
Janusz Bialek, a Senior Lecturer with the University of Durham, England, is the author of a transmission pricing methodology based on the concept of power tracking, and co-author of a novel approach to the design of robust and decentralized power system stabilizers based on the Lyapunov direct method. Dr. Bialek, who holds a Ph.D. from the Warsaw University of Technology in Poland, is author of over 40 technical papers and co-author of the textbook *Power System Dynamics and Stability*. His research has been funded by a number of institutions, including EPRI.

Rachid S. Cherkaoui received M.S. and Ph.D. degrees in electrical engineering from the Ecole Polytechnique Fédérale de Lausanne (EPFL) in Lausanne, Switzerland, and is now a Senior Engineer at the electric power system laboratory at EPFL. His research interests are in the areas of power systems simulation and optimization and distribution-system automation.

Alain Germond, who holds M.S. and Ph.D. degrees in Electrical Engineering from EPFL, has been a Full Professor at EPFL since 1985, and head of the Electric Power System Laboratory. His research interests are in the areas of advanced techniques for power system analysis and control. Dr. Bialek thanks the Royal Academy of Engineering and the EPFL for their financial support during his sabbatical leave in Lausanne.

Improving NERC Transmission Loading Relief Procedures

NERC's curtailment procedure can be significantly improved by curtailing more of those transactions with a bigger influence on the congested flowgates, and by curtailing multilateral packaged transactions as a package rather than on a bilateral basis.

Janusz W. Bialek, Alain Germond, and Rachid Cherkaoui

The recent modification by the North American Electric Reliability Council (NERC) of the transmission loading relief (TLR) rules addresses some of the concerns voiced by the community, but it retains some of the fundamental principles which are at the core of the dispute: the load-relief rules cause higher-than-necessary curtailments of transactions, and there is no effort to improve the economic efficiency of overall congestion management. In this article we provide geometric interpretation of the TLR rules and argue that they could be modified in a simple way so that the overall curtailment is less. Then we show how economic efficiency of the overall congestion management process can be improved by

allowing a multilateral, rather than bilateral, approach to congestion management.

I. A Review of NERC Procedures

Whenever interchange transactions threaten security of the system, some sort of procedure has to be applied to curtail them. NERC has been successfully using its TLR protocol for a number of years. With the increase of power transfers, there has also been a significant increase in the instances in which the TLR procedures are used. While in all of 1998 more than 300 TLRs were called, in the first 10 months of 1999 more than 400 TLRs were called, resulting in over 8,000

MW of power curtailment in the three-month summer period beginning June 1999.¹ The TLR rules have been widely criticized, however, for not attempting to optimize regional congestion relief, and for being cumbersome, inefficient, and disruptive to bulk power markets.² It has also been alleged that the TLR rules have been used in a discriminatory manner.³ NERC recently published a modified version which addresses some, but not all, of these concerns.⁴ NERC notes that a market mechanism may offer better ways of dealing with congestion management than does physical interruption of power flows, but asserts that it will always be necessary to have a nonmarket mechanism such as TLR in place to ensure that the stability of the grid is always maintained.⁵

In this article we address two concerns: that the transactions are curtailed more than is necessary, and that the overall congestion management process is not economically efficient. First, let us start with a short review of the curtailment rules, a full description of which can be found on the NERC Web site.⁴

All the interchange transactions between control areas have to be tagged and entered to the Interchange Distribution Calculator (IDC). Using this information and the power transmission distribution factors (PTDFs) also available from IDC, each of the security coordinators calculates the expected loading on the transmission facilities (so-called flowgates) within their security area. The PTDFs quantify the effect of a given trans-

action on a given flowgate. For example, if a transaction of 100 MW from A to B has the PTDF due to a given flowgate equal to 0.2, it means that 20 MW of the transacted power will flow through that flowgate.⁶ When a security coordinator identifies or is informed that a flowgate is about to exceed, or has exceeded, its operating security limit, he uses the TLR procedure consisting of a number of levels. Generally, the transactions are curtailed in inverse priority order. First, nonfirm transactions are curtailed in last-in, first-out order. If this does not bring about the necessary reduction in transmission flow, then firm transactions are curtailed, and so on through any other priority classes. In this article, we will only deal with the problem of how to curtail firm interchange transactions of the same priority.

