Two Visions of a Transactive Electric System[1]

Lorenzo Kristov, California Independent System Operator[2]

Paul De Martini, California Institute of Technology

Jeffrey D. Taft, Pacific Northwest National Laboratory

Conversations about distributed resurces, the transactive grid and the future of the electric industry abound with terms like decentralized, disruptive, distribution system operator and distribution marginal pricing. In this time of rapid dramatic change it is tempting to jump right to the sexy market design, utility business model, ratemaking and regulatory framework issues, and assume that the mundane operational realm where the laws of physics apply will sort itself out. As a result it is not surprising that most industry futurists have not noticed the emergence of two clearly distinct visions or paradigms for how a decentralized, transactive electric system with high penetration of distributed energy resources (DER)[3] could be designed to operate. This article starts from the operations perspective and describes the two visions in some detail, compares them and draws implications for key design and regulatory questions being debated today. Operations, in this context, means reliable operation of the high-DER physical electric system as a whole, from the balancing authority area to the end-use customer, and thus entails operation of both the distribution and transmission systems as well as the interfaces between them.

One vision centers around a centralized, whole-system optimization performed by the transmission system operator (TSO), who may also operate wholesale spot markets as an independent system operator (ISO) or regional transmission organization (RTO).[4]

---

[1] Accepted for publication in IEEE Power and Energy Magazine.

[2] Ideas presented in this article are the opinions of the author and do not reflect the views or policies of the California ISO.

[3] The term "DER" is used broadly here to include distribution-level devices connected on either the customer side or utility side of the meter, as well as aggregations of such devices to form virtual resources. It includes rooftop solar generation, energy and thermal storage, electric vehicles and charging stations, energy efficiency measures, demand response resources, and the communication and control systems that allow individual DER or aggregations of DER to provide services to the system as well as to end-use customers.

[4] The term TSO is used generically to apply to both non-ISO/RTO regions and to ISO/RTO regions which also feature wholesale spot markets. In non-ISO/RTO regions the TSO would typically rely on centralized cost-based decision making rather than organized spot markets and responses to market prices for short-term grid operation and balancing.

Under this model the TSO needs detailed information and visibility into all levels of the system, from the balancing authority area down through the distribution system to the meters on end-use customers and distribution-connected devices. The other vision involves a decentralized, layered-decomposition optimization structure, for which optimization at any given layer of the system only requires visibiity to the interface points with the next layers above and below, and does not need visibility to what's inside those other layers.[5] The TSO under this layered optimization paradigm would see a single virtual resource at each transmission-distribution interface[6] and would not need to be concerned with the individual DER or customers below the interface. This is not unlike the existing operational paradigm between balancing authorities, which primarily focuses on interchange flows between balancing authority areas.

Each of these visions can be used to characterize a different mature end state of the high-DER electric system.[7] The two visions are deliberately drawn as conceptually distinct to reveal key operational design choices and derive some observations to help inform today's policy, market design and system architecture discussions. The choice of which vision to aim for in any jurisdiction will have major implications for specifying the complementary roles and responsibilities of the distribution system operator (DSO) and the TSO, and consequently for the business model of the distribution utility. The choice will also imply different directions on questions like the value of distribution-level locational marginal pricing, the optimal uses of markets and controls to maintain reliable system operation, and the benefits of decentralization for enhancing system security and resilience. Perhaps surprisingly, both visions can fully support distribution-level peer-to-peer transactions, assuming the regulators adopt regulatory frameworks to enable such transactions.

---

5 J. Taft and P. De Martini, "Scalability, Resilience, and Complexity Management in Laminar Control of Ultra-Large Scale Systems," Cisco, 2012. http://www.cisco.com/c/dam/en/us/products/collateral/cloud-systems-management/connected-grid-network-management-system/scalability_and_resilience_in_laminar_control_networks.pdf

6 In ISO/RTO regions that operate locational marginal pricing (LMP) markets, the transmission-distribution substation is also a pricing node for spot energy prices.

7 Some participants in current industry debates may argue that this dual scheme omits designs that exchew all forms of system optimization in favor of complete reliance on autonomous responses to price signals. Our view is that proponents of such designs still implicitly assume some mechanism to calculate the needed price signals with a high degree of locational and temporal granularity. Some entity or mechanism must be running at all times to calculate the prices that align with grid conditions.

