Establishing Public Policy as a Primary Cause of Engineering Failure:
Did Market Deregulation Lead to the North American 'Blackout', August 14th 2003?

Christopher W. Johnson,
Glasgow Accident Analysis Group, Department of Computing Science,
University of Glasgow, Glasgow, G12 8QQ, Scotland, United Kingdom,
johnson@dcs.gla.ac.uk, http://www.dcs.gla.ac.uk/~johnson

Abstract:
On the 14th August 2003, a complex combination of latent problems and catalytic events led to a domino-effect in which 50 million people suffered some interruption to their power supplies. Losses have been estimated between $5-10 billion. It is, therefore, one of the most wide reaching and serious 'blackouts' in a national power distribution network. The causes of this infrastructure failure included technical issues to do with network capacity and the algorithms used to predict potential distribution problems. It also had managerial and human factors causes; these arguably included an over-reliance on automated monitoring systems. The infrastructure failure also stemmed from governmental and regulatory intervention, which led to problems in the operation of the energy market. The following paper applies accident investigation techniques to represent and reason about the complex interactions between these causes. In particular, we show how Violation and Vulnerability (V2) diagrams help to map out the competing arguments for and against market deregulation as a causal factor in the engineering failures that led to the blackout.

Keywords:
Accident analysis; root cause analysis; market deregulation; electricity distribution.

1. Introduction

The North American electricity network brings together some 3,700 utility organizations providing more than 320,000 kilometers of transmission lines. It provides more than 950,000 megawatts of generating capability for more than 100 million different customers. The complexity of this system has led to considerable investment to insure the resilience of energy infrastructure provision. Utility planners and regulators conduct regular studies to identify the interdependent causes that might defeat their 'defense in depth' approach. They also strive to provide backup generation and transmission capacity. In spite of these precautions, large portions of the Midwest and Northeast United States and Ontario, Canada, experienced an electric power blackout on 14th August 2003. The outage affected an area with an estimated 50 million people and 61,800 megawatts (MW) of electric load in the states of Ohio, Michigan, Pennsylvania, New York, Vermont, Massachusetts, Connecticut, New Jersey and the Canadian province of Ontario. Power was not restored for 4 days in some parts of the United States. Parts of Ontario suffered rolling blackouts for more than a week before full power was restored. The total costs in the United States range between (US) $4 billion and $10 billion. In Canada, gross domestic product was down 0.7% in the month of the failure, there was a net loss of 18.9 million work hours, and manufacturing shipments in Ontario were down (Canadian) $2.3 billion.

1.1 Competing Accounts of the Blackout

A joint US and Canadian commission was established to identify the causes of the power failure (US-Canada Task Force, 2004). The resulting report contains ten chapters and is well over 200 pages long. In contrast, the following pages demonstrate that graphical techniques can be extended from accident and incident analysis to provide a coherent overview of infrastructure failures. The intention is not to replace lengthy textual reports but to provide a more convenient road map to the interactions that occurred between many complex subsystems during this major interruption in power supplies. It is also important to stress that this incident has been the subject of considerable controversy. Federal (GAO, 2005) and State investigations (Massachusetts, 2005), commercial organizations (Delgado, 2005), pressure groups (Hughes, 2005) and media organizations (Hoggan, 2004) have all published alternate accounts. This has led Harris

(2005) to argue that "the August 14, 2003 blackout has become a Rorschach test in which every viewer interprets as evidence to support his or her concerns about the problems in today's electric industry". Many of the differences that distinguish these accounts stem from the authors' attitudes towards deregulation and the impact that it can have upon the reliability of complex systems. Some authors have argued that market competition cannot be relied upon to provide social goods, which including infrastructure reliability. In this view, public policy should focus not simply on reducing consumer prices but also on ensuring that the market will guarantee supply. Other authors have rejected the premise for this argument. They maintain that deregulation had little impact on the causes of the blackout. Political intervention in the market was a cause of, rather than a remedy for, infrastructure failure. Of course, this is an over-simplification. Many commentators hold positions between these two extremes. In this view, the blackout was not an inevitable cause of deregulation itself but by the particular mechanisms that were used to open access to electricity transmission networks across North America.

1.2 Structure of the Paper

The following sections begin by using graphical, accident investigation techniques to chart the immediate causes of the August 14th blackout. There were problems with the computer-based, monitoring systems that alert utilities and reliability organizations of potential transmission failures. There was inadequate maintenance of the transmission infrastructure that led to short circuits from changing power flows. These changes raised the core temperature of transmission lines that led them to sag in areas of the network that had become overgrown with vegetation. Using accident investigation techniques helps to map out the interactions between these different catalytic or triggering causes. However, the same diagrams can be extended back to identify the underlying impact that public policy had upon the infrastructure engineering. The closing sections of the paper show how these techniques help to reconstruct the different arguments for and against the impact of deregulation as a cause of the blackout. The intention is not to 'prove' that one side is correct in this on-going debate. In contrast, the intention is to cut through the rhetoric to look at the evidence that each side uses to support their claims about the engineering consequences of different forms of market intervention.

2. Catalytic Events

The catalytic events leading to the incident focus on the loss of monitoring software operated by a reliability supervisor known as the Midwest Independent System Operator (MISO). Their code supported State Estimation and Real Time Contingency Analysis functions. In simple terms, these programs used Monte Carlo techniques to assess the N-1 state of the network. Probabilities were used to estimate the set of worst possible conditions that might arise should a network component fail. These systems were only partially implemented and many of the managers regarded the software as undergoing a trial period of usage. Real time data flows were not being processed from all areas of the network and when an operator identified a discrepancy between the State Estimator and the actual state of the network, they disabled the software. After some period debugging the tools, they found that the discrepancy was caused by a failure in a sub-network that was not instrumented. The operator manually inserted the correct values to show that the sub-network had failed. He then checked that the software predictions agreed with the state of the network and went to lunch, without restarting the automated network monitoring function. In consequence, those agencies that depended on information about potential N-1 instability from MISO could no longer rely on an automated warning.

This summary should illustrate the complexity of software failures that contribute to the loss of national infrastructures. It should also be stressed that these software issues only formed a small part of the wider circumstances leading to the August 14th failure. In consequence, this paper argues that techniques from accident analysis must be extended to help us represent and reason about these problems. Figure 1 provides an overview of a V2 (violation and vulnerability) analysis. Dotted boxes represent events that lead towards a failure. For example, the following diagram records that MISO's State Estimation Software produces a solution outside bounds of acceptable error at 12.15. As we shall see this cannot be interpreted as a violation or vulnerability because the software functioned as it was intended by reporting a potential mismatch between the modelled state of the network and the various sensor readings that were being reported to the system. Solid boxes denote violations. These are events that contravene operating norms and procedures. Violations can be inadvertent. For instance, operators may not know about applicable rules and regulations. Similarly, system failures and communication breakdowns can make it difficult for

end-users to know whether or not they have violated a particular norm. Violations can also be deliberate and may, in some cases, be justified. This happens when, for instance, rules and procedures fail to take into account particular environmental conditions that would further jeopardize safety if operators were to follow them. For example, Figure 1 uses a violation symbol to denote the failure to restart the Real Time (N-1) Contingency Analysis (RTCA) after the State Estimator (SE) was enabled.

