QT-WP-003 032016

WHITE PAPER

Market Design for Congestion Relief With Energy Storage

Author: Ralph Masiello, PhD
Industry Advisor
Quanta Technology, LLC

Market Design for Congestion Relief With Energy Storage

Executive Summary

This paper explores how energy storage could be added to the portfolio of traditional transmission solutions when the driver for the investment is relieving transmission congestion. It demonstrates how storage can be operated so as to mimic the effects of those traditional solutions (such as transmission construction and line re-conductoring) to alleviate congestion. As energy storage is portable, it is more flexible than other solutions. The paper also makes an argument that storage is not necessarily a generation asset in this context but can also be considered a transmission asset, and as such should be developed via capacity markets or a regulated investment process as opposed to relying on merchant investment. The paper is focused on the transmission congestion relief application and not on other wholesale market applications which are generally understood to be merchant functions such as energy arbitrage and ancillary services. However, it is important to note that the full benefit of the technology is achieved through combined or shared applications resulting in the optimal return on the investments and benefits for the rate payers. In today's regulatory environment, storage is forced into the traditional Generation-Transmission-Distribution-Consumer buckets, with none fitting well the optimal use of the technology.

The congestion relief application is examined in terms of avoidance of scheduled thermal overloads – that is, ensuring that the dispatch schedules do not violate line limits; and in terms of N-1 contingency constraints. The latter is potentially a more attractive case economically if certain technical and standards issues can be overcome. Storage can be used to avoid line overloads in the scheduling and dispatch process as a resource that "transports energy in time" – shifts and reduces the peak on the transmission system. Storage for N-1 contingency relief is a "short term transmission reserve" that shifts the requirements for generation from being on line and scheduled to being a quick start reserve, which may be considerably less expensive overall. The storage needs to be more effective than the traditional alternatives to remedy congestion.

Storage is an asset characterized by high capacity costs and relatively low operating costs (depending upon the charge/discharge and self-discharge efficiencies of the technology) – much like transmission. The losses associated with the charging cycle and self-discharge are analogous to resistance losses in lines and transformers and to hysteresis or core losses in transformers. The capital costs of transmission assets are compensated via rate of return as a regulated asset or via cost recovery of payments under a PPA agreement. The operating costs – maintenance and losses – are either recovered as operating expenses (maintenance) or socialized in the market (losses), reflected in nodal prices, and recovered in wholesale energy costs inside total consumer billings. Different ownership / business models for storage as a transmission asset are discussed and the risk profiles of each described qualitatively. Approaches to the planning and evaluation of storage for N-1 contingency relief are explored.

In summary, we have examined the use of energy storage for congestion relief and made a case that this is an application better performed by a regulatory asset. As such, it becomes one of the portfolio solutions, (such as new transmission lines, re-conductoring to increase the line flow, FACTs or other T&D devices to redirect the flow) to allow for optimal management of assets to achieve reliability, safety and efficiency targets, as well as realize shared applications benefits of storage.

About Quanta Technology

Quanta Technology is independent business and technical consulting services company specializing in the electric power and energy industries. Our experts have years of practical working experience in the electric power industry, and are trusted to develop custom solutions to the industry's most challenging and complex problems. As an independent arm of Quanta Services, we are able to deliver comprehensive solutions that span the spectrum from management consulting, all the way through to engineering, procurement and construction.

Introduction

Storage improves asset utilization and displaces generation, transmission and distribution assets that are only utilized fully at peak periods. Improved asset utilization and the ability to smooth out variability make it valuable in facilitating renewable integration. The technical performance of some storage technologies (lithium ion batteries, flywheels, for example) make storage eminently suitable for ancillary services, especially regulation services.

However, with the exception of regulation services and possibly solar plus storage – the technology deployment is still in pilot and/or early adopter stages. Many key and promising technologies are still under development with manufacturing feasibility yet to be demonstrated. Most state regulators as well, are not well-informed or in a "wait and see" mode on storage. The lack of standards for safety and for interconnection is also a barrier.

