STUDY ON THE GENERAL DESIGN OF ELECTRICITY MARKET MECHANISMS CLOSE TO REAL TIME

YVES SMEERS

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

SCHOOL OF ENGINEERING [INMA] AND CORE

June 2008

1. INTRODUCTION: SELECTING A STANDARD

Competition is the instrument favoured by European Institutions to enhance the efficiency of the power sector. Competition requires trading, an activity that can be organised or spontaneously developed by the market. Experience and theory both suggest that some form of organized market is necessary for trading electricity. The reason goes back to the two special characteristics of this unusual commodity that have been most debated since the early discussions on the reform of the industry: electricity is not storable and it is difficult to transport. We are certainly well aware of these two important properties, but may not have fully drawn their consequences for the re-organization of the sector in Europe.

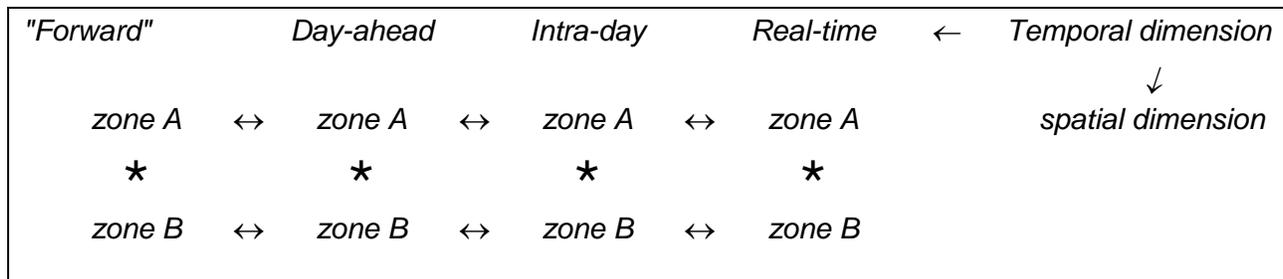Figure 1 gives a schematic description of the spatial and temporal dimensions of electricity trading.

"Forward"		*Day-ahead*		*Intra-day*		*Real-time* ← *Temporal dimension*
						↓
zone A ↔		*zone A* ↔		*zone A* ↔		*zone A* *spatial dimension*
★		★		★		★
zone B ↔		*zone B* ↔		*zone B* ↔		*zone B*

Figure 1: spatial and inter-temporal trading

The figure distinguishes four time stages: real-time, intra-day, day-ahead and "forward". It also refers to different "zones" between which trading takes place and hence introduces a spatial dimension. We assume in the rest of the discussion that a single organised market encompasses real-time, intra-day and day-ahead trading. In contrast we allow for "forward" trading to be organised or decentralised even though it is often admitted that even the "forward" electricity market should be at least partially organised. This over-encompassing trading platform does not reflect the current organisation of the European electricity market. Still we will use it as a reference in the rest of the discussion. We first briefly justify this choice either by invoking first principles or on the basis of existing experience.

1.1. On real-time trading

The sole existence of real-time trading is a first discrepancy between Figure 1 and the current organisation of the European market. The common view in Europe is that day-ahead, possibly extended to intra-day, but not real-time, are the appropriate periods for trading. Still, there are compelling physical and economic reasons for considering real-time as an effective trading period. The physical reason is that electricity is not storable and that real-time is the only moment when the commodity effectively flows from producers to consumers. This physical phenomenon has economic consequences. Electricity production and consumption need to balance instantaneously. The spot market therefore can only exist in real-time, which implies that all other trading activities are forward markets. Basic finance and economics teach that forward prices align somehow (by a non-arbitrage argument that we do not discuss here) on spot prices even when the spot market is essentially a residual market. This implies in the above scheme that both the intra-day and day-ahead markets are forward markets whose prices are driven by the real-time spot market. Wilson (2002) elaborates on the implications in terms of market design of this basic property of non-storability of electricity.

The idea of a real-time electricity market is not obvious. Because production and demand must balance instantaneously, the real-time market should also clear instantaneously. Ordinary markets cannot do that: traders, let alone brokers' clients cannot instantaneous clear markets. Only a "smart market", that is a market run by a computer program, can. "Smart market" is the name coined by Mac Kie-Mason and Varian (1993) in their analysis of the pricing of Internet access. The electricity sector has a long, but unnoticed, tradition of running smart markets. A well established tool of the industry, namely the security constrained optimal dispatch or a version of it is the instrument used today to clear real-time markets in many reformed power systems. It balances supply and demand and finds marginal values that can be interpreted as prices.

1.2. Spatial markets

The second fundamental property of electricity is that it is difficult to transport. Electricity flows in real-time and its transport is subject to different network constraints that must be satisfied permanently. The real-time electricity market is thus spatial, that is, the value of electricity is geographically differentiated as a result of the transportation constraints that are binding at that moment of time. The spatial differentiation of the price of electricity changes with the set of binding constraints. A network with ample excess capacities and hence few binding constraints

will not be split spatially very often but a market with tight bottlenecks in the grid will. The real-time market is thus a market of both transmission services and energy and the values of the service and the commodity are related. Again, Wilson (2002) explains the implications in terms of market design of this intimate link between energy and transmission.

In the same way as traders cannot clear a market in real-time, they cannot spatially arbitrage electricity, let alone arbitrage in real-time, at least if the intent is to account for the full possibilities of the grid. Electricity flows from one node to another by following different paths and it is the combination of all those flows on all those paths that may saturate the constraints of the network and change the arbitraging process. Again, a smart market is necessary in order to spatially clear supplies, imports, exports and demand. The optimal power flow or a version of it is again the appropriate tool for performing this spatial arbitrage. It balances quantities locally and globally; it also finds locational marginal values that can be interpreted as prices.

1.3. A reference trading platform

Figure 1 is based on these observations. It begins with a spatial real-time market for electricity and transmission and assumes that an instrument of the optimal power flow type clears this market, that is, computes quantities and prices. Forward markets complete the picture. Organized forward trading can, but does not necessarily, develop in risky markets; it all depends on the traded product. This also applies to electricity. In contrast with the real-time electricity market that emerged as a logical implication of fundamental properties of electricity, the forward electricity markets and their organization developed from practice. Given the need for a tightly organized real-time market that simultaneously clears energy and transmission, the question arose whether forward trading would spontaneously develop (or not) and whether it also requires a strong organization. The usual reasoning of finance that forward commodity prices are established on the basis of non-arbitrage and cost of carry principles does not fit here. Except for making it clear that the non-storability of electricity is incompatible with any cost of carry, theory offers no overwhelming advise on the possible organization of forward energy and transmission markets (see Wilson 2002). Different paradigms are thus possible, at least in theory.

1.4. Smart day-ahead and forward markets

It is convenient for the discussion to distinguish two stages in the temporal dimension namely day-ahead and ahead of day-ahead. Many reformed electricity systems developed a comprehensive set of forward markets that goes much beyond the organization commonly found

in standard commodity trading. The day-ahead electricity market indeed often developed into a smart market where the System Operator solves a security constrained unit commitment problem to determine forward prices. This is at least the practice in many reformed US electricity systems even though this view is not a universally accepted. Specifically, there is no well-established philosophy on the matter in Europe: the day-ahead market is often seen as a spot and not a forward market; ideas diverge as to its organization even if one progressively recognizes that it should be a smart market. The UK also offers an altogether different vision of the reform of the electricity sector. This report is about one of these day-ahead markets, namely the Flow Based Market Coupling (FBMC) foreseen for Central West Europe (CWE).

Day-ahead is not the end of the story. Smart markets can also be found in the "forward market" (ahead of day-ahead) where one could think that standard exchanges or even OTC markets would suffice. The reality is that many reorganized electricity systems also resort to smart markets to allocate long-term (well before day-ahead) financial transmission rights (FTR). The rationale is again technical and suggested by the physics and economics of electricity: for reasons of "revenue adequacy" that would be too long to explain here (see e.g. Hogan 2003) allocated FTR must indeed be physically feasible with respect to the capabilities of the network, something that only a computer program can guarantee. This feasibility condition imposes constraints on the forward market that drastically diverge from the standard hypothesis of an unbounded set of transactions that underlies the no arbitrage argument of finance. This divergence with standard finance assumption is itself a justification for a special organized market, even though it may create perturbations of the arbitrage process that are not yet fully elucidated today.

1.5. The reference platform in practice

This comprehensive and carefully structured set of organized markets emerged from academic analysis and real world implementation. Hogan provided the initial theoretical foundations in 1992 (Hogan 1992). He and others extensively developed these ideas later. These principles were first implemented in New Zeeland and in the pools of the East Coast of the United States (see Joskow 2007 for an historical discussion and IRC 2007 for the current status). FERC's eventually codified these ideas and experience in its Notice of Proposed Rulemaking (NOPR) on the Standard Market Design (SMD) (FERC 2002). The NOPR did not lead to mandatory requirements as the reform of the power system came to an almost complete halt in the US after the Californian crisis. But the expanse of the electricity systems operating under these principles

is impressive: close to 600 GW (UCTE capacity was slightly above 600 GW in 2006 (UCTE 2006)) operate today under an SMD type of organization in North America (IRC 2007). It is of interest, but also of concern, to note that the SMD architecture, which has now operated in various US systems since 1998 is much more demanding than what is commonly implemented elsewhere. Some, as the European Parliament, see this architecture as mainly administrative. General practice however shows that it works well and gives very satisfactory results in terms of competition. Recent studies (Mansur and White 2008) quantitatively confirm the good performance of these organized markets.

Many systems operate today under an organization similar to the platform of figure 1. New Zealand opened the way in 1996 with a real-time market; the pools of the East Coast of the United States followed suit and progressively added a day-ahead market. This market organisation then moved to Texas and other regions of the US. The case of Texas is particularly interesting. This system tried by different means to avoid resorting to the SMD but progressively and eventually adopted it. This success story is not a universal though. Not all reorganized electricity systems began with a real-time market or eventually adopted one. Nordpool's started with a day-ahead market as the rest of Europe is now trying to do. But the Nordic market was born in Norway, which, because of its considerable hydro basis is some sort of exception where electricity is storable as long as one has residual turbine capacity and reservoir volume. The need for a real-time market is then less pressing. It is only later, when Nordpool expanded to countries less endowed with hydro-resources that it tried to add an integrated real-time market, something that it found difficult to accomplish because it had not been foreseen before. The UK and Australia also began with a day-ahead market without any real-time market, but the UK completely modified its view when it moved to NETA and later to BETTA. The only organised market in the UK is real-time, the rest being left to bilateral transactions.

Except for the US, one thus finds that most reformed electricity markets in the world only adopt one or a few elements of the platform but do not develop the whole suite of markets of the SMD. Specifically, today's discussion in the Central West Europe (CWE) region (see ERGEG 2007, 2008) for the presentation of the regional initiative) concentrates on a particular organisation of the day-head market that has some SMD flavour, but still remains very far from it. Also, the intra-day trading and balancing markets in CWE are quite different from the day-ahead organization. This heterogeneity of design across time also diverges from the unifying philosophy of the SMD. It is worthwhile reflecting on whether adopting one piece of the SMD, significantly modifying it,

and connecting it with intra-day and balancing organised differently can effectively remove the barriers to trade that European Authorities often complain about.

1.6. The reference platform in the report

Figure 1 presents the skeleton organisation that we shall refer to later in the report. It contains the real-time, day-head and forward markets to which we add a physical intra-day market. We justify this addition as follows. First, European thinking on cross-border trade effectively includes a physical intra-day trading. The US SMD also introduced an intra-day virtual trading, which has some flavour of the European physical intra-day trading. A second reason is that intra-day trading comes closest to continuous trading, which is a common assumption in finance. Last, intra-day spans the interval from day-ahead to real-time, which naturally suggests revisiting the European thinking on balancing.

We do not go beyond the energy and transmission markets depicted in figure 1. Besides storability and transportation difficulties, electricity has other unusual characteristics that complicate trading and market developments. Reliability and the need to maintain a spinning reserve or the possibly insufficient incentives for investments in so-called "energy only" markets are some of these important questions. This paper does not venture into these areas; it focuses on the organization of cross-border trade in Central West Europe (CWE) and only looks at the energy and transmission markets.

1.7. The report

Regulators and TSOs of CWE are currently discussing a proposal (see Orientation study, CWE MC Project) to organize cross-border trade in CWE on the basis of a methodology referred to as Flow Based Market Coupling (FBMC). This paper concentrates on that proposal. The CWE market is largely bilateral but it also developed several Power Exchanges (PX) that clear in day-ahead. The duality of organized and decentralized markets is now common in reformed electricity markets. It is also the rule in the SMD where it only imposes certain requirements for the organization of the transmission market. FBMC is related to the nodal pricing system implemented in SMD but the details differ considerably. We can therefore draw on the experience of the SMD to anticipate possible difficulties with FBMC. Last FBMC only deals with the day-head market while the intra-day market and balancing system adopt a different organization. In contrast the SMD adopts a common design for day-ahead and real time. The

heterogeneity of market designs through time is another potential source of difficulties that the experience of the SMD can help elucidate. These are the themes of the report.

The text is organized as follows. The next section gives an overview of the situation that we summarize as follows: FBMC enables a potentially important step in the construction of the IEM but it is vulnerable to illusory "simplifications"; the crucial choice is how one represents the grid. Section 3 places the current developments in the history of the IEM: the introduction of a market coupling mechanism based on transmission capacities in part of CWE was a first major step as it introduced the necessary intimate pricing link between energy and transmission services. But this step remains by nature limited because of the inadequate transmission services that it involves. The FBMC proposal is about the introduction of more appropriate transmission services. As has now become common in the reform of the electricity sector the final outcome will however depend on the details of the implementation. Today situation is particularly delicate as the forthcoming penetration of wind power can dramatically exacerbate the consequences of an inadequate implementation of FBMC, let alone of the status quo. Section 4 introduces the technical material necessary to develop this argument. The section presents the main features of the flowgate model that underlies the FBMC proposal. It emphasizes the distinction between node-to-line Power Transmission Distribution Factor (PRDF) of the original flowgate model and the zone-to-line PTDF of the current FBMC proposal: the section explains that zone-to-line PTDFs require important assumptions about the pattern of injection and withdrawals in the zone. Section 5 turns to "critical infrastructures" that constitute another key feature of the FBMC proposal. The section offers two messages: first, critical infrastructures allow one to better identify the real causes of the limitations of cross-border trade and to mitigate these limitations; second the zone model of the grid may neutralize this potential progress. The rest of the paper takes stock of this discussion and explains the use of FBMC to construct a trading platform that spans the different stages from day-ahead to real time. We introduce the spatial dimension of the platform in section 6 and explain that the transmission services introduced in FBMC better represent the real possibilities of the infrastructure but are affected by the zone representation of the grid. Section 7 takes up the time dimension of the trading platform and first notes that all the considerations made up to this point in the report apply whether one is trading in day-ahead, intraday or real time. This implies that FBMC can provide a system supporting trading from day-ahead to real time. In the same way the dual theme of the report, namely the potential progress offered by FBMC and the potential damages resulting from the zone representation apply to all trading stages. But other remarks are also in order: (i) the difficulties created by the zone system in the day-each market are exacerbated when moving to real time; (ii) the separation of the PX

and TSOs that characterizes the EU organization is probably a further handicap at least if one compares it to the integrated ISO/RTO that successfully runs the reformed markets operating under the SMD in the US; (iii) there is no good reason for changing the organisation of the markets and in particular the respective roles of the PXs and the TSOs as one moves from day-ahead to real time. This discussion constitutes the (admittedly) lengthy background for the rest of the report in which we successively examine the role of FBMC in day-ahead, real time and intra-day trading. We elaborate on day-ahead trading in section 8. The OTC and PX markets constitute the two major components of trading. The PXs, together with the TSOs organize the coupling of the market in the FBMC. This operation involves a Consumer and Producer Surplus Maximization (CPSM) problem: we argue that this problem is affected by two difficulties: one is inherent to the zone system and can be eliminated by resorting to a nodal model; the other difficulty results from the trading of bloc orders and is inherent to machine constraints. We discuss these two questions in section 8. We prepare the extension of these considerations to real time and intraday trading in section 9. Section 10 particularizes the discussion of the trading platform to the real time market; section 11 does the same with balancing. Both conclude that except for the persistent problem of zones, FBMC could fit both types of trading. A short section 12 summarizes the main points covered so far. Section 13 briefly takes up the constraints placed by wind power on the trading system: the message is that intermitted sources justify developing FBMC in two directions: it should be spatially more sophisticated and cover the different trading stages. The last section gives a few indications on the construction of zones in case one is willing to depart from the one-country-one-zone principle that underlies the current FBMC proposal. Conclusions terminate the report.

2. THE CENTRAL WEST EUROPEAN MARKET IN CONTEXT

Many reformed electricity systems did not adopt the SMD paradigm. This market design is a US product even though it is also the only existing standard in electricity market architectures. Other markets have their idiosyncratic characteristics that make them difficult to compare. The reform of the European electricity industry is a case in point as all Member States approached market design differently. This diversity hindered the construction of the internal electricity market and led to the view that one should first try to integrate national markets on a regional basis. This resulted in the so-called Regional Initiative (ERGEG 2007, 2008). We concentrate on one of these regional markets namely the Central West European market (CWE), which consists of the three Benelux countries plus France and Germany. Besides allowing for bilateral transactions, four of these fives CWE countries also have an organized day-ahead market. These systems are today developing intra-day trading, but none of them is planning a real-time market, let alone an integrated real-time market. A cross border balancing mechanism is however foreseen after the implementation of the cross-border intra-day in 2009-2010. Priority has thus been given to the integration of the day-ahead markets, as these are the only ones operating on a significant scale today. This is the Flow based Market Coupling (FBMC) proposal.