Dotted ellipses represent contributory factors. These are a notational convenience and can be used to represent a series of events. For example, Figure 1 shows that the failure of the Bloomington-Dennis (B-D) Creek 230-kV line is not updated in the State Estimator. This might be represented by events showing that the fault was not updated at 12:11, 12:13, 12:14 etc. Solid ellipses represent vulnerabilities. These conditions threaten the safety of a complex system. The following diagram includes two vulnerabilities: the East Central Area Reliability Coordination Agreement (ECAR) data network did not provide data from the B-D Creek line as input to the State Estimator functions; MISO considered the State Estimator and Real Time Contingency Analysis systems still to be under development. The first condition is a potential vulnerability because it created the opportunity for mismatches between the software's network models and the actual state of the network. Operators were then forced to manually diagnose the source of the mismatch to the B-D Creek line. The second vulnerability arose because the State Estimator and Real Time Contingency Analysis seem to have played a central role in assessing the reliability of the network even though they were officially 'under development'. The triangle labeled 1 is a continuation symbol denoting that the diagram continues on another V2 figure.

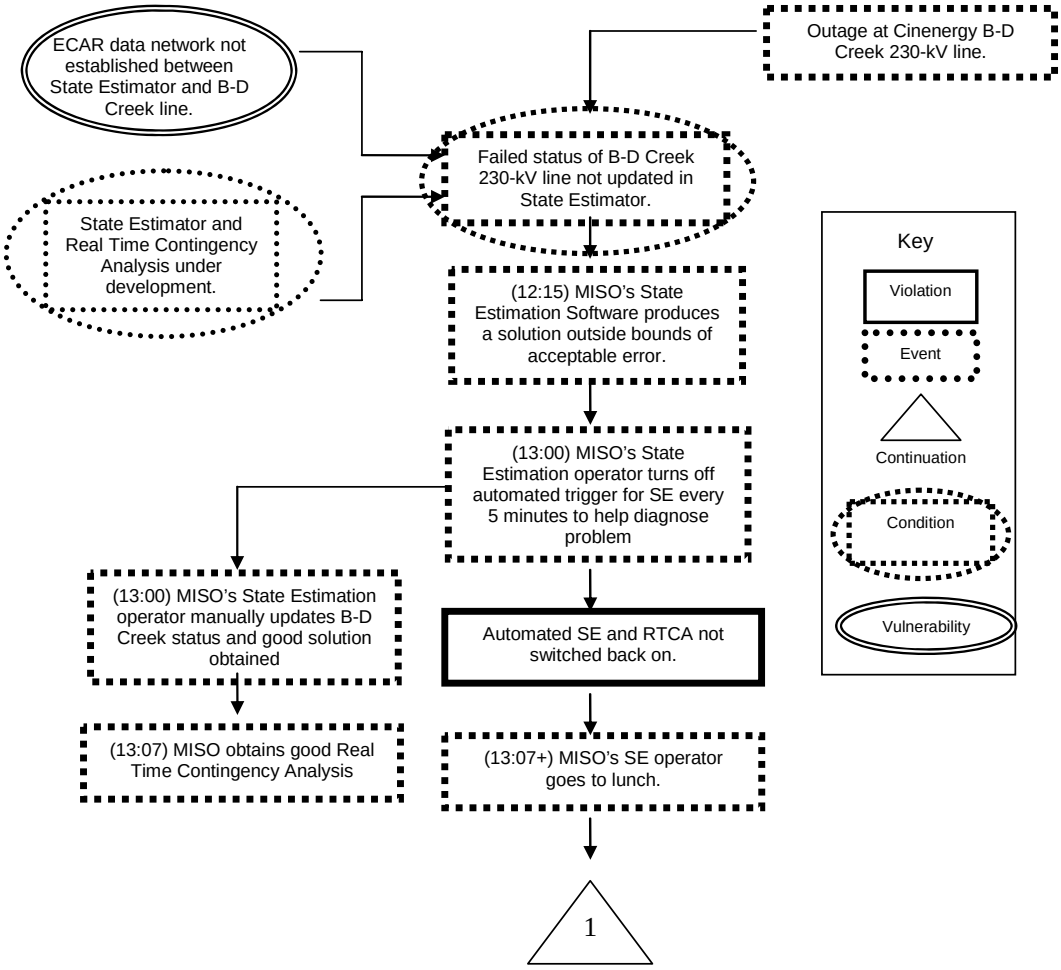

Figure 1: The Failure of MISO's State Estimator and Real-Time Contingency Analysis

Figure 1 illustrates the vulnerabilities that led to the blackout. As mentioned, these included the lack of sensor data from some areas of the distribution network. They also included the assumption within MISO that the state estimator/ real time contingency analysis were still under development even though, as we shall see, they provided warnings about the potential violation of N-1 criteria. The V2 diagram also shows the way in which the operator manually supervised the Real Time Contingency Analysis and State Estimator functions after they had accounted for the mismatch between the software model and the available network data. The operator forgot to restart the automatic State Estimator function that was intended to check the status of the network every five minutes. The lack of automated SE and RTCA functions made it difficult to assess N-1 reliability. This violated operating policies established by the North American Electric Reliability Council (NERC). This non-governmental body provides a forum for generation and transmission companies. The NERC 'N-1' guidance is intended to ensure that the network continues to operate even if a 'mishap' occurs. NERC Operating Policy 2.A—Transmission Operations states that "All CONTROL AREAS shall operate so that instability, uncontrolled separation, or cascading outages will not occur as a result of the most severe single contingency." In other words, the system must operate in such a way that it is resilient to a random failure involving the most 'critical' generation or transmission facility. It should also be possible for operators to assess the new worst contingency after the initial failure and to plan how to maintain the future reliability of this altered system. NERC operating procedures require that a network is returned to normal operation within 30 minutes of a contingency so that it "can once again withstand the next-worst single contingency without violating thermal, voltage, or stability limits" (US-Canada Task Force, 2004).