Now let us illustrate the TLR procedure using an example of three firm transactions: from A to D (transaction T_{AD}), from B to D (transaction T_{BD}), and from C to D (transaction T_{CD}). The values of transactions and their PTDF factors due to a flowgate X are given in **Table 1**. Let us assume that these are the only transactions

impacting the flowgate. The impacts of the transactions on the flowgate can be calculated by multiplying the transaction amount by the PTDF, and are given in the last column in Table 1. Note that the impact of transactions AD and BD is the same, as T_{AD} is three times T_{BD} but the PTDF of T_{BD} is three times the PTDF of T_{AD} . The total flow through the flowgate is equal to the sum of impacts, that is, 160 MW. Let us now assume that the security limit of the flowgate is 80 MW, which means that the overload is equal to $\Delta F = 160 - 80 = 80$ MW. The reduction of the flowgate flow by 80 MW is achieved by scaling down those transactions which have a positive impact on X (i.e., T_{AD} and T_{BD}) proportionally to their impact. The formula implicit in the NERC TLR procedure is:⁷

$$\begin{aligned} \text{CurtailedAmount} &= \text{Overload} \frac{\text{Impact}}{o(\text{PTDF}^2 \cdot \text{InitialTransaction})} \\ &= \text{Overload} \frac{\text{PTDF} \cdot \text{InitialTransaction}}{o(\text{PTDF}^2 \cdot \text{InitialTransaction})} \end{aligned} \quad (1)$$

When applied to our problem, the curtailments ΔT_{AD} and ΔT_{BD} can be calculated as:

$$\begin{aligned} \Delta T_{AD} &= \Delta F \frac{\text{PTDF}_{AD} \cdot T_{AD}}{\text{PTDF}_{AD}^2 T_{AD} + \text{PTDF}_{BD}^2 T_{BD}} \\ &= 80 \frac{0.1 \cdot 900}{0.1^2(900) + 0.3^2(300)} = 200 \end{aligned}$$

Table 1: Scheduled Firm Transactions

Transaction	Transaction Amount (MW)	PTDF	Impact
AD	900	0.1	90
BD	300	0.3	90
CD	100	-0.2	-20
Total	1,300		160

$$\Delta T_{BD} = \Delta F \frac{PTDF_{BD} \cdot T_{BD}}{PTDF_{AD}^2 T_{AD} + PTDF_{BD}^2 T_{BD}}$$

$$= 80 \frac{0.3 \cdot 300}{0.1^2(900) + 0.3^2(300)} = 200$$

As expected, both curtailments are the same and the total curtailment is equal to 200 MW + 200 MW = 400 MW. Note that although both transactions have been reduced equally (because their impact on the flowgate is the same), curtailment of AD reduces the congested flow by only $200 \times 0.1 = 20$ MW, while curtailment of BD reduces the flow by $200 \times 0.3 = 60$ MW. The total curtailment is equal to $200 + 200 = 400$ MW. The flow after the curtailment can be calculated as $(700 \times 0.1) + (100 \times 0.3) - (100 \times 0.2) = 80$ MW.

II. Concern 1: Curtailments Are Bigger than Necessary

It is clear from the above example that the TLR formula is inefficient, as it causes a bigger than necessary curtailment of transactions. Transaction AD is curtailed by the same amount as BD although it causes only a small reduction in the transmission flow. The question to be asked, then, is: What other formula would produce a smaller overall curtailment while being fair to both transactions? Unfortunately, both aims are contradictory and a compromise has to be reached.

Different options will be next considered, illustrated in **Figure 1**. The horizontal axis corresponds to transaction AD while the vertical axis corresponds to transaction BD. The

linear equation defining the line constraint is

$$0.1T_{AD} + 0.3T_{BD} - 0.2T_{CD} \leq 80, \quad (2)$$

(the flow must be less than 80 MW), or

$$0.1T_{AD} + 0.3T_{BD} \leq 80 + 0.2T_{CD} = 80 + 0.2(100) = 100 \quad (3)$$

The thick line in **Figure 1** corresponds to this constraint. The initial transactions correspond to the point of coordinates (900,300) and any curtailment must move this point onto the constraint line. The TLR procedure outlined above will move the initial point to position *NERC* of coordinates (700,100).