2.1. First step forward: adopting implicit auctions

Three of the day-ahead markets of CWE, namely those of Belgium, France and the Netherlands are coupled since November 2006 through what we shall refer to as a Transmission Capacity based Market Coupling (TCMC). TCMC is not equivalent to a full integration of the three markets but it makes some step in that direction. We disregard here (important) issues of lack of harmonization between opening and closing hours of different national markets. These questions do not require any particular methodological discussion, even though it is obvious that they should be treated properly in order to enhance integration (Frontier Economics 2007). More relevant for this report, TCMC differs from full market integration by its treatment of the network. Access to the grid has been discussed in Europe since the inception of the Florence Regulatory Forum and it is still the object of debate today in CWE. The current coupling of the Belgian, Dutch and French markets (the trilateral market) was a positive but limited step in the good direction; it anticipates FBMC and its discussion can thus help explain what is at stake.

The trilateral market is based on a representation of the cross-border trade possibilities in terms of Transmission Capacities (TC). This is depicted in Figure 2.
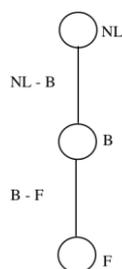
NL

NL - B

B

B - F

F

Figure 2: The network in the trilateral market

Each national market (or more specifically each national PX) is modelled as a node and the nodes are connected by capacitated links, also referred to as interconnections or interconnectors. Interconnectors in the trilateral market are entirely characterized by their Transmission Capacities (TC). The TCs impose bounds on the power flow through interconnectors. We shall sometimes refer to the view depicted in figure 2 as the copper plate description of the market: each node is seen as a zone with negligible internal congestion (the copper plate); these zones are linked by limited transmission capacities. This representation is intuitive and convenient but it does not correspond to the grid model that a TSO would use if the three systems were fully integrated in a single market. It is also a model that will be difficult to maintain when Germany is coupled in early 2009 (as planned) to Benelux and France. In fact we shall argue that the further construction of the European internal electricity market is partially conditioned by the abandonment of the copper plate description of the market. This statement should be taken positively. Even though the TCMC and its underlying copper plate view of the market are very far from allowing full market integration, they offered significant conceptual and practical advances with respect to other continental European systems as we discuss now. But a further step is necessary to move forward.

The TCMC essentially consists of an algorithm that iterates between the three national individual PX. The algorithm takes into account the "Transmission Capacities" between the systems and iterates until both the energy and TC markets clear. By so doing TCMC simultaneously finds energy and transmission capacity prices. We argued in the introduction that energy and transmission are intimately related in real-time. The usual non-arbitrage argument of finance suggests that the link that exists in real-time must also prevail in the day-ahead market. The TCMC recognizes this direct relation when it simultaneously computes day-ahead energy and transmission prices. TCMC implements what is referred to in EU jargon as an implicit auction. The recourse to implicit auctions in the implementation of market coupling is a positive logical

step in the integration of national markets. A rapid convergence of the prices in the three markets and a better use of the "interconnectors" was the immediate practical result of that move.

2.2. The second step forward: giving up Transmission Capacities?

Whatever its merits, TCMC is based on a representation of the grid that a TSO would never used if it were operating an integrated electricity system extending over Belgium, France and The Netherlands. The reason is that TCs give a very poor representation of the grid that TSOs are supposed to operate in the best possible way. The experience of FERC's Order 888 in the US shows that access to "Transmission Capacities" is not sufficient to enhance cross-system competition. The experience of Order 2000 also provides the alternative: one must abandon transmission capacities and give System Operators, whatever their form, the task of organizing a market of transmission rights that better fit the realities of the grid, while still clearing together with the energy market. The difference between the two approaches is thus a matter of traded transmission products: transmission capacities are not appropriate transmission products.

Flow Based Market Coupling (FBMC) introduces more adequate transmission products. We explained in the introduction that the transportation of electricity is subject to various constraints. The Flow Based model that underlies FBMC offers a reasonable representation of these constraints. Physical laws imply that a meshed grid is a single system where actions at some locations have an impact throughout the rest of the network. These impacts can be described by sensitivity coefficients (the Power Transmission Distribution Factors (or PTDF)) that we shall extensively discuss in the course of the report. TCMC neglects these impacts during market clearing; FBMC fully takes them into account. The existence of these impacts invalidates the representation of the grid of the figure 2 type where national systems are copper plates (congested free zones) linked by interconnections that can be used independently of one another (e.g. a flow on N-B has no impact on the flow B-F). The departure from TC and the adoption of a flow-based representation of the grid (PTDF and line capacities) is thus central to a better representation of the grid and hence also to a better access and pricing of the resources of the grid. We shall therefore extensively elaborate on the construction and use of the PTDFs as we proceed. The flow-based representation of the network through PTDF also departs from the underlying, physically erroneous, view that cross-border trade is a matter of cross-border transmission capacities. The introduction of FBMC after the inception of implicit auction is thus a second positive move in the right direction and a potentially quite important one.

While the construction of the internal electricity market can benefit from these two positive moves we shall also argue that it can be jeopardized by the particular construction of the PTDFs adopted in the current FBMC proposal. We just explained that FBMC abandons the view of national systems connected by limited interconnectors. We later explain that its particular implementation still retains the view of national systems even if it modifies the linkage between them. FBMC indeed sees a national system as a zone trading electricity with its neighbours at a single national price. By working with four price zones corresponding to the four existing PX, the proposed implementation of FBMC re-introduces at the trade level the zone model of the electricity system that it could abandon at the physical level by introducing node-to-line PTDFs. It does so at the cost of considerable assumptions that may eventually make FBMC unduly difficult if not impossible to implement.

2.3. Miscellaneous

FBMC appears in a minor mode in documents dealing with intra-day trading (Cegedel et al. undated) but remains absent from discussions of balancing. The experience of the SMD suggests that these different trading stages are related and should be based on similar descriptions of the grid. Similarly, financial reasoning suggests that the adoption of different trading organisations at different stages may create artificial arbitrages that are artefacts of the trading platform but do not correspond to economic realities. It thus makes sense to wonder about extending FBMC to these other trading stages in order to eliminate these artificial arbitrages.

Last, FBMC as well as TCMC encounter difficulties for treating bloc orders. This is a general issue that is part but not specific to the FBMC proposal. It is discussed in Section 8.4.

2.4. A summary of the recommendations

The above discussion delineates the content of this report. We first concentrate on the introduction of FBMC in the day-ahead market. We argue that this is a major and sound step because it establishes a strong relationship between energy and transmission prices using an adequate representation of the grid. Difficulties arise with the implementation of this principle. The Flow Based description of the grid is vulnerable to dubious simplifications. One of them is to adopt a coarse definition of the grid into zones and interconnecting lines. This view of the grid looks like a simplification but may in reality seriously complicate the problem. By equating zones and countries, even for large systems such as France and Germany, the FBMC proposal adopts

the most extreme version of these dubious simplifications. We therefore recommend implementing FBMC for the day-ahead CWE market but without resorting to illusory simplifications.

The second recommendation is to extend FBMC to intra-day trading. We argue that the tools used for day-ahead can be applied as such for intra-day if the description of the grid is sufficiently detailed. The organization of intra-day trading is even simpler than the one of the day-ahead market even though it will involve additional communication activities. The application of FBMC to intra-day trading will also imply a revision of certain ideas on the pricing of transmission in intra-day.

The next recommendation is probably more controversial: FBMC should also apply to real-time to construct an integrated real-time market that will substitute the current balancing system. Again the machinery used for day-ahead and intra-day trading can be used as such for clearing the real-time market.

We argue that it is urgent to adopt these measures. European policy is placing a lot of emphasis on the penetration of renewable energy, a significant fraction of which will be invested in wind. Because of its intermittency the penetration of wind energy will tax network resources, which will thus need to be managed with great efficiency. A good FBMC, based on a detailed description of the grid together with a single platform that allows trading energy in the day-ahead when wind forecast is poor and close to real-time when it is good can significantly contribute to that endeavour.

Finally we briefly tackle the question of zone definition in the Flow Based description of the grid. This is justified by our previous claim that the proposed version of FBMC is prone to difficulties: FBMC is an adequate concept, whose implementation can be distorted by inadequate simplifications. Little is known about zone definition as most systems that adopted a flow based approach in highly meshed networks have eventually taken the more drastic approach of moving to a nodal system where the grid is represented by its buses and lines. We briefly survey what is known.

3. THE REASONING OF FLOW BASED MARKET COUPLING

The spatial segmentation of the grid whether at the infrastructure or operations level, is currently limiting the integration of the European power market. Market integration requires a trading platform where agents can trade through space and time on a consistent basis. Regulation 1228/2003 introduces transmission capacity (TC) as the enabling concept of this platform. According to that view, the power system, or a part of it such as the CWE region, consists of a set of countries separated by interconnections of certain transmission capacities. This is the copper plate view: cross-border trade is limited by the congestion of these interconnections that occur whenever they are saturated, or in other words, when the flow on the interconnector reaches its TC. Regulation 1228/2003 does not cover congestion inside national borders. This should be taken care of by national TSO. Except for article 1.7 of the Congestion Management guidelines that obliges TSO to take the impact of national congestion remedying measures on neighbouring systems into account, national congestion is a non-problem for Regulation 1228/2003.

3.1. A first, but limited, positive step: TCMC

Network resources must be allocated among transactions. Stakeholders first introduced explicit auctions as a much desired "market based" method for allocating TCs. Explicit auctions recognized that transmission on existing infrastructures is a service that should have a market price. The introduction of explicit auctions created a specific market for TCs that clears separately from electricity. The recognition that TCs have a price was a useful move but the inception of implicit auctions to simultaneously determine the prices of energy and TC services was the real step forward. The result was the TC based coupling of the Belgian, Dutch and French electricity markets. Even though quite simple compared to the state of the art SMD, this coupling already permitted a significant step towards the integration of these three markets. It should be noted that these conceptual and practical advances have not taken place yet at the intra-day and balancing levels: there is no intra-day or real-time price for transmission capacities.

Notwithstanding the progress achieved with TCMC, it is now recognized that Transmission Capacities provide an inadequate representation of the grid that limits trading possibilities. TC's also have strange properties: for instance there is also no guarantee that transactions that fit within TCs at each cross-border level are globally compatible with grid capabilities. The result is that a right to use TCs does not guarantee the physical and financial security of the cross-border

transaction. This in turn induces TSOs to declare artificially low TCs in order to protect themselves against having to interrupt transactions. TSOs had recognized very early that the representation of the grid through TCs is inadequate. At the same time that several ETSO documents took considerable care defining TCs, others elaborated on the very strange properties of transmission capacities and explained that one had to be very careful when using them. TSOs did not explain what being careful implied. They instead introduced an alternative Flow Based (FB) description of the grids as early as April 2001, that is, when Regulation 1228/2003 had just been proposed and well before it came into force. Various reports later published by ETSO described the chaotic organisation of cross border trade that progressively developed from the allocation of TCs whether by market and non-market based methods. At the same time, the US literature extensively discussed similar issues and exposed the close relation between TCs and the ill fated "contract path" that ruled electricity transactions on both sides of the Atlantic. Attitudes were in many points similar in Europe and North America: while it was generally agreed that the notion of TC is flawed, adopting something else met considerable resistance, at least until the successful experience of PJM showed the way forward.

There is not much value in elaborating in detail on the fallacy of attempting to build a competitive electricity market on the basis of TCs, except for recalling once more that the problem arises from the trading of inadequate transmission services that do not represent the real possibilities of the grid. The implication is that progress can only come from the introduction of more adequate transmission services. This point will be developed in the rest of the paper; its discussion can be initiated as follows. As illustrated on figure 2, the description of the grid by transmission capacities relies, among others, on a double aggregation: nodes are grouped into zones and lines are assembled into interconnectors. The problem is that, except for the special case of radial grids, we do not know how to aggregate electric nodes and lines in a way that reasonably maintains the original properties of the grid in an optimal dispatch. We know that this aggregation is not possible in theory and should evaluate the errors of doing it in practice. Because of these difficulties, TSOs (i) restrict trading ex ante and (ii) remain unable to secure ex post the transactions that they accepted ex ante. One may also add that this organization gives contradictory incentives to TSOs: they cannot maximise TC as required by Regulation 1228/2003 without at the same time increasing the risk of having to curtail transactions and incurring penalties. Rational TSOs operating under a TC regime will thus restrict TCs. Wind penetration and the resulting generation intermittency will enhance these problems in the future. We thus take it for granted that a competitive electricity system in a highly meshed grid such as the one of CWE needs to be built on something else than TCs.

3.2. A second step: FBMC?

The TC representation of the grid is based on an aggregation of both nodes and lines. The proposed FBMC no longer aggregates lines and hence offers a better representation of the grid. But FBMC retains the aggregation of nodes into zones. Neglect, for the time being, the possible difficulties created by this latter aggregation and recall from the preceding discussion that TCMC improved on preceding designs by moving from explicit to implicit auctions of the infrastructure. We want to retain this progress in FBMC and conduct an implicit auction of the grid services offered by the flow-based representation of the network. This is precisely what the current FBMC proposal does. It offers two advances with respect to the explicit auctioning of TC that can potentially improve the functioning of the market: it retains implicit auctioning and applies it on individual lines. This improves on TCs but does not solve all problems: FBMC still relies on an aggregation of nodes by zones, a process that we no not really master well. We shall argue that an inadequate node aggregation in FBMC may seriously degrade if not fully destroy the potential for progress! We also argue that the penetration of wind power will make that possibility more and more crucial in the future. Nodal aggregation (and needless to say a combination of nodal and line aggregation) combined with higher wind penetration will indeed induce TSOs to further restrict cross-border trade in order to guarantee the security of the grid. An increased penetration of wind power implying a reduction of cross-border trade is not an acceptable goal for European policy! The following describes these points in more detail: it concentrates on the potential offered by FBMC, both for day-ahead and other trading stages and discusses how an improper aggregation of nodes into zones can destroy this potential. The report does not refer to TCs except when for the sake of clarification in the discussion. In the same way, the report does not mention explicit auctions, as these are not necessary for discussing FBMC.

4. THE PHYSICAL REALITY UNDERLYING FBMC: FLOWGATE

A power grid can be seen as a set of nodes linked by electrical lines. Injection and withdrawal of power take place at the nodes and power flows on the lines as a result. Physical properties known as Kirchoff laws relate injections, withdrawals and line flows. The first Kirchoff law states a simple conservation phenomenon: the sum of all incoming and outgoing flows at a node is equal to zero. This law is well understood and its interpretation does not cause any controversy. The second Kirchoff law is commonly known as the principle of least resistance. Power

distributes among the different lines of the network so as to minimize the resistance that it encounters. This second law is a common source of difficulties in the reform of the electricity sector.
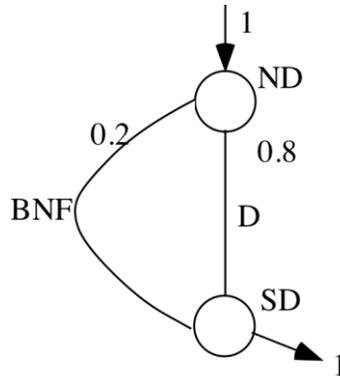

Figure 3: The illustration of second Kirchoff law

The implication of Kirchoff's second law is illustrated on Figure 3. Suppose a set of generators located at node ND and assume an important load centre located at node SD. Suppose that two lines BNF and D link the ND and SD nodes; suppose that line BNF offers four times as much resistance to power as D. The flow from A to B then distributes on the two lines in such a way that the flow on D is a four times as large as the one on BNF. Kirchoff laws are expressed mathematically by the so-called load flow equations. Roughly speaking, resistance is replaced by impedance; power decomposes into its "active" and "reactive" components; the former is directly consumed, the latter is necessary to insure the transmission of active power by controlling voltage. The active power is the one we are accustomed to; the reactive power is more difficult to interpret and is commonly, but by no means always (see FERC 2005), absent from discussions on the reform of the power sector. It is indeed common to limit the discussion to a linear approximation of the relations describing the flow of active power. This approximation is referred to as the DC load flow approximation. It is amenable to a representation of the grid through coefficients known as Power Transmission Distribution Factors (PTDF). The FBMC proposal adopts the PTDF representation. We illustrate its principle on figure 4 (taken from Chao and Peck 1998).

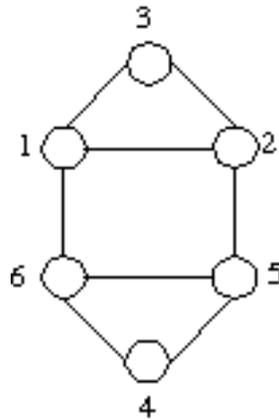### 4.1. The node to line PTDF as a description of the grid

Figure 4:

Figure 4 depicts a network consisting of six nodes and 8 lines. The three northern (1, 2, 3) nodes belong to a Northern Zone or zone I; the three southern nodes (4, 5, 6) belong to a Southern Zone or zone II. Lines 1-6 and 2-5 interconnect the Northern and Southern Zones. Lines are characterized by "admittance" (the inverse of the impedance) whose values are inserted as data in the load flow equations. Specifically and abusing language, Chao and Peck's example supposes that all lines have unit admittance (in the DC load flow approximation) except for the two tie lines 1-6 and 2-5 linking the Northern and Southern zones whose admittance is 2. The DC load flow approximation neglects reactive power and linearizes the non-linear relations describing the flow of active power around values expected to prevail during the relevant time period (e.g. hour 10 tomorrow). These linear relations are thus valid for flows that do not depart too much from these expected values. This linearization also amounts to neglecting losses.

