Durham University

Durham Energy Institute

Professor Janusz Bialek
tel +44 (0)191 3342505
email: Janusz.Bialek@durham.ac.uk.

Professor Phil Taylor
Tel 44 (0) 191 33 42524
email: P.C.Taylor@durham.ac.uk

17 November 2010

Project TransmiT – A Call for Evidence

Comments on behalf of Durham Energy Institute

Executive summary

The main thesis of this submission is that the current separation of the energy market from the transmission arrangements will become unsustainable with high penetration of renewables. We believe that the issues of transmission access, transmission charging and energy markets are inextricably linked together by physics of electricity transmission. Separating them for the reasons of convenience is bound to fail when congestion levels are rising rapidly. We suggest modifying the current market arrangements by introducing Locational Marginal Pricing (LMP) principles that have proved to be highly successful in dealing with congestion in a range of markets in USA. The main advantage of LMP mechanism is that it reconciles physics of electricity transmission with economics of pricing.

To implement LMP principles the existing GB market arrangements would require some modifications. All parties would be subject to energy-based forward transmission prices derived from bids and offers submitted to the Balancing Mechanism. Additionally, parties departing in real-time from their announced positions would have to pay nodal prices for the quantity that the party is short or long. The nodal prices would replace the current Balancing Services Use of System charge (BSUoS) as they would reflect the additional cost of changes in the dispatch pattern in the real-time market undertaken by the System Operator.

Transmission access arrangements would also have to change. The current practice of compensating curtailed-off generators means that they are effectively paid for doing nothing. This could be accepted as a price worth paying for simplifying market arrangements, encouraging competition and connection of new renewable plants when congestion levels were low, but is becoming increasingly unacceptable when congestion is high. Not only firm transmission rights make generators behind constraints oblivious to congestion but they would also actively encourage gaming. To address that, firm network access rights currently enjoyed by generators would be replaced by financial rights. The financial rights could be given free to incumbents, for a certain period of time, in order to overcome their resistance to any change, and they would be financed from congestion rents gathered through LMPs. Network costs would be recovered as at present by capacity-based TNUoS charges but, due to interaction with locational signals sent by LMPs, it would be necessary to establish how

strong the locational signal sent by TNUoS would have to be. One possible option would be for TNUoS charges to be locational but positive only, reflecting shared usage of the network.

Introduction

Current electricity market arrangements in GB aim to achieve the goal of separation of the energy market from the complexities of transmission, and specifically network access and congestion management. Project TransmiT also seems to maintain that separation by calling for evidence to review transmission charging and connection arrangements without mentioning the context of energy market arrangements. The main thesis of this submission is that the separation of the energy market from the transmission arrangements will become unsustainable with high penetration of renewables. The future network will become heavily congested at times of high wind making it necessary for the energy market to take into account transmission constraints. This issue is being dealt with by the industry in Transmission Access Review but the progress has been painfully slow. We believe that the issues of transmission access, transmission charging and energy markets are inextricably linked together by physics of electricity transmission. Separating them for the reasons of convenience is bound to fail when congestion levels are rising rapidly.

Physics of electricity transmission

It might be useful to start from reminding ourselves physics of electricity transmission as any market arrangements that disregard physical laws are bound to fail. In other words, for any pricing arrangements to be efficient and enhancing system security, they must reflect physical reality of the power system and its operation.

The two most important aspects of physics affecting pricing are the following:

- Because of the lack of significant storage capacity, the generation/demand balance must be maintained on a minute-by-minute basis. Hence the requirement for real-time energy pricing reflecting balancing actions taken by the System Operator (SO).
- The flow of power in individual lines is dictated by Kirchhoff's laws. In a meshed AC transmission network (such as that in GB), power flow in a line is the outcome of the overall pattern of generation and demand in the whole system. If any of the transmission lines is overloaded, the overall system generation/demand pattern has to be changed – this is referred to as re-dispatch or more generally as congestion management. Hence the energy prices should be locational, reflecting the value of generation and demand in different locations.

The main consequence of those two principles is that real-time pricing has to take into account not only the overall system generation/demand balance but also location of individual generators and loads. In other words the energy and the transmission markets have to be tightly coupled in a congested network.

Engineering and market-based solutions to congestion management

If the cheapest generation is scheduled without taking into account transmission limits, some of the lines may become overloaded and it is necessary to relieve overloading. A well-known engineering solution to congestion management is so-called Optimal Power Flow (OPF) optimisation tool[1] which optimally, i.e. at minimum cost, re-dispatches generators in a network to relieve congestion by using marginal cost characteristics of generators. OPF has

[1] Wood, A., J., Wollenberg, B. F. : "Power generation operation and control", Wiley, 1996.

been widely used since the late 1960s, well before the advent of liberalisation, by vertically integrated utilities which were able to control directly the generators they owned. Realisation by Schweppe et al[2] that a by-product of OPF, so-called shadow nodal prices, reflect the marginal cost of increasing generation or demand at any node, was a break-through that made it possible to introduce market-based power system operation. If generators and loads are faced with the nodal prices derived from their price bids, they are incentivised to adjust their generation or demand to the level required by OPF. In other words, market-based nodal prices achieve the same goal of congestion management as the direct command of generation by traditional vertically-integrated utilities. This way the prices are compatible with, and are integrated part of, optimal and secure power system operation. That methodology of energy pricing is referred to as LMP – Locational Marginal Pricing. It is important to emphasise that practically all utilities, including the vertically integrated ones, use OPF as an engineering tool to manage congestion. Hence the nodal prices are always readily available to use. Whether or not they are actually used depends on market arrangements in a particular system.