A. Minimum and Maximum Curtailment

Now let us pose the following question: What curtailment procedure would produce the smallest possible overall curtailment? The answer is that as the line flow is more sensitive to changes in transaction T_{BD} than T_{AD} , it is T_{BD} which should be curtailed while T_{AD} should not be curtailed at all. The corresponding curtailment of T_{BD} would be equal to $\Delta T_{BD} = \Delta F / PTDF_{BD} = 80 / 0.3 = 266.7$, and the result would correspond to point *min* in **Figure 1**. The position of this point is obtained by projecting the

initial point vertically on the constraint line. The overall curtailment of 266.7 MW is far less than the 400 MW obtained when NERC procedure was used.

Using a similar argument, the maximum overall curtailment is produced when only transaction T_{AD} is curtailed, that is, when the initial operating point is projected horizontally onto the constraint line to give point *max*. The resulting curtailment of T_{AD} is $\Delta T_{AD} = \Delta F / PTDF_{AD} = 80 / 0.1 = 800$ MW. Points *min* and *max* define the two extremes and any other curtailment point must lie between those two points. Generally, moving to the left on the constraint line from point *min* towards point *max* increases the overall curtailment.

Obviously, the minimum curtailment option is not really viable for two reasons. First, it is unfair, as it curtails only one transaction and second, it is unstable. Consider, for example, two transactions having the same PTDF. The minimum curtailment procedure would not be able to choose which transaction to curtail as both would yield exactly the same effect. However, the minimum curtailment procedure would

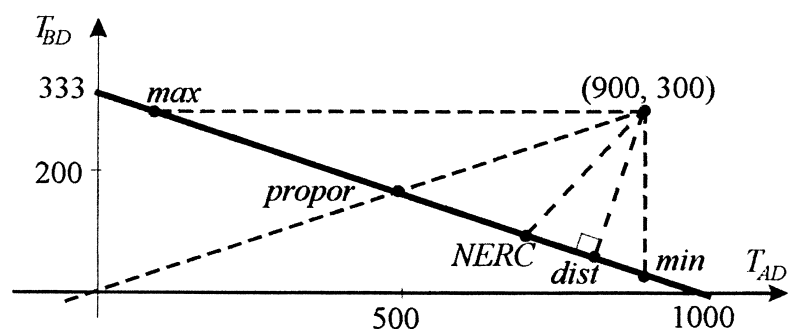


Figure 1: Different Options of Curtailment

serve as a benchmark against which all other procedures may be assessed. For example, the NERC rule results in 50 percent higher overall curtailment than the minimum.

B. Minimum Distance Curtailment

The other curtailment possibility is the minimum distance option, in which the distance between the initial operating point and the curtailed operating point, which must lie on the constraint line, will be minimized. Mathematically, this corresponds to the following problem:

$$\text{Minimize } \Delta T_{AD}^2 + \Delta T_{BD}^2 \quad (4)$$

(sum of the squared Cartesian coordinates of the distance),

subject to

$$PTDF_{AD}\Delta T_{AD} + PTDF_{BD}\Delta T_{BD} = \Delta F$$

(flow reduction = overload).

Solving this simple optimization problem gives the following result:

$$\begin{aligned} \Delta T_{AD} &= \Delta F \frac{PTDF_{AD}}{PTDF_{AD}^2 + PTDF_{BD}^2} \\ &= 80 \frac{0.1}{0.1^2 + 0.3^2} = 80 \end{aligned}$$

$$\begin{aligned} \Delta T_{BD} &= \Delta F \frac{PTDF_{BD}}{PTDF_{AD}^2 + PTDF_{BD}^2} \\ &= 80 \frac{0.3}{0.1^2 + 0.3^2} = 240 \end{aligned}$$

Note that the curtailments do not depend on the initial transactions but only on the PTDFs. The resulting operating point is shown in Figure 1 as *dist*, which is obtained by projecting the initial operating point at the right angle on the constraint line. This point lies to the right of *NERC*, which shows that the overall curtailment of 80 MW + 240 MW = 320 MW is less than that for the

NERC procedure (400 MW). However, it is still 20 percent more than the absolute minimum (267 MW). The minimum distance curtailment can also be viewed as unfair, as it does not take into account the transaction amount. For example, if two transactions had the same PTDF but one was three times larger than the other, both would be curtailed by the same amount despite their different impacts on the flowgate.