Consider a set of injections and withdrawals in this lossless grid. Choose a particular node (node 6) and refer to it as the hub or the imbalance node. Flows in the lines as determined by the DC-load flow approximation satisfy the following properties.

(i)   They are additive in the sense that the flow on a certain line (e.g. line 1 to 6) is the sum of the flows generated on that line by each nodal injection or withdrawal.

(ii) The flow in a line induced by a unit nodal injection or withdrawal is given by a coefficient (the Power Transmission Distribution Factor or PTDF) that depends on the node and the line. The PTDFs of injection and withdrawal are equal in absolute value but have opposite sign.

Table 1 reports the values of the PTDFs. In order to illustrate these figures, consider the top-left element of the table (0.25 in cell (1; L12)). It is equal to the flow in line 1-2 arising from a unit injection at node 1 with withdrawal taking place at the hub at node 6.

	L12	L23	L31	L16	L25	L45	L56	L64
1	0.25000	-0.12500	-0.12500	0.62500	0.37500	-0.12500	0.25000	-0.12500
2	-0.33333	0.16667	0.16667	0.50000	0.50000	-0.16667	0.33333	-0.16667
3	-0.04167	-0.47917	0.52083	0.56250	0.43750	-0.14583	0.29167	-0.14583
4	-0.04167	0.02083	0.02083	0.06250	-0.06250	0.35417	0.29167	0.64583
5	-0.08333	0.04167	0.04167	0.12500	-0.12500	-0.29167	0.58333	-0.29167

Table 1: PTDF of the example network

The FBMC proposal adopts a description of the grid in terms of PTDFs. As just discussed their computation relies on some simplifications and their use may therefore require ex post adjustments. The DC-load flow approximation neglects reactive power and hence does not allow for pricing it. This is taken care of in another part of the market design namely through ancillary services (see FERC 2005 for a discussion of the pricing of reactive power). Similarly, the DC-load flow approximation also neglects losses, which are thus also treated in a separate market. The linearization of the non-linear load flow equations of active power is another simplification. This can be taken care of by re-computing the PTDFs after a significant flow change.

This flow-based representation of the grid is known in the US literature as the flowgate model: each line is a flowgate. Flowgates are defined for each pair of node (b) and line (l): the node b to line l PTDF gives the flow in the line l resulting from a unit injection at node b and withdrawal at the imbalance/hub node. The capacities of the lines complete the description of the grid, which therefore consists of the set of node to line PTDFs and line capacities. The flowgate model provides a good local representation of the grid provided some (phase angle) conditions necessary for the linearization of the load flow equations are satisfied. These latter conditions are generally fulfilled for the European grid.

4.2. A delicate step: from node to line PTDF to zone-to-line PTDF

The major divergence between the original flowgate model and the FBMC proposal lies in the aggregation of nodes into zones. The FBMC proposal indeed does not describe the grid in terms

of individual nodes and lines. It aggregates the nodes into zones (e.g. nodes 1, 2 and 3 aggregated into zone I) and replaces the set of node to line PTDFs (e.g. $PTDF_{1,1\text{-}6}$) by zone-to-line PTDFs (e.g. $PTDF_{I,(1\text{-}6)}$). A zone-to-line PTDF therefore gives the incremental flow in the line induced by an additional unit injection in the zone (withdrawal always takes place at the balancing node 6). This amounts to moving from the description of the initial grid given in Figure 4 to the aggregate version shown in Figure 5: the lines retain their individual character but the nodes disappear in the zones.
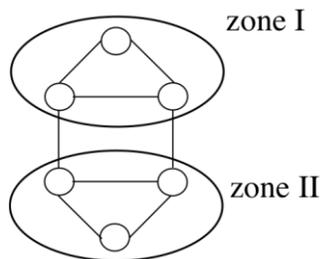

Figure 5: The FBMC representation of the grid

Note that the TC based representation of the grid goes beyond this aggregation and considers a representation of the network such as given in Figure 6 where the different lines linking the two zones are further aggregated into a single aggregate flowgate.
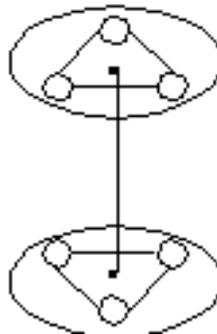

Figure 6: The TC representation of the Grid

It is well known that one cannot construct a coarse representation of the grid that is fully equivalent to the physical model where nodes and lines are individually identified. In other words, the grid of Figure 5 will never be equivalent to the one of Figure 4 and the grid of Figure 6 would give a still worse description of the reality. The construction of a reasonable aggregate network is obviously part of the problem of the implementation of the FBMC. We return to this

question in the next to last section but already note here that the FBMC proposal assumes that zones coincide with countries: all nodes of a country are aggregated into one zone. The main advance of the current FBMC proposal compared to the TCMC is thus that it does not aggregate lines and maintains an explicit representation of Kirchoff's second law by the use of PTDFs. Its main drawback is that it still aggregates nodes and by so doing, replaces node-to-line PTDFs by zone-to-line PTDFs. The relevant question is to assess the impact of the nodal aggregation. This is what we are turning to now.

4.3. Constructing a zone-to-line PTDF

The linearity properties of the flowgate model imply that the flow in some line accruing from a global incremental injection of 1 Mw in a zone is the sum, carried over all nodes of the zone, of the flows induced by the incremental injections at the individual nodes of the zone. It is obvious that the result depends on how the incremental injection of 1 MW decomposes into nodal injections in that zone. The FBMC proposal refers to these coefficients as Generation Shift Keys (abbreviated as GSK in the following) but does not elaborate much on their choice (see Section 3.2.4.3 of the proposal); it just explains that TSO will select the GSK on the basis of experience or particular rules. We illustrate the approach on figure 4 taken from Chao and Peck (1998). We interpret GSK in net generation, that is, after netting generation by withdrawals. The discussion can be readily transposed to the case without netting.

Consider the injection and withdrawals at the different nodes of the grid depicted on figure 4. Chao and Peck compute a perfect competition equilibrium on their network and arrive at injections respectively equal to 300 and 300 at nodes 1 and 2 and a withdrawal of 200 at node 3. The Northern Zone (zone I) therefore exports 400 to the Southern Zone (zone II). Suppose that these injections correspond to the observations used by TSO to construct the GSK. These are then equal to $\left(\dfrac{300}{400}, \dfrac{300}{400}, -\dfrac{200}{400}\right) = (0.75, 0.75, -0.5)$. Using the node-to-line PTDFs given in Table 1, one computes a Zone I to line (1–6) PTDF equal to

$$\mathrm{PTDF}_{I,(1-6)} = 0.75 \times 0.625 + 0.75 \times (0.500) - 0.5 \times 0.56250 = 0.5625.$$

Suppose a different generation pattern of 200 and 400 at node 1 and 2, still with a withdrawal of 200 at node 3 (Zone I still exports 400). The GSK become $(0.5, 1, -0.5)$ leading to a Zone I to line (1–6) PTDF equal to

$$\mathrm{PTDF}_{I,(1-6)} = 0.5 \times (0.625) + 1 \times (0.50) - 0.5 \times 0.56250 = 0.53125.$$

Taking now (200, 200, -200) as the pattern of injection and withdrawals (that is an export of 200), the GSK become (1, 1, -1) and one obtains a Zone I to line (1-6) PTDF equal to

$$\mathrm{PTDF}_{I,(1-6)} = 1.0 \times (0.625) + 1 \times (0.50) - 1.0 \times 0.56250 = 0.5625.$$

One may argue that a variation in the evaluation of $\mathrm{PTDF}_{I,(1-6)}$ of 5% as a function of the GSK is not much. For the time being we limit ourselves to conclude that zone-to-line PTDFs are no longer figures that depend on the sole physical characteristics of the grid. They also depend on the assumptions about the allocation of the flows in the zone, that is, on the GSKs. The FBMC proposal says very little about these assumptions. We come back to the question in several places later in the text.

5. THE PHYSICAL REALITY UNDERLYING FBMC: FROM LINES TO CRITICAL INFRASTRUCTURES

Security of grid operations was a major issue in the days of the fully regulated industry. Its importance is no lower today and it is likely to dramatically increase as a result of the penetration of intermittent sources like wind power. The N-1 and N-k criterion are standard instruments for guaranteeing the security of the grid: the criteria require that the system survives to a set of ex ante selected defaults, whether these are defaults on lines, generation capacities or a combination of both. The flowgate model of the grid allows for an immediate insertion of these criteria and therefore offers a particularly attractive way to tackle grid security. As an example, figure 7 depicts two contingent states of the illustrative grid that it presents together with the normal status of the network. The application of the N-1 or N-k criterion in the flow-based approach requires computing the PTDFs of the system in the different contingent states.

5.1. PTDF in contingent states

The computation of PTDFs in contingent states can be performed whether for node to line or zone-to-line PTDFs. We begin with the node-to-line PTDFs. Table 2 reports the results for the two contingent states. The results are respectively given for failures of lines 1-6 and 2-5.
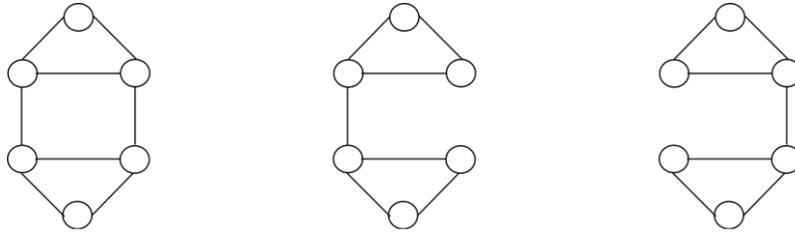
Figure 7: Normal and contingent states

	L12	L23	L31	L25	L45	L56	L64
1	0.66667	-0.33333	-0.33333	1.00000	-0.33333	0.66667	-0.33333
2	0.00000	0.00000	0.00000	1.00000	-0.33333	0.66667	-0.33333
3	0.33333	-0.66667	0.33333	1.00000	-0.33333	0.66667	-0.33333
4	0.00000	0.00000	0.00000	0.00000	0.33333	0.33333	-0.66667
5	0.00000	0.00000	0.00000	0.00000	-0.33333	0.66667	-0.33333

	L12	L23	L31	L16	L45	L56	L64
1	0.00000	0.00000	0.00000	1.00000	0.00000	0.00000	0.00000
2	-0.66667	0.33333	0.33333	1.00000	0.00000	0.00000	0.00000
3	-0.33333	-0.33333	0.66667	1.00000	0.00000	0.00000	0.00000
4	0.00000	0.00000	0.00000	0.00000	0.33333	0.33333	-0.66667
5	0.00000	0.00000	0.00000	0.00000	-0.33333	0.66667	-0.33333

Table 2 : PTDF of the example network in contingent states

We now compute the zone-to-line PTDFs when line (2-5) is down, and perform the same exercise as above with injection/withdrawals equal to (300, 300, -200), or GSK equal to $(0.75, 0.75, -0.5)$. The Zone I to line (1–6) PTDF is equal to

$$0.75 \times 1 + 0.75 \times 1 - 0.5 \times 1 = 1.0$$

One easily verifies that the result is identical when the generation shift keys are (0.5,1,-0.5). One can naturally object that the injections and withdrawals used in the computation of GSKs need to be adapted to reflect the capabilities of the network in case of a default. Specifically the pattern (300, 300, -200) implicitly implies an export of 400, which is not feasible when line (2-5) is down and the limits on lines (1-6) and (2-5) are both 200. One can easily verify that an injection and withdrawal pattern (200, 200, -200), with GSKs equal to (1, 1, -1) is export feasible. It will also give a zone-to-line PTDF equal to

$$1.0 \times 1 + 1.0 \times 1 - 1.0 \times 1 = 1.$$

In fact all feasible injection withdrawal patterns that saturate the remaining interconnection of capacity 200 give a zone-to-line (1-6) PTDF equal to 1 in this case.

5.2. Critical infrastructures

Knowing the PTDFs, whether bus to line or zone-to-line, it is possible to simulate the flow in a line implied by a given injection and consumption pattern. The relations for doing this immediately derive from the linearity properties of the PTDFs; they are given in the FBMC proposal (see Section 3.2.6) and need no further elaboration. The flow in each line is bounded by the capacity of the line. The PTDF based relations together with the line constraints on the flows in the normal and contingent states of the grid determine the feasible domain of injections and withdrawals. It is cumbersome and unnecessary to suppose that all lines could be overloaded in each contingency. The FBMC proposal therefore introduces the notion of "critical infrastructures"; these are equipments susceptible of being overloaded in normal and contingent states. Assuming that one can identify critical infrastructures and compute their PTDFs in advance, it is possible to limit the set of flow constraints and reduce the size of the problem to make it more manageable. The FBMC proposal leaves it to TSOs to determine the infrastructures that are "critical" in the normal state and relevant contingencies. TSOs are certainly best equipped to ex ante determine (before solving the market clearing problem) the lines that could become critical. We shall however argue that it is also reasonable to plan for a test that verifies ex post (after solving the market clearing problem) the validity of this ex ante choice.

5.3. Introducing a potential for progress

The notion of critical infrastructures is a direct by-product of the flow-based model of the grid. The introduction of critical infrastructures can induce a revision of all the reasoning on cross-border trade in Europe. The common wisdom indeed assumes that congestion in the grid mainly occurs on under-capacitated interconnections and that these limited infrastructures are the physical cause of restricted cross-border trade. The jurisprudence of European competition authorities illustrates this view: except for barriers to trade unrelated to this discussion, competition authorities essentially reason on insufficient interconnection transmission capacities to define the relevant geographic market in electricity. The experience of reformed electricity systems reveals that this view is misleading; congestion also occurs within domestic borders where it is often treated by countertrading (redispatching). It is also sometimes handled by artificially restricting declared transmission capacities, which means that limited interconnection capacities may simply express limited domestic capacities. A recent study conducted in ETH Zürich and in the Swiss TSO (Duthaler et al. 2008) confirms for UCTE what has been observed in many other systems: domestic congestions in contingent states are the real constraints on the system and they lead TSOs to artificially reduce transmission capacities. The mandatory Congestion Management guideline of Regulation 1228/2003 seems to be the only official EU document that today recognizes the possibility of restrictive practices of TSOs to solve their domestic problems and admits that cross border trade is not only a matter of interconnection capacities. The inclusion of critical infrastructures that possibly include domestic infrastructures, is thus of the essence for further progress. FBMC allows for this inclusion.

In order to understand the possible role of domestic infrastructures in limiting cross-border trade, consider the example of Figure 4 and the contingencies depicted on Figure 7. The observation of the normal state suggests that lines 1-6 and 2-5 are the sole critical infrastructures. A computation with the real data of the grid is necessary to prove this statement; this is very easily done here and it confirms that lines 1-6 and 2-5 are congested in Chao and Peck six nodes example when there is no line failure. However, the same calculation also reveals that other lines could easily become critical. The PTDFs from nodes 1 and 2 to line (1-2) shown in table 1 are respectively 0.25 and -0.33 in the normal state. Table 2 reveals that they are equal to plus or minus 0.66 when either line (1-6) or (2-5) is down. The flow in line (1-2) accruing from injections in 1 or 2 is thus twice as high when line (1-2) or (2-5) is down. Line (1-2) is thus indeed susceptible of becoming a critical infrastructure because of the need to account for contingencies. This conclusion on line (1-2) results from an ex ante computation performed for guaranteeing the security of the network. It will rarely be an ex post observation of the flow on the line because this line will only be saturated in the rather infrequent case of a contingency on

one of the interconnectors. But this possible saturation of line (1-2) is likely to be reflected in the TC between the two zones: limiting the flow on the interconnection reduces the total export from zone I to II and hence the flow on line 1-2 during contingencies. This in turn reduces the risk of congestion on that line in this contingent case. This mitigates domestic congestion problem and reduces countertrading costs.

The practice of "moving" domestic congestion to the border is in principle forbidden by the mandatory guidelines attached to Regulation 1228/2003. Experience reveals that the practice is real even in reputedly transparent and efficient systems such as Nordel. The already mentioned Duthaler et al. (2008) study confirms that domestic congestion plays an important role in the definition of the TC. The practice is rooted in changes of the PTDFs of domestic lines due to contingencies, a phenomenon that TSOs understand perfectly well but may appear complicated for external observers. The reasoning that leads to conclude that interconnector capacities limit the possibilities of cross-border trade may thus be intuitively attractive on first sight; it can turn out to be completely erroneous when TSOs artificially reduce TCs in order to remove potential congestions on domestic lines in contingent situations.

The Flow Based representation of the grid does away with transmission capacities and hence restores a correct vision of the congestion on individual lines. A possible congestion of line (1-2) (if this line has been declared a critical infrastructure) is now attributed to line (1-2) and not to an insufficient capacity of the interconnector. This reasoning also goes a long way toward correcting the misleading view of a grid consisting of national copper plates (low impedance infrastructure) connected by thin copper wires (high impedance infrastructure). The result is that all lines, domestic or cross border, can be congested. The introduction of critical infrastructures therefore completes a better representation of the grid through PTDFs with a better understanding of the restrictions to cross-border trade. This significant progress should be tempered by a word of caution: the FBMC proposal leaves to TSOs the possibility to manipulate line capacities by including the effect of remedial actions in their computation (section 3.2.2.1 "maximum allowable flow (Fmax)" in the FBMC proposal). This opaque statement signals that the reported capacities may not be the true capacities.

5.4. Limiting the potential for progress?

The extreme zone description of the grid adopted in the FBMC proposal is its first major drawback. Suppose, as argued above, that the TSO has found that domestic line 1-2 in zone I is critical and intends to take it into account for determining admissible injections and withdrawals.