The grand central optimization

The grand optimization vision is the logical extension of the wholesale market structure that exists in ISO/RTO areas today, but with much greater quantities and diversity of DER participating in the wholesale markets, including DER on both the customer side and the utility side of the meter, both individually and as aggregations into virtual resources. The structure of wholesale market participation by DER can remain much as it is today, with the ISO/RTO issuing dispatches and the DSO providing coordination services and utilizing qualified DER to support distribution system operations where feasible and economic.

Under the most likely version of this model – the Minimal DSO model – the TSO would see DER in its optimization as if they were located at the T-D substation, consistent with how the wholesale markets operate today. Under a more extreme version – the Total TSO model – the TSO's network model would include the distribution circuits and model the DER at their actual locations on those circuits.[8] The Total TSO requires such detail because spot market pricing or even direct operational controls rely on an accurate power system model including asset ratings, status and topography to perform real-time state estimation of power flows to calculate prices and control signals. In either case, the DSO would have to take on substantial new functions to coordinate the activities of DER on its system to maintain system reliability as the DER respond to TSO dispatch instructions while providing other services to end-use customers and possibly the distribution system. The distribution asset owner, who may or may not be the same entity as the DSO, would still be responsible for maintaining and operating the physical assets, analogous to the role of participating transmission owners in ISO/RTO areas today.

The grand optimization paradigm could be fully compatible with distribution-level markets, including markets for services DER can provide to the DSO to support reliable system operation or to defer investment in distribution infrastructure, as well as peer-to-peer transactions in which DER provide services to end-use customers or other DER.

---

[8] De Martini and Kristov use the term "Minimal DSO" or "Model B" for the version with DER modeled by the TSO at the T-D substation, and "Total TSO" or "Model A" for the version in which the TSO includes the distribution system in its optimization model and models DER at their actual locations. The authors provide several reasons why the Total TSO, although interesting as a concept, would be impractical to implement. See Paul De Martini and Lorenzo Kristov, "Distribution Systems in a High Distributed Energy Resources Future: Planning, Market Design, Operation and Oversight," Lawrence Berkeley National Laboratory series on Future Electric Utility Regulation, October 2015. https://emp.lbl.gov/sites/all/files/FEUR_2%20distribution%20systems%2020151023.pdf

Dispatch of DER for transmission-level and wholesale market services would come primarily from the TSO, however, because the DSO's operational role over DER would be limited to instructions or control signals needed to maintain distribution system reliability while supporting DER wholesale market participation. It may come as a surprise, but we will see that the layered optimization vision described next could also support the full range of DER transactions, so this aspect should not be a main distinguishing feature of the two visions.

The layered decentralized optimization

The layered optimization paradigm represents a substantial break from today's models of DER participation in the future grid and the wholesale market. Instead of numerous DER and DER aggregations bidding directly into the wholesale market and being scheduled and dispatched by the ISO, the DSO would aggregate all DER within each local distribution area (LDA), including DER that are aggregated into virtual resources by third-party aggregators, where an LDA is defined as the distribution infrastructure and connected DER and end-use customers below a single transmission-disribution interface substation or LMP pricing node. The "Total DSO" would then provide a single bid to the wholesale market at the T-D interface, reflecting the net aggregated needs and capabilities of all resources, customers and marketers[9] within the LDA to buy or sell energy and to offer capacity and ancillary services to the transmission grid.[10]

The TSO would then optimize its system only to balance net interchanges at each T-D substation, without requiring visibility into the distribution infrastructure or specific DER in any given LDA. When the ISO clears the market bid of the DSO and issues a dispatch or control signal based on that bid, the DSO determines how best to utilize the DER within the LDA to respond to the dispatch while coordinating other DER operations and grid needs to maintain reliability and meet customer demands. Thus the DSO would take on responsibility and accountability to maintain real-time supply-demand balance within each LDA, relying on internal DER as well as interchange with the transmission system or wholesale market.

Regions considering the layered optimization will typically feature strong customer interest in adopting DER and policies that support DER adoption. Supporting policies would include streamlined interconnection processes that are not prohibitively costly

---

[9] The term marketers is used here to denote commercial and institutional entities that buy and sell energy and related commodity and capacity services for resale.