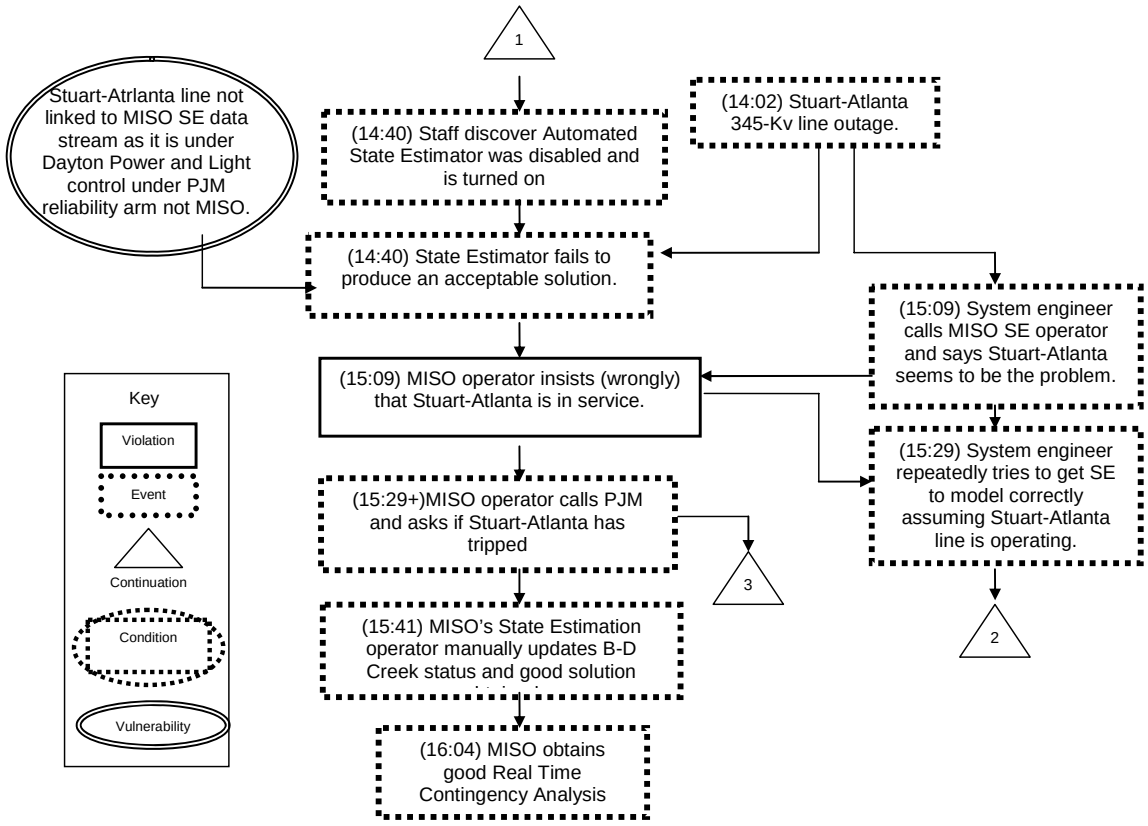

Figure 2: Restarting the Real Time Contingency Analysis

The continuation triangle at the top of Figure 2 extends events from the previous V2 diagram to represent what happened after the MISO operators discovered that the automated State Estimator and Real Time Contingency Analysis functions had to be restarted. The State Estimator failed to produce an acceptable solution, partly because there had been an outage on the Stuart-Atlanta 345-Kv line. A system engineer noticed the failure and called MISO. However, the MISO operator insisted that the line was working even though the State Estimator was failing to provide an accurate model. As in Figure 1, these events reveal

the vulnerability that is created by the lack of data input from portions of the transmission network. The MISO operator eventually is told of the failure by another reliability coordinating organization, PJM. All of these events explain why reliability monitoring resources were absorbed in trouble shooting the SE and RTCA systems while critical events in other areas of the network were not noticed.

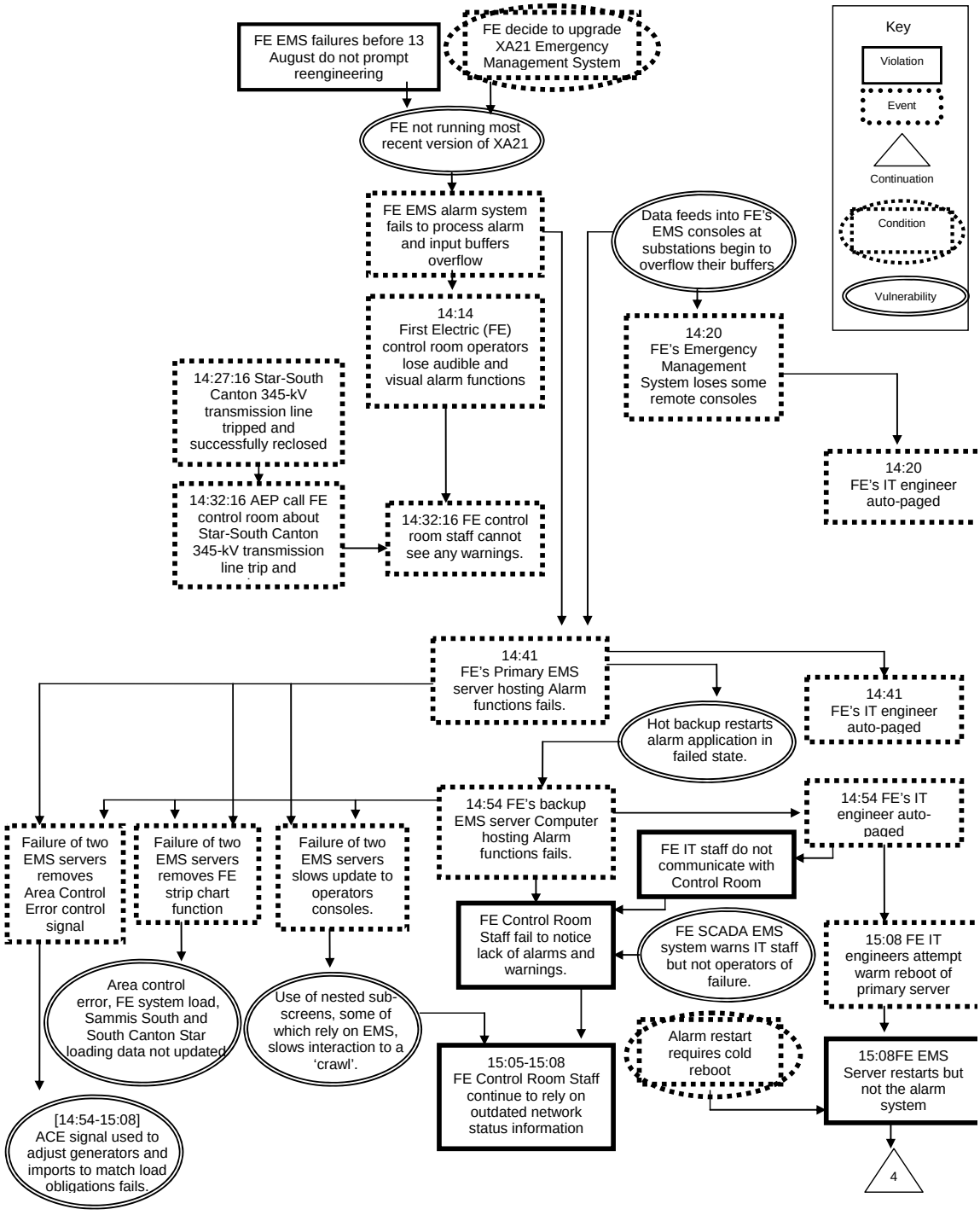

Figure 3: The Failure of First Electric (FE's) Emergency Management System

Figure 1 and Figure 2 sketch the violations and vulnerabilities that led to the loss of the MISO state estimator and real time contingency analysis tools between 12:15 EDT and 16:04 EDT. The governmental

