Change of paradigms in complexity and interdependencies of infrastructures: The case for flexible new protocols

Marija Ilic, EECS Department
MIT, Cambridge, MA
ilic@mit.edu

Abstract
In this paper a brief assessment of the technical, economic and regulatory paradigms underlying design and operations of complex infrastructures is provided. First, a description is given for the industry as it was. This is followed up by the industry in transition driven by technological advances and regulatory restructuring. It is concluded that a major qualitative change of paradigm has taken place. The paper describes the fundamental nature of this change, and its implications on the evolution of the infrastructure systems, their performance. Although the illustrative examples are mainly from the electric power industry, the change of paradigm described here can be found throughout the network type infrastructure industries, such as the airline industry, gas/fuel supply and delivery, and, more generally, transport industries. The paper identifies a major missing R&D piece well posed and flexible industry protocols for interactions among various newly evolved, unbundled sub-entities within each specific infrastructure, as well as for interactions among re-bundled entities across several infrastructures. It is suggested that these are essential for the infrastructure industries under the new operating paradigm to manage hidden complexities for predictable performance, technical and economic.

1. Background

It is not an overstatement that never before in their history have major transport infrastructures been presented by such tremendous new opportunities for innovation, and, at the same time, by the hidden threats related to these innovations. The overall process could be summarized as the one of going from a well understood top-down, hierarchical, relatively low risk design and operations mode to ultimately open-access transport network architectures mode with many distributed decision makers, with little explicit coordination , yet, at same time, capable of using inexpensive near real-time information technologies of various kinds in support their learning and adopting to the other entities within the systems of infrastructures.

A closer look into our knowledge of these two modes of interest quickly reveals that neither of the two are possible to design and operate for guaranteed performance. The reasons for this are many, and different depending on which mode is in place; yet, no set of rigid technical standards for designing and operating different components within the infrastructure is capable of ensuring guaranteed performance of the infrastructure (or systems of inter-dependent infrastructures). The state of the infrastructures is generically stochastic, which could be a result of small unintended fluctuations in supply/demand patterns, and/or larger un-intended changes of status of major equipment , and/or major natural disasters affecting the equipment status, and/or even intended attacks on the infrastructures (including terrorism). These stochastic status of the infrastructure status presents their designers and operators with a continuous challenge; the current state-of-the art in systems theory does not lend itself to designing and controlling the infrastructure systems for guaranteed performance; this, then, leaves the challenge to the human decision makers and their expert systems knowledge as the major means when the infrastructures are subject to significant random changes. The result is, all together, a best effort performance.

2. The old paradigm: Top down design and operations, and the resulting complexities

One could identify several major sources contributing to the fundamental complexities. To start with, a typical transport infrastructure could be characterized as a very large-scale network system, in which a change (disturbance) somewhere on the system affects everything else on the system. This makes the basic monitoring, data collection and data management into a useful information for decision making very challenging. Much R&D has occupied researchers over the years toward developing often infrastructure dependent data bases and methods for their use. Second, as a transport network gets exposed to a non-trivial disturbance, it no longer responds linearly; instead, depending on a scenario, a new equilibrium may not exist, or the equilibrium could be reached only by control actions which do not assume ‘‘expected linear response of the network to the disturbance and/or control at its nodes. We have witnessed many real-life examples in which major problems have been triggered by assuming a linear response; in the electric power industry the most obvious example relates to the first blackouts related to the voltage collapse experienced in France, Belgium, South Africa, and elsewhere; this process was triggered because the automated control devices had feedback whose logic was set for the assumed power network response to change, in this case reactive power demand an increase in demand was assumed to always lower the voltage, and vice versa. The actual nonlinear network response at an operating point far from the operating point for which the system was originally designed (including the controller logic) was just opposite, and the consequences were detrimental. It is relevant to observe that the voltage ‘‘collapse event did not take place right away, in other words a time critical event when no control could keep the network intact only occurred after a sequence of wrong quasi-static actions under the wrong assumptions. This points into the need for developing tools which adopt all the time under well understood system conditions, and if this is done, truly detrimental events could often be avoided. This observation becomes very important for the overall recommendation made in the conclusion of this paper; instead of having very rigid, non-adaptive standards, it is essential to have sufficiently flexible protocols for operating and designing the system to constantly respond to changes in an intelligent way, and keep the system away from the truly detrimental states. The lack of such adaptive methods is striking in current practices of typical infrastructures.