In today's regulatory environment, storage is forced into the traditional Generation-Transmission-Distribution-Consumer buckets, with none fitting well for combined or shared applications or fully realizing the potential of the technology. Storage has been considered a generation resource whenever it is used to provide market products normally provided by generators. Congestion relief has not been extensively studied, but in the context of the wholesale markets congestion and nodal pricing are driven by the supply side bids coming from generators and demand response (which is treated as a supply side resource).

This paper deals narrowly with the issues around whether storage assets whose primary purpose is to increase T&D asset utilization or capacity factors – i.e. to relieve congestion – should be considered T&D assets, which would then qualify for regulatory treatment as such under a regulatory cost recovery model. Relief from congestion and N-1 contingency congestion is shown via illustration with simple cases to be technically equivalent to a new transmission asset. The time arbitrage effects of shifting delivery of remote generation from off-peak to on-peak can be viewed as "transportation in time" with the effect of delivering inexpensive off-peak energy during on-peak periods and with all economic benefits accruing to consumers

Storage as a T&D Asset

The argument in favor of storage being a T&D asset has several points that are not found, so far, in the literature:

- Storage which is deployed to reduce transmission congestion can be operated so as to have identical effects on market nodal prices as a transmission asset, at least in terms of when it is used to “deliver” congested energy to load points. This is illustrated with some simple examples. This makes the point that it is a valid Non Transmission Alternative not dependent upon market dispatch for operations
- High efficiency storage (advanced batteries) falls into a general asset class where the marginal cost of “using” the asset is very low but the capital cost is high.
  - ◊ Operated as a market participant, it requires very high gross margins on its usage or production in order to recover capital costs. The industry has moved, to varying degrees in different market geographies, to capacity markets as a way to ensure that peak shaving / peak generation capacity is available for usage a small number of hours a year. This is an alternative way to enable these resources to recover capital costs.
  - ◊ Market distortions when zero marginal cost resources dominate the market at some hours of the day/days of the year are already visible and show up as negative pricing based on operator costs of not operating and/or the value of incentives (production tax credit, renewable energy credit, etc.) at risk if the resource does not participate. There has not been serious analysis made of market behavior in the presence of very high penetration of zero marginal cost resources and frequent periods where these resources are “on the margin” – but the likelihood of undesirable market behavior is real.
- The best way to operate storage used for congestion relief where the marginal cost of using the storage is very low is to co-optimize it in the market clearing/scheduling solution with the true marginal costs (operating losses; life cycle depreciation if applicable) as inputs to the market clearing overall cost minimization problem. The market operator is managing the delivery of energy across time as well as distance, as opposed to the market participant buying low and selling high. This eliminates the need for the storage operator to produce market bids for charging or discharging, eliminates the risk of mismatched charge-discharge cycles when the storage operator must bid. It also resolves a fundamental problem around the storage destroying the arbitrage “value” at high penetrations—which is good for the ratepayer, but bad for the merchant operator.
  - ◊ The arbitrage gains on the locational value of discharged vs. charged energy net of losses can then be socialized in the market, distributed to ratepayers as adjustments to net energy costs, and/or shared with the operator on a regulated basis that avoids double counting of benefits.
  - ◊ The storage operation schedule that maximizes energy arbitrage gains will in general *not* be the same as the storage schedule that minimizes overall system costs. Thus, a market-driven perfect bidding strategy by the storage operator is less efficient than a market co-optimized solution. This is true even if the storage operator has perfect and complete information (which would rarely be the case.)
  - ◊ This also avoids the hypothetical problem of a storage operator as a market participant that also owns generation resources and who might bid or even invest in the storage strategically so as to protect congestion revenues from favorably located generation. (So far there is not sufficient storage deployed to make this more than hypothetical, but it is another example of how markets and regulatory development are lagging technology development)

California’s mandate for 1.3 GW has been a huge driver for manufacturers, developers, and utilities. A precedent of storage as regulated asset has been implemented in California to promote deployment. For example, CPUC has allowed PG&E to own 50% (out of a total mandated 580 MW of storage in the footprint) and treat it as a regulated asset that could be a market participant as well.