report argued that disruptions to these systems "prevented MISO from promptly performing pre-contingency 'early warning' assessments of power system reliability over the afternoon of August 14" (US-Canada Task Force, 2004). In contrast, Figure 3 maps the events the prevented the First Electric (FE) generation and transmission company from accurately monitoring the state of their systems and those of neighboring utilities. Recall that MISO acted as the regional reliability monitoring organization for several of the utility companies such as FE. Most of Figure 3 focuses on problems relating to FE's Emergency Management System (EMS). Previous failures of this application had not prompted reengineering before August 2003. Instead, there had been a decision to upgrade the XA21 system. This decision had not yet been implemented and so FE were not running the most recent version of the application. The official report into the blackout suggests that this may partly explain why the EMS failed to process alarms from the network in the interval immediately before the blackout. At the same time, remote consoles at substations also began to experience buffer overflows. By 14:20, FE's IT engineer had been autopaged to help restore these EMS terminals. Meanwhile, the buffer overflow on the central EMS system prevented FE control room staff from receiving audio-visual warnings about network reliability. This proved to be significant because they could not confirm warnings from AEP when the Star-South Canton 345-kV line trips at 14:27. As we shall see, the lack of EMS warnings together with the failure of MISO's SE and RTCA applications gradually eroded the situation awareness of staff monitoring the generation and power distribution systems.

The loss of EMS remote terminals combined with central buffer overflow problems led to the failure of the primary server for FE's EMS application. Again, the official reports into the blackout provide limited details about these failure mechanisms. However, a 'hot backup' server was available and took over the EMS application while the FE IT engineer was again automatically paged to respond to the failure. However, the hot restart uses duplicate software and the buffer overflows are not resolved so the redundant server also fails in the same way that the primary EMS system. The FE IT engineer is again auto-paged with information about the failure of the backup server. They eventually attempt a warm reboot of the primary EMS server. This does not restart the alarm system, which would require a cold reboot.

Figure 3 shows a number of violations during this interval. Firstly, the FE IT staff do not pass on information about the failures to the control room staff even though they have received automated updates about the EMS server failures. Secondly, and partly in consequence, the control room staff fail to notice the failure of the Emergency Management System. They do not diagnose the failure from the absence of any warnings or alarms even though other organizations, such as AEP, have contacted them questioning whether particular lines have failed. The failure of the EMS servers has several further consequences. It sloes updates to the operators screens where EMS data can be nested within other displays that then cannot easily be refreshed. Hence the operators must rely on outdated information. The failure of the EMS servers also removed FE's strip chart function which provides an overview of critical data relating to area control errors and network loading. Not only does the server failure remove the presentation of data about Area Control Errors, the loss of the servers also prevents Area Control Error Signals which is a primary mechanism in the automated adjustments to generating and importing capacity so that available resources meet power obligations.

Previous paragraphs illustrate the complexity of the events and conditions that combined to create the content in which the North American blackout occurred. They should also illustrate the benefits of graphical modeling techniques, such as V2 analysis, which can provide an overview of these adverse events. It is possible to use the components of Figure 3 to trace the impact that vulnerabilities, such as delays in the decision to implement the update to the XA21 system, had on the course of this incident. It is also possible to trace the more immediate consequences of particular violations, including the failure of FE's IT engineers to inform control room staff of the EMS server failure.

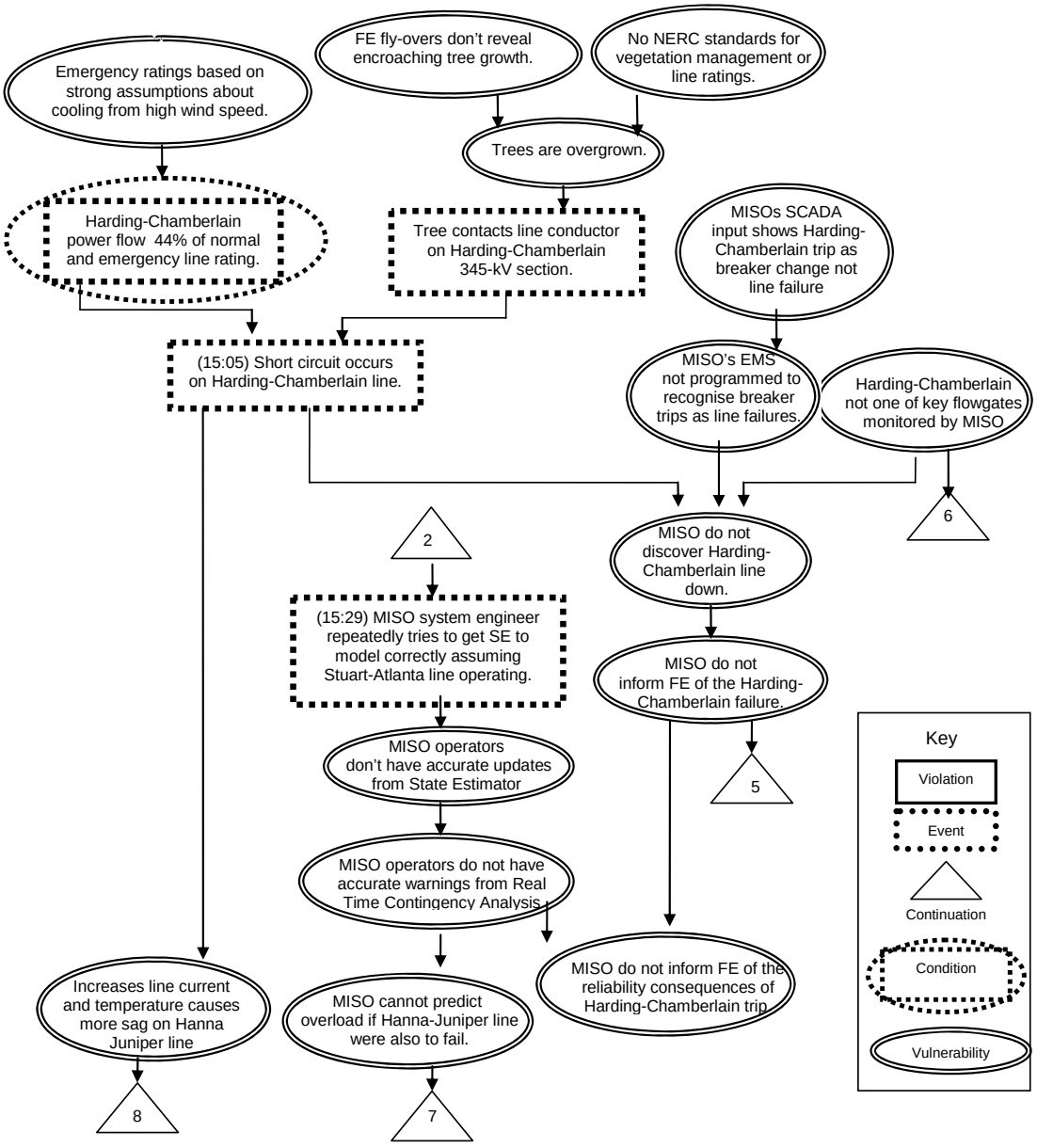

Figure 4: Lack of State Estimator and Flow Monitoring Functions Hides Consequences of Failures

Figure 4 continues the V2 analysis of the August 2003 blackout by analyzing the problems that MISO faced in monitoring and controlling the changing network conditions given the interruptions to the State Estimator (SE) and Real Time Contingency Analysis (RTCA) described in Figures 1 and 2. The diagram also describes additional network failures that began to destabilize the distribution network. As can be seen, there were no North American Electric Reliability Council (NERC) quality standards describing in detail how to manage the growth of trees and other vegetation. This is significant because growth over the Summer months creates the potential for short-circuits to interrupt line transmission through contact with overhead cables. First Electric (FE) used fly-overs to identify potential problems and after the blackout the governmental enquiries argued that these were insufficient to identify the potential problem sites. The likelihood of a short-circuit line failure is increased by the power loading on the cables; increased loadings result in increases in cable temperature that will eventually cause the cable to sag. Both the growth of vegetation and power loading may have been factors in the loss of the Harding-Chamberlain 345-kV line. Although the actual power flow was only 44% of the normal and emergency rating for this section of the

network, these upper bound estimates relied on assumptions about the effects of wind cooling on the cables that arguably could not be sustained.