Third, is very difficult to develop flexible methods for near real-time fault detection at the levels where the decisions are made; for example, in the electric power grids a fault could be geographically very remote from a control center where the supply is dispatched for the anticipated demand, for the assumed grid status. The state estimation is one of the basic tools which needs further development; it is already being used for fault detection, rather than monitoring of near real-time supply/demand patterns. Much more work is needed in this area.

Again, using the example of electric power networks, we suggest that it is very difficult to develop general methods for dealing with these fundamental sources of complexity. The R&D approach taken in this industry has been to develop the tools for data monitoring and processing, system control and estimation by drawing on unique characteristics of large scale electric power systems; for instance, under certain assumptions a response to a disturbance is localized, and the algorithms which take this into consideration are more effective than the ones which don t; under a wide range of operating conditions the electric power network responds as a monotonic network, and if the conditions under which this is true are monitored as methods under this assumption are employed for controlling the system, this adds tremendously to overcoming the fundamental complexity one started with; similar simplifications are possible when developing methods for state estimation, and researchers have over years attempted to develop these shortcuts to estimating a system fault, for example.

2.1 General reliance on expert knowledge of human operators and system designers

Based on the above assessment of the basic complexities, and the state of-the-art in dealing with these complexities, we suggest that it is almost impossible to rely on reliable models, and model-based decision tools for operating typical network infrastructures under stress (the network is subject to a relatively large change of status). It is therefore practically impossible to expect a guaranteed performance of an infrastructure under stress. (This is despite the fact that significant progress might have been made for particular classes of infrastructures.

Because of this, the current approach in practice has been to rely heavily on the human expert system actions in various emergencies. Operators generally do many off-line scenario studies, and prepare ‘‘nomograms , or a set of real rules based on this studies in case of a significant system fault. This approach is not sufficiently robust in cases when the underlying assumptions made by a system operator do not hold, such as localized response to a fault, expected monotonic network response to supply/demand at its node, etc. Therefore, the human operator is not in a position to guarantee technical performance in real-time if this exact scenario was not studied ahead of time. Moreover, the typical tendency has been to design conservative rules for operating in normal conditions, so that there is sufficient reserve to meet the demand under emergencies, without having to rely on near real-time decision making by suppliers and/or users. This conservative approach, referred to as preventive has led to relatively robust technical yet generally sub-efficient, low-risk performance. It is important to keep this in mind as one talks about the change of paradigm for the industry in transition, since the later is characterized by more dynamic decision making in which risk taking has a value in its own right.

2.2 Traditional policy/economic structure

It is interesting to look into the policy/regulatory framework behind the top-down technical design and operations. The two go hand in hand. The cost plus regulation in which large equipment capital investment is encouraged, and the cost recovery is fully guaranteed independent of the actual use of this equipment in operations, has been the basis for what one may consider an initial over-design for technical robustness. The economics of this had partial justification because of the economies of scale effects, which rest on the cost reduction by building larger units. The costs were paid by the users on an averaged basis, in which large industrial users generally ended up subsidizing smaller customers. Design principles were for the estimated demand growth and for peak demand, with little incentives to increase the load factor (difference between the average and peak load). The robust planning was made so that users get un-interrupted service even when one, or two major equipment components failed (except in the case of natural disasters, and/or intended attacks). Another relevant feature of planning was the ultimate responsibility to the native users, and cooperation with the neighboring subsystems in order to share burden of reliability when something major failed. The companies in charge of infrastructures had to typically present the case for new major investments, and the states would decide on behalf of customers what is acceptable future cost.