Barriers to Storage Adoption

Economic

- Battery costs
- Balance of System s costs
- Lack of easy to use/accurate valuation tools

Technical

- RTO & NERC position on N-1 congestion relief technical requirements
- Standard models and methods for storage in planning and operations

Regulatory

- Valuation buy-in
- Storage as G-T-D-C vs. storage as asset class
- Market vs. rate base functions

Storage for Congestion Relief as Compared to Added Transmission Capacity

Here we present a simplified example to illustrate that storage resources in the T&D network can be operated so as to have the same effect on market prices as a new transmission line when they are relieving congestion by charging "upstream" and discharging "downstream. "

Consider a generalized network as shown in **Figure 1.**

Line X from Bus A to Bus B is congested – let us assume that the least cost dispatch loads this line to its thermal long term rating. This means that the nodal price at B, C, D, F, and other nodes downstream will be higher than the price at bus A and that some generator downstream is dispatched out of normal merit order and setting that congestion nodal price. This relieves the congestion. It will also alter the flows on lines downstream and change nodal prices at buses C, D, E, and others downstream.

Figure 1

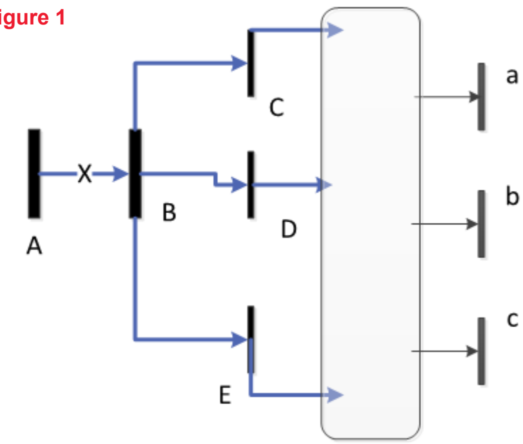
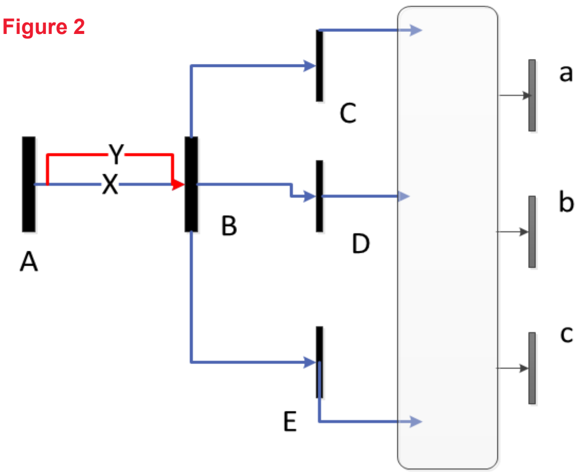


Figure 2



One way to relieve the congestion is to add transmission capacity – let us say, to parallel line X with a new line Y as shown in **Figure 2.**

If instead of adding line Y we add a negative injection at bus A equal to the flow on line Y from A to B, and a positive injection at bus B equal to that flow – we get exactly the same network conditions throughout (including nodal prices elsewhere in the network). The positive injection at bus B could be provided by a generator, but the negative injection at bus A could not be. However, a storage device at each bus charging (bus A) and discharging (bus B) would have the identical effect and would have the same impact on the network as the addition of line Y. This is shown in F**igure 3.**

BUT – the second, discharging battery does not have to be placed at bus B. We could put batteries at the ends of the lines from B to C, from B to D, and from B to E as we did with line Y – where the injections / withdrawals are equal to the changes in the flows on those lines as a result of line Y or of the two batteries. The "from" batteries will cancel out the discharging big battery at bus B and we end up with a situation as shown in **Figure 4** below. By continuing this superposition, we can get to the extreme case of a battery at each of the load take out points (or distribution busses or whatever) a, b, c (and other similar small buses at lower voltage) where the discharging of those batteries (or in some cases theoretically possible charging) is such that the line flows on the network, and the resulting bus conditions are what they are were line Y built.