[10] De Martini and Kristov, ibid, refer to this as variant C2 of the "Market DSO" model; in other contexts they have used the more descriptive label, "Total DSO."

and revenue opportunities for DER and DER aggregators on both the customer side and the utility side of the distribution system. These factors would in turn provide two conditions needed for successful implementation of the layered or Total DSO paradigm. First, the LDA would need sufficient liquidity and distributed resource diversity to enable the DSO to optimize the local system supply-demand balance. Second, the DSO would have to provide an open access distribution-level market that would aggregate DER offers to the wholesale market, obtain services from qualified DER to support distribution system operations, and enable peer-to-peer transactions within a given LDA and potentially even across LDAs. The layered paradigm thus requires a regulatory framework that will ensure transparency and non-discrimination in the DSO's planning, non-wires alternative sourcing, and operating decisions.

Insights from grid architecture

Grid architecture, for purposes of this article, is the application of the methodologies and tools of system architecture to the design of the future high-DER electricity system. As such, grid architecture insists on a whole-system view which places organizational structure questions, like comparing the two paradigms described above, into a larger context that includes energy and capacity market design, business models, regulatory frameworks, control engineering, communications and data management.[11]

Grid architecture offers some important insights which point to the layered optimization paradigm as the preferred design for the high-DER T-D interface. The first is the problem of tier bypassing, which occurs when two or more system components have multiple structural relationships with conflicting control objectives. In the electric system this can lead to behaviors that conflict with reliable system operation. Case in point, under the grand optimization paradigm, DER in the wholesale market have a market relationship with the ISO that bypasses the electrical interrelationship with the DSO that must respect distribution grid operational and safety considerations. This is reminiscent of the California ISO's original zonal market design, whose forward markets ignored intra-zonal constraints and cleared energy transactions that were not feasible on the grid and had to be unwound in real time. The layered paradigm precludes such tier bypassing.[12]

[11] The present article describes only a few key insights from grid architecture; readers interested in the more complete exposition should see JD Taft and A Becker-Dippmann, "Grid Architecture," Pacific Northwest National Laboratory, January 2015. http://www.pnnl.gov/main/publications/external/technical_reports/PNNL-24044.pdf

[12] See Taft and Becker-Dippmann, ibid, section 5.3.2.4.

A second key insight has to do with scalability, which means the relationship between the TSO and the DSO at the T-D interface can be replicated at lower or higher levels or layers of the system. For example, the layered paradigm can be applied to the interface between the DSO and a micro-grid connected to a distribution circuit, and yet again between the micro-grid and a "smart building" that is one of the components of the micro-grid. At each layer, the optimizing entity (TSO, DSO, micro-grid or smart building) only needs to manage its interfaces with the higher and lower layers without needing to be concerned with the specific components internal to those other layers. The value of such scalability is that it allows for a rigorous coordination framework, whereby the behavior of the components of the system can be coordinated to ensure predictable, reliable performance of the whole system.[13]

Distribution-level locational pricing

The idea of applying locational marginal pricing (LMP), common to today's ISOs and RTOs, to distribution systems is being discussed in the industry. There are two questions that bear on the grand central versus layered discussion. First, when would we want to implement an "LMP+D" regime where the price at a specific location on the distribution system equals the wholesale market LMP at the T-D interface plus a "D" factor that reflects distribution-system losses, congestion and other characteristics? Second, more generally what are the challenges, limitations and potential net benefits to developing a distribution-level locational pricing scheme for energy and capacity provided by DER and responsive end-use customers?

The simple answer to the first question is that the LMP+D approach may be appropriate for the grand optimization paradigm, but would not make sense for the layered optimization. LMP+D pricing is derived from the ISO/RTO market simply by extending the wholesale LMP at the T-D interface into the distribution grid to specific locations of DER and end-users. This could be done by modeling the distribution grid in the central optimization algorithm, but this would require an accurate electrical model of the distribution system and real-time state information, a complex and expensive enhancement that is not yet possible. Or, this could be done more simply by, for example, applying statistical pricing factors (such as loss factors) to determine the D adjustments appropriate to specific feeders or nodes on a feeder. A pricing method derived from the wholesale market intuitively makes sense for the grand optimization paradigm, where most DER are bid into and dispatched by the wholesale market.