It is interesting to look into the fine print of customers expectations in case of real major natural disasters and/or intended attacks. In this case infrastructure providers were not held responsible, it was the best effort attempt, hard to measure. From the point of view of missing concepts, the lack of systematic procedures to guarantee the minimal service is striking.

Finally, the cost plus regulation has not provided much in terms of incentives to make the capital intensive pieces of equipment more adaptive to the system needs, by providing them with near real-time flexible control, estimators, communications with the other parts on the system, and alike. The pricing/regulatory mechanisms had no direct way of providing incentives for dynamic decision making between investing into new technology, maintaining and improving the existing, or just using what is available. The lack of dynamic investments under uncertainties to guarantee that minimal level of performance even under natural disasters scenarios has been very serious. Part of the reasons for this situation had to do with the initial huge investment into very large pieces and equipment, based on macroeconomic type of static cost/benefit analysis in support of these investments.

Again, taking the case of the electric power industry, after major blackouts in the sixties and seventies, the question concerning the role of near real-time information on equipment status, exchange with the neighboring systems, etc had to be re-visited, since the lack of this type of data was closely linked to the causes of the blackouts.

This led to the establishment of Energy Management Systems (EMS) for each utility, and the strong progress in state estimation techniques, computer-based monitoring of the system and control. Even today, the near real-time information flow between the small(er) users and the system operators and designers is almost non-existent. As the operator of the transmission grid computes supply, this is only for the estimated demand at the large industrial (whole-sale) level. This is despite the tremendous metering, and information processing technology ready to be used. There is no economic incentive to go into real-time pricing in a regulatory setting based on macroeconomic thinking, which rests on spatial and temporal aggregation and averaging.

Much effort has taken place by the EMSs throughout the world to improve the wholesale level operations and design. This has taken place often despite the lack of the financial incentives; for example, France and Taiwan have implemented low cost, control/software intense so-called power system stabilizers on most of their power plants to make them more flexible when the system is subject to very large transients after major equipment outages. This relatively simple effort has practically eliminated the stability problem on these two systems!! Much more of this could be done, and, notably, the U.S. system lags some of the efforts of this type elsewhere.

For what is to come in introducing the new paradigm, one should differentiate current practices and issues at the highest level of the infrastructure (management of the EHV electric grid and large supply for whole sale demand, versus methods for small(er) loads and local power networks). In the following section we suggest that the change of paradigm is taking place bottom-up and that the links between many small users and their effects on the macro-performance of the infrastructure must be accounted for. The fine tuning of making more out of available hardware by means of software is not a often considered alternative in the current practice of large-scale electric power infrastructure, and, most likely in the other traditional transport infrastructures. The evolution of the infrastructure for reliable service has not been very systematic. The fine tuning of developing software for identifying how localized the effect of the large outage on the infrastructure is, to estimate the actual location of the fault, etc could add fundamentally to the improved performance of the existing infrastructures.

3. The evolving paradigm: Bottom-up distributed decision making, and the associated complexities

We start by recognizing that there has been a qualitative change in the basic supply chain management in several industries, such as electricity-, gas- and airline-- industries. Once quite capital intense, these industries have begun to rely on providing their products in a highly distributed, competitive way, with many small suppliers replacing a handful of very large suppliers, and in which the supply chain from these smaller-scale producers to the product users begins to vary significantly from the typically used macro-, wholesale level process. The supply chain, is, instead, highly distributed and e-commerce based. The e-commerce is typically based on open access rules, with much room of unbundling (within a single infrastructure) and re-bundling (across the infrastructures) of once well-understood products and services. For example, in the electric power industry, one could purchase electricity from one provider and delivery from another, yet it could purchase energy (electricity, fuel) and communications to his home as one re-bundled product.