The minimum curtailment option is not really viable for two reasons: it's unfair and it's unstable.

C. Geometric Interpretation of NERC Procedure

It is worth noting that the NERC formula (1) also has a geometric interpretation, as it corresponds to minimizing the distance between the initial transactions and the curtailed ones, with each of the squared Cartesian coordinates of the distance weighted by the inverse transaction amount:

$$\text{Minimize } \Delta T_{AD}^2 \frac{1}{T_{AD}} + \Delta T_{BD}^2 \frac{1}{T_{BD}} \quad (5)$$

(Cartesian coordinates of the distance weighted by the transactions),

subject to

$$PTDF_{AD}\Delta T_{AD} + PTDF_{BD}\Delta T_{BD} = \Delta F$$

(flow reduction = overload).

Solving this simple minimization problem yields the solution given by Equation (1).

D. Proportional Curtailment

Yet another possibility is to consider a proportional curtailment in which each transaction with positive PTDF would be curtailed proportionally to the transaction amount. The resulting curtailments are:

$$\begin{aligned} \Delta T_{AD} &= \Delta F \frac{T_{AD}}{PTDF_{AD}T_{AD} + PTDF_{BD}T_{BD}} \\ &= 80 \frac{900}{0.1(900) + 0.3(300)} = 400 \end{aligned}$$

$$\begin{aligned} \Delta T_{BD} &= \Delta F \frac{T_{BD}}{PTDF_{AD}T_{AD} + PTDF_{BD}T_{BD}} \\ &= 80 \frac{300}{0.1(900) + 0.3(300)} = 133.3 \end{aligned}$$

which corresponds to point *propor* in Figure 1 obtained by moving the initial operating point along a line crossing the origin. Point *propor* lies to the left of point *NERC*, which means that the overall curtailment of 400 MW + 133.3 MW = 533.2 MW is bigger than when the NERC formula was used.

E. Proposed Compromise Formula

Let us now consider what curtailment formula would provide a necessary compromise between fairness (i.e., sharing curtailments proportionally to the impact) and effectiveness (i.e., minimizing the total curtailment). The NERC formula weighted the Cartesian coordinates of the distance by the inverse transactions. This can be considered as a compromise between the minimum distance formula (point *dist*) and the pro-

portional formula (point *propor*) so that point *NERC* lies somewhere between points *dist* and *propor*, hence increasing the overall curtailment.

Considering a compromise formula in which the influence of the PTDFs (and therefore the effectiveness) is strengthened by weighing the Cartesian coordinates of the distance by the inverse impact of each transaction, i.e., by the product of the transaction and its PTDF, the minimization problem then becomes:

Minimize

$$\Delta T_{AD}^2 \frac{1}{T_{AD} PTDF_{AD}} + \Delta T_{BD}^2 \frac{1}{T_{BD} PTDF_{BD}} \quad (6)$$

(Cartesian coordinates of the distance weighted by the impacts),

subject to

$$PTDF_{AD} \Delta T_{AD} + PTDF_{BD} \Delta T_{BD} = \Delta F$$

(flow reduction = overload).

Solving this simple minimization problem gives:

$$\begin{aligned} \Delta T_{AD} &= \Delta F \frac{PTDF_{AD}^2 T_{AD}}{PTDF_{AD}^3 T_{AD} + PTDF_{BD}^3 T_{BD}} \\ &= 80 \frac{0.1^2 (900)}{0.1^3 (900) + 0.3^3 (300)} = 80 \end{aligned}$$

$$\begin{aligned} \Delta T_{BD} &= \Delta F \frac{PTDF_{BD}^2 T_{BD}}{PTDF_{AD}^3 T_{AD} + PTDF_{BD}^3 T_{BD}} \\ &= 80 \frac{0.3^2 (300)}{0.1^3 (900) + 0.3^3 (300)} = 240 \end{aligned}$$

The overall curtailment is 80 MW + 240 MW = 320 MW, which is, as expected, less than when the NERC formula was used. This result happens to be the same as when the minimum distance formula was used because the impacts of both transactions, and therefore the weights in equation (6), were the same. However, if the impact of T_{AD} is higher than the impact of T_{BD} , its weight is smaller and T_{AD} is curtailed relatively

more. This corresponds to a slightly higher overall curtailment than when the minimum distance formula is used. If the impact of T_{AD} is less than the impact of T_{BD} , the opposite is true.