Both nodes 1 and 2 belong to the same zone I which does not distinguish PTDFs from nodes 1 or 2: the question is thus whether combining nodes 1 and 2 in the same zone-to-line PTDF can distort the use of the critical infrastructures to construct the domain of admissible injections and withdrawals. If so, the zone aggregation would have reintroduced some of the inadequate representation of the network that characterizes TCs.

In order to see this, we now compute $PTDF_{I,(1-6)}$ when line 2-5 is down, using GSKs derived from flows that are compatible with the remaining export capacities of the grid. We noted that (1.0, 1,0, -1,0) was such a vector of GSKs; one can verify that (0.0, 2.0, -1.0) is also export admissible. The corresponding $PTDF_{I,(1-6)}$ obtain as follows:

$$(0.0, 2.0, -1.0): \quad 0.0 \times 0.0 + 2.0 \times (-0.66667) - 1.0 \times (-0.33333) = -1.0$$
$$(1.0, 1.0, \_1.0): \quad 1.0 \times 0.0 + 1.0 \times (-0.66667) - 1.0 \times (-0.3333) = -0.333333$$

These PTDFs therefore differ by 300% depending on the used GSKs. Such differences imply divergent approximations of the load flow model; they are not the result of physical characteristics of the infrastructure but are entirely due to the GSKs (that is, to assumptions), used for constructing the zone-to-line PTDF in the contingent states. One can submit that selecting GSKs in case of line failures is more difficult than in the normal state: the arbitrariness of the zone-to-line PTDF will thus be greater in contingent states.

A more intriguing result is observed by comparing zone-to-line PTDFs in normal and contingent states. We keep the original GSK vector (0.75, 0.75, -0.5), which is compatible with a 400 export from zone I in the normal state, and use the GSK vector (0.0, 2.0, -1.0) that is export compatible in case of defaulted line 2-5. One obtains for the two cases:

Default on line 2-5; (0.0, 2.0, -1.0): $0.00 \times 0.0 + 2.0 \times (-0.66667) - 1.0 \times (-0.33333) = -1.0$

Normal state; (0.75, 0.75, -0.5) : $0.75 \times 0.25 + 0.75 \times (-0.3333) - 0.5 \times (-0.04167) = -0.04167$

$PTDF_{I,(1-2)}$ is 25 times higher when line 2-5 defaults and the GSK vector is (0.0, 2.0, -1.0). In contrast we have seen that the comparison of the node-to-line PTDFs only reveals a doubling of $PTDF_{I,(1-2)}$ between the normal and contingent states. It is obvious that a PTDF that is 25 times higher on a domestic critical infrastructure in case of a contingency increases the chance of the TSO "finding" a congestion and hence drastically reduces the set of possible injections and withdrawals in zone I for security reason. Because domestic transactions may already require a significant fraction of the capacity of line 1-2 if $PTDF_{I,(1-2)}$ is 1, they will leave no room for international transactions if one wants to guarantee the integrity of the grid. This appealing

reasoning is erroneous though: the reality is that a $PTDF_{I,(1-2)}$ of 1 is fictitious and does not reflect what happens in the grid. The zone-to-line PTDF expresses the variation of the flow in the line when injections and withdrawals move proportionally to the GSK vector, here (0.0, 2.0, -1.0). But injection and withdrawals never move proportionally to GSK when clearing the market. The use of the GSK restricts trade compared to the real possibilities of the grid. We are back to artificial limitations on trade similar to those induced by TCs.

6. THE COMMERCIAL REALITY OF FLOW BASED MARKET COUPLING: THE TRADING PLATFORM

The flowgate/critical infrastructure model of the grid can be used to construct a trading platform. The European view distinguishes trade inside a zone (intra-zone trading) and between zones (inter-zone trading). A zone is a set of nodes where generators and consumers are located. Intra-zone trade takes place between the nodes of a single zone; inter-zone trade takes place between nodes of different zones. A Power exchange (PX) can be in charge of one or several zones as in Nordpool. More generally one could imagine that large countries such as France or Germany be split into several zones for network reasons even though there would only be one Power Exchange in each country. A PX clears the zone markets that it is in charge of and finds a zone price for each of them. This means that all the nodes of a zone trade at the same zone price. This implies that a PX could in principle clear the market at a different price in each zone. This perspective is economically quite rational and is in place in several regions of Europe (e.g. Nordpool and the Italian market) but it is not the common wisdom today in CWE, even if some are willing to accept it. We shall again invoke the trilateral market operating between Belgium, France and the Netherlands to explain these different ideas.

6.1. Case 1: TCMC and FBMC are identical

Let B, NL and F be the zones of the trilateral market (Figure 2). The system operates on the radial network depicted on figure 2. Let NL-B and B-F be the lines of that network. The network of figure 2 is a heroic simplification of the reality but a quite convenient support for the discussion: the transmission capacities between the zones coincide with the flowgates of the network. Taking F as the imbalance bus, the PTDF of flowgates NL-B and B-F are equal to 1. TCMC and FBMC are thus identical.

There is currently one PX in each zone of the trilateral market. Assuming one zone per Power Exchange simplifies the discussion but is in no way necessary. The current organization of the trilateral market is that a zone only trades through a single Power Exchange (there could be several PXs in competition in a zone). A PX receives all orders (bids and offers) of a zone. Assuming some net export (export minus import) from the zone, the PX can in principle clear its intra-zone market (see the discussion of bloc orders in Section 8.3 for the explanation of "in principle") at some zone price. Suppose this is done for different assumptions of the net export, one obtains a zone price for each net export: this is akin to an net export/import demand curve also called Next Export Curve (NEC) that links the zone price to the net export from the zone. Suppose that the NEC of the three PX are known, the TCMC operating between B, F and NL clears the inter-zone market taking TC constraints into account. This operation simultaneously finds the energy and TC prices and thus amounts to implicitly auctioning the TC. TC and flowgates are identical in the trilateral market because PTDFs are equal to 1. The trilateral TCMC is thus also a trilateral FBMC where one implicitly auctions the capacities of the critical infrastructures. This finding is only a particular case of a more general property: TCMC and FBMC are identical whenever the network is radial.

It is necessary before turning to the case of a meshed network where TCMC and FBMC are not identical to mention that the reasoning does not require all transactions to go through the PXs. Most of the trade today in Europe consists of bilateral deals concluded either directly or through the OTC market. They intervene as follows. Intra-zone bilateral deals completely bypass the PX. Inter-zone day-ahead deals go through the PXs of the relevant zones to get access to cross-border transmission. Energy in these deals is typically offered at price zero in the PX of the originating zone and bidded at a high price (e.g. the price cap) in the PX of the terminating zone. At clearing, the bilateral transaction receives the price of the originating zone from the PX and pays the price of the terminating zone to the other PX. The difference between these zone prices is the transmission price between the two zones. This transmission price is an outcome of the market coupling and is independent of the energy price agreed upon by the partners to the bilateral transaction. The trade through the PX has no impact on bilateral transactions other than this transmission price.

6.2. Case 2: TBMC and FBMC differ

Consider now the case of the network of figure 4 where PTDFs are no longer equal to 1 and hence TCMC and FBMC are not equivalent. An inter-zone transaction now requires a bundle of

transmission rights on critical infrastructures. One no longer allocates TCs but portfolios of transmission rights on line capacities. The PTDFs of the critical infrastructures in the different contingencies determine the composition of this portfolio. As discussed before, the value of these PTDFs depend on the GSKs. Specifically the amount of transmission rights on line 1-6 required by a unit transaction originating in node 1 of zone I and terminating in the hub is 0.04 in the normal state when the GSKs are (0.75, 0.75, -0.5) and 1.0 when line 2-5 is down and the GSKs are (0.0, 2.0, -1.0). The transaction will thus require 25 as many transmission rights on line 1-2 in the contingency state. This will undoubtedly limit the chances of this transaction getting access to the grid. This is especially true if one accepts cross border transactions only after accommodating domestic transactions as suggested by the discussion of section 3.2.8 of the FBMC proposal. The interesting point is that this large request for transmission rights on critical infrastructure 1-2 and hence the resulting limitation of access to the grid is artificial in the sense that it does not result from physical considerations. The contingent state that led to multiplying the demand for access by 25 would only have implied a multiplication of the PTDF by 2 if the FBMC had been based on a node-to-line description of the grid. The multiplication of the required transmission rights by a factor 25 originates from the use of zone-to-line PTDFs and the choice of the GSKs.

The example reveals that demands for transmission rights on an intra-zone critical infrastructure (line 1-2 is part of zone I) depend on the state of the network and can be quite different. This is justified by the physics of the grid. The example also shows that the recourse to zone-to-line PTDFs can exacerbate these differences and hence drastically reduce allowable transactions, because TSOs must plan for the worst case. Neither physics nor economics offer any justification to those increased differences. Part of the outcome of this worst-case calculation is artificial. It depends on the assumed GSKs in these different states of the grid. The common practice is to reduce intra-zone congestion by countertrading which is a costly operation. It is thus expected that a TSO will try to mitigate the impact of this factor of 25 either by selecting a different GSK or by countertrading. Both are quite possible and the FBMC proposal (section 3.2.2.1) explicitly foresees that line capacities can be adjusted to account for remedial actions. None of this is really documented further and it is thus impossible to really judge what is effectively proposed.

Intra-zone transactions also raise additional questions of transparency and possible complications for FBMC when the grid is meshed and hence TCMC and FBMC differ. Intra-zone transactions bypass FBMC but they still physically go through the grid. This means that they use

the critical infrastructures of the grid, not only those of the zone where they originate and terminate but also those of the other zones of the network. Economic rationality imposes that the use of lines, whether domestic, non domestic, or interconnectors, due to intra-zone transactions should either be charged somehow through the FBMC system or taken care of by countertrading. The latter proposal is infeasible: countertrading can relieve congestion but cannot fully eliminate the impact of loop flows; specifically countertrading in a zone will never eliminate the use of non-domestic lines by loop flows. The only economic rational is thus to charge intra-zone transactions for their use of domestic and non-domestic lines through the FBMC. The current FBMC proposal is also silent on the issue.

7. SPATIAL AND TEMPORAL ARBITRAGE

The above discussion applies whatever the time stage of the transaction, whether day-ahead, intra-day or real time. The description of the grid in terms of PTDFs and flowgates or critical infrastructures is static in the sense that it does not involve any dynamic property of the network (like stability). Except for the possible intrusion of countertrading measures in the computation of line characteristics, node-to-line PTDFs are entirely determined by the state of the grid (flow and topology) at any moment of time. This state evolves through time but these changes can be observed or deducted from measurements on the grid. It is thus possible to construct the node-to-line PTDF representation of the grid at any moment of time. The experience of US systems functioning according to the SMD paradigm shows that one can re-compute the PTDFs every five minutes. With this time scale, the PTDF description of the grid can thus also be used for day-ahead, intra-day or real-time trading.

7.1. Zone–to-line PTDFs may create difficulties

A FBMC trading platform that continuously re-computes node-to-line PTDFs as one moves from day-ahead to real-time is in line with the continuous trading model of finance that the reorganized power industry is trying so much to emulate. Unexpected flows or modifications of equipment status that require changing the topology of the grid are events that induce agents to adapt their energy and transmission positions. Continuously re-computing node-to-line PTDFs also permits tracking the capabilities of the grid with good accuracy. Flow Based Market Coupling, at least in the sense of an implementation in terms of node-to-line PTDF therefore offers a platform that applies from day-ahead to real-time provided one installs the adequate communication facilities to keep track of energy positions and grid status.

The situation is different with zone-to-line PTDFs. Zone-to-line PTDFs may not represent the real capabilities of the grid because the flows in the zone after market clearing do not necessarily match the GSKs assumed in their computation. TSOs can provide best estimate GSKs based on past observations of injections and withdrawals (these assumptions must also be common to all TSOs). In contrast, a well functioning smart market requires GSKs based on marginal injections and withdrawals taking place during the market clearing process. The difference is obvious in theory: the computation of zone-to-line PTDFs requires solving what mathematical jargon calls a "fixed point problem"; one needs the solution of the true (node-to-line) problem to compute the GSKs used to set up the approximate (zone-to-line) problem. The

construction of zone-to-line PTDFs in the FBMC proposal effectively supposes that TSO can reasonably guess in day-ahead the flows that will come out from market clearing in real-time. The common wisdom underlying the proposal is that this is possible because injection and withdrawals after market clearing should not vary too much from working day to working day or from weekend day from the same day in the previous weekend. If so, then the zone-to-line PTDFs probably offer a good representation of the grid. There is obviously no theoretical reason why this would be true; section 3.2.8 of the FBMC proposal hints that there may indeed be difficulties. Zone to line PTDFs may thus create difficulties in the day-ahead market.

The same difficulties appear in intra-day and real-time where the situation might still be more difficult to assess. These markets respond to random events and hence may have a less predictable character. One can certainly re-compute the zone-to-line PTDFs to account for observed positions as one proceed through the trading process; this allows for a good tracking of the current utilisation of the grid. But this does not give any hint on the changes in the utilisation of the grid resulting from changes of these positions. The difference between currently observed GSKs and forward GSKs is unimportant when the residual line capacities are large and there is ample room for additional flows. In contrast this difference is essential when the grid is intensely used and it is not clear whether a new transaction can be accepted or not. The adoption of node-to-line PTDFs is the only way to surely remove this drawback. Only experience will tell whether observed injections and withdrawals provide a good basis for setting GSKs in a way that permits a good simulation of the incremental flows in an already extensively used grid. The only message that can be sent today is that the explanation of possible difficulties in the implementation of FBMC should first be searched in the computation of zone-to-line PTDFs. These difficulties are likely to be more apparent in large countries such as France and Germany where the selection of GSKs extending over the whole territory is most demanding.

7.2. Interactions between Power Exchanges (PX) and Transmission System Operators (TSO) may also create difficulties

The trading platform is meant to facilitate the spatial and temporal arbitrage of electricity. Spatial arbitrage through market coupling, whether flow or transmission capacity based, involves energy and transmission services. It therefore requires the participation of PXs (for energy) and TSOs (for transmission services). Their interplay can take different forms that can raise governance difficulties. We mention in passing the possible conflicts arising from coordinating regulated (TSO) and unregulated (PX) entities. We instead focus here on the organisational implications of the technical aspects of FBMC. Some of them maybe directly linked to the distinction between the node-to-line and zone-to-line view of the Flow Based model.

7.2.1. From SMD: comparing with the integrated ISO/RTO

It is worth recalling at the outset that PX and SO are merged into a single entity (the ISO or RTO) in the SMD paradigm presented in the introduction. One can also recall that the PX and TSO were separated in the initial Californian market where the PX ignored, by design, the location of injections and withdrawals that the TSO had to later accommodate. The ISO or RTO of the SMD has essentially no stake in physical generation or transmission assets: its role is to run a complex auction and insure that all the necessary information for doing so is properly assembled. The ISO/RTO of the SMD runs a nodal system, which in a European context would mean the extreme case of a zone system: each node is a zone in the nodal system. The organizational divergence between the SMD's ISO/RTO running a nodal system and the pair of PX and TSO cooperating on a zone system in the EU is thus maximal.

A single RTO running a geographically extended nodal system also contrasts with the philosophy that led to the TCMC of the Belgian, Dutch and French markets, which is the ancestor of the foreseen FBMC. The acknowledged objective at the time was to retain as much as possible of the idiosyncratic aspects of each PX and TSO and to clear the market by organising iterations between these entities. The philosophy was also to retain the separation between PX and TSO (that prevailed in the initial Californian design but for quite different reasons) and to extend that separation of the trading and network functions to several PXs and TSOs. This requires an algorithm that iterates between these entities. We explained before that PXs construct net export/import curves (NEC) that clear the intra-zone market without considering the intra zone congestion that this may imply and that FBMC clears the inter-zone market by auctioning transmission rights on critical infrastructures. This creates an

organisational inconsistency: the intra and inter-zone markets both use the critical infrastructures and hence may induce congestions that the TSOs must accommodate; in contrast the PXs and the OTC markets are allowed to overlook their contribution to congestions when clearing intra-zone transactions. The RTO avoids that inconsistency: it clears the intra and inter-zone markets simultaneously taking all contributions to congestions into account.

The philosophy of the FBMC is to clear, at the inter zone level, zone markets that have already cleared at the intra zone level. This requires net export curves (NEC). Each PX is best equipped to individually construct NEC as long as there is no intra-zone congestion (equivalently if there is no critical infrastructure in the zone!). The operation can thus be decentralized to the PXs. In contrast it should be clear that the auctioning of the critical infrastructure capacities (or of the TC in a TCMC) must be centralized. It should also be clear that the computation of nodal or zone PTDFs (as the computation of TCs) requires an overall view of the grid and hence can only be undertaken at a central level. The spatial arbitrage of electricity through a system of individual PXs and TSOs therefore logically requires some hierarchical organisation of the tasks at a centralized level. The experience of the SMD and the geographical development of system like PJM confirm this finding. Stakeholders to the FBMC venture apparently acknowledge the need for this centralization. This is good news, even if unexpected in the overall context of the Internal Electricity Market where stakeholders have so far favoured a horizontal approach that resisted any centralization. This decentralization has led to the heterogeneity of national systems that makes it difficult today to harmonize opening and closing hours of national PXs! It is essential to recognize at the outset that FBMC requires some hierarchy of processes that cannot be handled by the sole horizontal iterations between national entities (the same applies to TCMC). We shall come back to that question in the discussion of bloc orders.

7.2.2. Evolving relations between PXs and TSOs through trading phases

This organisational discussion applies equally to all trading stages: one needs to manage congestion of critical infrastructures in real-time and prepare for this task ahead of real-time. Except for geographically segmenting the market when approaching real-time, one needs a central organisation that simultaneously clears energy and transmission over the whole CWE, from day-ahead to real-time. This requirement is not widely perceived today in CWE let alone in the rest of Europe. One admittedly recognizes the need for PXs and the TSOs to cooperate in the day-ahead market. In contrast, intra-day trading is largely seen as a TSO business and real-time operations are entirely within the hands of TSO and not even meant as a market activity.

