semantic. The natural form of the bids would be in term of prices at different locations for "receipts" and "deliveries." And an auction of this type would reduce to the form of the economic dispatch problem. Anyone who wished to use their capacity reservation could provide a nomination of this type, and exactly match their use y_k with their reservation t_k. Others would receive $p^t t_k$ for their capacity reservation and pay $p^t y_k$ for use of the system, with the net payment being $p^t(y_k - t_k)$.

An advantage of this financial interpretation is that we recognize that the actual economic dispatch consistent with the competitive market equilibrium can be defined and implemented without any direct consideration of the capacity reservations t_k. The financial payments can be left to a settlements system, without requiring the complications of real time trading.

This financial interpretation is the "contract network" perspective.[52] The capacity reservations could be implemented as a generalized system of "transmission congestion contracts" (TCC) which would amount to the right to collect the payment $p^t t_k$. A general description of a TCC could be any vector of net loads on the grid. The typical discussion of TCCs ignores the effects of losses and presumes that the vector describes transmission of a fixed amount of power from a source to a destination in the network. This special case for transmission of "x" MW

---

[52] W. Hogan, Contract Networks for Electric Power Transmission," Journal of Regulatory Economics, Vol. 4, 1992. See also W. Hogan, "Coordination for Competition in an Electricity Market," Response to an Inquiry Concerning Alternative Power Pooling Institutions Under the Federal Power Act, Docket No. RM94-20-000, March 2, 1995.

would be the vector

$$t_{TCC} = x(0,..,\ -1_{source},\ ...,\ 1_{destination},\ ...,0)^t.$$

The payment received under this balanced TCC would be $p^t t_{TCC} = (p_{destination} - p_{source})x$. Ignoring losses and with no congestion, the prices would be equal and the payment would be zero. As congestion appeared, the payment would be for the difference in congestion costs.

This form of a balanced TCC always sums to zero. However, there would be no necessity to impose this balancing requirement on each individual TCC. This would be especially relevant for TCC definitions that included reactive power, where losses would be high and individually balanced TCCs would be of no interest. All that would be required would be that the set of all TCCs would be simultaneously feasible and appropriately balanced. Ignoring losses and limiting attention to balanced TCCs, the payments could be defined only in terms of congestion costs relative to a reference bus. The aggregate physical balance to account for losses could be assigned to the reference bus, where congestion costs would by definition be equal to zero. Including losses presents no conceptual difficulties in the definition of the TCCs, but it does presume that (enough) market participants have TCCs with more inputs than outputs to guarantee simultaneous feasibility including the effects of losses on power flows and prices.

The generalized TCC would be denominated in the quantity of power receipt and delivery at various locations. This is similar to transmission from source to destination, intended to mean the actual flow of power, or at least specific performance on the locational delivery of the power. However, the TCC is not a contract for actual delivery of specific, identified power. The definition assumes that loads will be met either through actual delivery or through displacement. Hence, the actual power flows may be (very) different from the quantities embodied in the collection of TCCs. By contrast to a contract for physical flows, the TCC is a contract for payment of congestion costs defined as the difference in locational prices. These payments are designed so that the user is economically indifferent between meeting the load through actual delivery or through displacement and the effective trading or reconfiguration of the TCC. The result would be equivalent to a system of "physical" capacity reservations with efficient trading and reconfiguration necessarily organized through and by the system operator, but would be much simpler to implement. These tradeable rights are essentially financial commitments and provide an alternative and internally consistent definition of transmission capacity without suffering the defect of being unable to ensure the physical delivery in the actual dispatch.

Forward Contracts

The short-term locational prices for the actual dispatch give us p, the vector of prices for each bus. The contract administered by the system operator calls for a payment by the system operator of $p^t t_k$. For a TCC from a location with a low price to a location with a high price, the payment to the TCC holders would be positive, just compensating for the congestion and loss marginal cost differential in the price of transmission usage. In this case, $p^t t_k > 0$. However, nothing in the above development and interpretation of the capacity reservations or transmission congestion contracts guarantees this result. Implicitly, the above analysis assumed that the contract was of the type of a forward contract, an agreement to provide and take the transmission service. It is an "obligation," not just a "right." Since the network interactions can involve a large amount of displacement, the ability to honor one capacity reservation may depend on the actual use of another.

In another interpretation of TCCs, focussing on congestion only, the forward contract provides for transmission with zero congestion cost. If the cost of congestion were greater than zero, the TCC holder would receive a positive payment. If the cost of congestion between the locations were less than zero, the payment under the TCC would be negative. Of course, if the actual use of the system were the same as the TCC, and $y_k = t_k$, the net payment would still be $p^t(t_k - t_k) = 0$, no matter what the sign of the payment under the TCC.

In one case, the system operator makes a payment to the TCC holder. In the reverse case, the TCC holder would make a payment to the system operator, returning the negative transmission usage charge paid by the system operator. Hence, the TCC would not affect the dispatch or give the holder any control over the use of the transmission grid. In each case the holder of the TCC could perfectly hedge the congestion cost of transmission usage as though power had flowed according to the TCC but free of congestion cost.

In this sense, a TCC is analogous to a forward contract for the spot price of transmission congestion, with an exercise price of zero. If the spot price of transmission congestion were more (or less) than zero, the TCC would exactly balance the spot price payment for the quantity covered by the contract. This TCC could be traded in a secondary market and would provide a contractual mechanism for long-term pricing of transmission in a competitive, open access electricity market.

A necessary condition to guarantee revenue adequacy would be that the collection of TCCs or point-to-point capacity reservations be simultaneously feasible. For any market equilibrium, the payments under the system of locational marginal cost based prices for transmission usage and TCCs at the same equilibrium prices will always be revenue adequate. In the typical dispatch situation, there would be excess payments of profits, $\pi = p^t \Sigma(y_k - t_k)$, which would have to be distributed in some way to preserve the incentives for the system operator and the market participants.

Option Contracts

This interpretation of the point-to-point capacity reservations or TCCs as forward contracts or "obligations" is not the only, or even the most natural approach that comes to mind. In its discussion of the capacity reservation tariff, the FERC takes notice of this under point twelve of its principles: "In addition, the transmission provider also could offer an 'obligation' type of capacity reservation under which the customer would be required to use all of the capacity it has reserved." Apparently the intent, in the first instance, is to treat the point-to-point reservations as a "right" or an option contract, not an "obligation." The notion would be that the customer could choose to exercise the capacity reservation or not, at the customer's discretion.

In terms of the equivalent treatment under the TCC perspective, the option interpretation would be that the TCC holder would accept the payment $p^t t_k$ whenever positive, and the capacity reservation had value, but could elect not to make the any payment when $p^t t_k$ was negative, and the capacity reservation had become a liability relative to the spot market prices.

The practical difference between the two approaches would depend on the nature of the market and the degree to which prices for desired TCCs tended to change sign under different load and operating conditions. Many discussions of TCCs implicitly assume that the sign of the payments would always be positive. If true, then the distinction would be only a technical point. However, as will be discussed below, a useful system of point-to-point capacity reservations or transmission congestion contracts could involve circumstances where the payments under the forward type TCCs would be negative. If options are to be included, therefore, it is necessary to understand the implications.

The discussion of the forward type contracts provides the foundation for the expansion of the analysis to include options. The same set of system constraints apply, namely:

$$K(y) \le b \ ,$$

where y is the vector of net loads on the system. We obtain a competitive equilibrium with locational prices p, and all users pay the cost of transmission, $p^t y$.

Let $\{t_k^f\}$ be the set of TCCs of the forward type. At the market equilibrium prices, the payments to or from the system operator equal $p^t t_k^f$. Positive payments come from the system operator, negative payments are to the system operator.

Let $\{t_j^o\}$ be the set of TCCs of the options type. Only positive congestion payments are made by the system operator to the holder of the TCC. In the event that $p^t t_j^o < 0$, no payment would be made. This is equivalent to leaving the option unexercised whenever the congestion payments are negative or zero.

The feasibility for the TCCs of the forward type would be assured if and only if

$$\text{(Ff)} \qquad K(\textstyle\sum t_k^f) \le b.$$

Let $0 \le \alpha_j \le 1$. One definition for an option-type TCC of inputs and outputs, t_j^o, would allow the choice of a proportional use of some or all of the "optional" reservation. Hence, the exercise of the option would, in the event, amount to actual use of $\alpha_j t_j^o$. Given the nonlinearity of the loss balance equations, this would not be strictly feasible for different values of α_j. However, we could assume that small changes in marginal losses could be allowed for the feasibility test of the option type contracts, in effect dropping the loss balance equations from K. With this understanding, therefore, simultaneous feasibility would be guaranteed for the TCCs of both forward and option type if and only if

$$K(\sum t_k^f + \sum \alpha_j t_j^o) \le b, \quad \text{for all possible values of } \alpha_j .$$[53]

Evidently the set of feasible option type transmission congestion contracts or capacity reservations would be a strict subset of the feasible forward type contracts. Hence, the implied capacity of the network could be considered to be lower under the option contracts.

Using the notation $[z]_i$ to indicate the ith element of the vector z, this condition is equivalent to:

$$\max_{0 \le \alpha_j \le 1} \left\{ \max_i \left[K(\sum t_k^f + \sum \alpha_j t_j^o) - b\right]_i \right\} \le 0.$$

Because we can interchange the order of maximization operators over compact sets, this is equivalent to

$$\text{(Fo)} \qquad \max_{0 \le \alpha_j \le 1} \left[K(\sum t_k^f + \sum \alpha_j t_j^o)\right]_i \le b_i , \quad \text{for all i.}$$

Note the partial simplification, namely this optimization over α is done independently for each constraint i. In practice, we might ignore the loss balance constraints and evaluate only the other constraints assuming the losses would be made up at a reference bus or other loss supply points, apportioning the losses across the various TCCs.

Even with this simplification, the individual optimization problems in (Fo) may be hard to solve. In general, the elements of K are non-linear, often (but not always) convex, and not generally monotonic. Maximizing a convex function over a simple rectangle such as that defining the constraints on α_j must admit an extreme point solution, but except in special situations, this solution may be difficult to find.

The challenge will be in dealing with the potentially large dimensionality of K. Evaluating the feasibility of only the forward type TCCs in (Ff) amounts to solving a load flow problem for each contingency and then evaluating the various constraints. This is a difficult problem, but one that has been dealt with by a variety of means and for which there are workable algorithms and approximations.

---

[53] This observation for the special case of the DC-load model approximation can be found in Wu, F., P. Varaiya, P. Spiller, and S. Orren, "Folk Theorems on Transmission Access," forthcoming, Journal of Regulatory Economics, footnote 16.

Moving to the evaluation of the simultaneous feasibility including options in the optimization problem of (Fo) presents a significant increase in complexity. Typically there would be a large number of option-type TCCs, t_j^o, and a very large number of constraints. Evaluating each constraint in K requires solving a load flow problem, and optimizing over these constraint functions would require many evaluations. The combinatorics could be daunting. The problem is related to the problem of defining the transmission capacity on an interface, across all possible load conditions. The true answer may be difficult to determine, and the result might be a substantially lower capacity than normally assumed.

This individually proportional type of option is not the only approach. For instance, in its principles, FERC observed that "[i]ndividual PORs and PODs need not be 'paired' with each other." Although the proportional option might include multiple locations with inputs and outputs in t_j^o, and the locations are not paired, they are all tied to each other proportionally. Another approach that may be implied in the FERC principles would be to allow the individual t_j^o to be further subdivided into components, t_{jl}^o, with $t_j^o = \sum t_{jl}^o$. For instance, the subcomponents might have only one location, each representing a POR or POD. Then a more flexible option might allow combinations of these reservations, subject to the restriction that the combination is balanced, although unpaired. Hence, if $0 \leq \lambda_{jl} \leq 1$, we would have $t_j^{o*} = \sum \lambda_{jl} t_{jl}^o$ with the constraints that $e^t t_j^{o*} = 0$. For this definition of an option-type contract, the simultaneous feasibility requirement would imply:

$$K(\sum t_k^f + \sum \alpha_j \sum \lambda_{jl} t_{jl}^o) \leq b,$$

for all possible values of α_j and all values of λ_{jl} such that $e^t \sum \lambda_{jl} t_{jl}^o = 0$. Clearly if unpaired options t_{jl}^o of this more flexible type were feasible, then the corresponding aggregations $t_j^o = \sum t_{jl}^o$ would be feasible as proportional options. In this sense, the balanced unpaired options would be a subset of the proportional options.

Evaluating the test for feasibility for the unpaired options would be more difficult than for the proportional options. The counterpart to (Fo) would be

$$\begin{array}{ll} \text{Max} & [K(\sum t_k^f + \sum \alpha_j \sum \lambda_{jl} t_{jl}^o)]_i \leq b_i\ , \quad \text{for all } i, \\ 0 \leq \alpha_j \leq 1 & \\ 0 \leq \lambda_{jl} \leq 1 & \\ e^t \sum \lambda_{jl} t_{jl}^o = 0 & \end{array}$$

which promises to be a more difficult problem to solve.

Other definitions could be imagined. For example, we could consider defining a set of plausible equilibrium prices, P, and then anticipating that for each p contained in P, only those options for which $p^t t_j^o > 0$ would be exercised. Here we might require simultaneous feasibility only for each combinations of options that might be exercised for any price in P. And so on.

It is not clear how to characterize these progressively more complicated definitions of capacity reservations, or how to solve the optimization problems inherent in the feasibility tests. Fortunately, testing for feasibility of an allocation of capacity reservations or transmission congestion contracts is not something that must be done in real time. The initial allocation needs be done only once, and subsequent modifications to reflect long-term reconfigurations or modifications of the grid would be relatively isolated events. Hence, feasibility testing in practice may be able to accommodate a more complicated solution methodology. However, it is clear that seemingly innocuous assumptions, combined with network interactions, can lead to unanticipated difficulties. One implication is the need for making choices and being more explicit about what is intended, and what can be promised, in the definition of capacity reservations.

A Simplified Approximation

For the case of the proportional options, there might be acceptable approximations for solving the problem of maximizing over the combination of α_j for the most violated constraint value, or simplified approximations used in the definition of the point-to-point capacity reservations or TCCs. Here we consider the common DC-load model as an illustration to provide a further interpretation and analysis of the distinctions between forward and option type contracts in this context.

The DC-load model provides an approximate description of the transmission system. There are two primary simplifications. First, the DC-load model addresses only real power and ignores the reactive power component under the implicit assumption that there will be sufficient dispersed availability of reactive power to maintain system voltages. In some applications, the assumption is relaxed partly by imposing limits on real power flows to provide approximate protection to honor voltage constraints.

The second approximation is to linearize the nonlinear power flow equations and constraints. Interpreting now "y" as the vector of real power net loads, we pick an operating point and calculate the derivatives of the constraints relative to an operating point, say y^*, as $H = \nabla K(y^*)$. Then

$$K(y) \approx K(y^*) + H(y-y^*).$$

Hence, the approximate version of the constraints become:

$$Hy \leq b - K(y^*) + Hy^*.$$

This is the form of many optimal power flow and economic dispatch approximations. The discussion of the DC-load application often further simplifies by assuming that $y^* = 0$, and $K(y^*) = Hy^* = 0$.[54] Recognizing that the linearization could be, and is, done about a nearby

[54] For details, see F. C. Schweppe, M. C. Caramanis, R. D. Tabors, and R.E. Bohn, Spot Pricing of Electricity, Kluwer Academic Publishers, Norwell, MA, 1988, Appendix D, which applied this notation for H.

operating point, we adopt the revised notation for the simplified DC-load version of the economic dispatch problem:

$$\text{(c-DC)} \qquad \max_{y} \; B(y)$$

$$\text{s.t.} \quad Hy \leq b.$$

This simplified problem has only linear constraints, so the feasible set of net loads is convex. If the approximation is about the origin, marginal losses are ignored, and the approximate equilibrium prices differ only in terms of the congestion costs.

The matrix $H = \{H^{\omega}\}$ includes all the constraints across all contingencies, and $b = \{b^{\omega}\}$, the vector of constraint values across the same contingencies. The rows of H and b each identify a single constraint indexed by i. Typically there will be a very large number of potential constraints.

In this case, with the feasible set of net loads being convex, the equivalence between market equilibrium and the economic dispatch problem is exact. In other words, market equilibrium is always a solution of the economic dispatch problem, and the economic dispatch solution always satisfies the no arbitrage condition and is a market equilibrium. Revenue adequacy of the TCCs based on the spot prices, in this special case, is guaranteed by the convexity of the solution set.

The revenue adequacy condition is guaranteed by having all the forward type TCCs and the positive option type TCCs being simultaneously feasible. If this set were exercised and were simultaneously feasible, then under any optimal dispatch conditions the congestion payments would be revenue adequate.

The feasibility for TCCs of the forward type would be assured if and only if

$$H\sum t_k^f \leq b.$$

With $0 \leq \alpha_j \leq 1$, feasibility would be guaranteed for the TCCs of the forward and option type if and only if

$$H\sum t_k^f + H\sum \alpha_j \, t_j^o \leq b \quad \text{for all possible values of } \alpha_j \, .$$

As before, with $[z]_i$ as the ith element of the vector z, this is equivalent to

$$\max_{0 \leq \alpha_j \leq 1} \left\{ \max_{i} \left[H\sum t_k^f + H\sum \alpha_j \, t_j^o - b \right]_i \right\} \leq 0,$$

or,

$$\max_{0 \le \alpha_j \le 1} \quad [H\Sigma t_k^f + H\Sigma\alpha_j t_j^o]_i \le b_i, \quad \text{for all } i.$$

For each constraint, this would be a simple optimization problem which exploits the linearity of the functions. It must be that the solution is attained as

$$[H\Sigma t_k^f]_i + \Sigma \max(0, [Ht_j^o]_i) .$$

In the DC-load model, we know the sign of $[Ht_k^o]_i$, so this translates into the set of constraints

$$\text{(Fo-DC)} \quad [H\Sigma t_k^f]_i + \Sigma \max(0, [Ht_j^o]_i) \le b_i, \quad \text{for all } i,$$

which would establish the feasibility conditions for the forward and option type TCCs. For a given constraint, therefore, all the option type contracts with $[Ht_j^o]_i > 0$ go in the same "direction" relative to the constraint, and the guarantee of simultaneous feasibility requires that all option type contracts in the same direction be simultaneously feasible for that constraint.