MISO did not discover that the Harding-Chamberlain line had failed because their Supervisory Control and Data Acquisition (SCADA) system presented this as a change in the status of a breaker and not a failure of the entire line. Hence, staff may not have realized the full consequences of the short-circuit. The MISO Emergency Management System, not to be confused with FE's EMS, was not programmed to recognize breaker trips as a potential symptom of line failure. Finally, Harding-Chamberlain was not one of the flowgates that was routinely monitored by MISO. All of these factors prevented the reliability coordinator from detecting and diagnosing the line failure. This, in turn, prevented them from meeting the requirement to inform FE of any potential threats to system reliability. This was compounded by the lack of accurate updates from the State Estimator and the Real Time Contingency Analysis, mention in previous sections. The bottom portion of Figure 4 shows how the failure of the Harding-Chamberlain line increased current in the Hanna-Juniper 345-kV line. This would lead to a potential increase in core cable temperature in that area and make knock-on failures more likely. However, neither MISO nor FE were aware of the problem because they had not identified the initial fault on the Harding-Chamberlain line. The problems with the SE and RTCA applications also prevented MISO from identifying the potential N-1 hazards that could arise from the loss of Hanna-Juniper once Harding-Chamberlain was lost.

The V2 diagram in Figure 5 continues the analysis of the knock-on failures following the loss of Harding-Chamberlain. MISO lacked information about both the Harding-Chamberlain failure and the N-1 consequences from the loss of the 345-kV line. This prevented the reliability organization from warning FE about the potential problems that lay ahead. Recall that FE was also operating without full support from their Emergency Management System. The failure of Harding-Chamberlain increased the loading on Hanna-Juniper and this led to a further short-circuit even though this line was only operating at 88% of its normal and emergency loading. MISO's Supervisory Control and Data Acquisition (SCADA) system detects the overload and staff begin to contact FE to warn them. In the meantime, FE staff remain unaware of either loss. The failure of Hanna-Juniper after the loss of Harding-Chamberlain now places increased loads on the Star-Juniper and Star-South Canton lines. The utility company, AEP, and their associated reliability organization, PJM, recognize the increased loading on the Star-South Canton line. However, they were not alerted to the potential problem as part of a previous N-1 analysis because their contingency software did not draw sufficient data about the state of FE's lines. AEP attempted to reduce the load on Star-South Canton by asking the reliability coordinator, PJM, for Transmission Loading Relief. It can take more than an hour for such procedures to be implemented and usually they involve 25-30MW not the requested 350MW so the AEP request was delayed by repeated verification requests. The V2 diagram might also be extended to show that MISO, FE operators and other reliability organizations had NERC powers to take more extreme actions including re-dispatching generation, system reconfiguration or tripping load. However, limited situation awareness prevented all of the parties involved from realizing the need to take such emergency measures. AEP remained unaware that the loss of Hanna-Juniper had been the cause of the increased loading on the Star-South Canton line. Neither AEP not PJM realized that the Transmission Loading Relief would have had little effect on the course of the blackout as most of the loading was native to the immediate area.

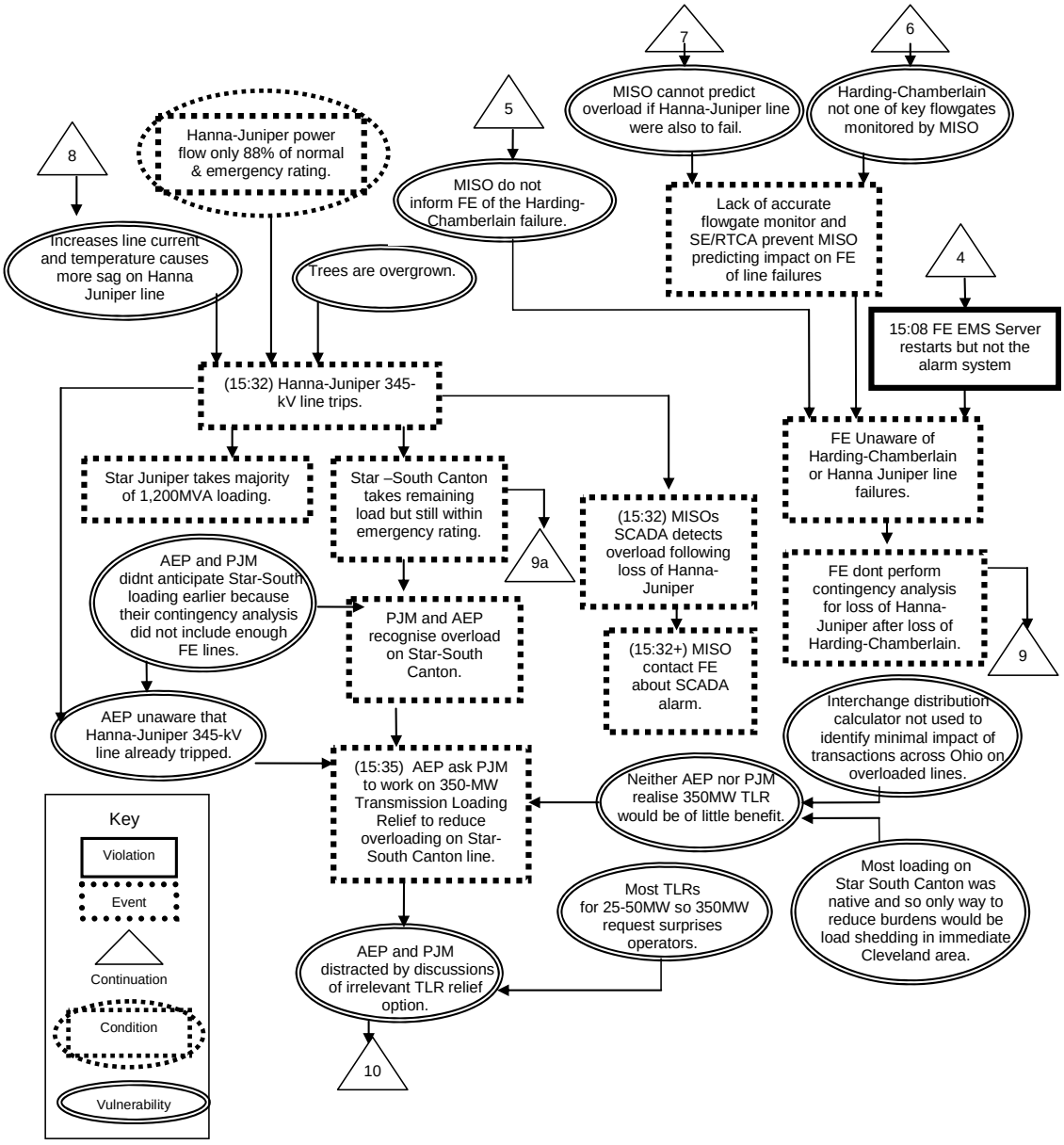

Figure 5: Consequent Network Failures and Attempts to Relieve Transmission Loading