These different views on the role of the PXs and TSOs confirm that these activities are not seen as part of a single trading platform. This violates basic economic principles and undoubtedly distorts arbitrage. This will create difficulties tomorrow except if they are anticipated today by defining a common organisation of the roles of PXs and TSOs in CWE from day-ahead to real-time. This is the discussion that we turn to now.

Tables 4, 5 and 6 present three views of the relations between Power Exchanges and Transmission System Operators across the different stages of the trading platform.

	Day-ahead (market coupling)	Intra-day (trading)	Real-time (balancing)
PX	interaction between PX	X	X
	interaction between PX and TSO	X	X
TSO	interaction between TSO (transmission capacities)	interaction between TSO (transmission capacities)	interaction between TSO (transmission capacities)

Table 4: the current organization

	Day-ahead (market coupling)	Intra-day (trading)	Real-time (balancing)
PX	interaction between PX	?	X
	interaction between PX and TSO	?	X
TSO	interaction between TSO (flow based)	interaction between TSO (flow based ?)	interaction between TSO (?)

Table 5: current interrogations

	Day-ahead (market coupling)	Intra-day (trading)	Real-time (balancing)
PX	interaction between PX	Interactions between PX	Interactions between PX

	interaction between PX and TSO	Interactions between PX and TSO	Interactions between PX and TSO
TSO	interaction between TSO (flow based)	interaction between TSO (flow based)	interaction between TSO (flow based)

Table 6: The proposed interactions

Table 4 depicts the current organisation. PXs and TSOs interact in the day-ahead market but PXs are absent from intra-day trading and balancing. This changing role does not reflect a market where stakeholders request intra-day trading services for adapting positions inherited from day-ahead trading to market innovations. The disappearance of PXs when moving from day-ahead to intra-day introduces a discontinuity that will distort arbitrage between these two trading stages. Without even invoking arbitrage, these different organizations will lead PXs and TSOs to organize day-ahead and intra-day trading differently, something that will create barriers, which competition authorities are likely to object to in the future.

Table 5 reflects the uncertain position of the current proposal: PXs and TSOs must cooperate to organize the day-ahead market coupling while the real time operations are recognized as a pure TSO matter. In contrast, the relation between PXs and TSOs in intraday trading in unclear. Cegedel et al. (undated) suggests that TSOs are aware of the problem and wonder about their taking on some trading activity or the PXs moving into the intra-day business.

Table 6 depicts a proposal that is the technically logical extension of the proposed FBMC into a single platform encompassing day-ahead, intra-day and real-time trading. Table 6 reflects the many points common to day-ahead, intra-day and real-time operations. The same flowgate model of the grid applies through time and PTDFs and line capacities are updated in order to reflect market innovation. This unifying view raises a fundamental problem however: the TSOs need the node-to-line view of the grid, possibly at all stages of trading but in any case in real-time. The PXs want to retain a zone view in day-ahead and for that reason may be unwilling to move into intra-day, let alone to real-time. At some point (possibly already in day-ahead) one needs moving from one to the other. This is the Gordian knot of European trading of electricity; it is both technical and organizational: PX and TSO are today different organizations in charge of different activities. Each will want to shape the system according to what suits it best: PX like zones and Transmission Capacities between zones; TSO must work with nodes and lines and prefer to remain within their zones (control areas).

7.3. Conclusion

The above discussion provides the methodological basis for the rest of the report. We embrace the principle of FBMC in the day-ahead market as a significant potential advance with respect to the use of TCs but expect difficulties because of the zone view, especially when zones are assimilated to large countries. The advantages of the FBMC with respect to TCMC are clear: the flowgate model better accounts for the real possibilities of the grid and the recourse to critical infrastructures helps guaranteeing the security of transmission services. The drawbacks of the zone implementation of the FBMC are also clear. The PTDFs that are at the core of the FBMC depend on GSKs that maybe inadequate. Significant difficulties may also arise because of the treatment of congestion on critical infrastructures inside zones, a subject that the proposal does not touch. These difficulties suggest that some aspects of the implementation of FBMC should be revisited. Very much like "market based" can mean quite different things that are far from equivalent, flowgate and Flow Based market coupling are fundamentally ambiguous notions in a zone system where the zones can be large and electrically heterogeneous. Specifically the introduction of critical infrastructures inside zones is a valuable addition but it will create implementation difficulties for managing intra-zone congestion and finding zone prices compatible with the profit seeking objectives of generators participating to the market and with the remuneration of the units involved in redispatching. This is discussed further in Section 8.2

The flowgate model applies equally well to day-ahead, intra-day and real-time. It thus seems natural to extend FBMC from day-ahead to intra-day and real-time and hence provide a unique trading platform that encompasses these different stages. This implies substituting the idea of a punitive balancing by a real-time market. Besides complying with the principle of continuous trading in finance, this view would also go a long way towards eliminating an important barrier to trade repeatedly pointed out by competition authorities. Note also that applying the same trading platform from day-ahead to real-time would facilitate temporal arbitrage and hence probably improve market liquidity.

These proposals are likely to encounter considerable organisational resistance. A main reason is that they question the assimilation of zones to countries. Large zone-to-line PTDFs are unlikely to be suitable in day-ahead but are almost surely inadequate as one moves from day-ahead to intra-day and real-time trading. Note that competition authorities are also likely to (erroneously) invoke the number of price zones as a sign of a segmented market. We simply mention this latter problem but do not elaborate on it here.

There might be other, more organizational problems. The extension of the idea of market coupling implies deep interactions between regulated and unregulated entities. It is difficult to imagine that this will not eventually cause difficulties. This problem is only part of the more general question of harmonizing both the rules of the PXs and the regulations of TSOs. Different market rules applied by different entities can only complicate market clearing. The problem becomes still more complex if one realizes that harmonization is also lacking between intra-day and real-time trading. Last, it is almost certain, in view of past experience in the internal electricity market, that some will object, at some stage of the integration process, to the creation of a higher level of responsibility in charge of performing certain computations. It is important to convey the message that not all computations can be decentralised in a way that fits purely horizontal organisational structures. It is also important that crucial locational information be exchanged between the PXs and the TSOs something that does not appear from the current proposal. Last but by no means not least, the language of Regulation 1228/2003 and its use of TCs will certainly facilitate the task of those who oppose moving ahead.

8. DAY-AHEAD TRADING IN FLOW BASED MARKET COUPLING

Figure 5 depicts a general organisation of FBMC. PXs operate price zones: the principal difference between the figure and the FBMC proposal is that we suppose that a single PX can cover several zones (as in Nordpool). Energy must clear at a single price in each zone. A large country should, for technical reasons, ideally decompose into several price zones, except if the network is sufficiently capacitated. Note that the requirement for a sufficient capacity is not limited to the domestic network as can be deduced from figure 3. We explained in section 4 that a fourth of the flow from ND to SD goes through the line BNF. Suppose that this line is itself composed of different line segments and some of them are congested. Then ND and SD should be part of two different zones, whether the line D is congested or not. The decomposition in two zones around ND and SD is not due to an insufficient line D capacity: bottlenecks on the non-domestic line BNF are the real cause. The reason is that grid is a single system and infrastructure limitations somewhere have implications elsewhere. Notwithstanding these physical and economic evidences, the current proposal seems to insist on one zone per country. Note also that competition authorities also assimilate a small number of price zones to a competitive market: the fewer the zone prices the larger the geographic market (and hence the smaller the concentration). This reasoning is extremely questionable and probably erroneous

except in a system with excessive transmission capacities: the experience of competitive systems operating under SMD is indeed that different nodal prices (and hence obviously different zone prices) is the norm, not the exception. The reasoning of European competition authorities would lead to the conclusion that each node of PJM is a separate market!
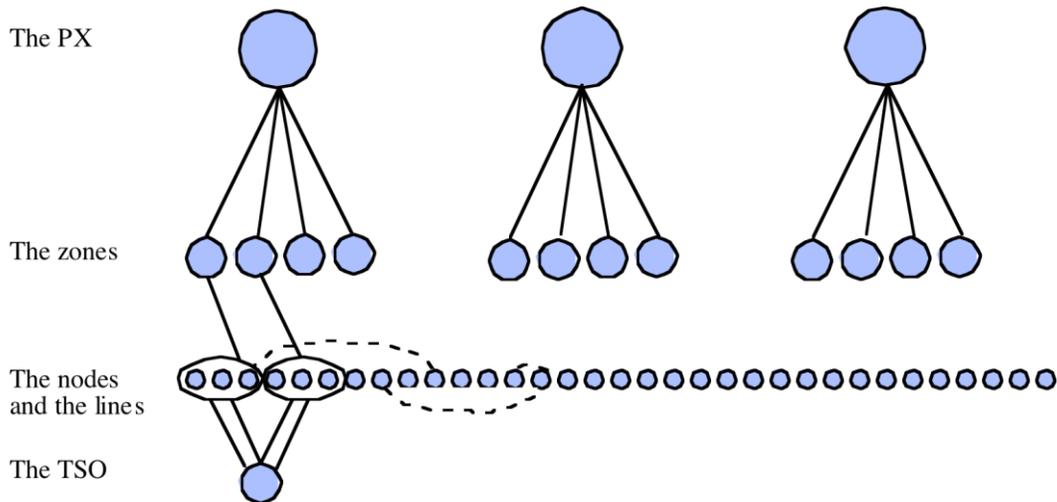

Figure 7

Price zones contain nodes that are linked by lines. PXs receive orders (offers and bids) from agents (generators, consumers and traders) located in a zone. The prices established in the PX can differ from those in bilateral transactions even though the non-arbitrage argument of finance suggests that they are likely to be closely related. PXs clear the market in individual price zone by running auctions, taking into account net exports from the zone (import from and export to other zones). Market clearing by a PX can normally be interpreted as the result of a consumer and producer surplus maximisation problem (CPSM) as depicted as in Figure 8. The FBMC proposal explicitly refers to that CPSM problem, which constitutes a pillar of the proposed architecture. Figure 8 is also classically interpreted as an equilibrium between supply (offer) and demand (bid) curves, with the equilibrium price at the intersection of both. The outcome of market clearing (for instance the intersection of the bid and offer curves at different hours) can mean that an order is only partially selected, possibly at a very low level in some hour. Market clearing is conducted under an assumption of net export from the zone. Changing this net export modifies the result of the clearing and hence the zone price. This gives a relation between net exports and the clearing price of the zone, which is in fact a net export/import demand curve (NEC). This curve normally shows a price that is increasing with the net export from the zone.

FBMC clears the inter-zone energy market formed by the set of PXs each represented by its NEC. This is obtained by another CPSM constructed over all the zones. The consumer and producer surplus maximization problem is constructed on the basis of the NECs of the PXs and subject to limitations in imports and exports imposed by the network. These limitations are established using the zone-to-line PTDFs and line capacities computed by TSOs. Bilateral transactions intervene in the process as follows; they bypass the PXs as long as they originate and end in a single price zone. Alternatively bilateral transactions between two pricing zones submit offers and bids to the PXs operating in the two price zones where they originate and terminate. The offer will typically be priced at zero and the bid at a very high price equal to cap. Bilateral transactions pay the difference of zone prices as transmission price. Long-term transmission rights also bypass the PX (but are known to the TSO). Their reserved transmission rights return to the coupling mechanism when they are not used.
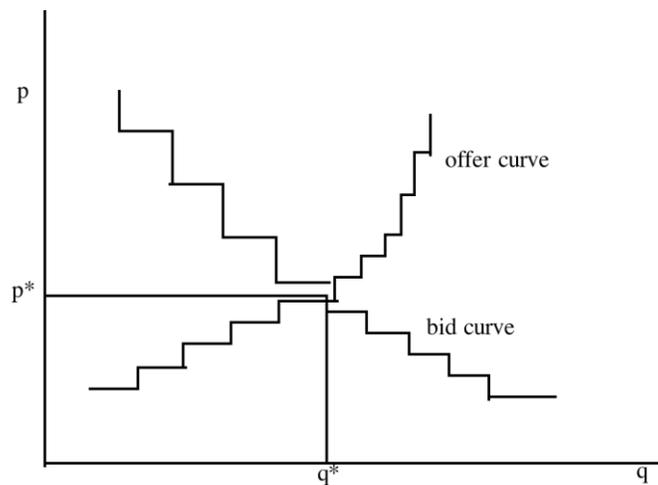

Figure 8: Producer and consumer surplus maximization

This organization raises different questions. We already extensively discussed the computation of the PTDFs and summarize this discussion in Section 8.2. Section 8.3 adds a few points to the question of critical infrastructures that was also presented before. Section 8.4 takes up the problem of bloc orders that has not been discussed so far. Before doing so we briefly introduce in section 8.1 a distinction between different degrees of market clearing that the CPSM problem might attempt to find. This distinction is briefly presented in the FBMC proposal.

8.1. Degrees of market coupling

The standard picture of market equilibrium and the associated classical Consumer and Producer Surplus Maximisation problem (CPSM) give a false impression of simplicity. The FBMC proposal refers to different degrees of market coupling that hide different degrees of complexity: market coupling can be "tight" or "loose"; it can also be "volume" or "price" based. We briefly review these notions.

By definition a tight market coupling accounts for the idiosyncrasies of the different PXs and inserts them all in the CPSM. A loose market coupling does not do so. It is presumed that that the loose market coupling requires some adaptation of the solution of the FBMC before its implementation by individual PXs. The principle is that volumes are accepted as such but prices need to be recomputed locally. It is also expected that this distinction would disappear with the harmonization of the PX that should hopefully intervene at some point.

The distinction between volume and price coupling appears methodologically much more important. By construction volume coupling occurs when only volumes determined centrally are transferred to the local PX. In contrast price coupling also requires that both the price and volume computed at the central level be implemented by the different PXs. The terms of reference of the FBMC proposal insist that the software be able to treat both volume and price coupling, as well as tight and loose coupling. We shall explain that this requirement is probably impossible to fulfil: volume coupling is always possible but may result in inconsistent prices. Price coupling, because it requires that prices be consistent, is sometimes impossible.

8.2. Computation of the PTDFs

We briefly recall and summarize our discussion of PTDFs. A node-to-line PTDF measures the impact on a line flow of a modification of nodal injection with respect to given flows and for a given network topology. Node-to-line PTDFs offer an adequate representation of the grid provided they are recomputed sufficiently often to account for changing flows. PTDFs always need to be recomputed when the topology changes.

A zone-to-line PTDF measures the impact of a change of zone injection on a line. The definition requires a given nodal profile of the injection, that is, its allocation into nodal injections and withdrawals. The zone-to-line PTDF therefore depends on this allocation, which is referred to as generation shift keys or GSKs. The dependence of the zone-to-line PTDFs on GSKs becomes more crucial with the size of the zone. The zone-to-line PTDFs are used for both computing the

loading of a line for an expected injection and withdrawal pattern, and the modification of that loading for variations of this injection and withdrawal pattern.

The FBMC proposal is based on zone-to-line PTDFs. We explained in Section 4.3 that it might be difficult to select appropriate GSKs: it is indeed not clear that GSKs should be the same for normal and contingent states of the grid and it is probably impossible to ex ante determine good GSKs for contingent states of the network. We also explained that the same GSKs are unlikely to apply through the different stages of trading (day-ahead, intra day and real-time). The recourse to GSKs is a key weakness of the proposal.

8.3. Contingencies and critical infrastructures

Both FBMC and TCMC clear the intra-zone market by solving a Consumer and Producer Surplus Maximization problem (CPSM). The objective function of the problem is identical in both models and is constructed from the NEC elaborated by the PXs. FBMC and TCMC differ by their representation of the network. TCMC requires that net imports balance and be TC feasible. FBMC imposes that net exports balance and do not overload the critical infrastructures. We already explained that zone-to-line PTDFs only give a good representation of the network under some conditions. One is easily met: PTDFs need to be recomputed sufficiently often so that node-to-line PTDFs give a good approximation of the load-flow equations. Others maybe more difficult to satisfy: GSKs need to correspond to the injections and withdrawals obtained from the market clearing process. Theory says that this requires to have already solved the market-clearing problem with the node-to-line PTDFs (in mathematical jargon, finding the GSKs is a fixed point problem). Practice may eventually tell that the problem is easier; it might also conclude to the opposite.

Suppose that these two conditions necessary for the validity of the PTDFs are satisfied. By construction, the CPSM will conclude on secured transmission transactions if the critical infrastructures have been properly selected. This implies that the lines not included in the set of critical infrastructures are not overloaded at the solution of the CPSM. The FBMC proposal indicates that individual TSOs ex-ante select the set of critical infrastructures included in the CPSM. Because of the numerous assumptions embedded in the zone-to-line PTDFs, we suggest to also introduce an ex post check of the loading of the lines, computed using node-to-line PTDFs (not zone-to-line PTDFs). This check is done after the first solution of the CPSM

problem. The CPSM can then be updated to include additional critical infrastructures if this is necessary. This ex-post test is easy to implement. We now turn to more complex issues.

8.3.1. Computational issues

The CPSM associated to the FBMC typically includes many more constraints than for the TCMC. The FBMC proposal explains that the number of critical infrastructures is of the order of 3000 before any pre-processing to reduce this number. Solving a CPSM problem with 3000 linear inequality constraints is normally not difficult for existing optimization software. The ex-post test on critical infrastructures mentioned above may increase their number but this would not create any computational difficulty as long as one remains within these orders of magnitude. In short the number of critical infrastructures is not a computational issue.