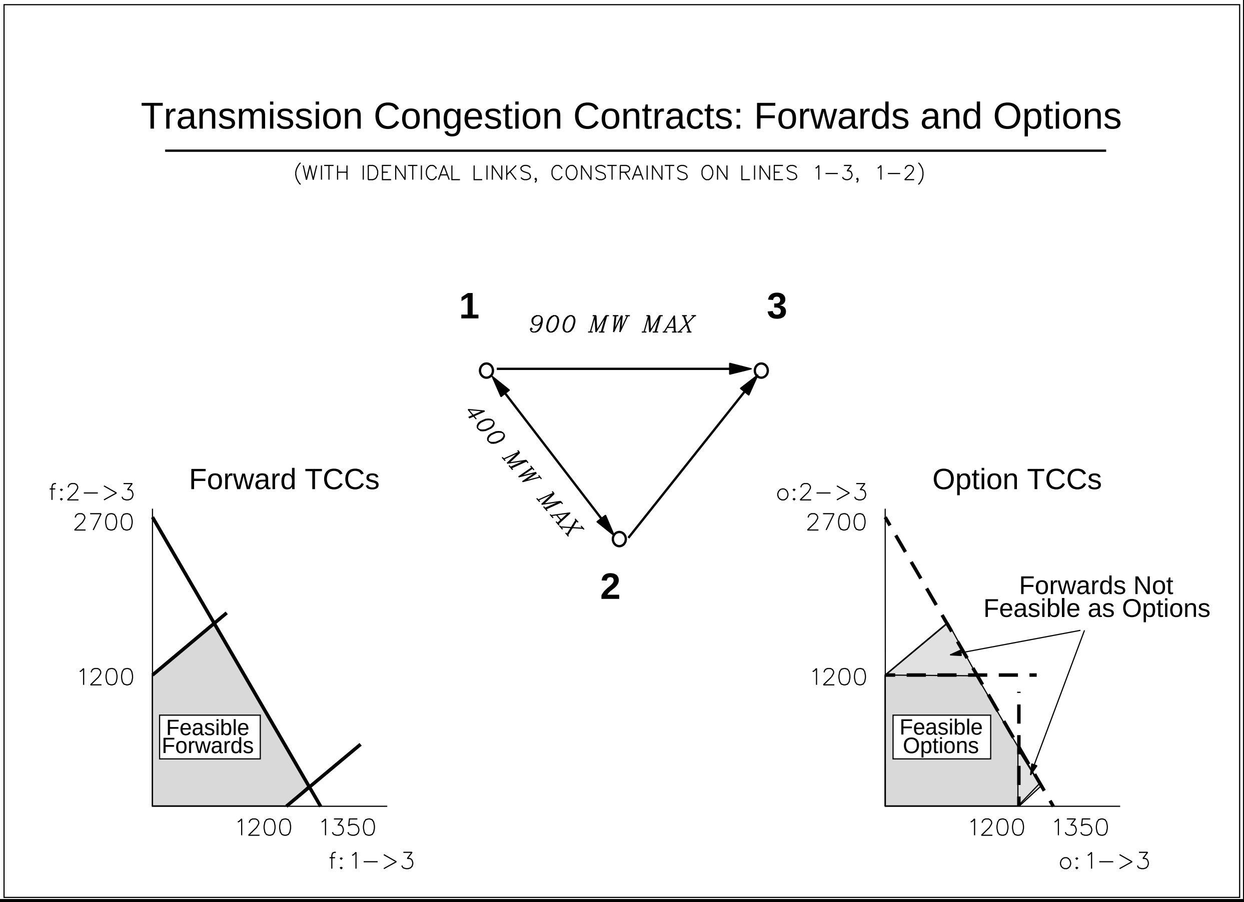
Hence, in the DC-load case, the set of constraints for the option form is not much more difficult, in principle, than the set of constraints for the forward form, and we could accommodate both types in the allocation.

The diagrams in Figure 7 illustrate the set of feasible forward and option TCCs for the DC-load approximation of a small network. For simplicity, we assume balanced TCCs and limit attention to contracts or reservations from bus 1→3 or bus 2→3. With identical lines, other than for the limits, two thirds of the power follows the shorter route, and one third of the power goes the longer way. With respect to the constraint on the line between 1 and 3, both types of contracts are in the same "direction." However, with respect to the limits on the line between buses 2 and 3, the two pairs of contracts are in opposite directions. Scaling the constraints and coefficients of H by the common factor of three, the combined constraints (Fo-DC) for both the forward and option type contracts include:

$$\begin{aligned}
2t_{13}^f + t_{23}^f + 2t_{13}^o + t_{23}^o &\le 2700,\\
t_{13}^f - t_{23}^f + t_{13}^o &\le 1200,\\
-t_{13}^f + t_{23}^f + t_{23}^o &\le 1200,\\
t_{13}^f,\ t_{23}^f,\ t_{13}^o,\ t_{23}^o &\ge 0 \quad .
\end{aligned}$$

If all the contracts or point-to-point capacity reservations were to be of the forward type, then the graph in the lower left of Figure 7 illustrates the possible combinations. If all contracts or point-to-point capacity reservations were of the option type, then the graph in the lower right of Figure 7 illustrates the possible combinations. Of course, the set of feasible option

Figure 7



contracts is smaller than the set of feasible forward contracts. In other words, the network has a lower capacity for option-type contracts than for forward-type contracts.

A mixture of forward and option type contracts would also be possible, with the range of feasible combinations falling somewhere in between the two graphs shown.

A Comparison of Transmission Forward and Option Contracts

Both forward and option transmission congestion contracts could exist in the system. They have certain common features and important differences. A brief summary of these features could provide some guidance into the questions that would need to be answered as part of the institutional design.

Equivalence with point-to-point capacity reservations. As discussed above, both types of transmission congestion contracts would be equivalent with the corresponding "physical" capacity reservations with trading organized through the system operator. Although the details of implementation and transaction costs would be different, most of the basic characteristics of this approach to transmission rights can be derived either from a formulation that describes the system starting with physical capacity reservations or with a system of strictly financial contracts.

Being able to trade capacity reservations would not be the same as being able to perfectly match an intended use of the system over different seasons and conditions. Capacity reservations and transmission congestion contracts could be developed to apply to different time periods. In the U.K., forward contracts for generation are written with different load profiles for each of the 2190 four-hour "Electricity Forward Agreement" (EFA) periods over the year. Similarly, both capacity reservations and transmission congestion contracts could be different for each EFA period. The various feasibility tests discussed above would then be applied separately to each period.

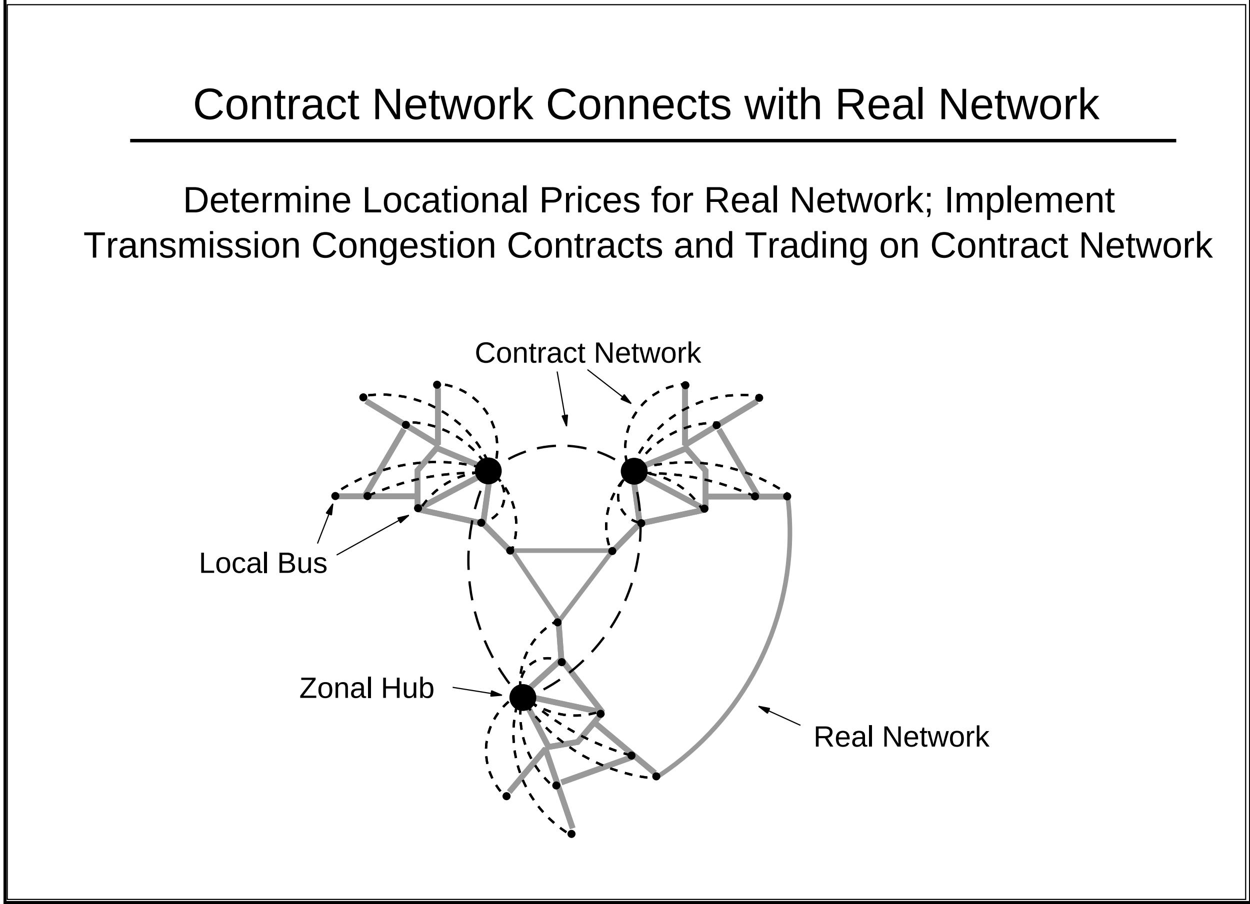
No effect on dispatch. Transmission congestion contracts of either the forward or option type would have no effect on the dispatch. The system operator would not need information on the contracts in order to solicit schedules and bids and then solve for the economic dispatch. In principle, the end result would be true as well for physical capacity reservations of the forward or option type, although under certain interpretations the system operator might have to be involved in the explicit trading and reconfiguration of the reservations on a real time basis. Hence the dispatch process would be affected, if not the result.

Different degrees of difficulty in testing for simultaneous feasibility. In awarding transmission congestion contracts that would apply over many periods where any combination of options might be exercised, there could be a substantial increase in the degree of difficulty of simply testing for feasibility between forwards and options. For a given configuration of the grid, the test of feasibility of the forward type contracts amounts to a calculation of load flows for the various contingencies. For the options contract, an optimization must be performed over these load flows for each constraint. This is possible in the case of the simplified DC-load approximation. It remains an open question as to whether the test could be implemented in the more general case of the nonlinear power flow equations considering real and reactive power.

Options would be a strict subset of the forward contracts. The set of capacity reservations that would be simultaneously feasible under an option type transmission congestion contract would be a strict subset of the forward type contracts. In effect, the option type contract forgoes part of the "expansion" of transmission capacity that comes from displacement of flows that have opposite impacts on transmission constraints. In practice, the actual use of the system, y, will exceed the transmission congestion contracts of the option type, by taking full account of the displacement that is implicit in the actual net loads.

Decomposition limitations. 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 8, the contract network could be different from the real network without affecting the meaning or interpretation of the locational prices.

Figure 8



The same simplification and decomposition could apply to the forward type transmission congestion contracts. A contract t_k^f can be decomposed into components, say t_{k1}^f and t_{k2}^f, with $t_k^f = t_{k1}^f + t_{k2}^f$. In the hub-and-spoke framework, if t_k^f could be from location A to location B, t_{k1}^f could be from A to the hub, and t_{k2}^f could be interpreted as from the hub to B. This form of reconfiguration would always be feasible, and because $p^t t_k^f = p^t t_{k1}^f + p^t t_{k2}^f$, the payments would be the same. These forward type transmission congestion contracts could then be traded independently.

The same decomposition would not apply to the option type transmission forward contract. Under the option, the payment would be $\max(0, p^t t_j^o)$. And for decompositions of the option type contract, we have often that $\max(0, p^t t_j^o) \neq \max(0, p^t t_{j1}^o) + \max(0, p^t t_{j2}^o)$. In other words, the components may have negative price differentials even though the original contract did not. Hence, the equivalence of the hub-and-spoke framework, or any other additive decomposition, would not apply to the option type contract.

Spontaneous market provision. The financial transmission congestion contracts are simply financial arrangements and operate much like other financial arrangements that arise naturally in markets. In principle, and at a price, the market could provide such financial hedges without participation by the system operator. However, the equivalence of tradable capacity reservations and the financial transmission congestion contracts provides a simple explanation of why it would be asking a great deal to rely solely on the market to develop an adequate collection of contracts independent of the system operator. From the physical perspective, it is clear that only the system operator could administer trading and reconfiguration of a system of physical capacity reservations and guarantee simultaneous feasibility. Relying solely on the market to provide such contracts would be equivalent to assuming that the system operator should not make available and administer a collection of capacity reservations.[55]

By the same token, only the system operator can administer a system of transmission congestion contracts that utilize the revenues from the equilibrium market prices to provide a long-term protection with no risk. The transmission congestion contracts of both the option and the forward type, administered by the system operator, would not prevent the spontaneous development of market alternatives, but transmission congestion contracts administered by the system operator would appear to be an essential element of a transmission access and pricing structure.

RELATED PRICING AND MARKET STRUCTURE ISSUES

This overview of the connections between capacity reservations and transmission congestion contracts emphasizes the link with opportunity cost pricing in the short run, and the use of locational marginal cost based prices to define opportunity cost for both electrical energy and transmission. This short-run focus provides the core of the connections and exploits the link between competitive market equilibrium and economic dispatch. In addition, there are other issues that must be addressed. Without fully developing each here, a short summary outlines the further pricing and market structure issues.

Scheduling and Balancing

Implementation of the short-run dispatch market could take many forms. In principle, the coordinated dispatch might be left to only the final hour, with all other commitments and

---

[55] Norway and Sweden together provide an interesting case study. The combined markets as of 1996 employ a common pool with locational differences in prices to reflect transmission congestion. Originally expected to be small, the difference in locational prices and resulting congestion costs emerged is a major economic issue, with participants unable to obtain a market hedge for locational differences. Moen, the Norwegian regulator, reports an effort to develop transmission rights, with the form not yet determined, to deal with the limitations on transmission capacity that increased in importance in the competitive market. (J. Moen, personal communication). A similar experience is reported in Chile. (G. Espinosa, personal communication)

trades developed through decentralized transactions, notifying the system operator of the schedules only in time to complete the final balancing. At the other extreme, with long lead times for changing the configuration of generation or the patterns of loads, it might be preferred to include unit commitment decisions over weeks or even longer periods.

In practice, most countries or regions adopt or recommended a procedure that falls somewhere in between these alternatives. The system operator accepts bids and nominations for scheduled dispatch, say for a day ahead, and determines an appropriate market clearing equilibrium and associated payment settlements. This schedule then defines a set of commitments for delivering and taking power in the short run (say hourly) dispatch. In the event, the actual dispatch will differ from the scheduled commitments, and appropriate balancing settlements would be arranged.

These connected scheduling and balancing settlements present no major difficulties, but there are a few points that need clarification to maintain consistent payments and incentives. The basic schematic appears in Figure 9. Participants in the market submit scheduling bids for the day ahead for both supply and demand. These bids may include start up costs, ramping rates and any of range of plant and load characteristics.[56] The system operator utilizes all this information to define a least-cost dispatch over the day that matches scheduled load and generation. The results is a set of market clearing prices (P) and quantities (Q) that define the schedule. In addition, the system operator arranges for any necessary reliability commitments, such as for spinning and standby reserves.

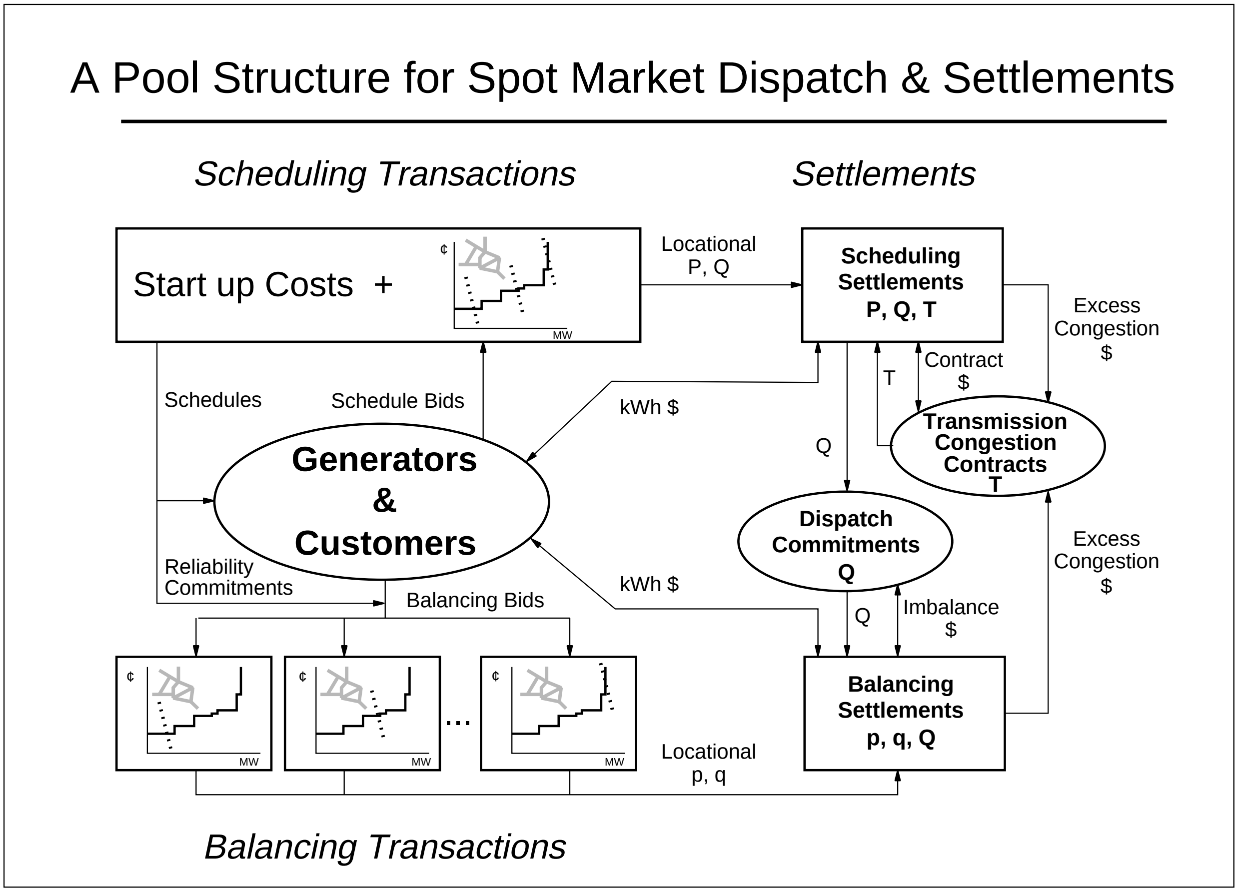
The equilibrium prices and quantities differ by location and represent immediate firm commitments. Based on these commitments, payments would be made through a settlements process. Conceptually these settlements could occur the minute the schedule is determined; in practice, the settlements would occur after the fact. However, to preserve the consistency of the incentives and payments, the settlements must be based on the scheduling prices and quantities. These settlements will include payments under long-term transmission congestion contracts (TCCs) shown by the symbol "T" in the schematic. Holders of TCCs would receive or pay the appropriate amounts of congestion cost differentials between locations for the contracted quantity under the TCC. Under some fixed sharing rule, the TCC holders, or the presumptive parties responsible for paying the fixed charges of the transmission grid, would also share in any excess congestion payments after settlement of the locational differences.

The schedule and the associated dispatch commitments (Q) would provide the reference point for the actual dispatch. In principle, bids in the scheduling market could be revised to create balancing bids for increments and decrements against the dispatch commitments. Again the system operator would find the least-cost dispatch, hour by hour, based on the actual conditions and the final balancing bids, as shown in Figure 10. The result would be an actual dispatch with associated equilibrium prices (p) and quantities (q) of Figure 9. These prices and

---

[56] NGC Settlements Limited, An Introduction to the Pool Rules, Electricity Pool of England and Wales, April 1993.