---

[13] Taft and Becker-Dippmann, ibid.

In the layered optimization, however, the DSO is balancing supply and demand within the LDA using the diverse mix of available DER within the LDA plus interchange with the transmission system. Thus the LMP at the T-D interface is only the price of imports and exports at the interface, and as such is only one factor affecting the cost of energy within that LDA. Under the layered paradigm, rather than pricing based on wholesale LMPs the pricing methodology should reflect the mix of resource types and customers within the LDA as well as the characteristics of the distribution grid itself, such that the marginal price at a location reflects the least cost of serving an additional kWh of load at that location. In particular the effect of the wholesale LMP in the locational price should go to zero continuously as the LDA's volume of net imports or exports approaches zero.

The second question goes to whether or in what circumstances some form of distribution-level LMP is the best way to facilitate DER operation and investment. There has been much debate over the years about whether wholesale LMPs provide effective incentives to finance generation or grid investment. In general it seems that generation investment requires long-term power purchase agreements and capacity payments, while transmission investment proceeds mostly through central planning and ratepayer funding. If the policy objective is to promote DER investment, there is no apparent reason to believe that distribution-level locational spot prices would have any greater success. Rather, the demonstrated value of wholesale LMPs has been to provide near-term scheduling and operating incentives consistent with grid conditions. By extension, then, this is what pricing on the distribution system should achieve.

If the goal is limited to near-term operating incentives, the design of a pricing paradigm to support reliable distribution system operation must address questions of spatial and temporal granularity. Control theory for large complex systems tells us that there are certain situations where markets, which rely on economic signals to entities that may voluntarily respond, are not sufficiently effective to maintain reliable system operation.[14] Such situations are typically defined by the rapidity of the required response or the signal refresh cycle, combined with the number of entities that receive and must respond to the signal.[15] Such situations are better managed via controls,

---

[14] The authors recognize that some parties to the transactive energy discussions assert that "getting the prices right" and sending the appropriate "prices to devices" would be effective in achieving economic efficiency and reliable operation of the system, without need for a centralized optimization or controls. The authors are not aware of any actual large-scale complex system in which this assertion has been successfully demonstrated.

[15] See JD Taft, P De Martini and L Kristov, "Market-Control Interactions and Structures for Electric Power Grids," forthcoming.

where the control signal elicits a hard-wired response that is predictable and consistent as long as the control system functions properly. In fact, it has been clearly demonstrated in all ISO and RTO markets on the bulk electric system that markets, controls, and integrated market-control structures[16] all have their places and contribute to the well functioning of the whole system. The present article is too brief to allow a fuller exploration of this topic; suffice to say that one must not assume that locational prices in themselves are an effective way, much less the best way, to coordinate the behavior of diverse DER on a high-DER distribution system to maintain reliable operation and meet the needs of end-use customers and other grid users.

Transactive energy markets

The GridWise Architecture Council offers the following definition of transactive energy:

> A system of economic and control mechanisms that allows the dynamic balance of supply and demand across the entire electrical infrastructure using value as a key operational parameter.[17]

The first observation to be drawn from this definition is that the wholesale markets in ISO/RTO regions are already transactive energy systems. The application of “economic and control mechanisms” to manage flows on the distribution system simply extends the scope of transactive energy to encompass DER and everything else below the T-D interfaces. The second observation is that the definition does not require the exclusive use of economic or market-based mechanisms. In fact, as discussed in the last section, a well-functioning transactive energy system would necessarily involve integrated economic-control structures. Spot market mechanisms such as LMP are just one option to determine economic value in such a system.

Thus we arrive at a transactive energy concept which, with high penetration of DER, encompasses the entire electric system from the level of the regional interconnection to the end-use customer meters and even the meters on individual devices located behind customer meters, as these too may participate in market constructs. Moreover, with regard to temporal granularity, the transactive energy system encompasses the entire

---

16 A well-known example of an integrated market-control structure is regulation service in an ISO/RTO market. The market optimization procures a target amount of generation capacity that has been certified to provide regulation. The procured generation receives a capacity payment in exchange for agreeing to receive and respond to four-second signals from the ISO/RTO’s automatic generation control (AGC) system.