The events and conditions in Figure 6 show how the various organizations involved in this start-up phase of the blackout gradually became aware of some of the problems that were affecting their systems. Before this, however, the Star-South Canton 345-kV lines tripped as a result of the increased loading created by the failures of the Harding-Chamberlain and Hanna-Juniper lines. The fact that the South-Star Canton lines crossed the boundary between FE and AEP may have increased the difficulty in coordinating vegetation control. However, the line was still only operating at 93% of its emergency rating. As can be seen in Figure 6, the loss of Star-South Canton had considerable implications as voltage levels began to degrade and flows increased on the 138-kV system towards Cleveland and on the Sammis Star line which remained the only 345-kV route into the city from the South. The failure of South-Star Canton lines forced a complete revision of the AEP and PJM contingency planning. They had worried about the consequences of a Sammis Star failure on South-Star Canton rather than the impact of the loss of South-Star Canton on the Sammis-Star lines. As the 138-kV system starts to trip both organizations begin to realize the extent of the emergency but cannot identify viable solutions.

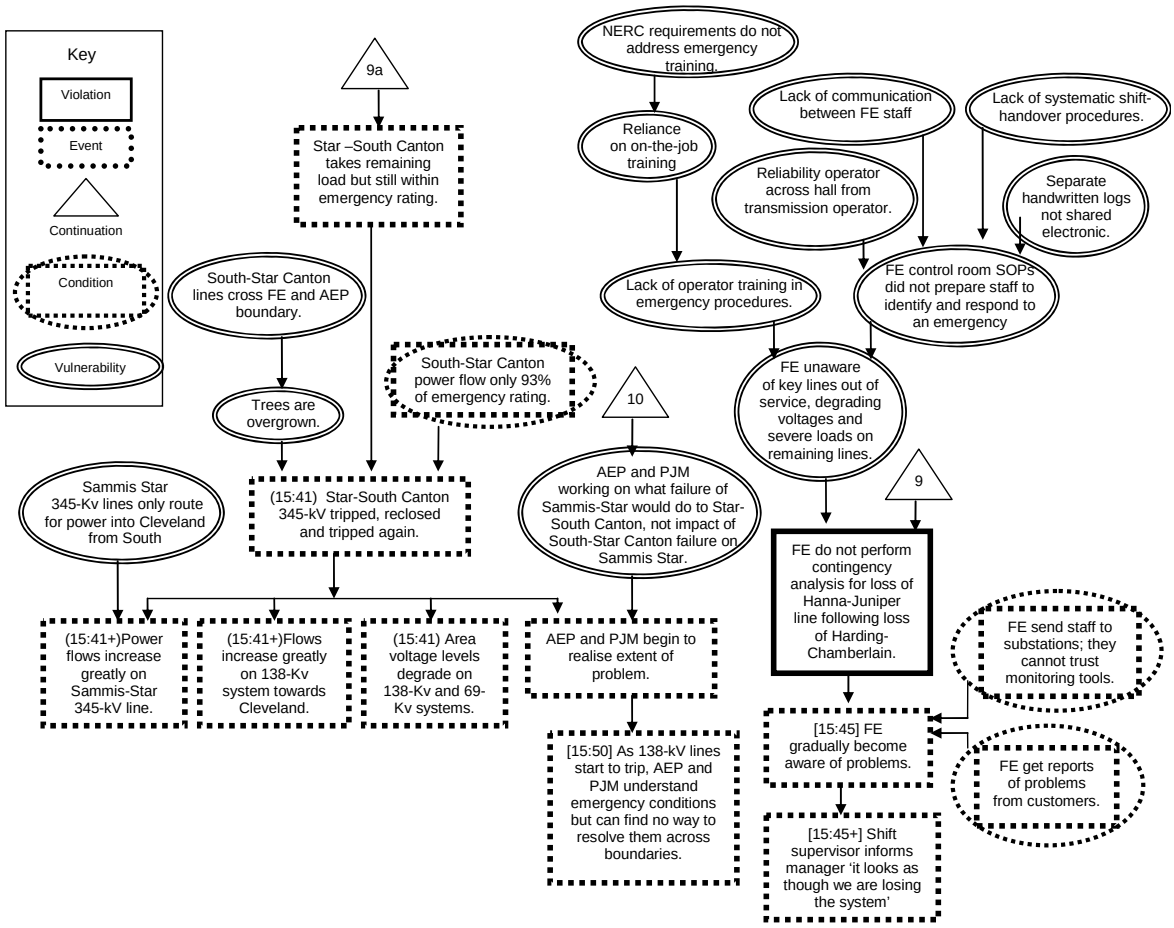

Figure 6: FE, AEP and PJM Begin to Realize They are 'Losing the System'

Meanwhile, FE are still trying to understand what has occurred. Operator situation awareness has been compromised by a range of factors including the lack of systematic shift handover procedures, the difficulty in sharing common logs across the organization to piece together colleague's observations, the lack of communication between key staff and the lack of training in emergency procedures. Figure 6 also captures the official report's observation that this lack or training and preparation may have stemmed from the limited guidance provided on this by the NERC. Calls from customers and the decision to deploy staff to substations rather than continue to rely on information from centralized monitoring systems gradually helped FE staff to understand the developing problems on the network. Eventually, after 15:45 the FE shift supervisor informs their manager that they may be 'losing the system'.

The loss of the Sammis-Star 345-kV line following the loss of South-Star Canton, Hanna-Juniper and Harding-Chamberlain led to weak voltages in Ohio and led to power flows that created a domino effect, triggering many subsequent line failures. One reason for this is that a particular form of 'zone 3' relay was widely used throughout the affected area. These were designed to trip in response to line overloads rather than to 'true' faults. In consequence, they helped to increase the speed of propagation in the cascading failures beyond the Cleveland- Akron area. The governmental investigation concluded that "the relay protection settings for the transmission lines, generators and under-frequency load-shedding in the northeast may not be entirely appropriate and are certainly not coordinated and integrated to reduce the likelihood and consequences of a cascade" (US-Canada Task Force, 2004). Brevity prevents a more detailed analysis of this cascade phase of the blackout. This is justified both by the similarities between the engineering issues in the domino effect and in the initial causes of the failure and also by the importance of

linking these engineering issues to the public policy decisions that arguably created the context in which the failures were likely to occur.