The extreme form of the zone system in the FBMC proposal may however make the CPSM model meaningless and its solution irrelevant. With four PXs and one zone per PX, the inter-zone CPSM has four variables (the net export) and a number of inequality constraints of the order of 3000. This problem is likely to be infeasible! Section 3.2.8 of the FBMC proposal suggests that this indeed already happens from time to time. The origin of the infeasibility can probably be found in the GSKs: flows that are feasible in the real network (that is with node to line PTDF) can turn infeasible in the model if the zone-to-line PTDFs are inadequately computed (other more worrying explanations like violations of security constraints in the real world can also be given). The proposed methodology assumes that the nodal injections and withdrawal patterns are proportional to the zone net import. This assumption is fundamentally unrealistic and unnecessarily constraints the CPSM. It also violates the logic of market clearing which is to find injections and withdrawals that correspond to the retained bids and offers, without supposing their pattern (the GSKs) given in advance. More fundamentally, this assumption may lead to erroneously conclude to the failure of the FBMC. This is what we referred to in Section 5 when explaining that the current FBMC proposal assumes GSKs that strictly speaking can only be determined appropriately by first solving the nodal FBMC. It is clear that the method finds neither volume coupling nor price coupling when the CPSM is infeasible. Supposing no structural generation and transmission deficiency, one can always handle volume coupling by returning to node-to-line PTDFs and the associated CPSM. As we now discuss, this solution will not necessarily be price coupling.

8.3.2. Pricing issues

We explained that the inception of intra-zone critical infrastructures in FBMC is potentially a major advance that permits identifying the real bottlenecks of the grid. But critical infrastructures raise pricing issues that are the direct consequences of the particular implementation of a zone system in the FBMC. The standard approach of zone systems is to resort to (intra-zone) countertrading or market (zone) splitting to eliminate intra-zone congestion. The economic logic of Flow Based systems is that energy prices differ on the two sides of a congested infrastructure. This holds whether the congestion effectively occurs in normal states or only potentially during contingencies. In other words, congested intra-zone infrastructures normally imply a zone splitting, that is, two different zone prices on the two sides of the congested infrastructure. The FBMC proposal retains a single zone and a single zone price and hence violates this principle. The standard alternative to market or zone spitting is to eliminate the congestion on the critical infrastructure by countertrading. But this implies that the critical infrastructure disappears from the constraints of the CPSM as a result of the countertrading activity. This modifies both the set of selected orders (the volume in volume coupling) and the transmission prices (and hence the PX prices in price coupling). The proposed FBMC also violates this principle since it retains the critical infrastructure as constraint. The pricing methodology described in the FBMC proposal is thus either inconsistent or at least inadequately documented when it comes to congested intra-zone critical infrastructures. Also methods for finding a single price on the two sides of a congested infrastructure increase the chance that this price might not clear the market: agents receiving this price may have an incentive to deviate from the orders (volumes) accepted by the CPSM. In the language of the proposal, the solution is not price-coupling. Economists refer to this as a problem of incentive compatibility. This situation is briefly alluded to in the preceding section. We shall take it up further when discussing bloc orders.

8.3.3. Governance issues

The FBMC proposal indicates that individual TSOs will select the list of critical infrastructures. TSOs obviously have the experience of their grid and hence are best qualified to do so. But the history of recent blackouts shows that TSO can also be wrong in their ex-ante evaluations. The ex-post tests proposed in the beginning of this section go beyond the sole control by individual TSOs. These tests are not very demanding and should not create serious organizational difficulties.

The computation of PTDFs demands a global view of the grid that extends beyond the jurisdiction of individual TSOs. This requires an entity that supersedes individual TSOs. In the same way, one needs a central entity to perform the CPSM subject to the constraints on critical

infrastructures. It seems that the idea of this central entity is accepted at the level of the day-ahead market. This should thus also not create too many organizational difficulties.

The main bottleneck will appear if the treatment of critical infrastructures in the zone system make it impossible to couple markets (volume and pricing), a possibility that the terms of reference of the proposal do not envisage The introduction of critical infrastructure is a key step forward but only a finer description of the grid can solve the volume coupling and mitigate the price coupling difficulties that the zone representation may create. Stakeholders will likely resist this solution and instead propose to reduce the role of critical infrastructures. This would be a major setback. The remaining alternative is to conduct an in depth investigation of a formal insertion of the recourse to countertrading to deal with intra zone congestion in the proposal. All these choices are fundamental and sometimes methodologically deep. The experience of the internal electricity market has shown that stakeholders have tended to shy away from difficult questions and decisions.

8.3.4. Conclusion

The use of critical infrastructures is a major advance that should improve the reliability of transmission services. But the adopted zone approach questions the correctness of the implementation. Here again, critical infrastructures modelled through node-to-line PTDF should not create any particular difficulty for the CPSM to select orders and achieve volume coupling. The FBMC proposal can easily be modified to do so. The price to pay for this modification is to accept that the CPSM conducted with node-to-line PTDFs does not give a single zone price. But the current FBMC proposal also suffers from this drawback: in contrast with what is stated in the terms of reference, the current approach cannot guarantee a single zone price (price coupling) that clears the market in case of intra-zone congestion except by resorting to zone splitting or countertrading. The FBMC proposal does not mention zone splitting and only very briefly alludes to remedial actions such as countertrading. Countertrading requires finding the right incentives for operators to keep machines available for that service. International experience shows that it is not always possible to achieve a single zone price by countertrading. The larger the zone is, the more serious the problem becomes. Only practice can tell whether there will be sufficient resources available for countertrading and how the service can be organized within FBMC. The whole model is at risk if a proper treatment of countertrading cannot be inserted in the proposal. It is advised to keep the relations between countertrading, incentive compatibility and zone-to-line PTDF in mind if difficulties arise in the implementation of the FBMC.

Decentralization of decision and implementation is a key concern in the internal electricity market. The construction of the Net Export Curves is the only element that can be decentralised to existing entities. The computation of the PTDFs, their inclusion in the CPMP and the ex post verification that all necessary critical infrastructures have been selected justify a central computation. This centralisation is limited to computation and data transfer. It does not involve any network asset beyond computers and communication equipment. It appears that a satisfactory agreement has been reached on the acceptance of this central organization, at least inside CWE.

8.4. Bloc orders: computational, pricing, and governance issues

Bloc orders probably constitute the technically most difficult part of the FBMC proposal. They raise both computational and pricing questions that may in turn induce organizational issues. Bloc orders, like critical infrastructures, require a centralized treatment. This seems well accepted today. They render the selection of orders in the CPMP computationally more intensive, but this obstacle can probably be overcome with current optimization tools. The real difficulty is that bloc orders, very much like intra-zone critical infrastructures may prevent finding zone incentive compatible prices of energy. As intra-zone congestion, bloc orders can make price coupling impossible. But in contrast with intra-zone congestion, moving to node-to-line PTDFs does not remove the difficulty. Incentive incompatible prices may later create organizational issues; this cannot be ascertained today but should be kept in mind. The critique applies to both the proposed FBMC and the existing TCMC.

Orders can be flexible or present a certain character of inflexibility in which case they are called bloc orders. Markets only consisting of "flexible" orders clear easily. As an example of a flexible order, take an offer of 1 MW in hour 10 to 11 at a price of 40 euros/Mwh. The offer will be entirely selected if the PX clears at a price higher than 40. It may be selected partially if this price is exactly 40. This offer is divisible if the PX can select a fraction of the offer, for instance 0.5 MW. Flexible bids and offers are easy to treat in the auction, their pricing is straightforward and they are quite amenable to the construction of the NEC that are the core of the trading platform whether in a FBMC or TCMC. The selection and pricing of flexible bids therefore fits quite well in a market coupling system (whether TCMC or FBMC) where PXs retain their individual identity. They are readily amenable to both volume and price coupling.

Bloc orders are transactions that involve some sort of inflexibility. The typical example is a sequence of offers that must be taken or rejected simultaneously. Referring to this case,

suppose an offer of 1 MW during hours 8 to 12 at a price of 40 euros/Mwh. This is a bloc order if accepting it implies taking the 1 MW during the four successive hours and remunerating the generator at a price that in average is no lower than 40 euro/Mwh. Another type of bloc order is an offer of 1 MW in hour 10 to 11 that, if accepted must be taken at a level of at least 0.5 MW. Bloc orders indirectly model machine constraints. They may cause computational and economic difficulties; their treatment requires a central operation that may induce organisational resistance, especially in case of pricing difficulties.

8.4.1. Computational difficulties

We explained before that market clearing in a single PX is normally interpreted as the solution of a standard consumer and producer surplus maximisation problem (CPSM). This is depicted in Figure 8. The figure also classically represents an equilibrium between supply (offer) and demand (bid) curves, with the equilibrium price at the intersection of both. The outcome of this market clearing process (for instance the intersection of the bid and offer curves at different hours) can imply that an order is only partially selected, possibly at a very low level in some hour. Alternatively, if two hourly auctions are run in sequence, the outcome may also imply that the same offer made for two successive hours is selected in the first hour but not the in the second one.

This does not cause any difficulty for "flexible" orders but would violate the constraints imposed by bloc orders. Bloc orders create computational difficulties in the sense that a standard auction process like the search for the intersection between the offer and bid curves is no longer suitable for maximising producer and consumer surplus and hence for solving the auction problem. The solution of the latter problem requires a much more complex process which involves either a heuristic, a special purpose algorithm (Lagrangian relaxation has been extensively used) or the recourse to (now standard and powerful) general purpose software solving a "mixed integer programming problem" (MIP). The FBMC proposal acknowledges the problem, even though the document is not clear about the exact solution that will be implemented.

8.4.2. Invoking the SMD

We indicated before that bloc orders are surrogates of machine constraints such as minimal technical output, minimal run or down time.... These constraints are obviously also present in

other thermal systems and in particular in those operated under the SMD. These systems bypass bloc orders and account for machine constraints by introducing special orders directed at these technical characteristics (e.g. offering a certain remuneration for starting up a machine). This dual approach (bloc orders vs. machine constraints) raises computational and economic questions. From a computational point of view, one can see bloc orders as reformulations of machine technical constraints. It is well known that reformulations of indivisibility constraints can modify the efficiency of MIP algorithms with the result that the same CPSM modelled through bloc orders or with an explicit representation of machine characteristics may solve quite differently. We are not aware of any computational study comparing the two approaches. In any case, it suffices to note here that the use of MIP in power auctions extensively developed these last years in various US systems operated under the SMD and that this experience shows that the computational difficulties can indeed be mastered. It is expected that the same will be true in the FBMC of CWE. But there exists other concerns as we discuss now

8.4.3. And adapting the experience to zone systems

Zone systems such as the proposed FBMC complicate the situation compared to SMD nodal systems. Zone systems normally rely on countertrading to remove intra-zone congestions (congestions internal to a price zone). These operations should ideally be conducted so as to minimize countertrading cost. This sub-problem of the overall CPSM problem is barely mentioned in the proposal. The CPSM must take into account that some of the machines will be requested for countertrading. Countertrading may seem a minor issue in CWE today and the common wisdom is that it will remain limited. This is possible but cannot be guaranteed: the experience of other systems shows that countertrading can become quite demanding in the sense that its cost can be high or worse, that there might not be enough machines available to perform the task. Both shortages of resources for re-dispatching or unexpectedly high re-dispatching costs have occurred in the early days (1997) in systems like PECO and more recently in ERCOT before both moved to the nodal organization. Insufficient re-dispatching resources have induced the Swedish TSO to curtain transmission capacities to Denmark in November 2005. The reasons go from structural (the composition of the generation system) to strategic (gaming) considerations but also involve market design (finding a single market clearing price). The penetration of wind is a structural element that may increase the countertrading activity and its cost in the future; it is not properly documented today. Gaming has been extensively reported at the time of the Enron saga and need not be elaborated upon. The market design aspect stems from the fact that machines are not only requested to remove line

congestion but to help find a single incentive compatible price, which is a more demanding objective. Needless to say one can expect these difficulties to also increase with the size of the zone: countertrading in a large price zone is likely to be more important than in a small zone except if the network is sufficiently capacitated. The current FBMB proposal works with very large zones.

To sum up, countertrading, if it develops (for instance as a result of wind penetration), raises at least three questions. First, it is not clear from the current FBMC proposal how the optimization of countertrading will be coupled with the CPSM. Second, because countertrading reduces the number of flexible machines that can be used in the auction, it may complicate the handling of bloc orders in the CPSM. Third the possible lack of re-dispatching resources may exacerbate the difficulties of price coupling. Very much like the FBMC proposal assumes that price coupling is always possible, it assumes that countertrading will always be possible and does not even mention the possibility of insufficient re-dispatching resources. The following elaborates on the pricing matter.

8.4.4. Pricing difficulties

The economic difficulties created by bloc orders seem less understood than the computational aspects in the FBMC proposal. European PXs are meant to clear at a single energy price in each zone, the set of these zone prices being incentive compatible with the volume exchanges determined by the global CPSM This is referred to as price coupling. The zone price is used to remunerate or charge all energy transactions going through the PX: the FBMC proposal does not foresee any other payment. In economic jargon, prices are linear (total payment is strictly proportional to the price) and anonymous (every agent sees the same price). Bloc orders may invalidate the existence of linear and anonymous zone prices. In other words, price coupling might be impossible for some patterns of offers and bids. The problem is not specific to the FBMC proposal; it also exists in TCMC where it is recognized through Paradoxically Rejected Blocs (PRB). These are bloc orders that should be accepted in price (they are profitable at the price found by the CPSM) but are "paradoxically" rejected in volume (they are not part of the solution found by the CPSM). The lack of market clearing prices also appears in the nodal systems operating under the SMD. They are here due to machine constraints like start up costs or minimal downtime (which are just the more direct formulation of the phenomenon creating bloc orders). We briefly elaborate on these points.

The question of the existence of a market-clearing price (price coupling) should be distinguished from the one of solving the CPSM (volume coupling). Bloc orders computationally complicate the finding of a mix of orders optimizing consumer and producer surplus (volume coupling) but they do not question the existence of that mix, at least if there is enough flexible generation capacity (one cannot guarantee that demand will always be satisfied if capacity is missing or only consists of inflexible plants). The CPSM will always conclude with a set of accepted and rejected orders possibly at some suboptimal solution. It will thus always allow for volume coupling. In contrast, the search for market-clearing prices may conclude that there exists no price that supports that solution. In other words, a linear and anonymous hourly energy price that induces generators and consumers to operate their plants as the CPSM demands may not exist. We briefly alluded to a similar phenomenon of lack of market-clearing price when discussing critical infrastructures: the absence of price coupling means that some agents may lack the incentive to comply with the results of the auction after learning the price that they will pay or receive in that auction.

The literature on the subject is limited but substantive (e.g. Hogan 2008, Hogan and Ring 2003, O'Neill et al. 2005, 2008 Scarf, H.E. 1990, 1994, Motto, A.L. and F.G. Galiana. 2002). The question appeared in the early experience of systems runs under the SMD. The cause of the problem is deep and was illustrated a long time ago on a very simple example in Scarf 1990. The remedy is to depart from single price auction and to turn to "multi-unit" or "combinatorial" auctions and to apply "non-linear prices". While the expression may involve a lot of jargon, the underlying idea is simple even though its implementation may be complicated. The principle is that payments and charges should not boil down to a single energy price but be accompanied by other payments. The most standard example is to remunerate generators for both producing energy and starting up or keeping up a machine (an uplift in SMD terminology). As just said the idea is simple but its implementation may not be; the discussion of non-linear prices goes beyond the scope of this report.

Notwithstanding its importance, economists tend to overlook the problem of non-linear prices in electricity: bloc orders or machine constraints imply what the jargon calls "non convexities", a phenomenon that economists dislike. Interestingly, optimization people also neglect the issue because it is a pricing question and they are rarely interested in prices. This double negligence has interesting impacts: the study conducted by London Economics (London Economics 2007) for the European Commission uses an engineering model ("Market Analytics" of "Global Energy Decisions") that effectively accounts for indivisibilities (non-convexities) at the machine level.

The study was conducted on five European electricity markets for the years 2003, 2004 and 2005 after an extensive data collection. The analysis is comprehensive but neglects verifying the incentive compatibility of the obtained prices. This led the Commission to claim that "…wholesale prices are significantly higher than would be expected on a perfectly competitive market…." This statement may be questioned: there may not exist any incentive compatible prices in a system that involves bloc orders and prices that are not incentive compatible "would (not) be expected on a perfectly competitive market". But a loose interpretation of the statement could still be useful in practice if the impact of bloc orders was not important compared to the situation where they would have been absent. But the study did not verify this point and hence says nothing about the possible incentive incompatibility of the prices that it finds. The report handled the computational issue of bloc orders (in that case by working directly on the technical constraints of the machines as in a SMD approach) but did not look at this economic consequence on prices.

The FBMC proposal is clearly aware of the issue but may not fully grasp all its consequences. The FMBC acknowledges the lack of market clearing prices when it refers to Paradoxically Rejected Blocs (PRB): blocs are rejected even though prices say that they should be accepted. But at the same time the authors of the FBMC assert that the software must always be able to find a price coupling solution. This is obviously not possible when market-clearing price do not exist. We mentioned the remedy above: use "multi-unit" or "combinatorial" auctions and apply "non-linear prices". We do not find any trace of such thinking in the current proposals.