There are many world-wide examples of successful implementation of LMPs and the economic benefits they bring. Examples include PJM[3], New York ISO, Texas, California, MISO and others. Mansur and White[4] have studied operation of PJM market and provided evidence that bilateral trading could achieve only 40% of economic benefits of LMPs. PJM serves about 51 million customers with the peak load of about 144 GW so the system is comparable with GB in terms of the number of customers but the peak load is over twice as high.

Current GB arrangements

The goal of the current GB market arrangements is to decouple the energy market from the transmission market. The energy market is based on the assumption that the transmission network has been optimally planned and therefore provides enough transmission capacity for trading. Hence energy prices are uniform, non-locational and do not signal congestion. The task of the System Operator (i.e. National Grid) is to provide system balancing, i.e. matching generation and demand, and to reconcile the energy market with the physical reality of network congestion by running the Balancing Mechanism (BM). Parties submit bids and offers to BM and National Grid chooses the most economic ones to manage congestion and overall system balancing. In effect BM is similar to the LMP market in that congestion management is based on bids and offers submitted by the parties but BM does not send strong locational price signals the way LMP market does. The main differences are as follows:

1. Prices emerging from BM (i.e. System Sell and Buy Prices) affect only those parties that are in imbalance (i.e. produce less or more than contracted at Gate Closure). In contrast, nodal prices in LMP market affect all the parties and trades.
2. Prices paid to the successful bids and offers are pay-as-bid, rather than proper marginal prices.
3. Generators have firm access to the network (upon payment of TNUoS charge) and are compensated if they are curtailed-off with the curtailment cost socialised among all the network users. In contrast, there is no notion of firm access in LMP market and generators are not compensated if they are curtailed-off.

---

[2] Schweppe, F. C., Caramanis, M. C., Tabors, R. D., and Bohn, R. E.: "Spot pricing of electricity" Kluwer Academic Publishers, Boston, MA, 1988

[3] www.pjm.com,

[4] Mansur, E. T., White M. W., "Market organization and efficiency in electricity markets" available at http://bpp.wharton.upenn.edu/mawhite/, 2009

This system has worked reasonably well so far because the cost of transmission constraints in GB tended to be relatively small. Pre-BETTA, Scottish generators were not allowed to contract exports to England above the level of the Interconnector constraint (Cheviot boundary) which means that transmission congestion was effectively hidden. However with BETTA in place, the Interconnector has become just another internal GB line with no special treatment in operational practice and the cost of constraints has been rapidly increasing year-on-year reaching now hundreds of millions of pounds. This increase is expected to continue and exceed £1bn as more wind generation is added in GB, especially given the DECC decision to implement Connect and Manage socialised access system that effectively oversells firm transmission rights. Congestion will affect not only the Cheviot boundary but will spread to different parts of the system. Even if the planned network reinforcements were to be implemented in time, it is still expected that the network will be heavily constrained at times of high wind. Otherwise the network would not be optimally designed – the incidences of high wind are relatively rare and it does not make economic sense to design a network which would serve rare events. One cannot plan a network which would rarely constrain the peak generation capacity of over 100 GW when the peak demand is expected to be only 61-71 GW.

Increased congestion will have serious consequences to system operation and will mean that maintaining separation of the generation and transmission market, with generators enjoying firm transmission rights, will become unsustainable. For one, compensation of curtailed-off generators means that they are effectively paid for doing nothing. This could be accepted as a price worth paying for simplifying market arrangements, encouraging competition and connection of new renewable plants when congestion levels were low, but is becoming increasingly unacceptable when congestion is high. Not only firm transmission rights make generators behind constraints oblivious to congestion but they would also actively encourage gaming.

There are many world-wide examples of examples what is happening when markets do not reflect physics of electricity transmission. The reasons why PJM, NYISO, Texas and California have adopted nodal pricing is that markets quickly learned to exploit discrepancies between physics and market arrangements (one well-known example is the infamous “inc-dec game”). Nodal pricing is the only known approach that coordinates electricity pricing with fundamental laws of physics. And if there is a conflict between the laws of physics and market arrangements, the laws of physics always win.