To illustrate this relationship, **Figure 2** shows by how much the overall curtailment is higher than the minimum for different curtailment formulas when the initial transaction T_{AD} is increased from 0 to 1,000.⁸ The initial transaction T_{BD} is adjusted in such a way that the line overload stays constant, which has the effect of moving the initial operating point along a line parallel to the constraint line in Figure 1. Consequently, the minimum curtailment and the minimum distance curtailments remain unchanged, so that the minimum distance formula always gives the curtailment of $320/266.7 = 1.2$ times the minimum.

Figure 2 shows that the pro-

posed compromise formula always gives a lower overall curtailment than the NERC formula. It also gives a lower curtailment than the minimum distance formula for most of the cases. Only when $T_{AD} > 900$ does the minimum distance formula give a slightly better result.

To illustrate better how the curtailments work, **Figure 3** shows the curtailments of T_{AD} and T_{BD} when T_{AD} is again increased from 0 to 1,000 MW and the line overload is kept constant at 80 MW. The solid lines correspond to the NERC formula while the dashed lines correspond to the proposed formula. The graphs clearly illustrate the fairness of the approach. Transaction T_{AD} is curtailed less using the proposed formula because the transaction's influence on the congested line is weaker.

Generally, with more than two transactions to be curtailed,

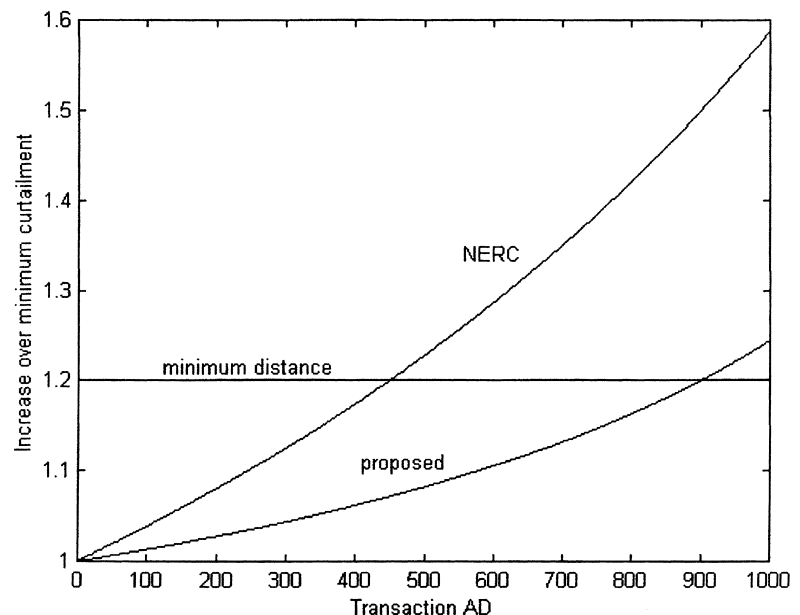


Figure 2: Increase in the Overall Curtailment over the Minimum as a Function of Transaction T_{AD} for Different Curtailment Options

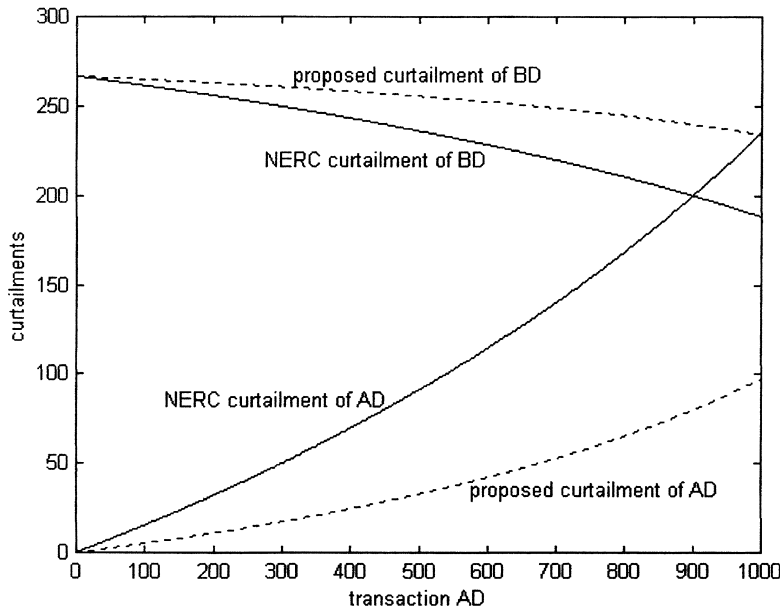


Figure 3: Curtailments of Transactions AD and BD when AD Increases

the proposed curtailment formula is:

$$\begin{aligned} \text{CurtailedAmount} & \quad (7) \\ &= \text{Overload} \frac{PTDF \cdot \text{Impact}}{o(PTDF^3 \cdot \text{InitialTransaction})} \\ &= \text{Overload} \frac{PTDF^3 \cdot \text{InitialTransaction}}{o(PTDF^3 \cdot \text{InitialTransaction})} \end{aligned}$$

The results of all the discussed options are compared in **Table 2**. Let us now interpret the results, looking again at Figure 1. As discussed before, the least-curtailment solution corresponds to the vertical projection of the initial transactions on the constraint line. As you go along the constraint line to the left, you increase the total amount of curtailment. Hence the next points, in order of increased curtailment, are *dist* (coinciding in this case with the proposed formula), *propor*, and *NERC*. The highest curtailment would occur if only transaction T_{AB} was curtailed. This corresponds to point *max*.

F. Coordination of Curtailments

The geometric interpretation of the curtailment procedures has an additional advantage, as it allows coordination between curtailment actions. Sometimes a problem may arise when the same transactions impact more than one congested flowgate. In this case the TLR rules applied to the interacting flowgates may result in contradictory commands. For

example, if the same transaction has the opposite effect on two congested flowgates (i.e., its PTDFs have opposite signs), curtailing it to alleviate constraint on one flowgate will actually worsen the congestion on the other flowgate. Even when the impact of a given transaction is positive on both flowgates, the curtailment commands due to each flowgate will be different.

The solution to this problem is simple. Optimizing the curtailment could be achieved by minimizing the total weighted distance between the initial and the curtailed transactions. In other words, the curtailment would minimize function (7) subject to interacting line limits, where the summation would be over all the transactions impacting the interacting congested flowgates. This task could be performed within the Interchange Distribution Calculator module, as IDC has all the necessary information. Computationally, this is a standard minimization task and should present no serious problems.

Table 2: Different Curtailment Options for $T_{AD} = 900$, $T_{BD} = 300$

	ΔT_{AD}	ΔT_{BD}	$\Delta T_{AD} + \Delta T_{BD}$	Excess Above Min (%)	T_{ADcurt}	T_{BDcurt}
Min	0	266.7	266.7	0	900	33.3
Min dist	80	240	320	20	820	60
Proposed	80	240	320	20	820	60
NERC	200	200	400	50	700	100
Proportional	400	133.3	533.3	100	500	166.7
Max	800	0	800	200	100	100

III. Concern 2: Achieving Higher Economic Efficiency

The main problem with any TLR rules which are purely administrative is that they are not economically efficient. The curtailment is grounded on a purely technical basis without taking into account its economic effect. As NERC stated, the congestion pricing methodology should "ensure that the generators that are dispatched in the presence of transmission constraints must be those that can serve the system loads at least cost, and limited transmission capacity should be used by market participants that value that use most highly."⁹ The most popular solution to this problem is bid-based locational marginal pricing (LMP). The question is therefore: Can any purely administrative TLR procedure, which is not based on a centralized bid mechanism, satisfy the goal of achieving economic efficiency? We claim here that even without using a bid-based system, a substantial improvement can indeed be achieved.

Our proposed solution to this problem is based on the observation that the NERC procedure is aimed at bilateral trades, although congestion management is by its nature a multilateral problem, i.e., the optimal solution usually involves a coordination of two or more parties. In our example, congestion relief could be achieved if trade AD is re-arranged so that the delivery of power from A is reduced and instead delivery of more power from C is scheduled. As the PTDF of T_{CD} is negative, this would create a counterflow

and relieve the flowgate. However, the NERC procedure is bilateral and would treat the packaged trade as two bilateral trades. Thus, if the flowgate was still congested, the NERC procedure would curtail the trade from A without taking into account that it is a part of a packaged trade. This problem has been extensively discussed by Rajamaran and Alvarado in the October 1998 issue of *The Electricity Journal*.¹⁰ They showed that re-

The NERC procedure is aimed at bilateral trades, although congestion management is by its nature a multilateral problem.

packaging penalizes the trader who undertakes the re-packaging while benefiting the competitors. To solve this problem, Rajamaran and Alvarado suggested a bid-based system, where each trader submits incremental/decremental price bids for both the supply and demand of the trader's transaction. This information would be used to determine centrally the optimal re-dispatch of transactions.