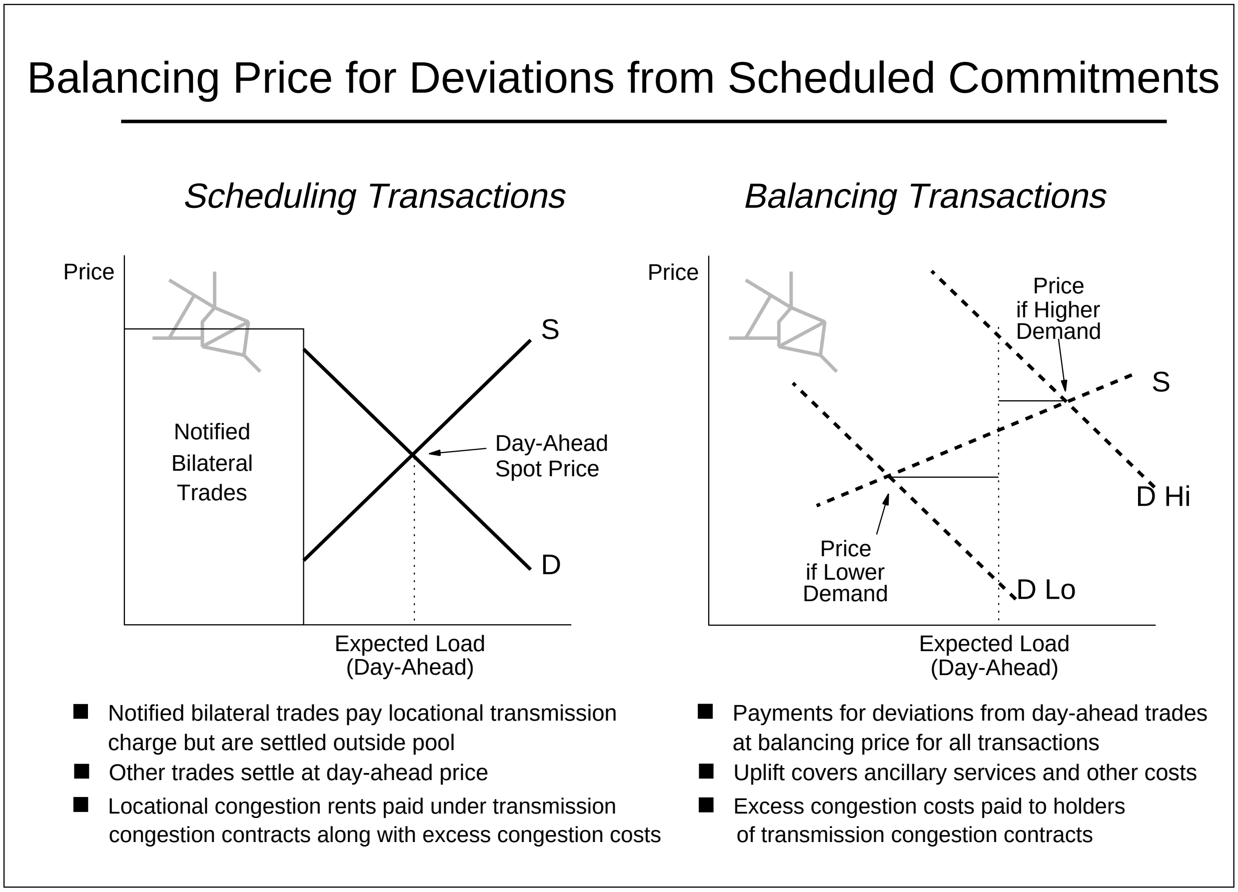
Figure 9



quantities would differ to a degree from the schedules, with the "imbalances" (q-Q) settled at the balancing price of "p". In this balancing settlements, the dispatch commitments are conceptually similar to the TCCs that apply in the scheduling settlements. And just as for the TCCs, after settling all the imbalances at the market clearing price "p", there could be some excess congestion payments that would be disbursed to the users and not kept by the system operator. In part, this excess would be used to reduce user payments for overhead and ancillary services, not shown in the schematic of Figure 9, or rebated again to those who bear the fixed costs of the grid.

The precise treatment of the excess congestion rentals is not important, other than to disburse them to the users and not the system operator in a way that creates no incentives for an inefficient dispatch. What is important, however, is to settle the scheduling markets and balancing markets with their own internally consistent prices (P,p) and quantities (Q,q). This is required, for example, to avoid the initial over-the-day gaming problems created in England and Wales by settling based on scheduled prices (P) and actual quantities (q).

Figure 10



Approximations

The dispatch and pricing model based on economic dispatch could accommodate a great deal of the real system operations and pricing. In principle, the opportunity cost prices of real and reactive electric energy, transmission, capacity reservations, and energy imbalances would be provided automatically with the equilibrium locational marginal cost based prices. Modest extensions to include spinning reserve and other ancillary services would be required, but could be consistent with the basic approach.[57] In some cases, this broader set of prices could also reflect and attribute locational marginal costs. In others, the costs of ancillary services would not necessarily be attributable to individual transactions, and would be collected as part of an average cost uplift applied to all uses.

The static optimization framework simplifies the reality of a dynamic system that will involve interactions over time. Ramping rates and other dynamic limitations will enter the

---

[57] Trans Power New Zealand Limited Transmission Pricing 1993, February 1993.

problem. The usual approach to dealing with these dynamic interactions amounts to a compromise between extending the horizon of the optimization to cover several periods of short-run dispatch and developing rules which interpret the optimal solution in terms of a sequence of static approximations for pricing purposes.

Similarly, start up costs, minimum running rates, and other non-convexities appear in the real economic dispatch problem. The unit commitment problem, which considers these complications, could be included under the generic label as economic dispatch. To the extent that these features are important, there may be no set of equilibrium prices that support a fully decentralized competitive market. However, it would be possible to solve the economic dispatch problem and the come as close as possible to finding a set of locational prices that support the equilibrium solution for the customers. Any deficit in the costs of individual generators or loads and the revenues under the locational prices could be allocated on average as part of a general uplift.[58]

In the United States, at least for the foreseeable future, a single system would not be likely to have control over all interconnected systems. Hence, there must be operating and pricing rules that extend to the flows between and among system operators. In practice, this will mean conducting both dispatch and pricing using simplified representations of the external systems. Depending on the coordination and dispatch rules, a consistent set of ex post pricing approximations would be applied.

In real systems, system operators must deal with all these problems, and it is impossible to provide an *exact* solution to the economic dispatch problem. If the real solution is close enough, however, then the actual dispatch can be interpreted after the fact as optimal, with the prices calculated accordingly. The discrepancies can be minimized, and the resulting prices applied as the best approximation of the locational price representation of opportunity costs.[59]

Best Efforts

The theory of capacity reservations and transmission congestion contracts applies for a given configuration of the transmission grid. The complexity of network interactions implies that even for this simplifying assumption, defining and measuring transmission capacity is a challenge. In reality, the transmission grid is more like a living entity with constant changes in configuration. Explicit modifications of the grid may be relatively rare and easy to accommodate.

---

[58] NGC Settlements Limited, An Introduction to the Pool Rules, Electricity Pool of England and Wales, April 1993.

[59] The perspective of consistent approximation is implicit in every realistic pricing scheme. For an explicit examination of the implications, see Ring, B. J., "Dispatched Based Pricing in Decentralized Power Systems," University of Canterbury, PhD. Dissertation, 1995.

However, lines are constantly being removed or brought back into service, if only for routine maintenance, the capacity of individual facilities depends in part on changing temperature and weather conditions, and so on. Providing the controllers of transmission the incentives to keep the network available and schedule maintenance will present its own challenges.

In testing for simultaneous feasibility of capacity reservations there may be available only an approximate solution. Therefore, it might be that the assignment is not, in all circumstances and under all conditions, actually feasible. Hence, the guarantee of the capacity reservations or payments for any of the form of TCCs would not be absolute. For example, in some cases we may not have enough money to honor all the best efforts contracts, and these receive a pro rata share of the rents after we pay for forward type TCCs. The situation may be different in different hours, with (most) hours showing an excess of congestion payments over transmission congestion contract obligations. However, in some hours this revenue adequacy may not hold, due to the approximations. If the deficit is temporary, the surplus in many periods could be used to fund the deficit in the few. However, if the deficit is persistent, there may be a necessary caveat to apply revenue sharing across all the types of contracts.

Presumably, it would be desirable to minimize the possibility of deficits -- promising more capacity reservations than can be delivered. This may imply both extra effort in carefully defining the meaning of the capacity reservations and transmission congestion contracts, and conservative determination of the evaluation of simultaneous feasibility of the allocations. The best mix is a matter that is not yet well understood.

Fixed Charges and System Expansion

The locational opportunity cost pricing system does not pay for the grid. The fixed charges of the grid, reflect investments that, in principle, have been made to avoid paying the opportunity costs of congestion and more expensive generations. Because of economies of scale and scope, furthermore, typically the fixed charges would be greater than the congestion opportunity costs after the investment was made. It follows that fixed cost recovery for grid investments would be through a combination of access charges and long-term contracts. The capacity reservation or transmission congestion contracts would be the embodiment of the long-term guarantee that would be obtained in exchange for the fixed charge payments. Transmission congestion contracts would provide a way to untangle the network and convert short-run opportunity cost prices into long-term transmission arrangements.

Allocation of embedded costs and transmission rights for the existing system presents a major transitional challenge. However, one implication of the creation of a system of capacity reservations, is that this challenge must be met. To support a competitive market, and to meet the test of comparability, the old implicit allocations of rights must be made explicit. There will be many issues to resolve in this transition, and many ways to allocate costs and benefits without distorting the market. Looking forward, however, the market framework would condition system expansion decisions.

Within the contract environment of the competitive electricity market, new investment occurs principally in generating plants, customer facilities and transmission expansions. In each case, corresponding contract-right opportunities appear that can be used to hedge the price uncertainty inherent in the operation of the competitive short-run market.

In the case of investment in new generating plants or consuming facilities, the process is straightforward. Under the competitive assumption, no single generator or customer is a large part of the market, there are no significant economies of scale and no barriers to entry. Generators or customers could connect to the transmission grid at any point subject only to technical requirements defining the physical standards for hookup. If they choose, new customers or new generators would have the option of relying solely on the short-run market, buying and selling power at the locational price determined as part of the half-hourly dispatch. The system operator itself makes no guarantees as to the price at the location. It only guarantees open access to the dispatch at a price consistent with the equilibrium market. The investor takes all the business risk of generating or consuming power at an acceptable price.

If the generator or customer wants price certainty, then new generation contracts could be struck between a willing buyer and a willing seller, possibly through the intermediation of the aggregators and brokers. The complexity and reach of these contracts would be limited only by the needs of the market. Typically, we expect a new generator to look for a customer who wants a price hedge, and the generator defers investing in new plant until sufficient long-term contracts with customers can be arranged. The generation contracts could be with one or more customers and might involve a mix of fixed charges coupled with the obligations to compensate for price differences relative to the pool price. But the customer and generator would ultimately buy and sell power at their location at the half-hourly price.

If either party expects significant transmission congestion, then a transmission congestion contract would be indicated. If transmission congestion contracts are for sale between the two points, then a contract could be obtained from the holder(s) of existing transmission congestion contracts. Or new investment by the grid owner could create new transmission congestion contracts. In the case of transmission investment, economies of scale and network interactions loom large, unlike the case assumed for generation. Hence, because of economies of scale it is expected that for any given transmission investment there will be a material change in the market prices through reduced congestion rentals. In addition, the network interactions will create many potential beneficiaries.

In a market driven grid expansion process, these facts typically would require that any transmission expansion be organized by a consortium of transmission investors who negotiate a long-term contract that allocates the fixed cost of the investment and the corresponding allocation of new transmission congestion contracts. The grid owner, as a regulated monopoly, builds the lines in exchange for a payment that covers the capital cost and a regulated return. The grid owner does not make transmission investments without long-run contracts signed by willing customers who will pay the fixed costs and recover any future congestion revenues. The system operator participates in the process only to verify that the newly created capacity reservations or transmission congestion contracts are feasible and consistent with the obligation to preserve the existing set of contracts on the existing grid. Unlike in the traditional definition of transmission transfer capacity, which can be ambiguous, there is a direct test to determine the feasibility of any new transmission congestion contracts for compensation, while protecting the existing transmission congestion contracts, and the test is independent of the actual loads. Hence, incremental investments in the grid are possible anywhere without requiring that everyone connected to the grid participate in the negotiations or agree to the allocation of the new transmission congestion contracts.[60]

Grid expansion and pricing would continue to present a need for regulatory oversight, but the existence of workable capacity reservations or transmission congestion contracts would substantially simplify transmission investment decisions. Economies of scale and complex network interactions would continue to create incentives that would not be wholly compatible with decentralized decisions in a market. This need to address network expansion as an integrated problem leads to a continuation of the expected need for something like the Regional Transmission Groups (RTG). An RTG would be needed to review the operating reliability standards and evaluate the impacts of proposed transmission expansions. However, this evaluation need not extend to a central decision on the need or cost responsibility for transmission expansion. The users of the system who are buying and selling electricity without a complete hedge through transmission congestion contracts would face the short-term market clearing price. In the face of transmission congestion, the locational prices provide the proper incentive for investment in transmission facilities. Investments should be made when justified by the savings in congestion costs. Those who are prepared to make the investment would obtain the associated transmission congestion contracts. The role of the states, the RTG and the Commission, therefore, would be to review requests for transmission expansion, examine the compatibility with the companion request for new transmission contracts, and ensure an open process for all to join in developing combined transmission investments recognizing the interactions in the network. The regulator would be responsible for enforcing a requirement for existing transmission facility owners to support expansions and reinforcements at a traditional regulated cost that recovered the incremental investment, and then to assign the corresponding transmission contracts. If no coalition of grid users were able to agree to pay for a grid expansion that appears to be beneficial for the system as a whole, any interested party could propose a project and an allocation of its costs among those grid users who would benefit. Regulatory procedures, similar to those used now, would determine whether the project should go forward and how its costs should be allocated to those expected to benefit from the effect on future locational pool prices, with the payers granted rights to compensation to assure that future

[60] For an investigation of a number of problems, and some solutions, in dealing with transmission expansion, see James Bushnell and Steven Stoft, "Electric Grid Investments Under a Contract Network Regime," Journal of Regulatory Economics, Volume 10, 1996, pp. 61-79. See also, J. Bushnell and S. Stoft, "Grid Investment: Can a Market Do the Job?" The Electricity Journal, January/February, 1996, pp. 74-79.

congestion does not rob them of the benefits they are paying for.[61]

The transmission congestion contracts, once created, would no longer need any special regulation. Although investments in the transmission grid would by lumpy and would require the cooperation of the owners of existing facilities, the transmission congestion contracts would be divisible and freely tradable in a secondary market. This secondary market would provide a ready source of transmission hedges that would serve as an alternative to system expansion. The price of the transmission contracts should never rise above the long-term expected congestion opportunity costs or the cost of incremental expansion of the grid. In this way, the unregulated market for transmission congestion contracts would emulate the broad outlines of the Commission pricing policy. Transmission contracts would be obtained at the lesser of opportunity costs or incremental costs. Holders of existing transmission rights, converted into the appropriate transmission congestion contracts, would pay embedded costs but not opportunity costs. Those using the transmission grid without holding transmission congestion contracts would pay opportunity costs but not any embedded costs other than the costs of any stranded assets that would be collected from all users.[62] Most important of all, the long-term transmission market could be more like a market, relying as much as possible on the incentives and forces of competition, limiting the role of planning and regulation to address the unavoidable interactions in the transmission grid. Investment decisions would be made at the initiative and with the agreement of those required to bear the cost.

Allocating Initial Transmission Contracts

An estimate of the available transmission capacity would be needed and defined only for the combined set of point-to-point contracts. The explicit requirement would be that this set of contracts be simultaneously feasible if no other loads were on the system. There would be no requirement that actual power flows would follow the contracts. Only that the power flowed or the opportunity cost payments followed the terms of the contracts.

Under any scheme, there will likely be a need to allocate the "rights" to the existing transmission grid. For the point-to-point contracts as defined, two broad options are available. First, some contracts could be allocated to the existing rate payers, or their local distribution companies, who are responsible for the fixed costs of the grid. Since the definition depends only on simultaneous feasibility, and does not require any forecasts of future loads, any of a number of allocations would be possible. For instance, the power injections and withdrawals from a previous year would define a collective set of feasible contracts, and would guarantee the ability to repeat what was done last year. These rights might be different for different periods, and

---

[61] D. Garber, W. Hogan, L. Ruff, "An Efficient Electricity Market: Using a Pool to Support *Real* Competition," The Electricity Journal, Vol. 7, No. 7, September 1994, pp. 48-60.

[62] These stranded assets might include transmission sunk costs that could not be recouped through a sale of transmission rights.

could be scaled up to allow for some growth in load.

Second, others who wanted additional transmission contracts as outlined above could submit bids for such contracts to be awarded in an auction. The simultaneous evaluation of the bids would be a problem that is a variant of the least cost dispatch solved by the system operators, and would result in an award of incremental transmission contracts at market clearing prices. The revenues would be applied to reduce the basic transmission fixed charges. In effect, any residual embedded costs would be collected from ratepayers in the same way as other stranded assets.

Again the details need to be developed, but there is nothing especially complicated, in principle, when compared with other approaches that face the reality that whatever the definition of transmission rights, there will have to be an identification and explicit allocation of the rights to the existing system.

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 another demand for continued regulatory oversight. An advantage of the market model with opportunity cost pricing 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.

Opportunity cost pricing and open access help mitigate market power. The system operator 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. Hence, locational marginal cost pricing would reduce market power relative to the alternatives and make the exercise of market power more transparent.

The two ends of the policy spectrum for dealing with the remaining market power are regulation and divestiture. At the regulatory end, firms with a high concentration of generation may be subject to a form of continued cost-based regulation designed to prevent any abuse of monopoly power. For the obvious reason, as a long-term solution, this is an unattractive approach that would be by definition inconsistent with the competitive market. At the other end of the spectrum would be a policy of requiring divestiture of generation into a sufficient number of competing entities. In the U.K., where concentration of generation ownership has produced the expected behavior inconsistent with competitive pricing, the regulator has embraced this divestiture strategy as the principal tool for mitigating the effects of market power. However, the divestiture approach has its own limitations, including what might be the strong objections of the existing utilities who will dispute the existence of market power or argue the inability to exploit what potential power that may exist.

In the middle of the spectrum are many lesser options that could be implemented within the pool-based model and should be explored further. For instance, contracts adopted for a transition period can dramatically alter the incentives of generators with market power. In effect, a long-term power contract at a fixed price transfers the beneficial interest in the plant from the owner to the customer, leaving the generator with the incentives to control costs and maximize the economic use of the plant. This is easy to achieve in the pool-based model and is exactly what happened in the U.K. during the early days of its electricity restructuring. The generators were fully contracted and they behaved like competitors. Only when the generation contracts began to lapse did behavior turn strategic and pricing begin to deviate from the competitive norm.[63] Similar generation contracts could be fashioned in the United States and implemented as contracts for differences, perhaps as part of a larger strategy for recovery of past investment costs. A closely related transition approach would call for a medium-term period of continued cost-of-service mechanism with performance based rates for existing generation, allowing new generation investments to enter as competitive investments and eventually removing even the existing plants from regulation.

Absent contracts for the sale of the power, incentive contracts could be structured to insulate the operators of generating plants from the control and interests of the owners of the plants.[64] Generation owners could contract out plant management and bidding, with the incentive payments for the plant geared only to successful operation of the plant in a competitive framework, not to the profits created by strategic behavior that exploited market power. It would be an easy matter for regulators to monitor the terms of such contracts, with the result that the plant operators should perform in the same way as under a divestiture but without the difficulties of actually forcing the sale of the plants.