3. Public Policy and Failures of Infrastructure Engineering

Previous paragraphs have used V2 diagrams to identify the conditions and events that led to the August 2003 blackout. These involved many different systems, including MISO's SCADA, Flowgate monitoring, EMS, SE and RTCA applications. They also included FE's XA21 EMS primary and backup servers, their Strip Chart systems, Area Control Error signaling, remote EMS terminals as well as the East Central Area Reliability Coordination Agreement Data Network. The V2 analysis referred to operators from utilities such as FE and AEP as well as their reliability organizations, MISO and PJM, and national bodies including NERC. Although the previous analysis has shown the immediate events leading to the blackout, it has not explained how or why these vulnerabilities began to affect national infrastructures. The following pages, therefore, extend our use of diagrammatic technique to examine the way in which public policy helped to create the context for these failures.

In order to understand the causes of the August 14th 'blackout', it is important to introduce the structures that governed the generation and distribution of electricity in 2003. The North American Electric Reliability Council (NERC) was established in 1968, as a result of the Northeast blackout in 1965. It is a non-governmental, organization relying on peer pressure to establish the standards that ensure reliable generation and distribution. This creates potential confusion; public policy is defined to be guidelines or rules that results from the actions or lack of actions of governmental entities. As a non-governmental organization, the NERC does not fall within this definition. However, we would include it within our analysis of public policy because the self-regulatory nature of the organization arguably reflects the lack of direct governmental involvement in the industry.

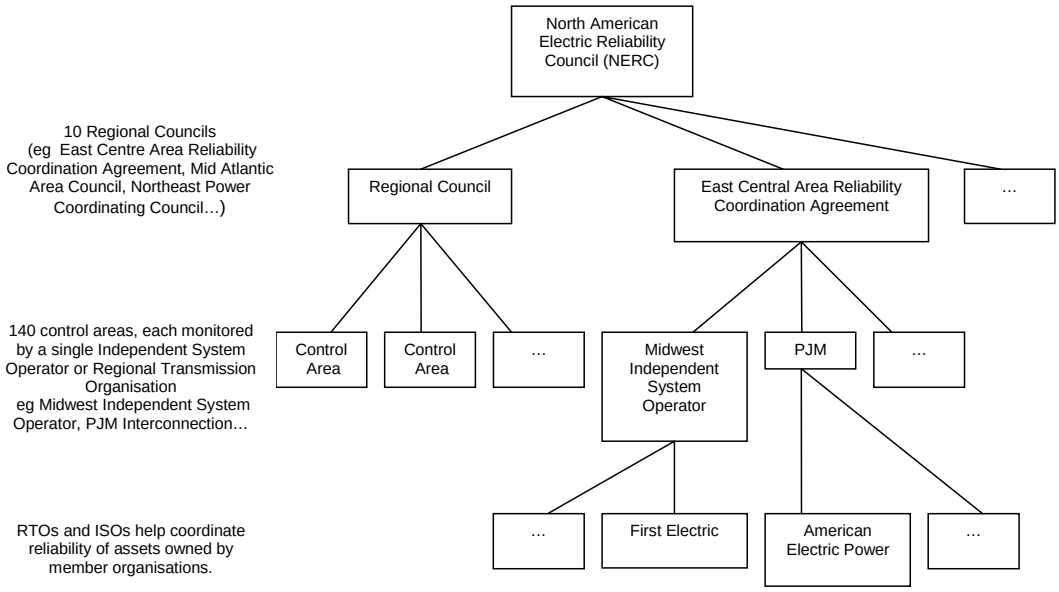

Figure 7: High-Level Overview of the Organizational Structure Prior to August 2003

Figure 7 shows how the NERC is composed of ten regional reliability councils. Three were affected by the August 14th failure: East Central Area Reliability Coordination Agreement; Mid-Atlantic Area Council and Northeast Power Coordinating Council. These regions can be broken down into a total of 140 'control areas'. Each control area is, typically, monitored by a single Independent System Operator (ISO), or Regional Transmission Organization (RTO). Five RTOs/ISOs were directly affected by the August 14 blackout: Midwest Independent System Operator (MISO); PJM Interconnection (PJM); New York Independent System Operator; New England Independent System Operator and Ontario Independent Market Operator. Each of the RTOs/ISOs manages the real time and 'day-ahead' reliability of the bulk power system within their areas. They do not own transmission assets but direct the operation of assets

owned by their members. The blackout originated in two control areas. The first was in northern Ohio where FirstEnergy (FE) operated. The second was to the South under American Electric Power (AEP). These organizations were monitored by their reliability coordinators, Midwest Independent System Operator (MISO) and PJM.

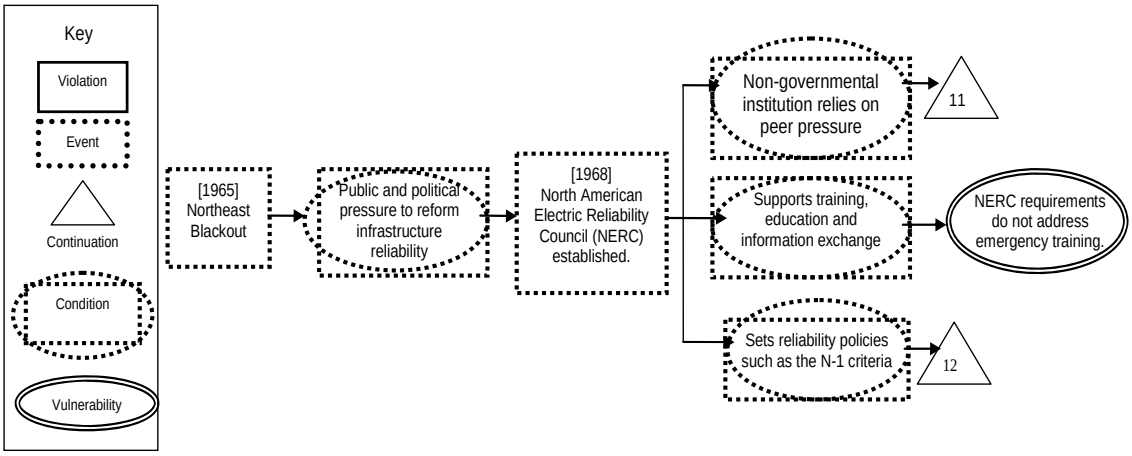

Figure 8: Conditions Created by the Development of NERC

NERC coordinates the development of tools to enhance infrastructure reliability, including data exchange systems. Their objectives include maintaining a balance between generation and demand. They are also concerned to limit the thermal heating that takes place in network components, including distribution lines, from dynamic power flows such as those described in earlier paragraphs. Previous sections have also described NERC's 'n-1 criterion'; operators must assess the new worst contingency after an initial failure so that they "can once again withstand the next-worst single contingency without violating thermal, voltage, or stability limits" (US-Canada Task Force, 2004). Figure 8 extends the use of Violation and Vulnerability diagrams to illustrate some of the conditions that were created by the establishment of the North American Electric Reliability Council (NERC). This can be demonstrated by showing the manner in which these conditions influenced the immediate events that have been identified in previous diagrams. For instance, the NERC's role in supporting education and information exchange amongst utility and reliability organizations can be linked to the lack of requirements for emergency training. Figure 6 identified this omission as an important vulnerability that, arguably, left the FE operators unprepared to identify and diagnose the failures that occurred during the afternoon of 14th August 2003. As can be seen from Figure 8, it is relatively easy to represent the relationship between this high-level objective of the NERC and particular events during the blackout, represented in Figure 6. However, the relationships between public policy and engineering failures are not always this simple. The two continuation symbols in Figure 8 denote that the non-governmental nature of NERC and their associated role in establishing reliability policies, such as the N-1 criteria, form part of a more detailed analysis that will be presented in the following pages.