17 GridWise Architecture Council, “Gridwise Transactive Energy Framework, Version 1.0,” January 2015, p. 11.

electric system life cycle from long-term resource and infrastructure planning to micro-second control mechanisms to maintain frequency and voltage. Figure 1 provides a whole-system transactive energy model for the electricity grid.

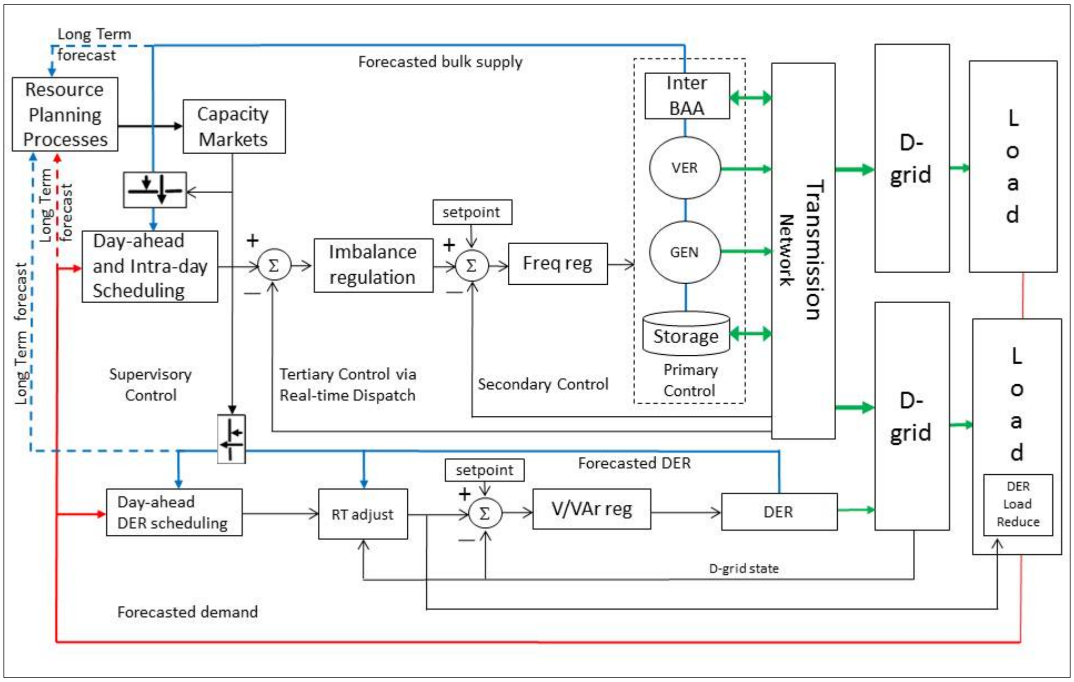

Figure 1 Whole System Transactive Energy Model

The final observation in this section is that peer-to-peer transactions, which are generally posited to be a mainstay of a transactive energy future, can be fully realized under either the grand central or the layered decentralized optimization paradigm. DER and third party DER aggregators could provide services to the transmission grid, the distribution grid, end-use customers and other DER under either model. The only difference will be whether the DER interact directly with the TSO for wholesale market transactions and for obtaining transmission service to support inter-LDA transactions, or engage in such transactions indirectly via the distribution-level markets operated by the Total DSO. Most importantly, the layered optimization or Total DSO paradigm, which is preferred from the perspectives of grid architecture and control theory, will not inhibit or limit the ability of DER to participate in economic transactions, given a regulatory framework that is designed to support such transactions.

A regulatory framework for a transactive energy system

The transactive energy system envisioned in the preceding sections will require several new regulatory framework elements. First, the layered optimization does require sufficient DER penetration and market liquidity within the LDA to enable the Total DSO to balance the local system. For these conditions to develop there must be broad interest by customers and DER developers, supported by an enabling regulatory framework governing that provides for open access in a manner comparable to the federal rules for ISOs and RTOs. Open access and transparency principles must apply to the interconnection process, distribution infrastructure planning, real-time operating procedures, rules for aggregating wholesale market bids from DER within each LDA into a composite bid for the LDA as a whole, and rules for disaggregating and conveying instructions back to the appropriate DER when the ISO market clears the composite bid.