These examples illustrate the possibility of remedies that might avoid either extreme end of the system. Before even these remedies may be needed, however, further consideration should be given to diagnostics that could reveal abuses of any market power. Again the pool-based model would simplify the regulator's use of such diagnostics to track the performance of the generators. Even if generators have market power, they may not use it, because it would be easy to detect. For instance, it may be possible to design data monitoring schemes that could effectively uncover abuses of market power along the following outline. In the absence of transmission constraints that involve loop-flow effects among the plants owned by a single firm, attention would concentrate on possible market abuses with existing power plants; entry provisions should be sufficient to assure competition with new facilities. For the existing plants,

---

[63] R. Green, "Britain's Unregulated Electricity Pool," in M. Einhorn (ed.) From Regulation to Competition: New Frontiers in Electricity Markets, Kluwer Academic Publishers, Boston, 1994, pp. 73-96.

[64] This idea was suggested by Michael Schnitzer.

there is a great deal of data that could be used to provide reasonable estimates of the capacity and operating costs of the plants. With this information, and the transparent prices of the system operator's dispatch, exercise of market power to capture monopoly profit would be revealed by three simultaneous conditions:

> **Price Above Operating Costs.** The market price must be high enough to contribute to monopoly profits. If prices are at or below the operating costs of the plant in question, there is no profit, hence no monopoly profit. The plant may not be running, or if dispatched it would have no impact on the market price.
>
> **Output Below Capacity.** The exercise of monopoly power in the pool-based market requires restricting output in order to support higher prices. If the plant is running at full capacity, high prices above operating costs may generate large profits, but these profits reflect scarcity and not use of market power. Scarcity prices should be paid and charged to provide the right incentives in the market; but scarcity prices apply only when the applicable plants are offered and running at full capacity.
>
> **Significant Affiliated Output.** The profit from monopoly bidding and pricing is captured on the commonly owned output which enjoys the benefits of the higher prices created by the restriction. Hence the person restricting output must have other output sufficient to demonstrate a higher profit captured because of the use of market power.

These conditions should be easy to monitor with the information available from the system operator's dispatch. If any of these conditions fails to hold, and there are no significant loop-flow effects, then there is a natural explanation of the market outcome that differs from use of market power. Since the exercise of market power should be the concern, not the simple fact of concentration of ownership, this outline of the elements of a possible diagnostic suggests a policy that could be followed before remedies need be applied. This would be a variant of light-handed regulation. The regulator would monitor the dispatch results for existing plants. As long as the three conditions did not exist simultaneously for a significant number of hours a year, bidding behavior would be accepted as consistent with market competition. Otherwise, the search for remedies would be on. Given the nature of the likely remedies, the result might well be that behavior would be competitive and no remedies would be required.

In the presence of significant loop flow effects, with the ability of generators to manipulate the interactions, it would be possible to profit from market power without meeting these three tests. The analysis of market power in the face of significant transmission constraints is a subject for future research.[65]

---

[65] Judith B. Cardell, Carrie Cullen Hitt and William W. Hogan, "Market Power and Strategic Interaction in Electric Networks," Harvard Electricity Policy Group, September 25, 1996 (Revised October 15, 1996).

SUMMARY

The contract network approach and its key elements -- access to essential facilities including the wires and pool dispatch, use of short-run marginal-cost pricing, and reliance on long-term contracts to provide economic hedges rather than specific performance -- are not far from actual operations or proposed reforms in other systems. The FERC proposed capacity reservation system moves very far in this direction, posing a point-to-point transmission reservation definition that does not depend on a decomposition and tracking of the actual flows. Decentralized trading of these capacity reservations would not be enough to support a competitive market, but this trading could be coordinated through the system operator. Opportunity cost pricing would define the trading criterion. Because of the strong and unavoidable network interactions, opportunity cost pricing for transmission and energy cannot be separated, but they would arise naturally as the result of bidding and economic dispatch offered by the system operator. Locational prices would define the opportunity costs for energy bid through the spot market, and the difference in locational prices would define the opportunity cost for transmission scheduled in addition to the spot market transactions. Then tradable point-to-point capacity reservations with opportunity cost pricing for unused or overused amounts would be functionally and financially equivalent to transmission congestion contracts. Conversely, transmission congestion contracts would be functionally and financially equivalent to the tradable point-to-point capacity reservations, would be easier to manage, and would fully support the competitive market while being fully consistent with the actual use of the transmission grid.

APPENDIX I
Capacity Reservation Open Access Transmission Tariffs
FERC's Fifteen Proposed Principles[66]

A capacity reservation tariff might have terms and conditions very much like those for point-to-point service in the Final Rule tariff. These would need to be modified to accommodate former network service customers. It is premature to specify detailed terms and conditions of capacity reservation service in advance of the comments and technical conference. However, we propose certain general capacity reservation tariff principles for comment.

1. Purpose of reservation service

Transmission products and services should be provided on an open access, comparable basis. In order to ensure comparability, transmission service should be nominated and reserved on a non-discriminatory basis. Transmission for wholesale sales of electric energy should be made available on an unbundled basis.

2. Basic service concept

All firm transmission service would be reserved, and all reserved service would be firm service. Reservations of transmission capacity should permit the customer to receive up to a specific amount of power into the grid at specified [Points of Receipt] and to deliver up to a specific amount of power from the grid at specified [Points Of Delivery], on a firm basis. Individual PORs and PODs need not be "paired" with each other. The customer's capacity reservation would be the higher of either (1) the sum of the reservations at all PORs or (2) the sum of the reservations at all PODs. All nominations for a capacity reservation would be evaluated using the same standard; for example, the utility could apply a feasibility criterion that states that the grid must be able to accommodate the scheduled use of all capacity reservations simultaneously.

3. Use of capacity reservations

A customer with a capacity reservation could use the reservation to deliver or receive any type of power product (such as firm or non-firm power). That is, use of the capacity reservation should not be restricted to particular power products. Any such restriction would be inconsistent with unbundling. This would allow the capacity reservation holder to combine transmission and power products in any way that satisfies its needs.

4. Applicability to all customers

---

[66] Federal Energy Regulatory Commission, "Capacity Reservation Open Access Transmission Tariffs," Notice of Proposed Rulemaking, RM96-11-000, Washington DC, April 24, 1996, extract of pp. 20-25.

Capacity reservations for all firm transmission service would be made under the [Capacity Reservation Tariff], including reservations nominated on behalf of the transmission provider's bundled wholesale and retail customers. This would make it possible to allocate capacity and costs comparably among all transmission users. This would not require the unbundling of the transmission component of bundled retail rates or affect state authority with regard to the rates, terms, and conditions of service to bundled retail customers.

5. Application of penalties for overuse

Any charges for exceeding capacity reservations should be non-discriminatory. If a CRT penalizes use in excess of reserved amounts, these penalties should be applied comparably to all reservations. Any dispositions of penalties assessed against the utility for violating bundled retail capacity reservations would be under the state commission's ratemaking authority. If penalties are not authorized by the state commission's ratemaking authority, the Commission would not authorize recovery of such penalties from other transmission customers.

6. Standard for accepting nominations

A nomination for a capacity reservation would be accepted if the transmission provider determines that it can be reliably accommodated without infringing on other firm reservations. If transmission capacity expansion is needed and approved by state siting authorities, a nomination should be accepted if the nominating customer is willing to pay its appropriate share of the cost of the expansion.

7. Non-firm transmission service

In addition to reserved firm service, transmission providers would offer non-firm transmission service. Non-firm service could be provided from transmission capacity not scheduled by customers with reservations or from capacity that is not previously reserved. Non-firm service would be allocated to the highest valued use by opportunity cost pricing as described in the Open Access Final Rule or by some other pricing consistent with the Commission's Transmission Pricing Policy Statement.

8. Open season for new facilities

We would anticipate requiring a transmission provider to publicly announce its plans for capacity expansion projects to allow market participants to reserve capacity. Participants would pay an appropriate share of the costs of the project. All market participants would be treated comparably in securing additional transmission capacity reservations when the grid capacity is expanded.

9. Cost allocation and pricing

The fixed costs of the transmission network would be allocated among reservation holders on the basis of their capacity reservations. Rates would be designed to recover these costs and would be revised from time to time to reflect changes in the level of fixed costs or changes in reserved amounts. In this way, transmission providers would have an opportunity to fully recover their fixed costs. Transmission providers would be expected to propose specific mechanisms for recovering fixed costs from transmission customers.

10. Standardized products and priority protocols

Just as the Commission has required under the Open Access Final Rule tariff, the CRT would offer standardized transmission products and services, defining reserved and non-reserved transmission service and setting reservation priorities and curtailment protocols. This would reduce uncertainty and facilitate the trading of any transmission capacity in a secondary market. Such trading can be an important tool in price discovery and risk management.

11. Service modifications

Customers with a capacity reservation would be allowed to modify their capacity reservations at no additional charge if the modification can be accommodated without infringing upon any other firm capacity reservations. Modifications should not result in the customer's capacity reservation being exceeded. Modifications could include reallocation among the customer's already specified receipt and delivery points or reallocation from existing to new receipt and delivery points.

12. Scheduling flexibility

Customers with capacity reservations would be given the option of scheduling (using) less than their full capacity reservation at each POR or POD. In addition, the transmission provider also could offer an "obligation" type of capacity reservation under which the customer would be required to use all of the capacity it has reserved.

13. Reassigning reservations

Customers would be allowed to reassign their reservations to other entities eligible to take service under the CRT at no additional cost, subject to certain limitations, such as those in the Open Access Final Rule point-to-point tariff provisions.

14. Opportunity Cost Pricing

Opportunity cost pricing would still be an option under a capacity reservation service. Under a CRT, a holder of a capacity reservation would not pay opportunity costs for use of its own capacity when the utility encounters a transmission constraint; instead, it would be eligible to receive opportunity cost payments if it did not use its full capacity reservation across the constrained interface. In contrast, a customer seeking a capacity reservation or using non-firm service might have to pay opportunity costs.

15. Planning obligation

Each market participant would be responsible for planning its own transmission needs. The transmission provider would not be responsible under Federal rules for planning the CRT nominations of others, even relatively small customers. Transmission providers, of course, would be free to enter voluntary arrangements to perform this task, or they may be required to do so under state laws. The Commission would consider approving negotiated rates and conditions between a small customer and a transmission utility that reflect different risks accepted by each party when one plans for the other.

APPENDIX II
Competitive Electricity Market Pricing

INTRODUCTION

Examples of pricing in networks illustrate the issues and the use of least-cost dispatch with accompanying transmission congestion contracts under the pool-based model.[67] Pricing in a competitive electricity market is at marginal cost. The many potential suppliers compete to meet demand, bidding energy supplies into the pool. The dispatchers choose the least-cost combination of generation or demand reductions to balance the system. 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 pricing are intimately connected. The FERC has outlined objectives for transmission pricing that would be compatible with a competitive market.[68] 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. These fundamentals provide the building blocks for an energy and transmission pricing system that addresses the several requirements of the FERC outline.

SHORT-RUN TRANSMISSION PRICING

A system operator (SO) 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 SO provides balancing services, adjusts for transmission constraints and charges for transmission usage. The SO would match buyers and sellers in the short-term market. The SO 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.

---

[67] These examples illustrate the elements of locational marginal cost pricing. They are adapted from 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; W. Hogan "Transmission Pricing and Access Policy for Electricity Competition," The Harvard-Japan Project on Energy and the Environment, February 1996.

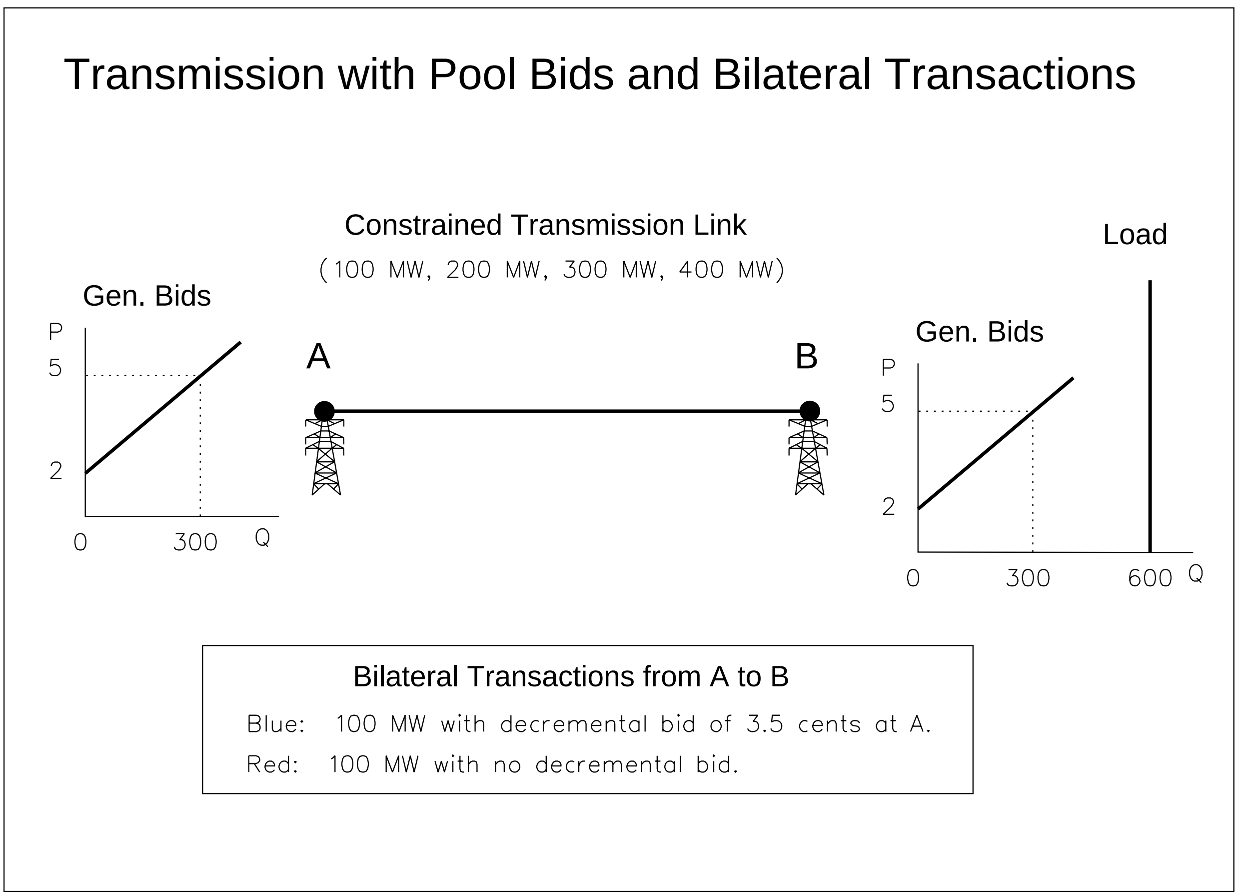
[68] Federal Energy Regulatory Commission, "Transmission Pricing Issues," Staff Discussion Paper, Inquiry Concerning the Commission's Pricing Policy for Transmission Services Provided by Public Utilities Under the Federal Power Act, Washington, DC, June 1993, pp. 7-8.

The responsibility of the SO 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.

Transmission Pricing Examples

A set of examples can illuminate the treatment of spot-market transactions and bilateral transactions, under the SO'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.

Figure 11



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 lines, as shown in the accompanying Figure 11, 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.

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.

- 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 SO 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 SO 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 SO reduced both pool and Blue transactions. There is no artificial bias induced by the SO 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 SO reduced both pool and Blue transactions. There is no artificial bias induced by the SO 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 SO reduced both pool and Blue transactions. There is no artificial bias induced by the SO fulfilling the directives of the economic dispatch.

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 SO	cents (x1000)	0	450	600	500

The net spot-market payments that are made to and from the SO are summarized in the accompanying table. Note that the cases of transmission congestion include net payments to the SO. 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.

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 the argument above, that the SO 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 SO 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 SO. 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).

Failure to offer a bid-based economic dispatch will return us to the complications and fictions of the contract-path world of old, and the many artificial arbitrage opportunities that create profit by creating confusion. This would not be good public policy.

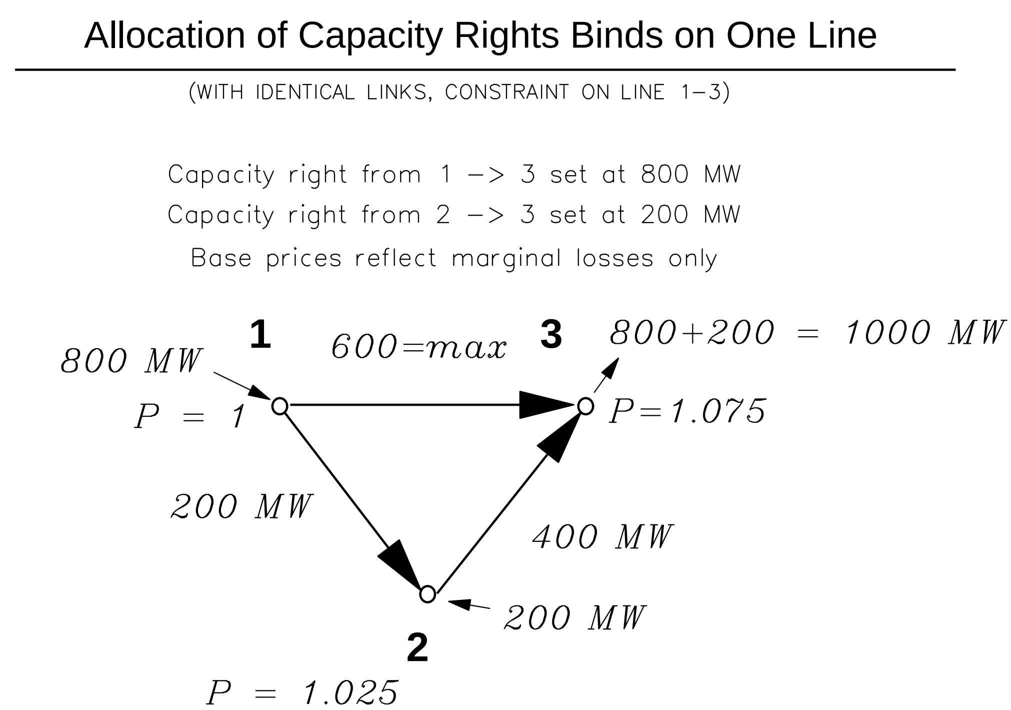
CONTRACT NETWORK EXAMPLES

For purposes of further illustration, consider the case of a three bus network with identical lines and identical thermal limits on each line. A three bus network is the minimum case needed to observe the network interaction effects of loop flow. Here we use the DC-Load approximation for real power only, and ignore contingency constraints. Reactive power and contingency constraints can be included without changing any of the fundamental points examined here.[69]

An alternative base case model and an allocation of congestion contracts are shown in Figure 12. Here we assume that the desired transmission congestion contracts are for 800 MW from bus 1 to bus 3, and 200 MW from bus 2 to bus 3. Or, an equivalent definition is that the customer at bus 3 has the contract to purchase 800 MW at bus 1 and 200 MW at bus 2. The simultaneous allocation of these contracts is feasible, but it does hit the thermal transmission

[69] W. W. Hogan, "Contract Networks for Electric Power Transmission," Journal of Regulatory Economics, Vol. 4, No. 3, September 1992, pp. 211-242.