Moreover, the line between products and services is becoming less pronounced than in the past. As we are all well aware, the electric power industry, and several other transport infrastructures, are currently undergoing restructuring. The qualitative transition to the industry characterized by many independent decision makers as a result of several major changes, such as:

- Regulatory changes, where once fully regulated monopolies are undergoing both a) vertical functional/corporate separation into several different sub-entities (suppliers, delivery businesses and (groups of) consumers are basic results of vertical separation , and b) horizontal restructuring from the hierarchical structure in which each subsystem had well understood responsibilities to the local (native) customers, and very little responsibility to the customers outside its franchise area, to an open access structure, in which any product supply/demand is required to be services (for example, delivered) without differentiating local from the outside users. In addition, these regulatory changes have plowed the way to competitive, rather than cost plus pricing of products and services.

-Technology progress which has resulted in smaller-scale cost effective producers (such as small, distributed generation in the electricity industry). Moreover, with the influx of powerful communications/data systems (Internet, in particular) it has become plausible to have much distributed, near real-time information to support individual decision makers as they decide to sell and/or purchase power. Most recently, there has been real progress in load control technology [ ]. On a service (delivery) side there has been major progress in developing distributed network flow control, through primarily Flexible AC Transmission Systems (FACTS) technologies.

We suggest in this paper that one of the real problems has been lack of basic R&D necessary for designing supply/demand electricity markets and delivery (transmission grid, and local distribution) markets in support of meaningful evolution of the electric power network characterized by certain regulatory, economic and technological features leads to an efficient and reliable industry as a whole, although much is operated and planned in a highly distributed way by the separate functional/ corporate sub-entities. The industry has been moving forward by trial and error (California case) without ever being able to predict its outcomes.

To start with, it could be shown that the current practices described in Section 2 are fairly satisfactory with respect to technical performance (supplying reliable electricity to the customers). It is also somewhat tempting to a theoretician to think that this is an unacceptably conservative practice with regard to economics of electricity provision This temptation is one of the starting reasons for new R&D : in order to prove that there is something potentially more efficient, one has to introduce such tools.

Particularly disturbing has been lack of methods to differentiate among the technology candidates and chose the most effective technology (hardware and/or software). One of the major objectives of this project is to explore potential regulatory/pricing and technical frameworks capable of valuing the right technology. The question of technology choice becomes very rich and critical in light of information revolution, and almost unlimited Internet and other data networks. The availability of almost free information about the status of each customer, or groups of customers, makes it very tempting to think about the customer as an active decision maker, who, in its own right begins to shape the requirements of once unidirectional power supply chain.

The traditional static equilibrium thinking when designing engineering tools and/or static equilibrium thinking when assessing pricing/policy rules are simply not applicable to the industry in transition. The industry in transition needs sufficiently flexible information flow protocols between the evolving new entities as well as decision making tools for the entities themselves, to catalyze evolution of once fully regulated, large scale, macro level averaging-based industry to the competitive, smaller scale industry in which economies of scope , risk management and dynamic decision making under various uncertainties become dominant factors determining success of various entities.