Concerns about nodal pricing

Common concerns about nodal pricing are that it is complex, non-transparent and that the lack of firm transmission rights would create undue uncertainty for the generators. In fact nodal prices are calculated according to a simple optimisation function that minimises system cost given transmission constraints and bids, offers or schedules submitted by generation and load. This allows the resulting prices to reflect physics of electricity transmission. This transparent algorithm provides market clearing prices for all and thus values properly generation and demand at a given location. Thus also small generators can fully benefit from the market: The generation or demand response will be called whenever the nodal price exceeds the submitted bid and will receive the market clearing price. Thus it suffices to submit cost reflective bids in order to maximise revenue from competitive behaviour and the need for complex market research is limited.

Firm and financial access to the network

Adoption of nodal pricing would mean that generators lose firm access to the network which they currently enjoy. This is bound to be met with protest for obvious reasons – turkeys do not vote for early Christmas. However there are financial hedges available (so-called Financial Transmission Rights, FTRs), which are widely used in any LMP market, and which ensure financial (rather than physical) transmission rights to the network by refunding any transmission charges a user would have to pay. FTRs are funded directly from market proceeds (so-called congestion rents) so they do not require any additional funding mechanism. Hence the opposition from incumbents to move to the new LMP-based regime could be overcome by giving them tradeable financial transmission rights for a certain period of time. There is a vast body of literature analysing the benefits of financial, rather than physical, access rights. They are more flexible, tradeable, and they facilitate efficient exit/entry decisions. The value of transmission rights would also give an indication where transmission reinforcements is needed and would therefore minimise the risk of investing in stranded transmission assets.

A possible GB model

We suggest that the following possible LMP-based model of combined energy and transmission market arrangements could be developed. The model would be a modification of the current NETA/BETTA contracts-based model[5]. The generator-load pairs would submit, as at present, their physical notifications with price bids and offers to the System Operator (SO)[6]. After Gate Closure, National Grid would run the optimisation program (so-called Security-Constrained OPF - SCOPF) to resolve constraints, i.e. issue redispatch instructions based on submitted bids and offers and calculate forward nodal prices valid for the next trading period. The first part (redispatch decisions) would be basically unchanged from what is happening now; the new element is calculation of forward nodal prices that would be charged to <u>all</u> system generation and load. Those nodal prices would reflect the marginal cost of increasing generation/demand at a given location. We understand that National Grid's is procuring computational tools that could be adapted for that purpose. Such tools are readily available on the market.

All the transmission-connected generators and loads would be therefore charged forward nodal prices emerging from SCOPF. As GB market is contracts-based, a transmission charge paid by a generator-load pair would be equal to the nodal price difference between the two locations. This means that if there is no nodal price difference between the locations, that pair would pay no transmission charge. If there is a difference, the generator-load pair would pay or receive a transmission charge equal to the nodal price difference. Charging the same price to generators and loads would ensure that embedded generation is treated the same way as the transmission–connected generation.

Parties departing in real-time from their notified positions would have to pay the real-time nodal price, emerging from the Balancing Mechanism, for the quantity that the party is short or long. Positive and negative deviations would be treated symmetrically, thus avoiding

---

[5] Bilateral transactions tend to form a majority of trades in majority of LMP-based markets

[6] Ideally, the market should run more often than the current half-hour in order to achieve more efficient time resolution to resolve constraints. In PJM the market runs every 5 minutes.

discrimination against small actors. This would replace the current Balancing Services Use of System charge (BSUoS) as it would reflect the additional cost of changes in the dispatch pattern in the real time market undertaken by the System Operator.

We note that Ofgem has recently rejected locational BSUoS charges on the basis that two-day ex-post charging does not meet applicable charging objectives[7]. The experience across liberalised US markets, where they have eventually adopted LMP, was that a consistent and integrated market design for real-time, intraday and day-ahead trading avoids opportunities for gaming and provides a robust framework for participation of demand side response.

Interaction between short-term nodal price signals and long-term access charges

The LMP-based pricing covers only the short-term network costs, i.e. the cost of congestion management (and possibly losses). The long-run costs (i.e. investment cost of installing transmission assets) would have to be recovered, as at present, by Transmission Network Use of System (TNUoS) charges. TNUoS charges would have to remain capacity-based in order to avoid affecting dispatch decisions. Currently TNUoS charges are strongly locational and are calculated using a variant of marginal MWkm method. As LMP-based short-term pricing would also provide a strong locational signal, it would be necessary to study to what extent TNUoS charges should be locational, reflecting different usage of the transmission depending on the location, and how strong that locational signal would have to be.

One important aspect of TNUoS charges is whether or not the charges should be positive and negative. Currently generators deemed to relieve the load on the system, i.e. creating so-called counterflows, enjoy negative TNUoS charges so that effectively they are paid for being there even if they do nothing. If LMP-based energy prices send a strong locational signal, a case could be made for TNUoS charges to reflect shared network usage only, i.e. for the charges to be only positive, and therefore for the TNUoS locational signal to be much weaker. Transmission charging has been the subject of intense academic research over the last 20 years and a number of different approaches have been developed that could achieve that purpose.

I hope you will find our comments useful and we would be happy to discuss them in detail.

Yours sincerely,

---

[7] PJM settles transactions hourly