Figure 12

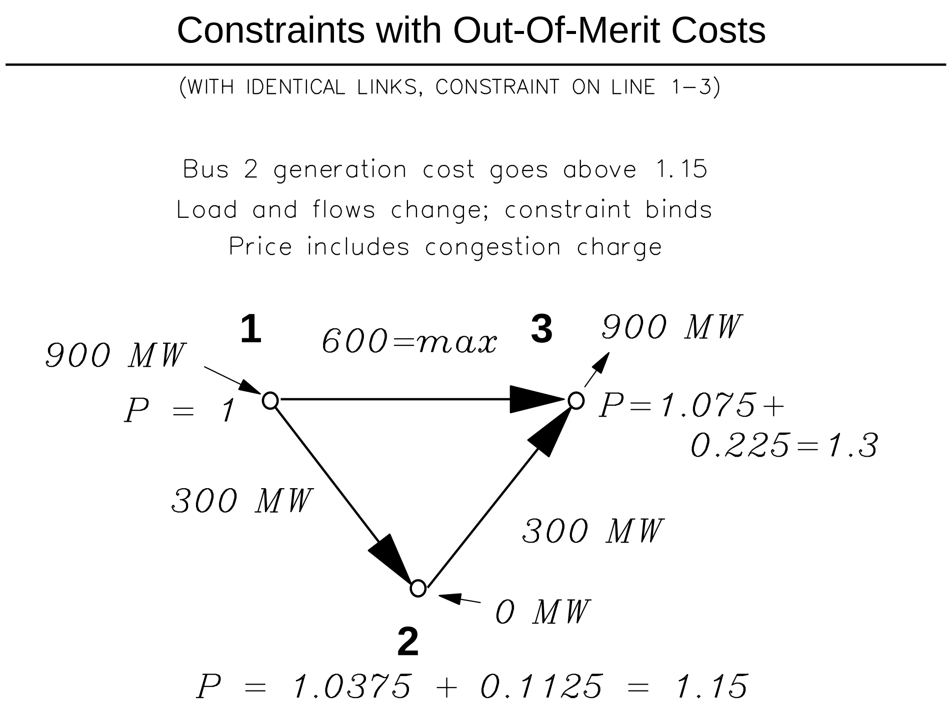


constraint of 600 MW on the line between bus 1 and bus 3.

In Figure 12 the prices calculated for this dispatch are shown relative to the price at bus 1. In this instance there is no congestion and the prices cover only the cost of generation at bus 1 and the marginal cost of losses. In this simplified case, the equilibrium required is that the marginal losses are linear in the flow on any link and are the same along any parallel path. Hence the marginal loss of one additional megawatt from bus 1 to bus 3 is 0.075, whether by the path 1→3 or via 1→2→3. There is no additional congestion cost, and hence there is no payment from the pool under the congestion rental contracts. Equivalently, the customers at bus 3 buy their 800 MW from bus 1 and 200 MW from bus 2, just as specified in the transmission congestion contracts.

Of course, a change in the economics of generation could induce transmission congestion with the associated differences in prices across locations. In Figure 12, it was economic to generate power at bus 2 and the actual economic dispatch is the same as the dispatch with simultaneous use of the allocated congestion contracts. In Figure 13, the assumed conditions change with an increase in the running cost of power at bus 2 and the need to use expensive

Figure 13



generation at bus 3. If the gross load at bus 3 is still 1000 MW, then part of the load must be met with local generation, which costs 1.3, including a congestion rental of 0.225. At this price for bus 3 and with these loads and flows, the price at bus 2 is determined by the equilibrium conditions of optimal economic dispatch. The easiest way to verify the equilibrium prices is to assume that an additional 2 MW of power could be supplied at bus 2. This would allow a reduction of 1 MW at bus 1 and an increase of 1 MW delivered to bus 3 that could displace the expensive generation at bus 3. The flow from 1→2 would reduce by 1 MW, with a corresponding 1 MW increase along 2→3. The flow along 1→3 would still be at the limit of 600 MW, and total losses would be the same. Hence the net savings would be 1 unit at bus 1 and 1.3 units at bus 3, for a total of 2.3 units. This implies a price for the incremental generation of 2 MW at bus 2 of 2.3÷2 = 1.15, which is the equilibrium price. Along with the marginal losses, this total opportunity-cost price implies a bus 2 congestion price of 0.1125.

By assumption, the cost of the generating plant at bus 2 is above 1.15, so the plant does not run. Now all the power transmitted is generated at bus 1, and only 900 MW can be transmitted. The thermal constraint of 600 MW on the line between bus 1 and bus 3 is binding. All the users of the grid pay or are paid these prices for the actual dispatch. In addition, the holders of the point-to-point transmission congestion contracts receive payments from the pool operators.

The resulting payments are shown in Table III. Hence the owner of the 800 MW contract from bus 1 is paid the congestion rental difference from bus 1 to bus 3 of 0.225 for the full 800 MW, requiring a payment from the pool of 180. Likewise, the owner of the 200 MW contract is paid the difference in the congestion rental between bus 3 and bus 2, or 0.1125, for the full 200 MW contract and a payment of 22.5. Both users actually buy power from the pool at bus 3 for the price of 1.3. For the customer with the 800 MW contract to purchase at bus 1, the payment of 180 is the total value of the congestion price differential between bus 1 and bus 3. And for the customer with the contract for 200 MW at bus 2, the payment of 22.5 is the total value of the congestion price differential between bus 1 and bus 2.

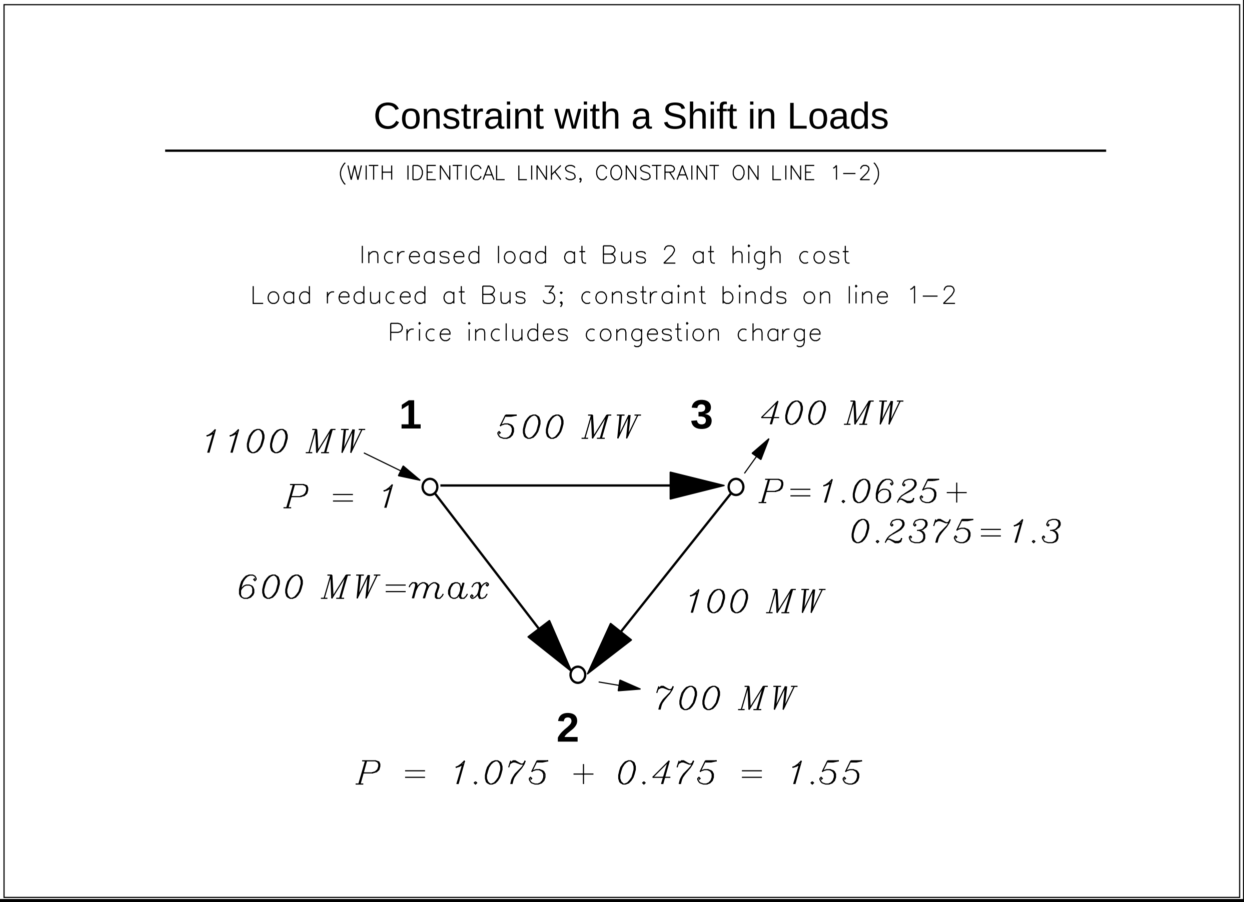
By purchasing 200 MW at bus 3 at a price of 1.3 and then applying the transmission congestion payment of 22.5, the holder of the 200 MW transmission contract can in effect purchase 200 MW at the price at bus 2 and pay only the marginal losses to move the power to bus 3. Although the actual dispatch is different than the simultaneous allocation of congestion contracts, the payments to the congestion contract holders provide the guarantee in effect that the congestion contract holders can purchase power at the price of power at another location. This holds true even if no power was actually generated at that location, as here for bus 2. Furthermore, specific performance to actually generate and transmit the 800 MW and 200 MW according to the congestion contracts would not be feasible under this economic dispatch. Only by foregoing the advantages of the economic dispatch, and increasing total costs, could the specific plants be used for specific customers. The transmission congestion contracts guarantee the economic value of the transmission, but do not determine the actual flows.

Table III: Congestion with Out-Of-Merit Costs

Congestion Contracts	Q (MW)	Congestion Price Difference	Direct Receipts	Excess Share	Excess Rentals	Net Receipts
1->3	800	0.225	180	80.00%	0	180
2->3	200	0.1125	22.5	20.00%	0	22.5
Total			202.5		0	202.5

The example in Figure 13 finds the pool dispatcher collecting congestion rentals from the actual users and paying the same rentals to the owners of transmission congestion contracts. Because the same transmission constraint limits both the actual dispatch and the initial allocation of transmission congestion contracts, there are no excess congestion rentals. All the congestion revenue collected is required to compensate the holders of the point-to-point transmission congestion contracts.

Figure 14



An alternative case is shown in Figure 14. In this case the economics of load and dispatch have changed significantly. Power still costs 1.3 at bus 3, but now the net load is reduced to 400 MW. There is a big net load at bus 2, and the equilibrium power cost there is at 1.55. The relatively cheap generation at bus 1 is used at the level of 1100 MW, which causes a shift in the flows. Now the dispatcher has no problem with a thermal limit on the line between bus 1 and bus 3, but the line between bus 1 and bus 2 has reached a similar thermal limit at 600 MW. This transmission constraint induces the indicated bus prices and congestion rentals.

To verify the equilibrium price at bus 2, given the prices at buses 1 and 3, note that 1 MW of additional load at bus 2 would be met by an increment of 2 MW in generation (or an equivalent reduction in load) at bus 3 and a reduction of 1 MW of generation at bus 1. The net change in generation cost would be 2*1.3 -1 = 1.6 units. The flow on line 1→2 would be unchanged, but the flow on line 1→3 would be reduced by 1 MW and the reverse flow along 2→3 would increase by 1 MW. Since these lines have different marginal losses, there would be a net reduction in losses of 0.05. Hence, the total increase in cost for an additional MW at bus 2 would be 1.6 - 0.05 = 1.55, the price at bus 2.

Again the pool pays or is paid the short-run prices for power at each of the locations. And again the pool makes payments to the holders of the point-to-point transmission congestion contracts. The summary of the various payments appears in Table IV. For the customer with the contract for 800 MW between bus 1 and bus 3, the congestion differential is 0.2375 and the total payment from the pool is 190. This allows the contract holder to purchase 800 MW at bus 3--with some of that power necessarily generated by plants located at bus 3--pay the price of 1.3, and after netting out the payment of 190 from the pool, effectively purchase the 800 MW at bus 1 and pay only marginal losses to move the power to bus 3.

Similarly, the customer with the contract for 200 MW from bus 2 to bus 3 can purchase 200 MW at bus 3 at the price of 1.3. However, this price is lower than the price at bus 2, and the difference in congestion rentals is now negative, at -0.2375. Under the terms of the point-to point contract, this customer must make an added payment of 47.5 to the pool. When coupled with the purchase of 200 MW at bus 3, this is equivalent to purchasing 200 MW at bus 2 at 1.55 and then moving to bus 3 paying only the marginal losses (in this case the marginal losses also would be negative between bus 2 and bus 3). The final effect is as promised under the transmission contract of the customer at bus 3 to purchase 200 MW at bus 2.

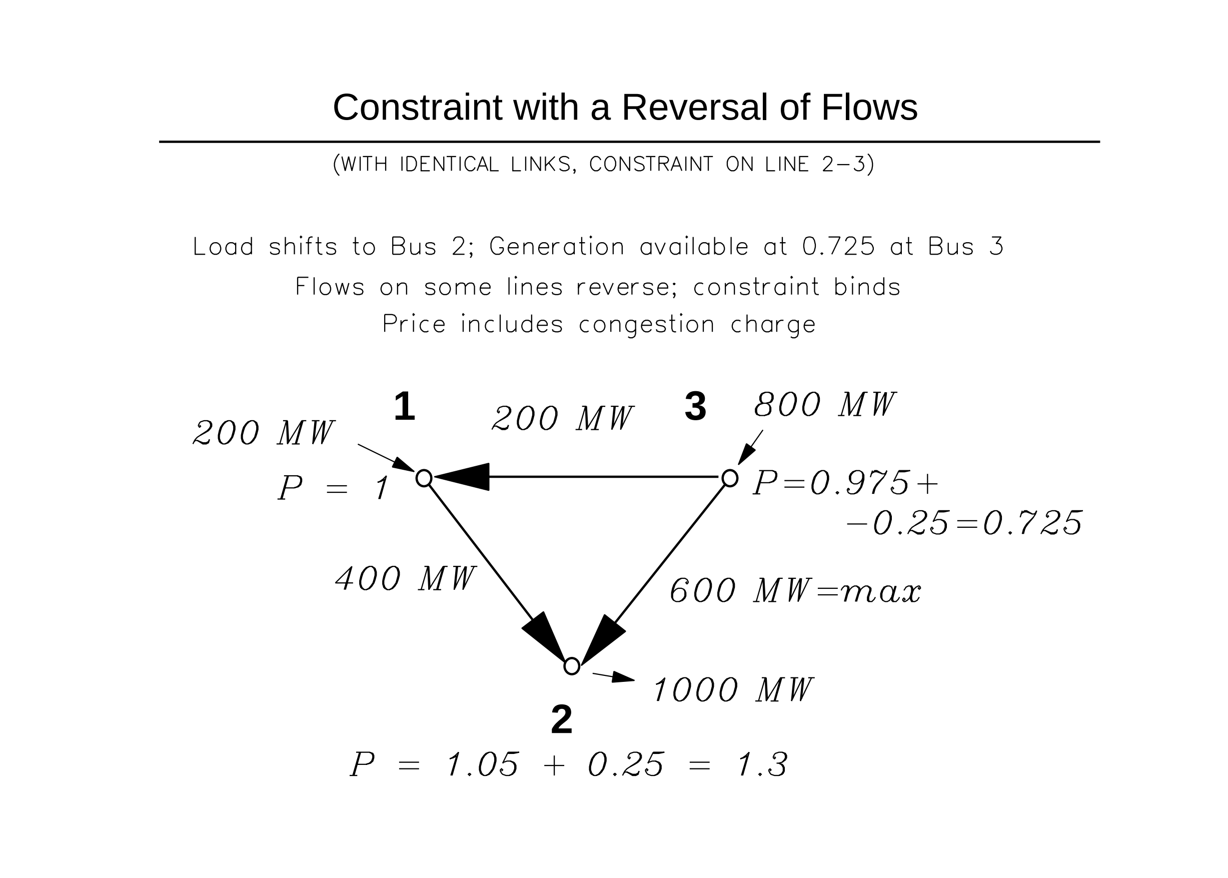
Table IV: Constraint with a Shift in Loads

Congestion Contracts	Q (MW)	Congestion Price Difference	Direct Receipts	Excess Share	Excess Rentals	Net Receipts
1->3	800	0.2375	190	80.00%	228	418
2->3	200	-0.2375	-47.5	20.00%	57	9.5
Total			142.5		285	427.5

In this case of a shift in loads, with a new transmission constraint binding, the pool can make all the necessary payments to the holders of the point-to-point transmission congestion contracts, but the payments out amount to only 190 - 47.5 = 142.5. However, the total for the congestion rentals paid by the users of the grid is 700*0.475 + 400*0.2375 = 427.5. There remain excess congestion rentals of 285. Assuming that the fixed charge payments are proportional to the transmission congestion contracts, one way to dispose of these excess congestion rentals would be to pay them out to the transmission congestion contract holders in the ratio of 80 to 20. Hence the transmission contract holder from bus 1 would receive an additional payment of 228, for total receipts of 418. Likewise, the transmission contract holder from bus 2 would receive 57 from the excess congestion rentals for total receipts of 9.5.

In Figure 14 there is a shift in load and economics, and one of the transmission contract holders, for power from bus 2, is required to make addition congestion payments to the pool. With enough of a change in the loads and transmission flows, it is possible that everyone with a transmission contract holds them in the reverse direction, and in this case the payments

Figure 15



under the sharing of excess congestion rentals take on added importance. For example, consider the conditions depicted in Figure 15. Here the economics of the dispatch and load have changed even more dramatically compared to the initial allocation of transmission congestion contracts. Now there is low price at bus 3 and a net input of 800 MW, and the higher price is at bus 2 with a net load of 1000 MW. The flows on the links from bus 3 are now reversed.

The prices at the buses include a positive congestion component at bus 2 and a negative congestion impact at bus 3, all relative to bus 1. As before, to verify the equilibrium price at bus 3, given the prices at buses 1 and 2, note that 1 MW of additional load (or an equivalent reduction in generation) at bus 3 would be met by an increment of 2 MW in generation at bus 1 and a reduction of 1 MW of generation at bus 2. The net change in generation cost would be 2*1 - 1.3 = 0.7 units. The flow on line 2→3 would be unchanged, but the reverse flow on line 1→3 would be reduced by 1 MW and the flow along 1→2 would increase by 1 MW. Since these lines have different marginal losses, there would be a net increase in losses of 0.025. Hence, the total increase in cost for an additional MW at bus 2 would be 0.7 + 0.025 = 0.725, the price at bus 3.

Once again, the users of the grid pay or are paid according to these short-run marginal cost prices. The pool collects the payments and, in turn, makes the necessary payments to the holders of the transmission congestion contracts. In this case, both the customers with congestion contracts to bus 1 and those with contracts to bus 2 face negative congestion rent differentials. Hence the customer with congestion contracts of 800 MW from bus 1 sees a differential of -0.25, and makes a total additional payment to the pool of 200. With the purchase of 800 MW at bus 3 at the price of 0.725, these combined payments are equivalent to a purchase of 800 MW at bus 1 and then moving to bus 3 at the cost of marginal losses.

For the customer with a 200 MW contract to bus 2, the congestion price difference is -0.5, and the direct payment to the pool is 100. These payments from the contract holders to the pool add to the total congestion rentals collected by the pool from the actual users of the grid, who pay -800*(-0.25) + 1000*0.25 = 450. In all, as summarized in Table V, there are 750 units of excess congestion rentals. As before, these congestion rentals could be distributed according to the share in the fixed cost allocation. In the present example, this would provide a payment of 600 to the customer with transmission congestion contracts of 800 MW from bus 1, for net receipts of 400; and 150 for the customer with contracts of 200 MW at bus 2, for net receipts of 50.