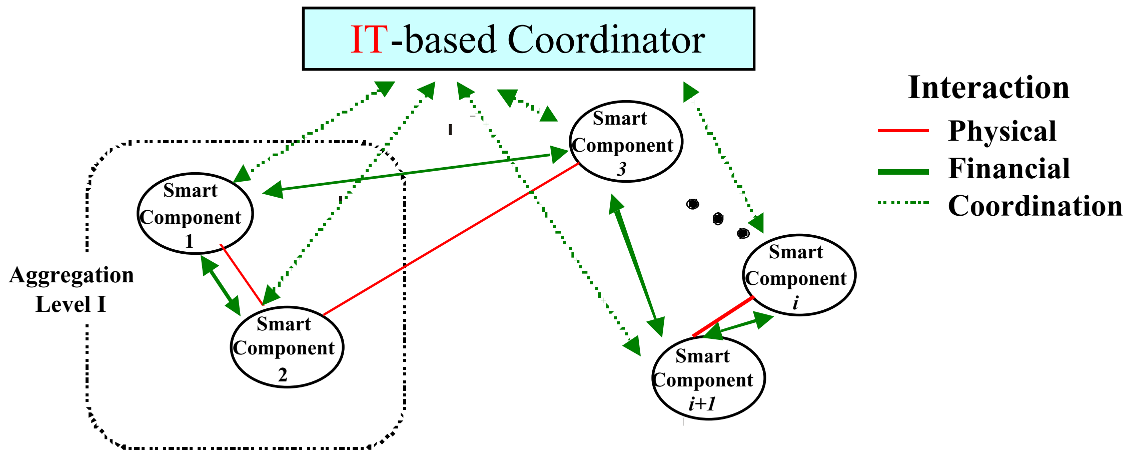
4. Managing infrastructure complexities: The case for new flexible protocols

To summarize the material in Sections 2 and 3 above, we are presented by two qualitatively different paradigms for designing and operating the infrastructure systems and their interdependencies. One is top-down, and the second one is bottom-up. It was argued that neither of the two approaches are likely to lead to a silver bullet answer. They both have complexity problems. The first paradigm is less risk dynamic, better understood, the second promises interesting new opportunities in which distributed decision makers re-bundle for value across infrastructures.

In this situation, of particular importance is the interplay between financial, IT and technical signals all interacting over various time horizons and various uncertainties infrastructures are presented with over time. It is essential to understand that for the infrastructures in transition one should not take a rigid top-down approach in which the feedback of system users is minimized. At the same time, given unique temporal and spatial challenges in managing the transport type infrastructures, one needs to provide some minimal coordination to reflect the status of the entire system . This could be technical and/or financial and/or IT type coordination, all working in an interactive flexible way. Figure 1 here shows the basis for evolution of intelligent transport infrastructures, in which complexity is managed by various coordinating interactions. We stress that these are not remnants of the old paradigm rigid standards.

Figure 1: Basis for possible protocols

Basic Problem of Interest



¥ Two Question

— Optimizing Performance at Component Level subject to System Imposed Constraints

— Satisfying System-wide Performance Criteria

Much R&D is needed to develop these protocols for managing complexities. The technical and economic foundation for such protocols in the evolving power industry could be found in [4a]. The borderline between designing markets, technical and/or IT structures is no longer as pronounced as in the traditional industry. None of the signals are exogenous, they all become mutually endogenous and interactive.

At the same time, the IT open access protocols to implement this coordination are necessary as well. For an example of this, see [4b]. These are initial efforts in this direction, and much R&D is needed.

References:

[1] Many conversations with Professor Mariann Jelinek , during her stay as an NSF Program Director for Industrial Innovations, 2000; Professor Paul Kleindorfer, Wharton School, University of Pennsylvania, 2000; Professor Ingo Vogelsang, Boston University, 2000; Dr. Paul Werbos, NSF, 2000.

[2] Jelinek, M., Ilic, M., ‘‘Strategic Framework for Electirc Technologies , Proceedings of the NSF/DoE/EPRI Workshop on Complex Interacitve Networks, 2000.

[3] Ilic, M., ‘‘Power and Data Networks: How Common are Their Basic Problems , EPRI/DoE Porgram on Complex Networks, April 2001.

[4a] Ilic, Marija, ‘‘Technical and Economic Foundations for Designing Protocols in the Competitive Electric Power Industry , Power Systems Computations Conference, July 2002 (submitted)

[4b] Ilic, Jovan, Intelligent Energy Systems, Inc. (IESI) ‘‘Implementation of Protocols for the Competitive Electric Power Industry , Power Systems Computations Conference, July 2002 (submitted).