Figure 3

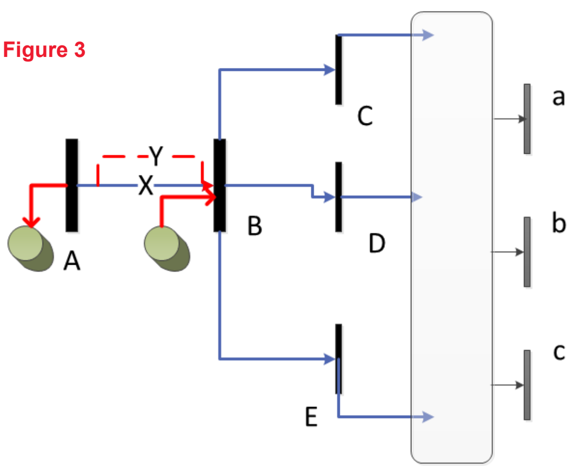


Figure 4

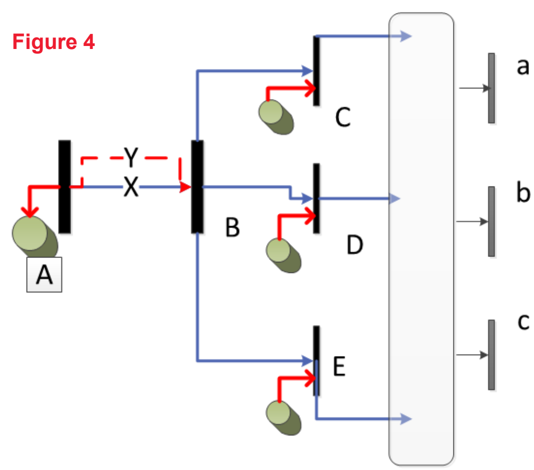
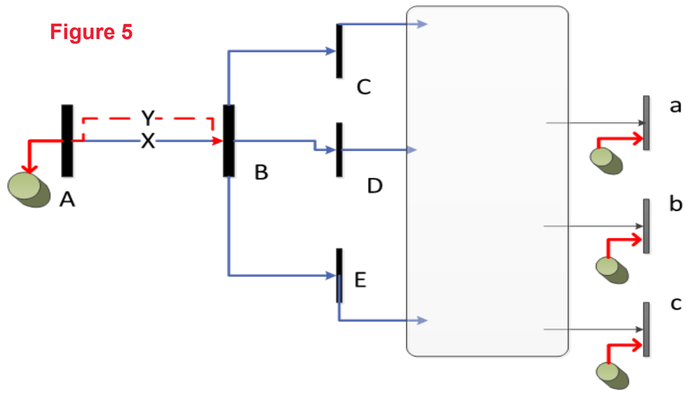


Figure 5



For every different set of base conditions on the network and nodal prices as a result of line X congestion – there will be different downstream outcomes and the discharging values of the downstream batteries will be different. But it is sufficient to size the batteries for the worst case outcome and then operate them so as to duplicate the effect of line Y on the system.

This argument shows that logically the storage devices can be operated so as to mimic the congestion relief of adding a transmission line.

Line Y, if added, alters the dispatch and the nodal prices. The batteries have exactly the same effect so have no adverse effect on market participants, note. Thus if the transmission utility builds the batteries instead of line Y – it is logically and market impact equivalent.

When X is not congested – the addition of Y does not change nodal prices at all so is neutral to the market. The operations of the batteries to discharge and charge in the reverse of what they did when X is congested should have neutral impact as well.

The above statement is valid only if the size of the batteries and the time duration of the original congestion are such that the batteries can be discharged/charged during uncongested periods without adding to overall congestion. Said differently – if the line in question is congested 24 hours a day, then the batteries cannot be operated consistent with the above argument, and in fact storage is not a solution at all. It will only make economic sense to deploy and operate storage consistent with the overall congestion relief economics. If the line id congested more than a few hours a day, storage is not likely to be an economical solution given today's storage costs.