Table V: Constraint with a Reversal in Loads

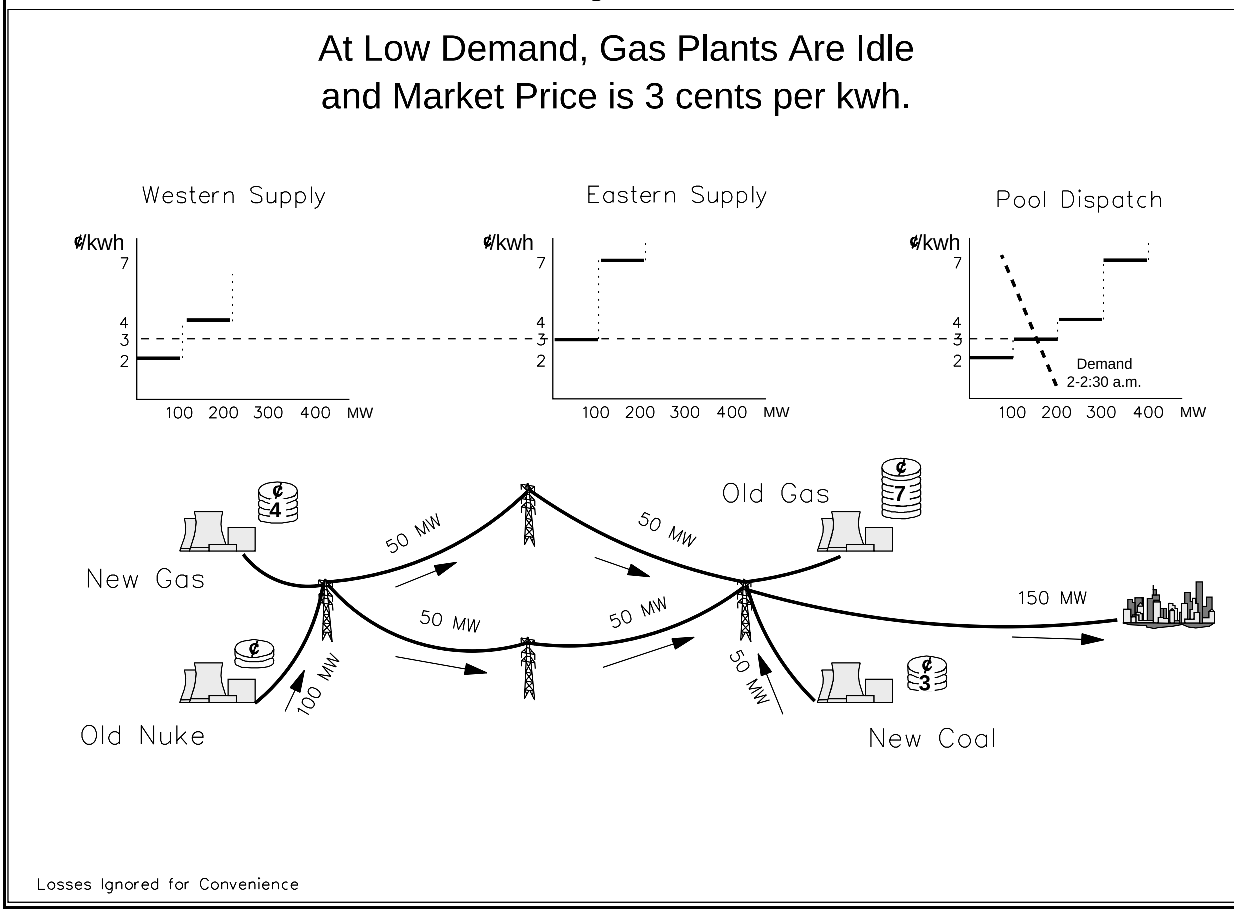
Congestion Contracts	Q (MW)	Congestion Price Difference	Direct Receipts	Excess Share	Excess Rentals	Net Receipts
1->3	800	-0.25	-200	80.00%	600	400
2->3	200	-0.5	-100	20.00%	150	50
Total			-300		750	450

In all cases, the net effect of economic dispatch, marginal cost pricing, and assignment of transmission congestion contracts is to collect congestion costs from the actual users of the grid and pay the congestion costs to those who bear the burden of the fixed charges. The transmission congestion contracts based on price differences compromise two forms. First, point-to-point transmission congestion contracts can be offered which provide the economic equivalent of a customer at one location always having the effective contract to buy delivered power at the cost at a distant location plus the marginal losses. Second, to the extent that there are excess congestion rentals, these rentals can be distributed according to some agreed formula. In aggregate, the congestion rentals paid are always adequate to honor the point-to-point transmission congestion contracts, and sometimes there can be additional rents that could be dispersed according to a sharing formula. In some instances, the congestion payments under the point-to-point contracts can be negative, but only when the economics of the dispatch have switched to provide the contract holder, who has access to cheap generation, the money from an operating margin through the pool dispatch that can fund the congestion payments. The transmission congestion contract is analogous to a futures contract which provides a perfect hedge for the cash market in transmission. The pool dispatcher and operator of the settlements system is taking no financial risk in providing these price guarantees, and the actual dispatch is not constrained by the transmission congestion contracts. The dispatcher always has the freedom to provide the most economical generation possible given the current costs, bids, and system constraints.

Economic Dispatch on a Grid

The three bus and three line example is the minimum network needed to illustrate the effects of loop flow and the impact of locational prices in a network. The results for the three link loop can be quite different than those found for a single transmission line or a radial connection. Analogies built on the case of a single line can be misleading. However, the analysis of the three bus case extends to more complicated networks, with one additional and important amendment. In the three bus case, it may be easy to fall into the trap of assuming that transmission congestion contracts are connected with individual lines between buses, since there is no difference in the geography of point-to-point definitions when every bus is connected to every other by a direct line. In a more complicated network, the transmission congestion contracts can be defined quite separately from the map of the individual lines.

Figure 16



Consider the simple market model in Figure 16, which will serve as the starting point for a set of a succeeding examples for a grid that move 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.

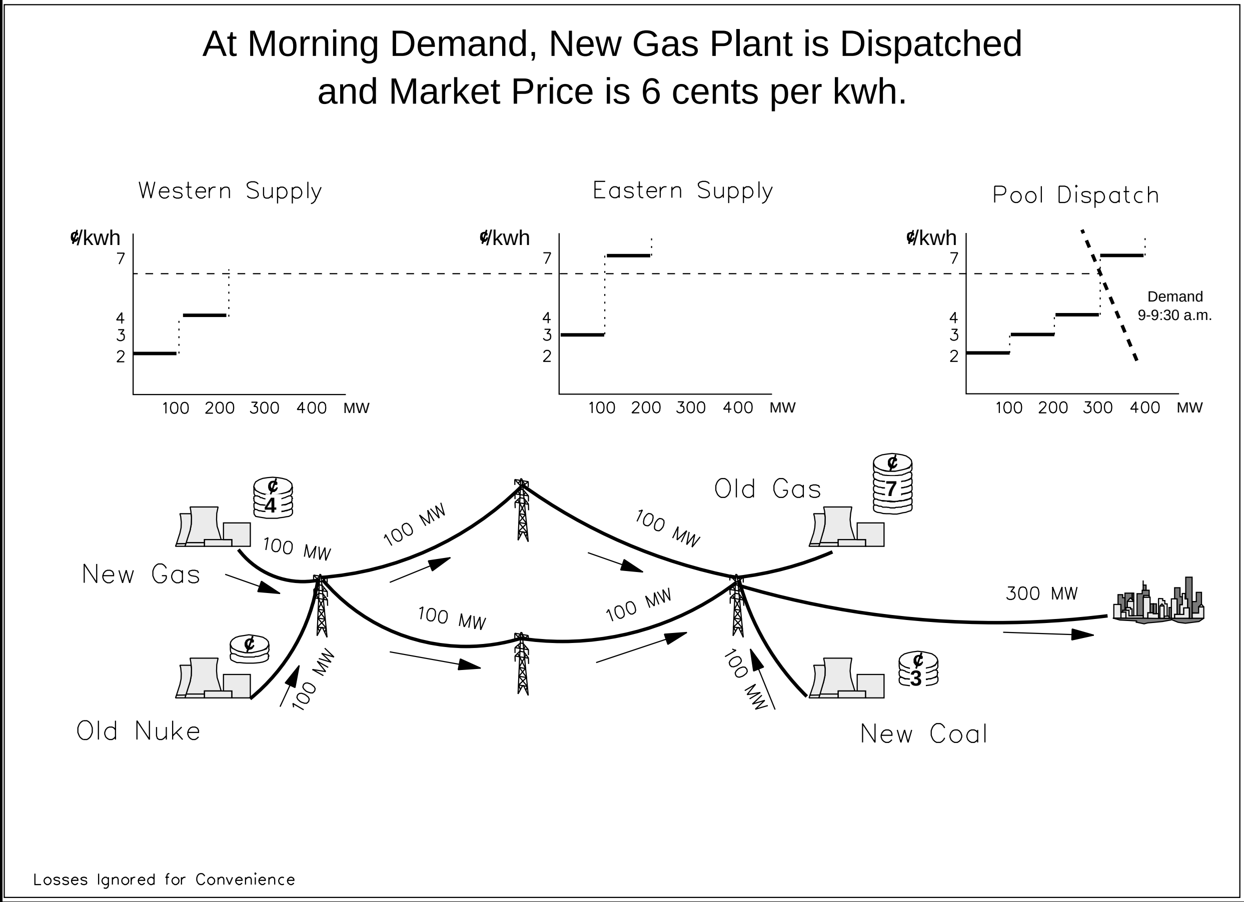
Under low demand conditions, as shown in Figure 16 for the early hours of the morning, the supply curves from the two regions define an aggregate market supply curve that the pool-based dispatchers can balance with the customer demands. The aggregate market supply curve stacks up the various generating 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 16, 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 17. 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

Figure 17



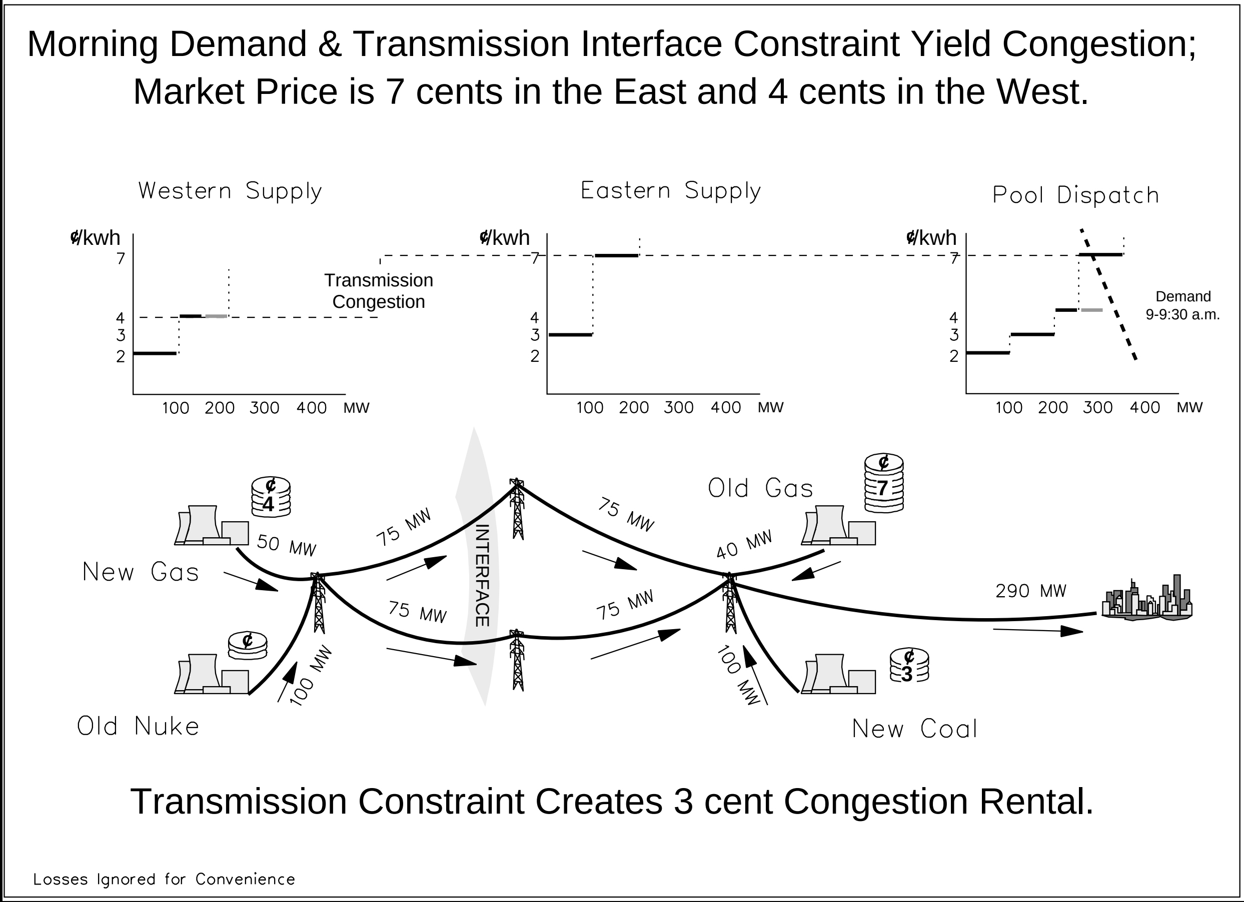
they are willing to pay for electric energy at that hour. This short-run demand bidding defines the demand curve which allows the system operator to raise the price and reduce consumption until supply and demand are in balance. In Figure 17 this new balance occurs at the point where the market price 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 17 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 18

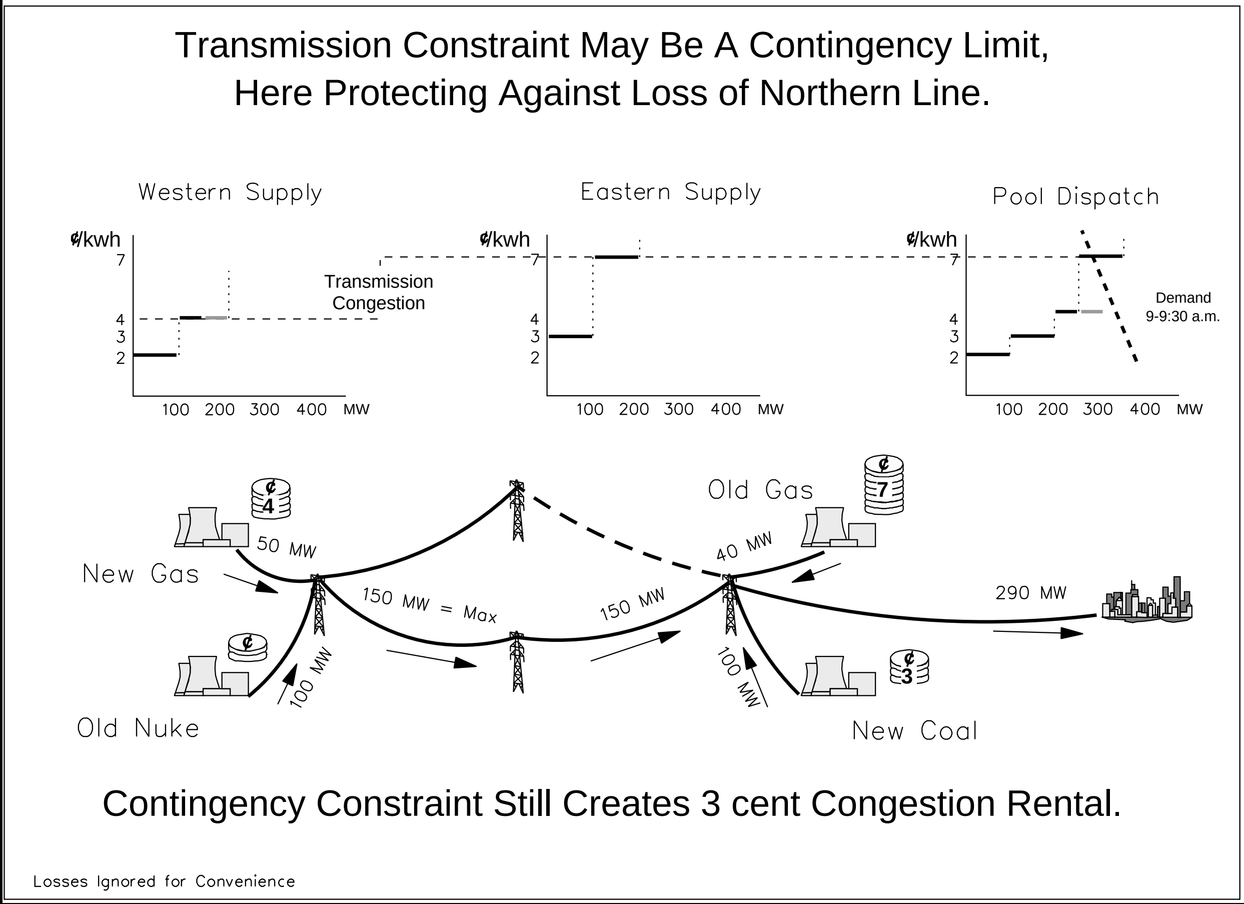


As shown in Figure 18, this transmission constraint has a significant impact on both the dispatch and market prices based on short-run marginal costs. In Figure 18 the level of demand from the city in the east is assumed to be the same as in the case of Figure 17. However, now the pool-based dispatcher 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 19



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 18, with 75 MW on each line, but the pool-based dispatchers must dispatch in anticipation of the conditions in Figure 19. 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 19 even though the flows are as in Figure 18. 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.

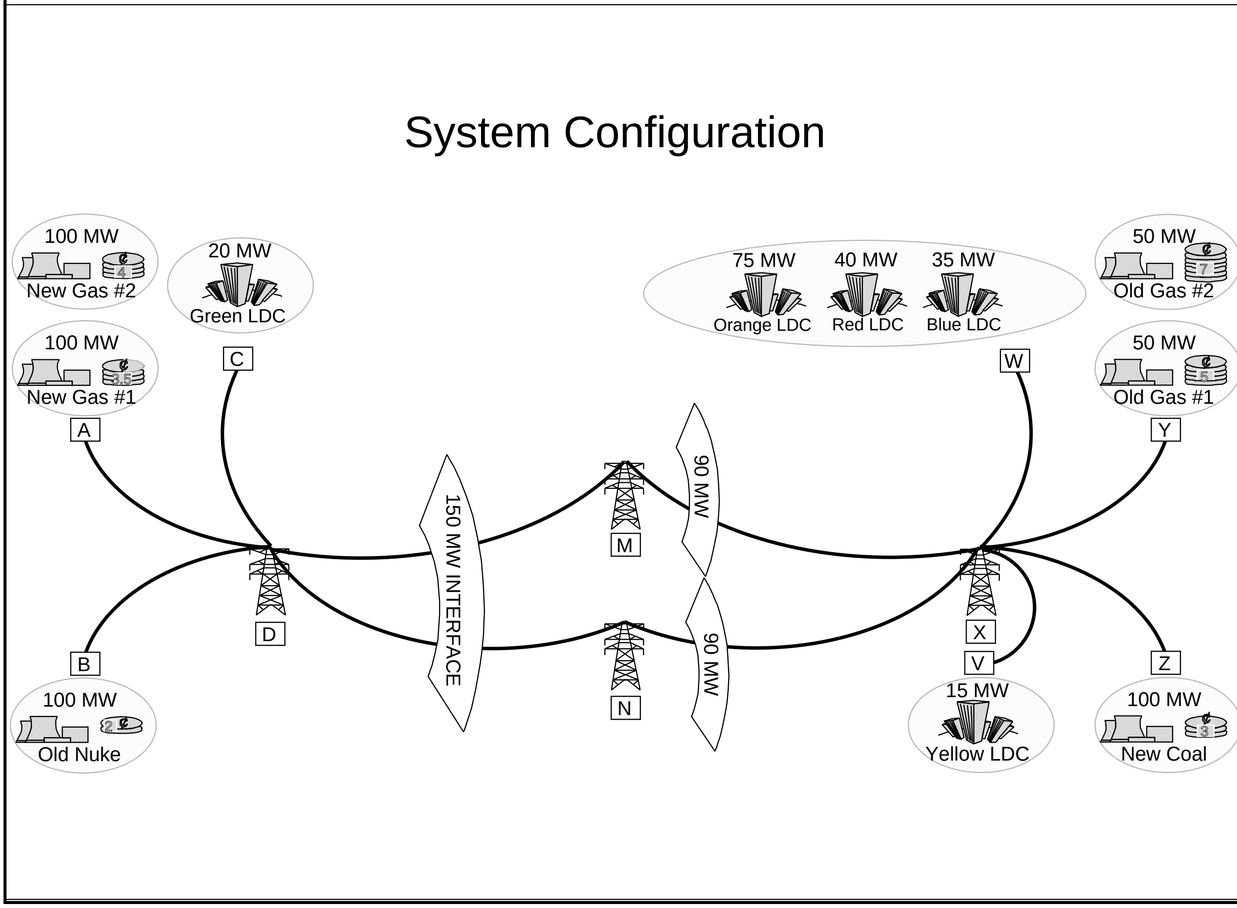
Buying and selling power at the competitive market prices, or charging for transmission at the equivalent price differential provides incentives for using the grid efficiently. If some user wanted to move power from east to west, the transmission price would be negative, and such "transmission" would 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 16 through Figure 19 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. 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 20. These generators 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

Figure 20



Blue LDCs at W.