It is not our aim to dispute the bid-based approach to congestion management. Practical implementation of such a system to all interchange transactions, however, would encounter a number of

problems, both political and technical.¹¹ Without going into details, let us just say that any centralized solution is likely to be met with a deep suspicion. Also, there are many examples that bid-based congestion management may encourage gaming. Before all those problems are overcome, we would suggest a simple practical solution. Additionally, even if a bid-based approach is implemented, it will always be necessary to have a non-market mechanism such as TLR in place to ensure that the stability of the grid is always maintained.

We suggest that the NERC approach to transactions management be changed to reflect the multilateral, rather than bilateral, nature of the problem. To do this we propose that each trader should be allowed to submit a balanced schedule with multiple generation and delivery points. Moreover, interchange transactions between control areas should be treated in the same way as the native loads—that is, each control area would tag and enter its internal schedule to the Interchange Distribution Calculator. The importance of that will be discussed later. In the case of congestion, the security coordinator would apply the TLR procedure as proposed earlier in this article, but to the packaged rather than bilateral trades. Hence, if congestion is expected which would affect a transaction from A to B, the trader could arrange for a packaged trade whereby supplying B is achieved partly from A and partly from another source. How this is achieved is the trader's business.

When congestion still occurs and the package is curtailed, how the individual generations and demands within the package are curtailed is again the internal trader's business. Obviously, he would do it in an optimal way from his point of view, thus introducing some efficiency into the process. Ideally, following a curtailment, the traders should be allowed to arrange additional packaged trades, provided they do not overload the flowgates. It is this secondary round of repackaging which should make the resulting re-dispatch equivalent to the optimal dispatch.¹² All the necessary information for this approach to work is already available in the IDC module.

IV. Advantages of the Proposed Approach

Let us now discuss the advantages of the outlined multilateral approach. First, efficiently packaged trades which do not overload the flowgates are encouraged. This puts the onus on the traders to find the most cost-effective solution to the congestion management problem. The important thing, however, is that all the cost-benefit information is private to the traders and does not need to be provided to a central body (as in centralized optimal re-dispatch). It is important to realize that the suggested multilateral trading model achieves better efficiency if the cost and benefit functions cannot be explicitly written down as functions of generation and consumption as required for the centralized approach. This

also stimulates development of creative market solutions.

With NERC's current bilateral approach, there is scope for a lot of inefficiency because the societal benefits of each transaction are not coordinated. With the proposed multilateral approach, the transactions making up a bundle are optimal within a bundle (otherwise they would not be put together). Also, any curtailment decisions imposed on a bundle are

NERC's current bilateral approach inefficiencies because the social benefits of each transaction are not coordinated.

executed in an optimal way within the bundle. Hence, any scope for inefficiency arises only between the bundles. Theoretically, the traders should sort out those efficiencies between themselves and arrive at the optimal solution in the second round, but in practice, due to the size of the problem and time constraints, some inefficiencies could remain. These inefficiencies, however, would be far less than in the bilateral approach. In the extreme, if one marketer arranged all the interchange transactions, this would be equivalent to a centralized optimal dispatch.

Even without the secondary

round, which may be difficult to organize in real time, an improvement in economic efficiency is achieved through combining the individual trades in a bundle. Note that those who are keen not to get curtailed will organize a bundle and get it accepted. Those who are indifferent will get curtailed. Hence, economic criteria would be indirectly introduced without introducing a bid-based system. Conceptually, the proposed multilateral trading model goes in the same direction as FERC's *Order 2000—Regional Transmission Organizations*. Grouping smaller trades into bigger ones is bound to increase the overall efficiency.