It is important to restate, at the risk of redundancy, that the possible non-existence of a market clearing does not originate in the use of a Flow Based, Transmission Capacity or nodal model of the grid. It is rooted in bloc orders and eventually in machine constraints that is, in generation technology. There is nothing one can do about it: the larger the fraction of a given market covered by bloc orders is, the more difficult it is to find market clearing energy prices. But the choice of one or another network model may exacerbate the problem. The principles are simple: First, enlarging the set of flexible machines (not involved in bloc orders) participating to the energy market should facilitate the finding of a market-clearing price: The architects of FBMC cannot do much to the structure of the generation system and hence to the existing capacity of flexible machines. But they can act on the number of zones and on the recourse to countertrading. The zone system reduces the space of possible prices (one price per zone and few zones) and countertrading to eliminate intra-zone congestion reduces the number of flexible machine available to set prices. This discussion may look theoretical but it is not. On 28

November 2005, the Swedish TSO found itself confronted with a price space that had reduced to nothing: it could not find machines to achieve a market clearing price for its zone and had to reduce the transmission capacities with Denmark in order to solve its zone problem. PJM gives the other facet of the picture. The extension of the geographic coverage of the market did indeed simplify the problem of dealing with machine indivisibilities (in Europe, bloc order). But PJM uses non-linear prices to "price couple" the nodes!

8.4.5. Governance issues

The initial intent of Market Coupling was to integrate markets through an iterative process between TSO and PXs without requiring any additional centralised layer of activity. Bloc orders invalidate that objective by requiring a centralized MIP treatment of the CPSM. The solution of this MIP problem cannot be found by interactions between PXs and TSOs at a purely horizontal level. FBMC, as TCMC, requires a centralized level of computation to treat bloc orders. This could have raised organisational difficulties; they are apparently solved but it is worthwhile keeping the possibility of their resurgence in mind. We indeed saw before that critical infrastructures also require a central computation. Bloc orders and critical infrastructures combine to progressively lead to an ISO like organization where an asset free entity runs an auction to clear the market and processes the flow of information necessary for doing so. The concept of an ISO is difficult to convey in Europe. Moreover this ISO like organization has something special: it is constructed on the basis of the NECs established by the PXs in each zone and not on the basis of the nodal individual orders. As argued throughout the report, this difference is far from trivial and might create difficulties with very large zones. The experience of the internal energy market therefore suggests being very cautious: any solution that requires a centralized treatment of TSO functions can always bounce back, especially if this function is an integrated ISO like function and it encounters technical difficulties.

8.4.6. Conclusion

Bloc orders add to the difficulties created by critical infrastructures and already mentioned in Section 8.3. But bloc orders raise deeper new questions in the sense that they cannot be resolved by simply moving from a zone to a nodal system. These difficulties originate in indivisibilities in the functioning of generators. The authors of the FBMC proposal allude to the problem when they distinguish volume and price coupling and mention PRB. It is not clear however that they are fully aware of the importance of the question when they require that the software be able to find both volume and price coupling. Volume coupling can always be done;

price coupling may be impossible in some cases at least as long as one does not resort to more drastic measures such as non-linear prices. The proposal does not mention non-linear prices.

The experience of the Internal Electricity Market suggests that technical difficulties are commonly treated by ad hoc methods. The combination of volume coupling difficulties due to the zone system (see section 8.3) and price coupling intricacies due to bloc orders (this section) may lead to the decision to reject the whole FBMC approach or to downgrade it to the point that it is not better than the TCMC. It is thus important to recall that most of the difficulties created by both critical infrastructures and bloc orders are well understood today and that their solution is known.

9. INTRA-DAY TRADING AND BALANCING: PRELIMINARIES

We argued that the proposal to organize the day-ahead market on the basis of node-to-line PTDFs would be a major step forward but that the zone-to-line PTDF approach significantly complicates matters. Working with zones identical to countries may even question the validity of the whole endeavour. Theory and practice both suggest being cautious: theory explains that the choice of the GSKs in the proposed zone system requires in principle to solve a fixed point problem, an idea that is completely missing in current documents. The preliminary tests reported in the FBMC proposal reveal feasibility difficulties that may very well be related to this choice of GSKs. Practice reveals difficulties with many zone systems that, like ERCOT, tried a similar zone based flowgate model but finally opted for a nodal system. But the move to FBMC and the introduction of critical infrastructures is a potentially sufficient important step forward to continue the discussion and explore the possible extension of the approach to the other phases of trading namely intra-day and real-time. We pursue by largely assuming a node-to-line FBMC. This may not be strictly necessary, but it certainly simplifies the argument.

Today's approaches to day-ahead, intra-day and balancing operations are mutually inconsistent. The objective of intra-day trading is to allow agents to update positions inherited from day-ahead trading in response to innovations in the market. This is in line with the standard financial model of continuous trading. This updating process requires compatible organizations of the day-ahead and intra-day markets: specifically the traded commodities and transmission property rights must be identical throughout the evolving market and traded under similar conditions. In contrast, observation reveals striking differences between the architectures proposed for day-ahead and intra-day trading. A first major discrepancy is that transmission is no longer priced in intra-day-trading but given free on a first-come first-served basis as long as there remains available capacity. This purely quantitative allocation (not price driven) creates a true discontinuity between the pricing of transmission in the day-ahead and the intra-day markets that spoils any meaningful arbitrage possibility between them. There is no obvious reason for this discontinuity and one finds none in the existing documents. On the contrary, one may easily imagine reasons why this discontinuity of organization is counter-productive. Agents that have to pay to procure transmission in day-ahead but can get it free in intra-day will have an incentive to wait till the intra-day market. They are not certain to find that transmission capacity in the intra-day market but they have an incentive to try. There is no economic reason to give free a capacity that has not been allocated in the day-ahead market if the demand for that capacity increases in the intraday market. The real value of the capacity is the one that prevails in real-time when the

market physically clears. It is not the value in a forward market, which according to standard finance principles should be geared by the value in real time. This whole financial logic is lost if one starts allocating capacity free of charge between day-ahead and real time. One should also note that waiting for the intra-day market to procure a resource that it is artificially made free until one hits its availability limit is not gaming as stated in the TSO report (Cegedel et al. (undated)): it is normal economic conduct. Economic agents behave according to the incentives that they receive from the designers of the system: they are not supposed to correct the errors of the designers.

An analogous reasoning can be made for the respective architectures of the intra-day market and balancing operations. While one admits that intra-day can continue till close to real-time, the arrangement is completely different in real-time: one allows for trading in intra-day but tries to prevent it in real-time. This is again a discontinuity that prevents arbitrage between intra-day and real-time. But the discontinuity is here intended: the acknowledged objective is indeed to induce agents to remain in balance by penalizing, more than pricing, imbalances. In other words the objective is to prevent agents from moving part of their procurement to real-time. Again, there is no good economic reason for this. There is certainly no economic reason for making balancing more expensive than its marginal cost and hence for discouraging procuring electricity in real-time. For the same reason as in the day-ahead market, there is no economic reason for not trying to reduce the cost of balancing by arbitraging balancing resources between systems, that is, for trying to organize cross-border balancing in the same way as one tries to organize cross-border electricity trade.

The comparison between the day-ahead and intra-day markets and the balancing system leads one to conclude to three different organizations that prevent temporal arbitrage. These multiple arrangements violate the finance view that day-ahead, intra-day and real-time are just different steps of a single trading process and hence require a single trading platform. This view has already been exposed in the introduction and we only recall its principle here. Because of non-storability, the physical trade of electricity only takes place in real-time, which is thus the only true spot market. The other markets are forward markets that trade derivatives products maturing in real-time on the spot market.

Electricity is a special commodity and there might be technical reasons to depart from this standard view of finance. But the experience of the SMD in the US tells us otherwise. The SMD organizes the day-ahead and real-time markets in the same way. There have been attempts to proceed otherwise for instance by adopting a coarser representation of the network in the day-

ahead than real-time market (Kamat and Oren 2004). This attempt is of particular interest here: the idea was to be zone based in the day ahead and node based in real time. This was tried and later abandoned in ERCOT. But attempts to work with different views of the grid have never gone as far as introducing a full discontinuity between the day-ahead and real-time markets. Extending this SMD experience to the inception of the intra-day market, one shall conclude that there is also no reason why day-ahead, intra-day and real-time should be organised differently if the same market design can apply to all steps.

It may be useful before concluding to try to identify the logical flaw (the deep reason) that lead to the discontinuities in the proposed organizations of day-ahead, intra-day and real-time in Europe. The three organizations differ in form but also on a fundamental principle: there is no pricing of transmission in the organization of the intraday-day or real time trading. Transmission is implicitly priced at zero in intra-day and at infinity in real time as long as there is no cross-border trade of balancing. In other words, the fundamental link between energy and transmission that results from the basic properties of the electricity, and took so long to be admitted in the day-ahead market is here again abandoned. There is no logical reason for doing so: recall that the basic reasoning is that the link between energy and transmission prices is established in the real time market and extended by standard economic and finance arguments to the other markets. There is also no practical reason for not keeping this link between energy and transmission at all stages of trading. The experience of the PJM shows that one can re-assess grid characteristics every five minutes for a system of more than 8000 nodes and use the information to simultaneously produce electricity and transmission prices at that frequency. This implies that one can cast day-ahead, intra-day and real-time trading in the same architecture. Using PTDFs and updating them very frequently to adapt the representation of the grid brings us very close to the finance model that looks at the price formation process as the result of continuous trading. The rest of our discussion is based on this finding: one should reconcile the current design to the standard view of finance by moving the flowgate model to intra-day and real-time trading.

10. Making balancing a real-time market

Competition authorities rightly see the current organisation of balancing as a barrier to entry. The argument is simple and compelling: an entrant with a limited generation resource and customer basis may find it difficult to remain in balance. This is the barrier: it can be more or less deterring. If balancing is charged at more than what it costs or its costs is higher than it should be because of the balancing arrangement, then the barrier is reinforced. This obviously raises the question of whether the objective to remain in balance is justified. Remaining in balance obviously simplifies the task of the TSOs. But this simplification should not go as far as charging imbalance more than what it costs or should cost if it had been provided by an efficient TSO.

The observation of the existing situation suggests, but does not prove, that the cost of balancing is higher than what it should be, and this for a trivial reason. The same reasoning that suggests that cross-border arbitrage in the day-ahead market is globally beneficial (increases Consumer and Producer surplus) also suggests that cross-border arbitrage should be beneficial for balancing. Similarly, the same reasoning that suggests a market based approach for the allocation of "transmission capacities" also suggests a market-based approach for pricing balancing resources across borders. The very question is whether the same instruments that are today used for organising cross-border trade (that is Transmission Capacities) in the day-ahead market can also be used for a market-based approach to cross-border trade of balancing resources. The answer is probably no: the computation of TCs is cumbersome, it is not doable in real-time and even it was, updated TCs would not permit simulating the pricing impact of flows resulting from real-time exchanges of imbalances. The German system illustrates today's different views on day-ahead and balancing. While operators insist on a single zone in the day-ahead market, they maintain four balancing zones in real-time. They thus limit their operation to their own control areas. The reason is probably that operators can arrange for extensive countertrading to eliminate congestion in the day-ahead markets while this is becoming much more difficult in real-time when working with Transmission Capacities. Specifically one cannot meaningfully price balancing in real-time on the basis of TCs. Can PTDF help?

10.1. Adapting FBMC to balancing

Balancing is a zone notion in Europe: any plant of adequate flexibility in a zone can contribute to balancing. Congestion management is taken care of in another function. Suppose we abide to the principle: the question is whether we can establish a cross-zone trade of balancing services

using PTDFs for congestion management purposes. The answer depends on whether one is using node-to-line PTDFs or zone-to-line PTDFs

It should be clear that it is possible to use a FBMC platform based on node-to-line PTDFs to organize a real-time market of balancing resources. Node-to-line PTDFs can be recomputed on short notice, taking into account current injection and withdrawal positions and the topology of the grid. Node-to-line PTDFs are then equivalent to an approximation of load flow equations that can deliver real-time electricity and transmission prices for any pattern of imbalance and use of balancing resources. The statement requires some qualifications though.

We first note that the CPSM used for clearing the day-ahead market is also applicable to clear the market of balancing resources and real-time transmission services. The model must run faster because one is operating in real-time but its task is also much simpler. Bloc orders are not part of the balancing market (one can hardly balance with indivisible orders). The CPSM will thus run with a given set of bloc orders. Computational difficulties (the use of a MIP algorithm) mentioned in Section 8.4.1 disappear as well as the pricing difficulties discussed in Section 8.4.1. The optimisation is conducted with "flexible resources" (flexible machines) for a given time segment and there is no price coupling difficulty in a nodal system when machines are committed and one only needs to decide on their operations level (which is the case in real time). The experience of the SMD confirms this view. These systems indeed began with the sole real-time market and hence did exactly what a real time balancing market would do. We thus know that it can be done and that node-to-line PTDFs are adequate instruments to do it.

The nodal aspect of this solution is also its drawback in today European discussions. The FBMC proposal is based on zones and not on nodes. The representation of the network is constructed with zone-to-line PTDFs and not with node-to-line PTDFs. Moving from one to the other requires assumptions on the distribution of injections/withdrawals in a zone into nodal injections/withdrawals (the GSKs). One may argue (possibly unconvincingly) that it is possible to find good GSKs in the day-ahead market; one cannot submit this type of argument for real-time when the mix of balancing resources is changing quickly. We reasoned before that there is no objection to using node-to-line PTDFs for selecting orders in CPMP. A zone system can select orders and hence conduct volume coupling using node-to-line PTDFs. This will produce nodal price coupling as a by-product. The issue is the problem of incentive compatibility when moving from the nodal to zone prices. We did not conclude on this question in Section 8.4.4 because of lack of information on the zone price computation process. We do not conclude here either: many things can be explored but they are beyond the scope of this report.

10.2.Governance issues

Balancing is currently a non-market activity run by individual TSOs: existing documents available from ETSO's web site reveal little harmonization and almost no integration among TSOs. There are plans and there is some progress, but all this remains quite vague at best. Transforming the current balancing system into a real-time market will thus create considerable organisational resistance that one cannot hope to overcome in the near future. As discussed, there are also real technical difficulties due to the quality of the zone-to-line PTDFs in day-ahead and real-time. It is thus reasonable at this time to concentrate on FBMC in day-ahead while maintaining, in parallel, a shadow activity that looks at the implications of transposing these developments to balancing and real-time trading.

Making balancing a real-time market also raises the question of the involvement of the PXs. The logic of continuous trading suggests that PXs should be part of the real time market but the current organization does not readily indicate how this can be done. The question of the interaction of TSO and PXs in real time can thus also be included in that shadow activity.

11. THE INTRA-DAY MARKET

Intra-day trading is meant to allow agents on the market to adapt their position in response to innovations after the closure of the day-ahead market. There is no physical intra-day trading in systems operating under the SMD paradigm as agents can arbitrage between the day-ahead and real-time markets. But the SMD system includes a virtual intra-day market that is purely financial and does no imply modifications of physical positions. The objective of virtual intraday trading is to facilitate arbitrage though time, a theme that should be familiar by now. Physical intra-day trading is certainly necessary in today's European markets that have no real-time markets. Even though the economic logic would be to begin by developing the real time market, there are pragmatic reasons to give priority to intra-day markets. First, stakeholders are demanding an intra-day market much more than a reform of balancing. Second the intra-day market is currently in development and still in flux. In contrast, balancing is an already established system that is implemented in a rather heterogeneous way in the different Member States. Transforming and harmonizing an existing heterogeneous system is organizationally much more demanding than gearing a system still in development. Introducing FBMC in the intraday market therefore offers the opportunity of a learning experience that may mitigate

resistance against a future reform of real time activities. The third reason will be invoked in the next section: the penetration of wind and the progress in wind speed forecast will benefit from the development of a sophisticated intra-day trading. The last reason looks academic but has by now become a recurring theme in the report: intra-day trading approximates the paradigm of continuous trading in finance; it hints at a globally more consistent system.

11.1. Adapting FBMC to intra-day trading

The development of intra-day trading in Europe is presented in some ETSO documents that can be downloaded from the organization's web site. Cegedel et al.(undated), offers a recent analysis of the integration of the different intra-day systems. The authors suggest, but do not assert, a FBMC based architecture. This suggestion is sound: intra-day trading is the follow-up of day-ahead trading and it is quite reasonable to organize both systems on the same basis. In fact the whole discussion of the day-ahead market conducted in this report applies to intra-day with some simplifications. Specifically the representation of the grid in intraday trading should also be based on PTDFs because this representation of the grid allows one to track and forecast the impact of changes of injections and withdrawals on line utilization. One might argue that zone-to-line PTDFs are acceptable close to day-ahead but the question of moving to node-to-line PTDFs mentioned for balancing cannot be avoided when approaching real-time. As before we leave this question unanswered. In any case and possibly controversial in view of some statements of Cegedel et al., the adoption of the Flow Based model implies that energy and transmission should be priced at the same time and that transmission should not simply be offered for free up to available capacities. The same system that charges for transmission services in the day-ahead market should apply in intra-day. This condition is required for arbitrage. The CPSM used for clearing the day-ahead market can also be used in the different steps of the intra-day market provided one updates the data after the closure of the day-ahead market. There will be simplifications in the CPSM as one should prevent or limit modifications of bloc orders as one moves close to real time. The computation of the PTDFs can be updated to account for changing positions in the energy market and the possible modifications of the grid topology.

11.2. Governance issues

Even though the idea of developing intraday according to the same lines as day-ahead appears in Cegedel et al., that report reveals a much less structured thinking than the current FBMC proposal. It is indeed easy to analyze the existing FBMC proposal and pinpoints areas of

concern; in contrast it is very difficult, except for the mentioning of a possible reliance on the FBMC architecture, to decipher what Cegedel et al. really intends to do. The first objective is thus to clarify ideas.