- Load in the West, consisting of a Green LDC at C.

- Interface constraint of 150 MW between bus D and buses M and N.

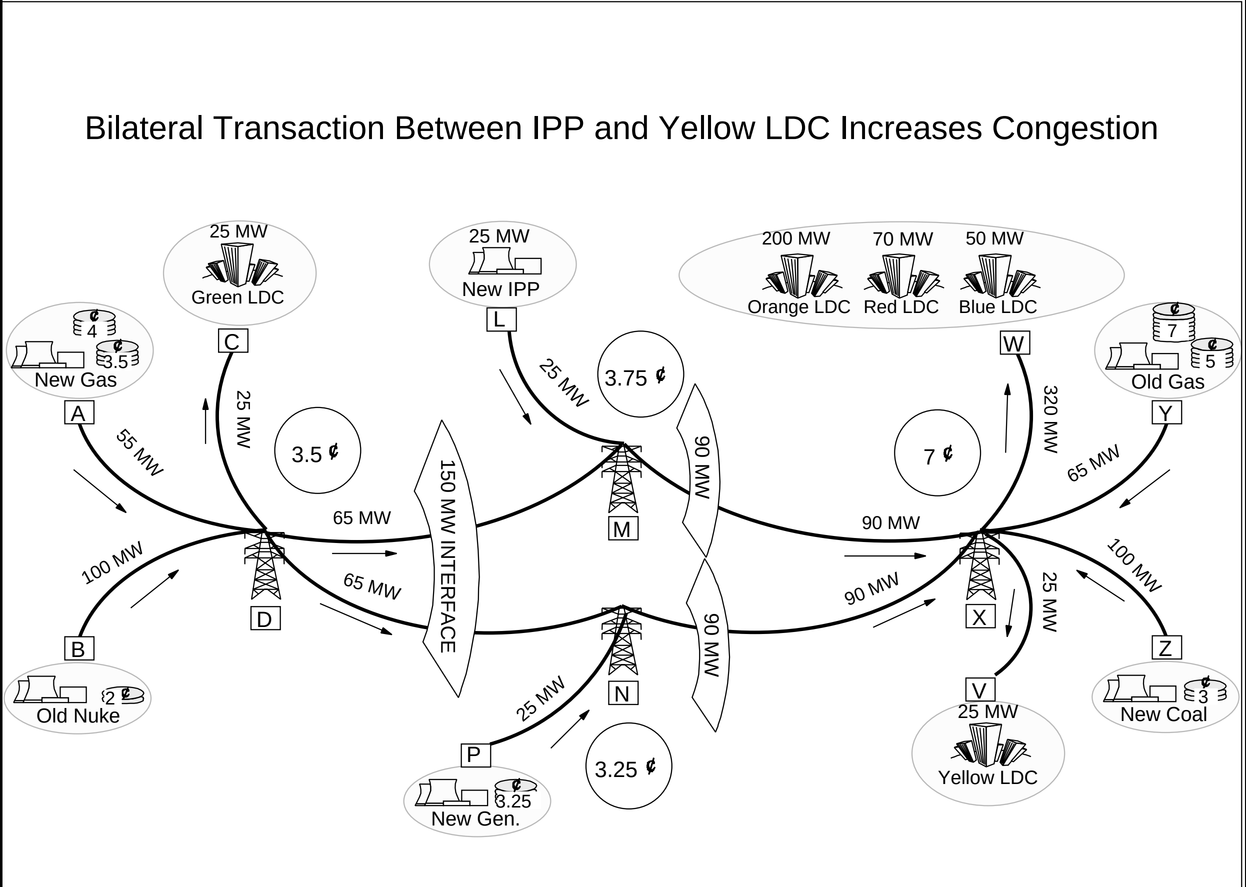
- Thermal constraints of 90 MW between M and X and between N and X.

- The New Gas and Old Gas generating facilities each consist of two generating units whose marginal costs of production differ.

Loads in Figure 20 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 21 at bus "P." An IPP at bus "L" has bid in a must run plant at 25 MW, having arranged

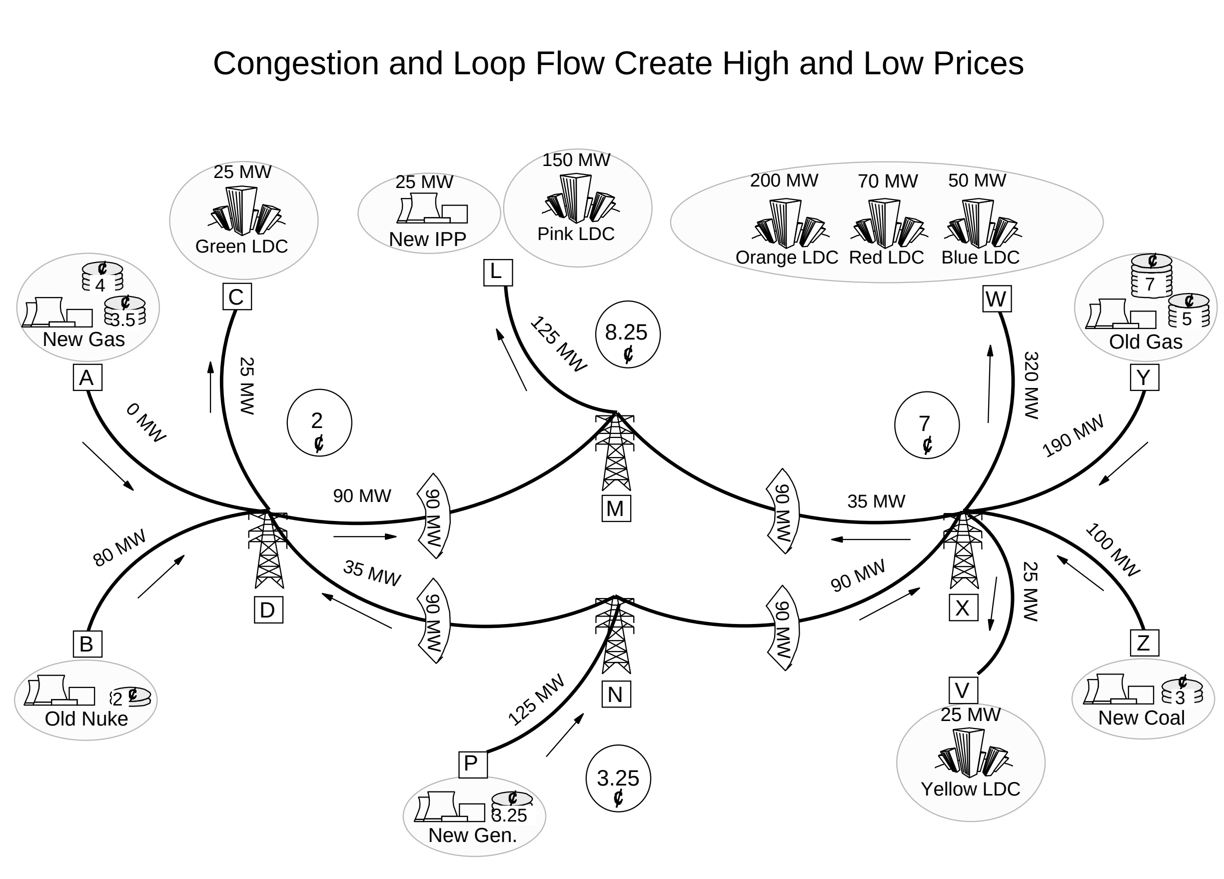
Figure 21



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". 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 22, 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".

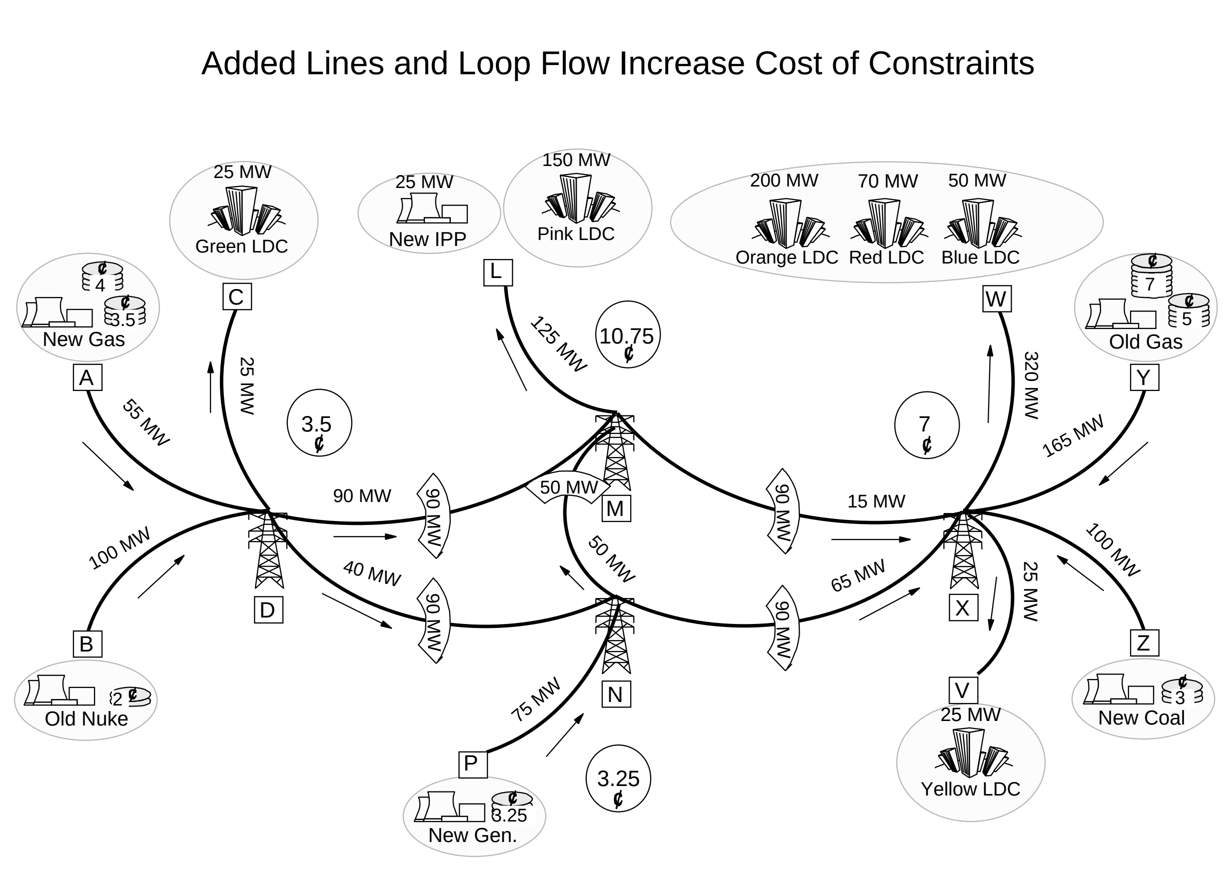
Figure 22



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 added to the network in Figure 23, 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 22 to $19,912.50 in Figure 23. 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 24, 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

Figure 23



elsewhere and providing benefits to the overall system.

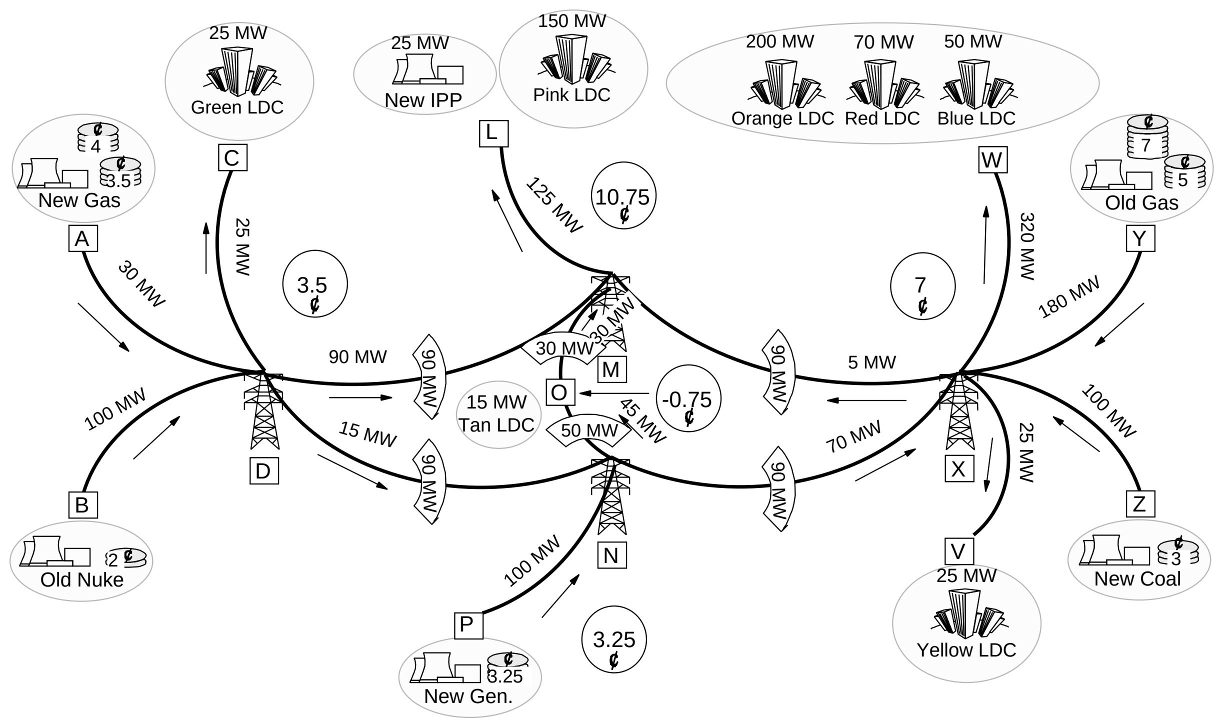
This final example, therefore, illustrates and summarizes the types of interactions that can develop in a network with loop flow. Power can flow from high price nodes to low price nodes. The competitive market clearing price, equivalent to the marginal costs for the least-cost dispatch, can include simultaneously at different locations prices higher than the cost of the most expensive generation and lower than the cost of the cheapest generation source.

Zonal Versus Nodal Pricing Approaches

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. 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

Figure 24

A Tight Constraint from Bus O to Bus M Yields a Negative Price



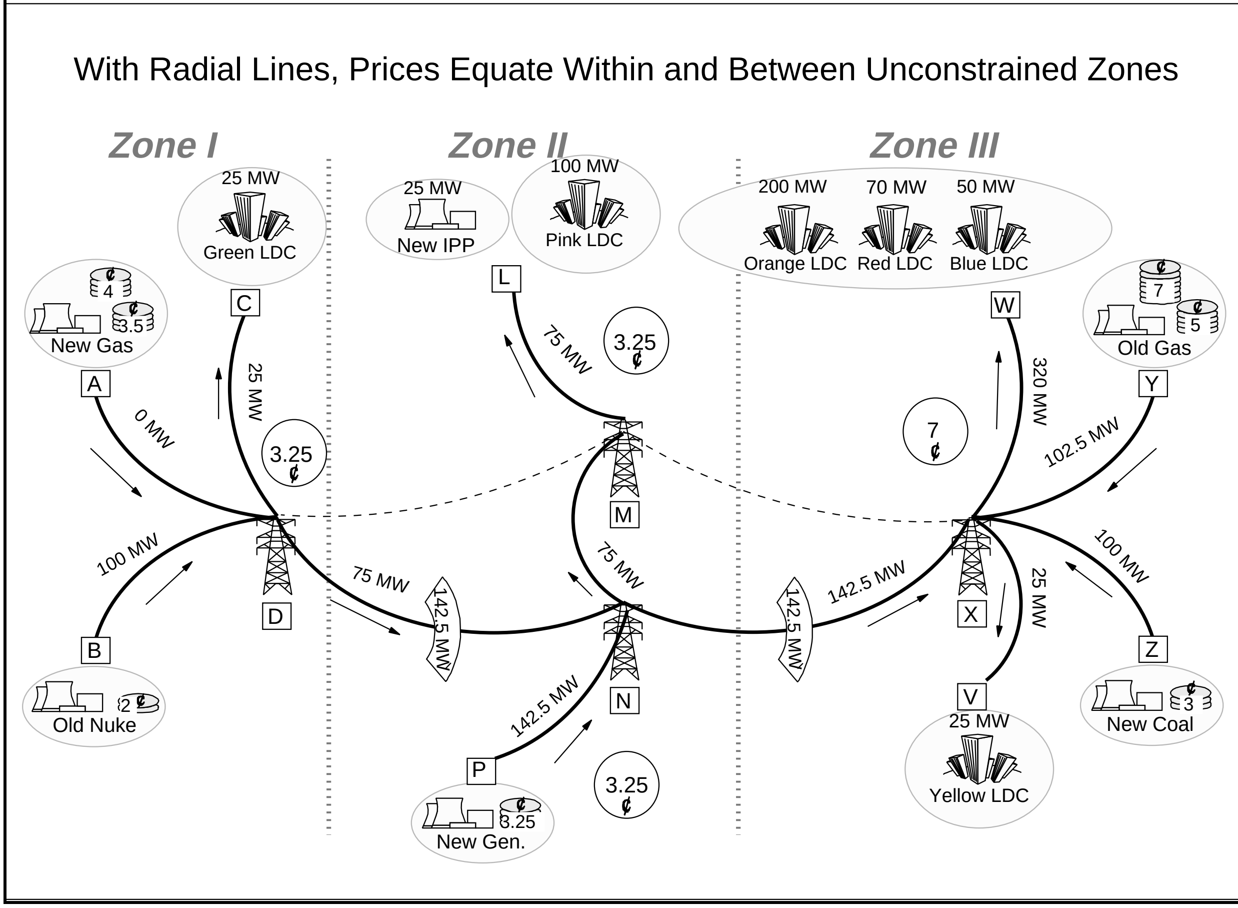
complexity. However, in the presence of real constraints in the actual network, the zonal approach may not be as simple as it might appear without closer examination.[70]

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 25. The network has been constructed so that there are only radial connections. With strictly radial connections, locations within and

[70] 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, October 8, 1996.