In today's market, the merchant storage operator has to schedule the charging (as a load) and bid in the discharging (as a generator) before the congestion dispatch results are known – so the storage as a load/generator model will definitely NOT mimic the addition of line Y and will be suboptimal in a dispatch sense.

But – if the storage devices are **transmission assets** and the market operator can co-optimize their use in the market as part of the congestion dispatch – then the optimal solution will be incorporated in the dispatch. In this model the losses of the charge-discharge cycle are part of the co-optimized dispatch as would be the line losses on line Y.

Considered as a Non-Transmission Asset (NTA), storage would have to be shown to be more economical and reliable than simply building the line Y, or more feasible if siting issues make building line Y problematic.

The Case for Market Co-Optimization of Storage to Minimize Congestion Costs

The above logic assumes that the market operator would schedule the batteries so as to exactly mimic the network behavior of an equivalent transmission line. In theory, this could be done, and could be extended to include more than one hypothetical new line. In practice, it would seem absurd – far better that the market operator schedule the batteries to minimize costs overall which in the perfectly efficient case is precisely the same as operating the batteries to minimize congestion costs.

This is in contrast with the storage operator as a market participant who will operate the device so as to maximize energy arbitrage profits. Let us assume that the source of the upstream generation is all renewable generation – with zero marginal costs, and that the source of the downstream generation is (a) a small but efficient combined cycle (CCGT) unit and (b) a collection of peakers. If the peakers normally set the congestion price (assuming zero for the renewable generation), then the storage operator would deliver only enough power such that the last remaining peaker still set the marginal price. If enough power were delivered so as to displace the last peaker, then the much more efficient CCGT would set the price and the arbitrage value of the first tranche of stored energy would drop to the CCGT price. The storage operator would never deliver so much power as to drop the downstream price to near zero and eliminate the arbitrage gain altogether - that is not in the storage operator's interest at all. This logic, carried out, implies that the storage resource would have to be a QF with constrained operations and be under regulatory rate recovery instead. Once sufficient storage were deployed to "nearly but not quite" displace the peaker, any prospective storage investor would quickly realize that additional storage deployment would destroy the arbitrage value and make the investment unattractive. This is the same logic that wind farm investors go through in performing curtailment studies before investing – to see if congestion will render the additional generation less valuable.

When Congestion is a Result of N-1 Contingency Dispatch

Now let's consider **Figure 6.** In it the N-1 contingency is the outage of line Y. Let's assume that the lines tying buses B and E together are strong enough that for all intents and purposes lines X, Y, Z are inherently in parallel, and let's further assume that the thermal ratings of X, Y, and Z are identical. In order for the network to withstand an outage of any of X, Y, Z – the post contingency flows on the lines in service (X and Z after an outage of Y) must be within ratings. So if the rating is "*R*" then the post contingency transfer from A to the network at B, E must be < 2*R*. The pre-contingency flow must have been 2*R* therefore, so that only 2/3 of the thermal rating was in use.

Another solution which may be cost effective is to have 0.5*R* of fast storage available downstream – located at buses B, C, D, E variously. As shown in **Figure 7,** this storage is there to provide temporary post-contingency loading relief in the event of an outage of line Y. The two batteries shown are connected to the buses B and E, and normally are not charging or discharging. On an outage of lines X, Y, or Z the batteries immediately begin to discharge at a power level of 0.5*R* and thus avoid any overload of the two lines left in service. The batteries only have to sustain this level of discharge until the combustion turbines downstream can be started and brought into service to provide the same congestion relief as would have occurred under "normal" N-1 contingency dispatch.