Very much like for real time, but possibly with a higher sense of urgency, it is thus reasonable at this time to introduce a shadow activity that would systematically explore the implications of adopting FBMC principles to intraday trading. The main danger perceived from reading Cegedel et al. is that an intraday specific way of thinking develops creating difficulties of harmonization not only between day-ahead and intraday but also among geographic systems in the future. Needless to say the need for a common thinking between day-ahead and intraday also applies to the relations between PXs and TSOs.

12. SUMMING UP

It appears that the recourse to a Flow Based representation of the grid makes it possible to integrate the trading platform into a single scheme going from day-ahead to real-time, and passing by the intra-day market. The driving theme comes from the idea of continuous trading in finance. Agents trade on the basis of their current knowledge in the day-ahead market and update their positions as they learn more about demand, production, network...In so doing they take forward positions that are settled at the price of the market at the time. At each moment of time the market clears at the result of a CPSM producer model. Volume coupling will be possible throughout. Questions remain about price coupling. The trading of bloc orders should stop or at least be drastically reduced at some point in the process. We certainly know that bloc orders cannot play any role in real-time; but their trading should stop before real time as changes of accepted bloc orders introduce drastic variations of market clearing prices and sometimes lead to the disappearance of these prices and the impossibility to price couple markets. This question is relevant irrespective of the chosen market coupling approach, whether Flow Based or Transmission Capacity Based.

The other issues are direct consequences of the application of the zone system and already appear in the day-ahead market. It is not clear how congestion on a critical infrastructure located inside a zone intervenes in the zone pricing. It is also not clear how one shall compute zone-to-line PTDFs as one comes close to real-time. Note that we mention these difficulties with the view that they have to be solved; difficulties should not be used as argument for bypassing real

problems that need to be tacked seriously if the Internal Electricity market is itself to be taken seriously.

13. INTERMITTENT SOURCES

The European Union has embarked on an ambitious renewable policy. This accompanies a much-needed objective of GHG emissions reduction. Renewable energy sources are intermittent and hence not fully dispatchable. Nuclear, coal and gas units equipped with CCS are GHG free (or almost free for CCS) and there remain issues of flexibility. The dispatching of the generation system is thus expected to become more difficult as we progress towards these renewable and GHG objectives.

Intermittent sources pose a problem of their own that can dramatically benefit from progress achieved in the market architecture. It is today impossible to forecast wind speed a day in advance. This implies that wind generations will not be known, or at least very imperfectly known in the day-ahead market. But forecast improves as one moves forward in time and becomes quite good a few hours before real time. This should normally imply active trading as one moves from day-ahead to real time and information on forthcoming wind generation is progressively known. The idea of a single trading platform going from day-ahead to real time is thus of the essence. Wind also has its own geographic characteristics, as the bulk of the plants are generally located along coasts with new installations progressively developing offshore. This will have network implication with the consequences that the spatial dimension of trading will also become important. These characteristics suggest a sophisticated trading activity where one shall need to account for a geographically differentiated forecast of wind speed that becomes more precise as one moves from day-ahead to real time. This also means that the price signals and the trading activity will need to account of these spatial and time differentiations. Considering the current FBMC proposal and the intended wind penetration prospects, it is hard to believe that one shall be satisfied in the future with a single day-ahead price of electricity prevailing from northern to southern Germany.

The experience of the ISO/RTO functioning under the SMD paradigm in the US has shown that these organisations have so far adapted very well to wind penetration, even though accommodating much smaller market share that those foreseen today in Europe. This US

experience emphasises the importance of price coupling for sending the right dispatch signals and producing the right incentives for controllable machines to generate in a way that properly adapts to wind variation. The old command and control would have dispatched controllable sources taking into account the many sources of dispersed generation. The reorganized system proceeds through price incentives and it is essential that prices exist and be correct. Specifically the continuum between the successive phases of the market from day-ahead to real-time and the insistence on price coupling vs. volume coupling especially closer to real time when wind forecast become good are here of the essence. Coupling markets through prices determined at a local PX level (and not at the central level) might be problematic if these are not strictly "price coupling" (that is incentive compatible) and induce decentralized generators to deviate from volumes determined by the CSMP. The continuous update of prices by the CSMP through time is also of the essence as one moves from day-head to real-time. The experience of the SMD based systems has led to the following conclusions: it is easier to accommodate intermittent sources in a geographically larger system (PJM) and the move from zone to nodal system also facilitates the inclusion of wind (ERCOT). Both these experiences and theory suggest that the "simplifications" currently introduced in the FBMC planned in CWE will backfire with the growth of wind power. Specifically the German claim that one can manage a high penetration of wind power by countertrading is at best intriguing, except if one heavily invests in the grid.

It is thus suggested to revisit the proposal and the comments made in this report in the context of a high penetration of intermitted sources. It is suggested to assess the extent of the needed countertrading and the consequences of the zone system in view of the high wind penetration.

14. ZONE SELECTION

The FBMC proposal departs from the notion of TCs and moves to flowgates. This step is most welcome. The standard Flow Based representation of the grid models a line flow as a linear function of injections/withdrawals at nodes, with corresponding opposite operations (withdrawals/injections) at the hub/balancing node. The FBMC proposal modifies this definition by modelling line flows as a function of injections/withdrawals in zones with opposite operations at the hub/balancing node. This modified definition assumes a distribution of the injections/withdrawals at the different buses of the zone. This is done using Generation Shift Keys (GSK). The proposal understandably does not give much information on GSKs and leaves it to the experience of TSOs to determine them. While the recourse to PTDFs and critical

infrastructures is an important step in the right direction, the impact of the zone-to-line PTDFs remains a major uncertainty of the proposal: the validity of the network model (the zone-to-line PTDFs) entirely depends on the extent to which the GSKs chosen ex ante by the TSO are close to those that will emerge from the solution of the CPSM. If the discrepancies are too large the flow pattern found by the CPSM will (i) not correspond to the best use of the grid (and hence the trading platform will not make the best of the existing resources) and (ii) not guarantee the security of the operation (because the underlying network model and hence the flow forecast are wrong). This problem created by the use of zones can be approached in different ways, none of them being guaranteed to succeed but all of them certainly worth trying.

14.1. An ex ante determination of the zones

The zones appearing in the FBMC proposal are full countries. Suppose this is not required and one can admit zones smaller than countries. One can then attempt to construct zones by grouping nodes with similar node-to-line PTDFs. The task does not appear insurmountable: the FBMC proposal estimates of the order of 3000 critical infrastructures. One can thus associate a vector of 3000 PTDFs to every node. Grouping the nodes in zones can be done by a cluster analysis of the nodes with respect to the similarity of their vectors of PTDFs. It is not certain that this operation will succeed but it will provide useful information. If successful (e.g. French nodes can be grouped in five zones) then it is possible to construct a better model of the grid in terms of zone-to-line PTDFs that is less dependent on the choice of the GSKs. If unsuccessful, then one knows that the current approximation of the load flow equations of the grid is not good and something else should be tried.

14.2. An ex post adaptation of the generation shift keys

The ex ante construction of the GSKs is the main difficulty of the current approach, as these may not correspond to the flows coming from the solution of the CPSM. The usual remedy in this type of situation is to proceed by adapting the allocation coefficients ex post and redoing the computation (this amounts to solving the aforementioned fixed point problem). This usual remedy is not guaranteed to converge and the treatment of bloc orders does not facilitate this convergence. Needless to say the convergence is likely to be better if the zones have been carefully selected ex ante. The updated computation of GSK can be interpreted in different ways: it can take place in one shot with the view of solving the already mentioned fixed-point

problem: one then tries to find GSKs compatible with the volume coupling solution. One may also see it as taking place through time, for instance at a new clearing taking place after nomination when more information on the location of injections and withdrawals is available

14.3. Zone splitting

We noted several times that the use of node-to-line PTDFs does not complicate the selection of orders in the CPSM (volume coupling) but that the real difficulty is to find market-clearing prices (price coupling). A reasonable approach is to split zones when circumstances justify it, that is, when it appears that one cannot achieve price coupling. Unfortunately, this solution is likely to be rejected, given the current philosophy to work with fixed zones. There might still be a way out but it requires further investigations. The true solution is indeed to resort to non-linear prices. This may be considered more acceptable than zone splitting.

Note that the price coupling operation step is crucial whatever the fraction of the market going through the PXs. As the comparison with the SMD shows, the market coupling system is the sole operation capable of finding market-clearing prices at the scale of the whole market. This process is thus driving the whole cross border arbitrage. It is thus of the essence that all effort be made to go beyond the sole volume coupling if one wants to avoid creating inadequate incentives, that is, volume transactions not supported by prices.

15. CONCLUSION

The current FBMC proposal is a potentially fundamental step in the right direction. But its pricing scheme maybe insufficiently understood and the extreme zone system adopted may lead to fatal difficulties. We briefly summarize the forward steps at the same time as we mention possible shortcomings.

15.1. Market coupling and price coupling

The day-ahead market operates an implicit auction in the FBMC proposal. This means that it simultaneously clears the energy and transmission markets. This is good as both prices are fundamentally related in the economics of electricity. The selection of retained orders (which is always possible) and the computation of market clearing prices (which might not always be possible) are done centrally on the basis of information exchanged between a central agency and the PX and TSOs. The central agency essentially runs a market-clearing algorithm and controls the exchange of information. The recognition of the need of a central processing is also a major step forward. But admitting the need for centralization is so unusual in the history of the Internal Electricity Market that one should remain wary that it might backfire if and when serious difficulties are encountered in the implementation.

Market clearing is performed on the basis of energy prices (euro/Mwh). There is no other payment in the proposal. These prices apply equally to all: in the jargon, they are linear and anonymous. It is known both from economic theory and the practice of existing reformed systems that there may not always exist anonymous linear prices that clear the market when some of the traded products are bloc orders. The proposal recognizes the difficulty and treats it by the distinction between price and volume coupling. This distinction acknowledges the problem but does not solve it; the consequence is that it is unrealistic to impose that FBMC be always able to find a market clearing energy price (price coupling), except if market clearing is to be interpreted in a weak sense. The true solution to this problem is to resort to non-linear prices that complete the energy prices by side payments: only non-linear prices can be guaranteed to clear the market when there are bloc orders. The discussion of this solution goes beyond the scope of this report. The lack of market clearing price is sometimes invoked as allowing some PX (in this case EEX) to comply with national law that requires prices to be computed domestically. This is well acknowledged but does not contradict our point; our argument is that this legal requirement does not remove the inherent difficulty of a market that lacks a clearing price. This signals

inadequate incentives, which the history of reformed electricity market suggests have often led to negative unintended consequences.

The zone character of the proposed system is at the origin of a related pricing problem as the experience of former zone systems reveals. The FBMC proposal adopts a representation of the grid by zones and suggests that zones may cover whole countries. It is well known that congestion can persist inside zones with the implication that they will need to be relieved by countertrading. Countertrading requires resources (machine or demand side) that are therefore no longer available for submitting orders to the PX or for contracting on the OTC market. These resources are in general flexible: withdrawing them from the market and reserving them for congestion management makes it more difficult to find energy only market-clearing prices. The problem may become crucial with wind penetration where managing congestions by countertrading will demand more flexible resources. One should recall that the experience of reformed power systems has revealed that it is sometimes impossible to find zone prices because of lacking countertrading resources. This complication of the zone organisation of the market adds to the bloc orders to make the finding of energy only market clearing prices more difficult. In contrast with the bloc orders that inherently require non-linear prices, one can here always facilitate price coupling by resorting to more zones.

15.2. Critical infrastructures

The FBMC selects orders that comply with the capacities of critical infrastructures. These are defined both for normal and contingent conditions. Congestion is thus no longer the sole matter of interconnections but can also play an important role inside zones. The inclusion of critical infrastructures is a very positive step. But the zone character of the proposal leads to a definition of the PTDFs that is likely to create difficulties. PTDFs no longer express the sensitivity of the flow on a line with respect to nodal but to zone modifications of injections or withdrawals. This in turn requires an allocation of the zone injections/withdrawals among the different nodes of the zone. This allocation may not be correct, which implies that the modelling of the critical infrastructures in the CPSM is itself incorrect. The problem can be remedied to some extent. One shall first note that there is no need to use zone-to-line PTDF for selecting orders; in other words, volume coupling can be performed on the basis of a zone-to-line as well as on a node-to-line representation of the grid. In this latter case, the CPSM is solved with a good description of the grid guaranteeing that the accepted orders will be secure. The problem remains to find zone market clearing prices when there is congestion inside the zone; it is here compounded by the

need to distinguish whether it is the exact model (based on node-to-line PTDFs) or the simplified problem (using zone-to-line PTDFs) that does not have zone market clearing prices. We argued that one might try to use cluster analysis to arrive at a good definition of the zones, that is, one where the PTDFs depend less on the GSKs. We also explained that one could try to update the GSKs in case of given zones. If successful, these remedies will give zone market clearing prices. If not, the only solution to arrive at price coupling is to split zones (which helps but does not guarantee price coupling) or resort to non-linear prices.

15.3. Extending FBMC to intra-day and real-time trading

The current organisation and thinking about intra-day trade and balancing is not satisfactory. There exist major discontinuities between the three steps and these are inconsistent with both the financial model of continuous trading and the practice revealed by systems operating under the SDM paradigm. FBMC offers an opportunity, starting with the day-ahead market to cast the two other trading phases under a common framework that relies both on the same CPSM and flowgate model. This common framework will become more and more urgent as Europe accommodates a high penetration of wind energy.

◢ ◢ ◢ ◢

REFERENCES

Cegedel Net, Elia System Operator, EnBW Transportnetz, E.ON Netz, RTE, EDT Transport S.A., RWE Transportnetsstrom, Tennet TSO, undated. Cross border intraday market.

Chao, H.P. and S.C. Peck, 1998. Reliability management in competitive electricity markets. *Journal of Regulatory Economics*, Vol. 14 (2), 189-200.

CWE region, Implementation Study
http://www.belpex.be/index.php?id=32&tx_ttnews[tt_news]=45&tx_ttnews[backPid]=1&cHash=3f25d90566

Dijkgraaf, E. and M. Janssen, 2008. Is there a Belgian wholesale electricity market? Erasmus Competition and Regulation Institute, Erasmus University Rotterdam, available on line on http://www.ecri.nl

Duthaler, C., Emery, M., Anderson, G. and M. Kurziden, 2008. Analysis of the use of PTDFs in the UCTE Transmission Grid. To be presented at the Power System Computation Conference, July 14-18, 2008, Glasgow.

FERC. SMD NOPR, 2002.

FERC. 2005 Principles for efficient and reliable reactive power supply and consumption. Staff report Docket no. AD05-1-000, Federal Energy Regulatory Commission, February 4, 2005.

ERGEG 2007 ERGEG Regional initiatives annual report 2007. Available on line http://www.energy-regulators.eu,.

ERGEG. 2008 ERGEG Regional initiatives annual report 2008. Available on line http://www.energy-regulators.eu, .

Frontier Economics. 2007 *Harmonization of gate closure times.* Frontier Economics, Ltd., London, 2007.

Gribick, P.W., Hogan, W.W. and S.L. Pope. 2007 Market-clearing electricity prices and energy uplift, December 31, 2007. Available on line on http://ksghome.harvard.edu.

Hogan, W.W., 1992. Contract networks for electricity power transmission. *Journal of Regulatory Economics*, 4(3) 211-242

Hogan, W.W., 2003. Transmission market Design. Available on line on http://ksghome.harvard.edu.

Hogan, W.W., 2008. Electricity market design: Coordination, pricing and incentive. Available on line on http://ksghome.harvard.edu,

Hogan, W.W. and B.J. Ring, 2003. On minimum uplift for electricity markets. Available on line on http://www.ksg.harvard.edu.

IRC, 2007. Progress of organized wholesale electricity markets in North America. ISO/RTO Council,

Joskow P.L 2007 Lessons learned from Electricity Market Liberalization, Available on line from http://econ-www.mit.edu/faculty/pjoskow

Joskow P.L 2006 Markets for Power in the United States: An Interim Assessment, *The Energy Journal* 27(1), 1-36

Kamat, R. and S. Oren, 2004. Two settlements systems for electricity markets under network uncertainty and market power. *Journal of Regulatory Economics,* 25(1), 5–37, 2004.

London Economics, 2007. Structure and performance of six European wholesale electricity markets in 2003, 2004 and 2005. London

Mac-Kie-Mason, J.K. and H.R. Varian, 1993. Pricing the Internet. Memorandum 20/1993, , Department of Economics, University of Oslo.

Erin T. Mansur and Matthew W. White, Market Organisation and Efficiency in Electricity Markets, Discussion draft, October 3, 2008
http://bpp.wharton.upenn.edu/mawhite/papers/MarketOrg.pdf

Motto, A.L. and F.G. Galiana, 2002.. Equilibrium of auction markets with unit commitment: the need for augmented pricing. *IEEE Transactions on Power Systems,* 17(3), 798-805.

O'Neill, R.P., Sotkiewicz, P.M., Hobbs, B.F., Rothkopf, M.H. and W.R. Stewart Jr., 2005. Efficient market-clearing prices in markets with non convexities. *European Journal of Operational Research,* 164(1), 269-285.

O'Neill, R.P., Sotkiewicz, P.M. and M.H. Rothkopf. 2008 Equilibrium prices in power exchanges with non-convex bids (in review).

O'Neill, R.P., Fisher, E.B., Hobbs, B.F. and R. Baldick. Towards a complete real-time electricity market design (to appear in *JRE*).

Scarf, H.E., 1990. Mathematical programming and economic theory. *Operations Research,* 38, 377-385.

Scarf, H.E., 1994. The allocation of resources in the presence of indivisibilities. *Journal of Economic Perspectives,* 8(4), 111-128.

Wilson R. B., 2002. Architecture of power markets. *Econometrica*, Vol. 70(4), 1299-1340.