Figure 25

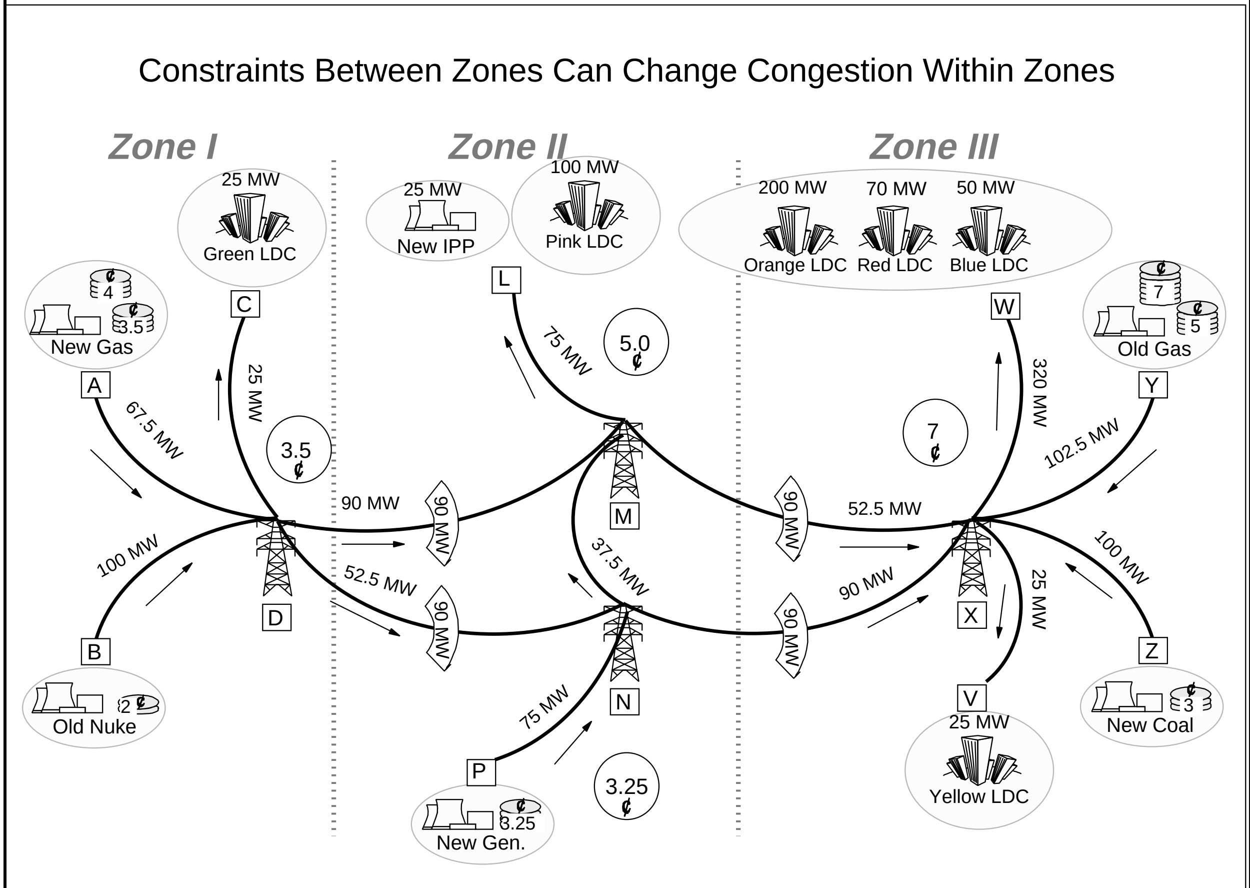


between unconstrained zones would have a common price. Hence, aggregation of locations offers an apparent simplification by reducing to a few distinct zones. This motivation from a typical radial examples leads to the assumption that in general there could be areas in a real network that would have the 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.

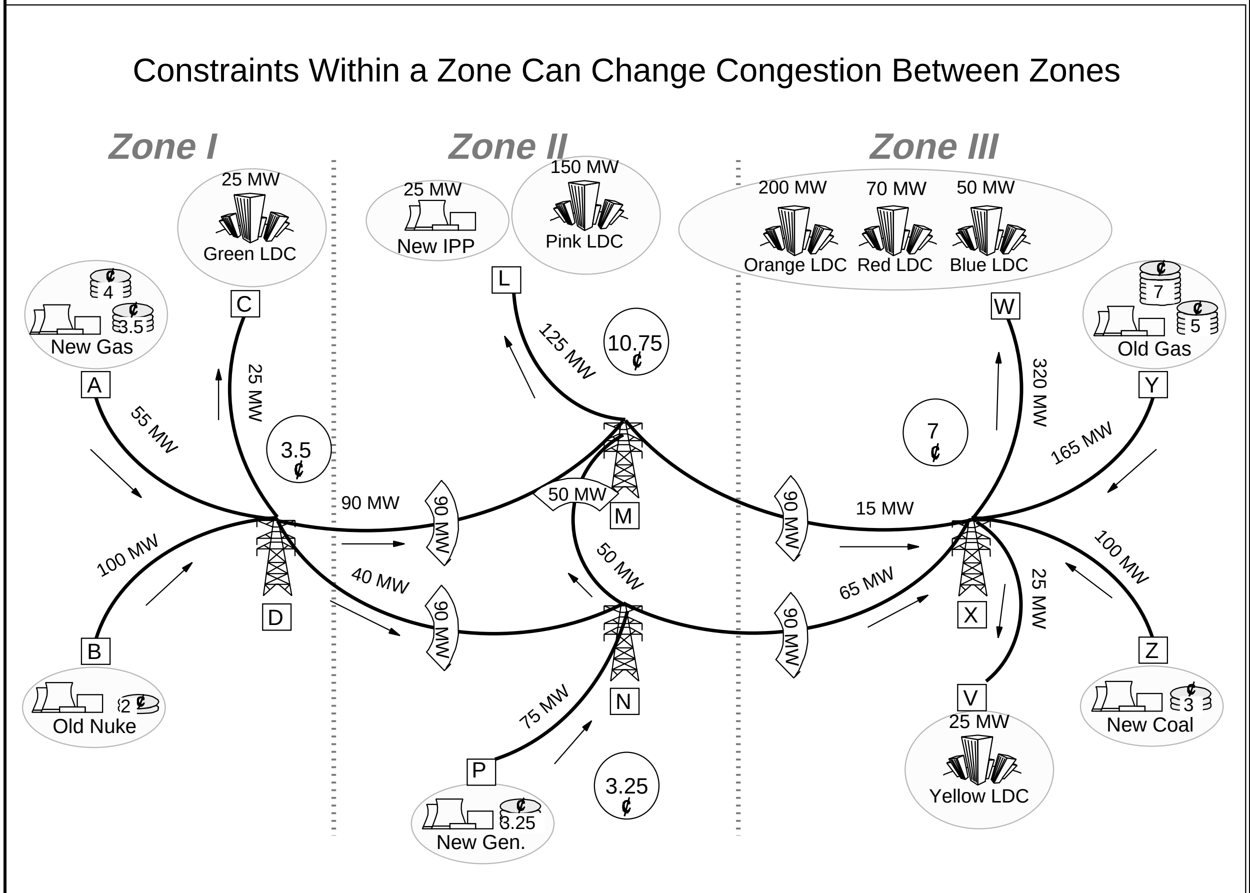
Second, the definition of a zone, which appears easy in the case of a radial network, becomes more problematic in the case of a more realistic network with loop flows. The radial examples can be a poor guide to thinking about interactions in networks. For example, it is often argued, or assumed, that 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 26



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 26. 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 25. 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 27



A symmetric result appears in Figure 27 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.

Transmission Congestion Contracts

The congestion rental received by the system operator provides the key to defining property rights in the transmission grid. In the face of transmission constraints, prices will be more volatile and it will not be possible for a generator to provide guaranteed price stability in the form of a long-term contract with a customer. Furthermore, customers and generators will have an incentive to expand the grid only if they recognize and believe that after making an investment in the grid there will be some protection against any future congestions costs. A simple way to define the property right and provide this guarantee is to assign the congestion rental not to the system operator but to the holder of the transmission congestion contract. In the transmission constrained cases of Figure 18 or Figure 19, the congestion contracts for a total 150 MW of power might have been held by customers in the city, or by generators in the west. In either case, this right is defined only as the right to collect the congestion rental. The generators and customers would not control the use of the grid. The system operator would determine the efficient pattern of use through economic dispatch. The system operator would collect the congestion payments from the actual users of the grid and pay them in turn to the holders of the transmission congestion contracts.

With this definition of transmission congestion contracts, it is an easy matter for generators at Old Nuke and New Gas to arrange long-term contracts that provide price stability for customers in the city. For example, the owners of Old Nuke may have acquired power contracts for 100 MW, and signed long-term contracts that guaranteed to provide power delivered to the city at a price of 5 ¢/kWh. In the case of low demand as in Figure 16, the short-run price is only 3 ¢/kWh, which customers pay and generators receive through the system operator. Separate from the system operator, the customers pay Old Nuke the difference of 2 ¢/kWh owed under the contracts. If demand shifts to the higher case in Figure 17, the market price is 6 ¢/kWh, and again the customers pay and generators receive this short-run price through the system operator. In this event, the generators separately pay the customers the difference of 1 ¢/kWh required under the long-term contract.

When transmission constraints bind as in Figure 18 or Figure 19, the price paid by the customers to the system operator is 7 ¢/kWh, and the price received by the generators from the system operator is 4 ¢/kWh. If the generators own the transmission congestion contracts, then the system operator pays the generators an additional 3 ¢/kWh which allows the generators in turn to pay the customers the 2 ¢/kWh difference agreed to by contract. The owners of Old Nuke are always making an operating margin of 3 ¢/kWh, and the customers are always receiving the equivalent of 5 ¢/kWh electricity. Likewise, if the customers own the transmission congestion contracts, the customers receive the 3 ¢/kWh from the system operator and in turn pay 1 ¢/kWh to the generators. Again the owners of Old Nuke are always making an operating margin of 3 ¢/kWh, and the customers are always receiving the equivalent of 5 ¢/kWh electricity. Furthermore, in either case the system operator ends up with no transmission congestion rentals; the system operator serves only to pass through the congestion costs from the actual users of the grid to the holders of the property rights in the economic interest of the transmission grid.

This system of transmission congestion contracts and payments back and forth may seem unnecessary and cumbersome in the case of the simple system of Figure 16 through Figure 19. After all, couldn't the pool-based dispatchers in effect assign the generation from Old Nuke to the long-term customers in the city? In principle, this specific performance model--assigning particular generation to particular users--is possible in this simple case, but it does not generalize into the more complicated reality of an interconnected grid with many different sites of load and generation, real and reactive power, thermal and voltage limits, and multiple contingencies. The examples in Figure 21 through Figure 24 illustrate the difficulties attendant to the network interactions. It is impossible in a real system to meaningfully assign any particular sources and destination of electricity, and attempts to do so can only serve to compromise the efficiency objective of maintaining an optimal dispatch which may require only partial use of plants in constrained regions, violating the assumptions of specific performance. However, the payments of congestion rentals from the system operator to the holders of point-to-point transmission congestion contracts do generalize to the more complicated case, and allow optimal dispatch for efficiency while accommodating long-run contracts for price differences and congestion rentals, contracts that provide both stability and the essential protection of investment in the network.

These point-to-point price protection transmission contracts defined in alternative equivalent ways, with various advantages for implementation and interpretation. For example:

- **Difference in Congestion Costs**. Receive the difference in congestion costs between two buses for a fixed quantity of power.

- **Purchase at a Distant Location**. Purchase a fixed quantity of power at one location but pay the price applicable at a distant location.

- **Dispatch with No Congestion Payment**. Inject and remove a fixed quantity of power without any congestion payment.

The total quantity of these transmission congestion contracts can be defined for a given configuration of the network, and the congestion contracts guaranteed for any pattern of loads in the network. In a real system, the transmission congestion contracts would respect all the constraints in the grid, including contingency constraints for thermal limits on lines and voltage limits at buses.

The point-to-point congestion contracts can be provided by the pool and allow protection of new investment in generation and transmission. The pool takes no risk in offering these transmission congestion contracts because the revenue collected from users paying at short-run marginal cost is always great enough to honor the obligations under the congestion contracts. Under certain circumstances, the revenues collected by the grid will be greater than the obligations on the point-to-point congestion contracts, and there will be "excess congestion rentals." For example, if there are no point-to-point contracts assigned, then none of the congestion rental would be paid out under such contracts. In the more interesting cases, if point-to-point congestion contracts are assigned up to the limit of some transmission constraint, it is possible that some other network constraint will be binding in the actual dispatch. In these cases, there will be more than enough revenues to honor the rental or purchase rights, and the pool will face the added task of allocating the excess congestion rentals.

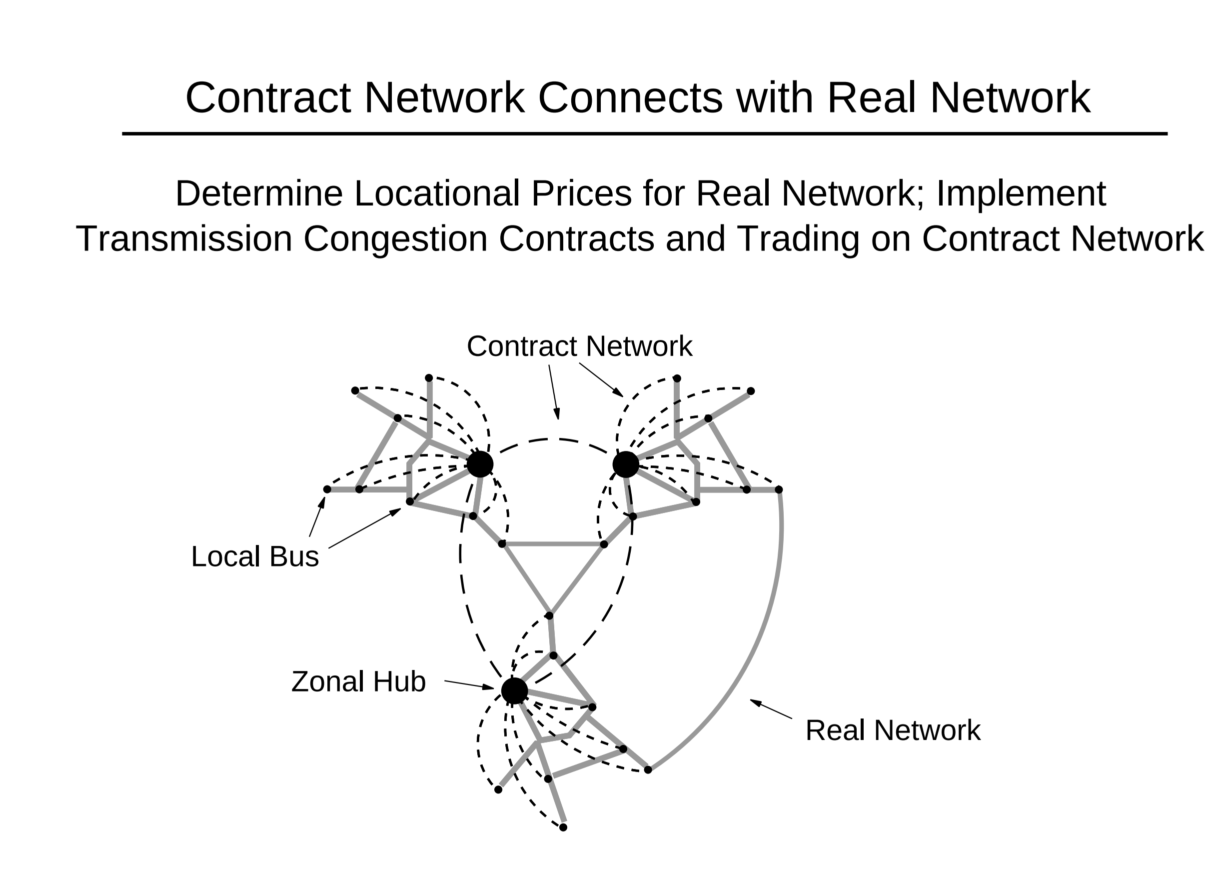
The natural assignment of any excess rentals is to those who are paying the fixed charges. There is no single rule for allocating the rights to the excess congestion rentals. The rule examined here is to share the revenues according to the share of the fixed charges. Hence, the owners of the grid who commit to pay the fixed charges have access to two types of well defined and tradeable transmission congestion contracts: the point-to-point rental contract and a share in any excess congestion rentals.

As shown in Figure 28, the contract network must anchor to the same locations, but the point-to-point contracts can follow a very different geography. Market hubs can arise and be included, with the contract connections in the network following a configuration convenient for contracting and trading. The separation of the physical and the financial flows allows this flexibility with the congestion revenues always sufficient to cover the obligations under transmission congestion contracts, no matter what the resulting pattern that appears in the least-cost dispatch.

To illustrate this conclusion for a more general network, consider again the example network in Figure 24. For this example, consider the extreme case where the market has elected to use bus "O" as the market hub, with transmission congestion contracts all defined relative to this hub. Generators may have the contracts to get to the market at "O". Customers may have similar contracts to get from the market at "O" to their own locations. The individual transmission congestion contracts may embody flows which would never be individually feasible, especially given the limits on the lines connecting "O" and the unusual conditions in this extreme case. As long as the collective flows under the contracts would be feasible, however, the congestion costs collected will always be sufficient to meet obligations under transmission congestion contracts, even in this extreme case.

Here Table VI offers with two feasible sets of transmission congestion contracts (TCCs) for a hub at "O". If the only inputs and outputs in the system consisted of those with TCC 1 at 180 MWs in at "D", 30 MWs in at "M", 30 MWs in at "N", 60 MWs out at "O" and 180 MWs out at "X", the flows would be feasible even though the individual contracts appear to require flows that are not feasible. Similarly for TCC 2. With these feasible transmission

Figure 28



congestion contracts, alternative dispatch cases illustrate the impact of the changing congestion payments.

Table VI: Transmission Contract Allocations		
From-To	TCC 1 (MW)	TCC 2 (MW)
"D-O"	180	160
"O-X"	180	160
"M-O"	30	10
"N-O"	30	70

The results summarized in Table VII capture the outcomes for economic dispatch with the only change being three different assumptions about the load at bus "L" in Figure 24, ranging from 0 MW to 150 MWs, with all other conditions the same. The table shows the corresponding prices at key buses, ignoring losses, and the associated collection of congestion rents. These rents are compared in turn with the obligations under the point-to-point contracts of the two sets of transmission congestion contracts. In the case of no load at "L", the congestion payments amount to $6300. Under the TCC 1 the obligation is also $6300; under TCC 2 the point-to-point obligation is $5750, leaving an excess of congestion payments to be disbursed through a sharing formula.

Table VII: Transmission Congestion Payments

Load at "L"	Bus Prices ¢/kWh					Total Rents $	TCC 1	TCC 2
MW	"D"	"M"	"N"	"O"	"X"			
0	3.50	3.75	3.25	3.50	7.00	6300	6300	5750
50	3.50	5.58	3.25	4.15	7.00	6300	6138	6084
150	3.50	10.75	3.25	-0.75	7.00	10950	1650	1650

As the load at bus "L" increases, dispatch reconfigures and prices change. Power flows are different than in the transmission congestion contracts. However, in every case and for each set of TCCs, the congestion rentals equal or exceed the obligations under the point-to-point contracts. The transmission congestion contracts can always be honored, no matter what the pattern of load. In some instances, there will be excess congestion rentals to disburse, but transmission congestion contract holders will always be able to hedge power contracts without requiring physical transmission rights and without compromising the least-cost dispatch.

- end -