A second key regulatory element will be to allow states to regulate Total DSOs and DSO-operated markets within the territories of their jurisdictional distribution utilities. These distribution-level markets will include sales for resale which are mostly subject to federal jurisdiction under the Federal Power Act. Federal and state policy makers seeking to develop DSO-operated transactive markets should consider how this provision might be modified legislatively to allow state regulation of transactions that are completely within a single LDA and thus do not rely on the transmission system. This could enable states to regulate Total DSO transactive markets and thereby facilitate their wider adoption nationally.

A third key element involves developing shared accountability for reliability between the TSO and the Total DSO. If the TSO optimizes only to the net interchange at each T-D interface while the DSO performs supply-demand balancing within each LDA, as per the layered paradigm, then the regulatory framework for the DSO must assign responsibility and accountability for reliable service to the end-use customer in a manner that goes beyond reliable distribution wires service. This may play out differently in different states. One possible scenario is where the DSO is also a load-serving entity with provider-of-last-resort responsibility, as is the case with most restructured distribution utilities today. Another possibility is that the DSO is a pure wires company, like the investor-owned utilities in Texas, and as such may be the entity that enforces supply adequacy requirements on load-serving entities within each of its LDAs. This could lead to a new resource adequacy paradigm based on bifurcated jurisdiction: the FERC-jurisdictional ISO/RTO enforces resource adequacy for load-serving entities in its footprint commensurate with their shares of net system demand measured at the T-D interfaces, while the state-jurisdictional Total DSO enforces resource adequacy for load-serving entities commensurate with their net load within each LDA.

The way forward

To date, industry discussions about the future transactive energy system with high DER implicitly assume, at least for ISO/RTO regions, a version of the grand central, Minimal

DSO paradigm. This would be a straightforward extension of how demand response participates in wholesale markets today. Probably for this reason, and for lack of a well specified alternative, its implementation challenges and potential shortcomings have not been seriously probed. But, implicit unquestioned assumptions in the design of a complex system like electricity can have disastrous consequences at the scale of DER adoption we can anticipate based on historical trends. It is therefore crucial for the industry to address the question of how best to design the functional roles and responsibilities of a DSO vis-à-vis the operator of the transmission grid and wholesale markets around the T-D interfaces. To that end the authors have presented the layered decentralized model as an alternative to the current trajectory, and offered reasons why the alternative is preferable and can facilitate the successful evolution of the high-DER transactive energy electric system.

That said, there are some areas requiring further work which this article only has space to mention. These have to do with the evolutionary path to the layered structure. In many ISO/RTO areas expansion of DER participation in the wholesale markets is already underway. The authors do not suggest an abrupt shift of the operational paradigm, but we do urge an abrupt shift of thinking to view and then to shape the present trajectory as a transition to a future layered optimization structure.

The pace of the transition will vary, depending on the rate of DER growth among other things. Different regions and states – even different cities and counties in a given utility service territory within a state – will approach the transactive future at different rates based on their customer mix, climate zone, natural resources, geography and public policy goals. Still, there should be common elements and strategies based on the laws of physics governing the movement of electricity through wires and transformers, the tools of grid architecture and control theory, and widely-held goals such as reliability, system security and resilience, efficiency and affordability.

As a concrete example, the staging of distribution infrastructure investment needs to be considered within a holistic framework that includes market design, telecommunication and control strategies, business models, economic incentives and public policy goals. The sensing and control needs of a grid with low penetration of DER and little or no market activity are much simpler than those of a system with greater amounts of DER some of which provide operational services to the DSO. And the needs of the latter are simpler in turn than those of a system to support peer-to-peer transactions among DER. Similarly, the design of markets at distribution level must be linked directly to the policy goals of the jurisdiction. If the goal is to incentivize DER expansion, are locational spot prices an effective strategy or is longer-term revenue certainty required?

In the electricity system the above questions are all intertwined. To ignore the holistic discipline of grid architecture at this time of rapid change is to risk massive stranded investment in market and operational systems that work poorly, delay or even derail beneficial system evolution, and fail to achieve desired societal goals. The value of adopting the layered decentralized optimization as the end-state operational paradigm at this time is that it grounds industry transformation and transactive energy discussions in physical reality and thus provides a foundation for addressing the entire constellation of regulatory, market design, business model and infrastructure investment questions.