Another advantage of the multilateral approach is that it would implicitly assign added value to the counterflows, i.e., transactions which have negative PTDF and therefore relieve the congestion. Under the current mechanism, such transactions do not receive any benefit (apart from not being curtailed when congestion occurs). Under the proposed approach, such counterflowing transactions would have an added value from the point of view of a trader who tries to arrange a bundled trade to transmit power over a congested flowgate. Hence, they would expect to be additionally paid. This would encourage generation investment in the areas where it is most needed.

The proposed approach would also solve another problem. Because transmission providers currently do not have information about their neighbors' on-system transactions to serve native load, they are unable to predict the impact of potential TLR events.¹³ In the face of this uncertainty, transmission pro-

viders often make overly conservative, but inaccurate, assumptions which unnecessarily reduce the amount of transmission capacity available to the market. If, as suggested here, the transactions to serve the native load were treated in the same way as an interchange package, the neighbors could assess the impact on their own system. Also, any curtailment procedure would apply to the native transactions in the same way as to interchange transactions. This would be in full agreement with the principle of non-discriminatory third-party access.

The proposed multilateral approach to congestion management would also solve one more problem indicated in FERC *Order 2000*.¹⁴ Apparently, there are some rampant manipulations of the TLR by so-called "shell game" when the transmission providers are able to dispatch their resources to manipulate their exposure to TLRs while customers cannot. This is done by designating fictional sources and sinks and treating one transaction as two separate transactions. Obviously, the proposed multilateral approach would put an end to such practices.

One of the possible criticisms of the multilateral approach is that it favors big marketers at the cost of small traders who do not have the muscle to arrange a bundle. The answer to this is that, from the societal point of view, cooperation between the market participants is necessary. Under the old paradigm, this cooperation was achieved by enforcing the will of the traditional vertical utility. Under the new market paradigm, this cooperation can

be achieved by arranging bundled trades. If a trader does not wish to cooperate he will be swept aside by market forces.

V. Conclusions

This article has argued that the existing NERC curtailment procedure can be significantly improved. First, on a strictly technical level, the curtailment rule can be changed by curtailing more of those transactions with a larger

One possible criticism of the multilateral approach is that it favors big marketers at the cost of small traders.

influence on the congested flowgates. This should result in a lower overall curtailment. The geometric interpretation of the TLR provides a basis for coordination of curtailment of transactions impacting interacting flowgates.

Second, to improve economic efficiency, multilateral packaged transactions should be curtailed as a package, rather than on a bilateral basis. This in itself should result in a more efficient outcome. Additionally, following a curtailment a second round of packaged trades should be allowed which would bring the solution closer to the optimum. All this would signif-

icantly improve economic efficiency of the congestion management without introducing a centralized bid-based system. ■

Endnotes:

1. FERC, *Order No. 2000: Regional Transmission Organizations*, issued Dec. 20, 1999, at 18.
2. *Id.*, at 34.
3. *Id.*, at 36, 43, 47, 52, and 335.
4. NERC *Operating Manual*, Appendix 9C1—Transmission Loading Relief Procedure, Version 2. Available at <http://www.nerc.com>.
5. *Supra* note 1, at 342.
6. The PTDFs are calculated using an approximated linear network model, and thus neglect the effect of transmission losses and system nonlinearities.
7. The TLR pro forma performs the same calculation using a spreadsheet, but the result is the same.
8. Transaction AD cannot be greater than 1,000 MW, as this would require making T_{BD} negative in order to observe the line limit.
9. *Supra* note 1, at 526.
10. R. Rajaraman and F. Alvarado, *Inefficiencies of NERC's Transmission Loading Relief Procedures*, ELEC. J., Oct. 1998, at 47.
11. This problem has been a subject of extensive discussion in the literature. See, e.g., Alan E. Rosenberg, *Congestion Pricing or Monopoly Pricing?* ELEC. J., April 2000, at 33.
12. See the seminal paper by F. F. Wu and P. Varaiya, *Coordinated Multilateral Trades for Electric Power Networks: Theory and Implementation*, ELEC. POWER AND ENERGY J., 21(1999), 75–102. The paper proves that a decentralized multilateral market achieves the same optimal dispatch in the presence of constraints as a market based on price bids and centralized economic dispatch (i.e., when locational marginal pricing is used). All the cost-benefit information is private and is used only for the purpose of negotiating contracts between willing parties.
13. *Supra* note 1, at 39–40.
14. *Id.*, at 52.