Figure 6

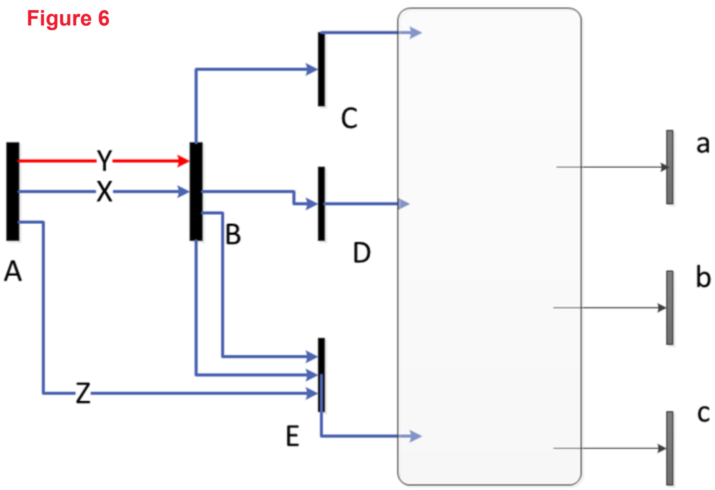
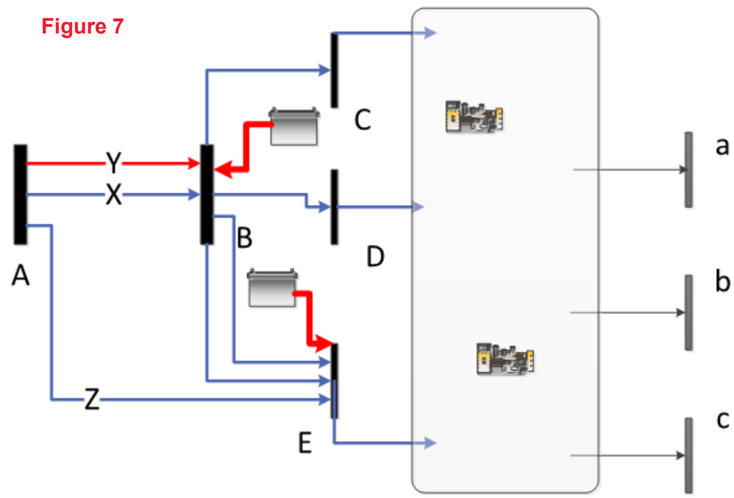


Figure 7



If the total demand that needed to be served was 2.5*R* and it all had to be served via X, Y, Z – then we have congestion of 0.5*R.* And to go back to the simplistic situation described in the first discussion above, let's assume that the generation upstream of "A" is hydro and that the generation downstream of "B" is CT based – so we have a congestion cost of 0.5*R* * $CT where the $CT is the marginal price of generation from the most expensive CT required to get collective downstream generation up to 0.5*R*. But that price will apply to all the load – to 2.5*R.*

One solution to the congestion cost is to build a much more efficient generation capability downstream. And the same logic about merchant investment decisions applies here—inventors will be leery of reducing the congestion price and thus the value of the new generator. But siting issues and potentially fuel availability often act to prohibit this option. A second solution is to build a new transmission line paralleling X, Y or Z to add capacity; or to repower/reconductor XYZ to increase ratings. Siting and technical feasibility can make these choices difficult or can delay them for years.

This application can harvest the congestion relief benefits across all the hours of the year that the lines are congested, and do so with very low operating costs, just as with the addition of a new transmission line. The cost of the batteries has to be weighed against the cost of additional transmission capacity to see what the relative economics are.

From an overall reliability viewpoint, the additional transmission line is the bullet-proof most robust solution. If batteries and quick start generation are used instead, the grid operator has to be confident that the batteries will begin to generate when demanded, and that the quick start generator will start – so come capacity rating factor has to be applied to these resources. The batteries and associated power electronics are fast, reliable, and can be tested as frequently as desired. The generators would be subject to the same regimes as quick start reserve units.

The inverter technology and the appropriate battery technology would be able to respond to a line outage about as rapidly as normal protection devices would, so speed should not be an issue. The battery technology would also have to be chosen for ability to sit fully charged for long periods with minimal self-discharge. (As is the case with Li-Ion batteries today)

Added to the congestion cost savings of this scheme, there would additionally be an emissions benefit as the downstream generation would not be operating. For urban areas where congestion is a constant or near constant condition, this is significant.

In this case, the N-1 congestion relief is a *transmission* application. There are zero market based ongoing revenues available under today's product definitions from delivery of stand-by energy services or ancillary services. So there is no market mechanism to incent a storage developer to deploy batteries for N-1 congestion relief. Additionally, the most natural location for the storage units could well be inside the existing substations. If that is the case, then the transmission utility has ownership.

There are several possibilities for "acquiring" fast storage for N-1 relief. In all cases, the grid operator and NERC would have to approve the scheme, including any direct automatic control schemes.

One mechanism is for the transmission utility to deploy, own, and operate the storage resource as a regulated asset. The grid / market operator would procure the quick start reserve capability from downstream conventional generation.

A second possible scheme is for the grid operator to conduct a capacity auction for fast storage in conjunction with quick start generation for N-1 congestion relief. Because the siting would likely be quite specific, this is practically the same as the transmission utility conducting a competitive procurement for equipment supply/installation, with possibly different terms of contracting as opposed to the lifetime of a rate-based asset.

A third possible scheme is for the market operator to define a new ancillary service – “instantaneous reserve” targeted at inverter/battery resources; and procure it as a market service on a site specific basis. The market would have to co-optimize across procuring this locational instantaneous reserve vs. scheduling downstream energy and paying a congestion price. This mechanism would allow the market to continue to apply nodal pricing that reflected the locational reserve cost as opposed to socializing it as a transmission investment. In theory informed investors would make decisions to enter this market based on the expected revenues vs the capital and operating costs. Note that if the storage were distributed further downstream, it would look like a near instantaneous load reduction or load shedding on a net basis with the same congestion benefits. As such it could be owned/operated by the utility or procured from aggregators or others.

Treatment of Inevitable Time Arbitrage

There is still a fundamental difference between the storage resource and the transmission line as means to relieve congestion. The new transmission line relieves congestion by increasing instantaneous delivery capacity - it increases peak transportation capacity. The storage resource relieves congestion by improving transmission capacity utilization – shaving the peak transportation demand by displacing energy delivery in time. Note that it is not altering total energy demand or total deliveries (except for operational losses as with the transmission line itself) – it is shifting peak deliveries in time.

This may provide a distinction between storage as a peak shaving-time arbitrage resource and storage as a congestion relief locational arbitrage resource. If the only benefit of storage is the “buy low sell high” time arbitrage accomplished by shifting demand from peak to off peak, then arguments can be made that it is akin to Demand Response and should be treated as such for regulatory jurisdiction and asset treatment purposes.

However, a counter argument can logically be made to this view of peak shaving. New transmission forces generation resources to compete with other distant generation resources that could not compete based on available transportation. Storage forces expensive generation at peak to compete with less expensive generation off-peak. So a different market treatment of the time arbitrage cold be to deliver the inexpensive off-peak generation to the market on-peak at the off-peak price, adjusted for the costs associated with storage. That is the “time transport” equivalent of what new transmission does for “distance transport,” and assures that all the benefits accrue to consumers. When sufficient inexpensive generation is delivered on-peak, the peak prices are depressed and all consumers benefit—exactly as with transmission.

Ownership of the Asset vs. Ownership of the Stored Energy

So far we have discussed the “merchant storage” and the “regulated transmission storage” model. In the former the merchant operator has the risk of capital recovery for the physical asset as well as the risk (and potential gains) on the time arbitrage of energy or the revenues from market services provided. In the latter the ratepayer effectively has the risk of investment prudence and the market has the control of time arbitrage with all benefits accruing to ratepayers. There is a third way that exists in gas transmission today.

The RTO and the utility could determine the need for storage capacity for congestion relief and or peak shaving and then arrange for the deployment of storage via a NTA process, possibly a capacity auction.

The physical asset is then a rate based asset either utility owned or developer owned sold to the utility as a PPA. The ownership of the energy and scheduling of the storage is distinct and is operated as a merchant function available to market participants – not unlike the use of gas storage in transmission lines. The merchant trader then schedules the charging and discharging via bilateral agreements or via bids to the market. If the most attractive use of the storage is for N-1 contingency relief via an “instantaneous reserve” product then the merchant trader would make a tactical decision to offer that service preferentially over time arbitrage or the market operator could co-optimize this as is the case today with other ancillary products. Storage becomes an adjunct to transmission that market participants can use to balance variable and lower cost energy with demand, or to manage delivery and transmission capacities.

The argument so far has been that the storage resource should be treated as a zero-marginal cost dispatchable or controllable resource under the complete scheduling and operation of the market / grid operator – not a market participant. That is to say, storage for congestion relief should not be operated as a merchant asset, nor should the decision to invest be made by a merchant operator seeking asset recovery and net profit based on the locational and time arbitrage. Consequently, the decision to deploy the storage asset should be the result of a planning process that leads to identification of the need for the congestion relief capacity and a determination that storage is a viable alternative to traditional transmission technologies.

The decision between a rate-based utility asset vs. a QF merchant asset comes down to cost and risk. Under utility ownership, the technology/need risks are directly with ratepayers. Under merchant ownership, those risks are translated to higher costs for financing and faster capital recovery.

To the extent that storage for congestion relief is a critical reliability asset, and one which cannot easily be replaced by a different storage resource in a different location, utility operational control is desirable. That argues for utility ownership.

Another factor in the ownership question is around cost recovery for the initial planning and engineering. Determining the best locations for storage for congestion relief and the sizes of the storage will depend upon several factors:

- How many lines/transformers are affected by a given contingency or set of contingencies
- What nodes afford the best opportunities to reduce contingency overloads for a set of contingencies if that is desirable
- Siting issues for large batteries at or adjacent to the selected substations – space, safety

The engineering studies to determine storage sizes and locations are non-trivial and will ultimately require RTO/ISO validation and approval in terms of market benefits.

Performing an analysis of storage for N-1 congestion relief today in terms of a cost benefit analysis versus other transmission and non-transmission alternatives will require some new developments in planning and production costing simulation tools. Most planning and production costing tools today do not handle the co-optimization of storage in the markets comprehensively. The one or two that do will not consider storage for N-1 congestion relief as no market product for this other than generation re-dispatch exists today for these tools to simulate. Once the tools are able to simulate this new function – referred to above as "instantaneous reserves" then a portfolio optimization capability is required to determine the optimal locations and sizes of storage against a set of N-1 contingencies to be mitigated. In other words, this is a non-trivial problem. A determination of favorable cost benefit in some high profile / high congestion cost cases would be a first step in persuading the industry to take a serious look at this alternative and the necessary control and market infrastructure to exploit it.

Conclusions

We have examined the use of energy storage for congestion relief and made a case that this is an application better performed by a regulatory asset. As such, it becomes one of the portfolio solutions, such as new transmission lines, FACTs and other T&D devices, to allow for optimal management of assets to achieve reliability, safety and efficiency targets, as well as realize shared applications benefits of storage in an optimal way.

A merchant asset will inherently be deployed and operated so as to maximize the improvement in economic surplus for itself and to minimize the amount that accrues to the ratepayer. In particular, it is possible to entirely destroy the arbitrage value of the storage resource used for congestion relief, so that the market nodal prices are not a good incentive mechanism for storage operation . A regulatory asset whose costs are recovered as such will not have this attribute.

Storage is a class of asset that is characterized by high capital costs and low or negligible operating costs. As such, forcing it into a "Generation" asset class where market participation dictated by marginal cost is a bad fit, and at high penetration may be counter-productive. The next logical step in discussing this new concept is to develop a cost benefit analysis of storage for N-1 contingency congestion relief.

For More Information

To learn how Quanta Technology can help in understanding how storage can assist with performing N-1 Contingency Congestion Relief, in evaluating alternatives and performing market product design, please contact us by visiting: www.quanta-technology.com/contact-us

Quanta Technology, LLC
4020 Westchase Blvd.; Suite 300
Raleigh, NC 27607
+1 (919) 334-3000
www.quanta-technology.com