It is important to stress that the industrial structures described above had emerged through a process of change in the years before the blackout. In 1978, the U.S. Congress passed the Public Utility Regulatory Policy Act (PURPA). The aim was to encourage investment in newer, more efficient technologies and, in consequence, to lower costs. These regulations enabled new entrants into the market to sell energy to utilities without many of the reliability obligations that governed established companies. Traditionally, companies had been vertically integrated within particular regions where they owned and operated generation, transmission and distribution. PURPA companies could sell power without necessarily providing guarantees about continued service provision. There was no assumption that they would invest in, for instance, the transmission infrastructure. Utilities that housed a PURPA generator were also faced with the costs of paying for these contracts which in the short term might be more expensive than native generating capacity. Delgado (2005) argues that this, in turn, reduced the utility's earnings and jeopardized

their own future investment. These moves to open access to the energy markets continued in 1996 with Federal Energy Regulatory Commission (FERC) Order 888. New industry participants, known as energy marketers, gained access to the distribution grid under the same conditions as the utilities provided for their native generating loads.

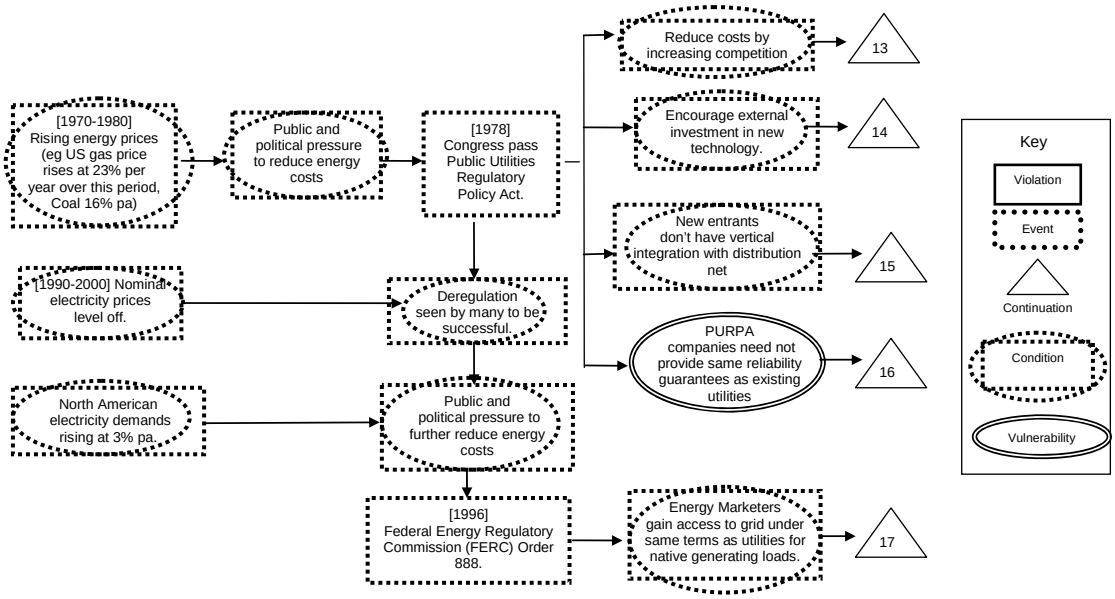

Figure 9: Changes to the North American Electricity Market Structure

Figure 9 uses the V2 notation to characterize some of the changes in public policy. As can be seen, most of the consequences of FERC Order 888 and the PURPA legislation are shown as conditions. They led to a series of events that were intended to 'reduce costs by increasing competition', to encourage 'external investment in new technology', to ensure that 'new entrants didn't have the same vertical integration with the distribution network' and to provide access to the gird 'under the same terms as utility's native generating loads'. The impact of these conditions on the causes of the blackout will be discussed in later diagrams that include the associated continuation symbols. However, Figure 9 also represents one consequence of the PURPA legislation as a potential vulnerability; 'PURPA companies need not provide the same reliability guarantees as existing utilities'. This illustrates the way in which V2 diagrams capture the subjective viewpoint of the analyst. As we shall see, subsequent diagrams can show the manner in which this condition directly led to events in the blackout. However, it would equally be possible to argue that the lack of vertical integration created similar problems. The key point here is that by distinguishing some conditions as potential vulnerabilities, these diagrams help to express and, therefore, expose the analysts' perspective on particular public policy issues. This is important because there are several competing views about whether Order 888 and PURPA created the necessary conditions in which the blackout was likely to occur. If analysts identify the condition that 'PURPA companies need not provide the same reliability guarantees as existing utilities' as a potential vulnerability then they must explain the mechanisms by which this led to particular engineering failures. Conversely, if an analyst contradicts this argument then they must provide evidence to undermine these mechanisms. In either case, it is essential that rhetorical arguments about the positive or negative impact of public policy should be grounded in the engineering details of the infrastructure failure.

The creation of 'open access' tariffs for the distribution network had a profound impact on the wholesale energy market. Energy marketers were able to trade power over increasing distances in response to pricing changes. Similarly, the utilities themselves began to trade power regionally both to gain revenue from their generation capacity and also to obtain additional power at lower costs. These trades created some of the preconditions for the August 2003 blackout as it became increasingly difficult for control area operators, such as MISO, to predict and resolve congestion problems. Power was increasingly traded on an ad hoc point to point basis. In response, a series of Transmission Loading Relief (TLR) procedures were drafted.

However, as mentioned before, the costs of implementing these agreements fell heavily on the utilities who owned the transmission system while many of the benefits accrued to the energy marketers who joined the market following the PURPA and Order 888 changes. It has also been argued that most of the regulatory attention was focused on ensuring fair access rather than on ensuring system reliability. For instance, the NERC elected a fully independent board of trustees in 2001 replacing the utility CEOs who had traditionally held these positions. This further distanced the owners and operators of the transmission systems from key regulatory issues. Previous sections have described the voluntary nature of NERC rules, the high costs associated with its requirements intensified the pressure to evade compliance as utilities were faced with increasing competition and with caps on their rates. The US-Canadian task board recognized that "recent changes in the electricity industry have altered many of the traditional mechanisms, incentives and responsibilities of the entities involved in ensuring reliability".

Figure 10 continues the V2 analysis of public policy influences on the August 14th blackout. As can be seen, many of the conditions created by the establishment of the NERC, by Order 888 and by PURPA resulted in four further vulnerabilities. The lack of NERC sanctions led to increasing violations of this non-governmental agency's requirements (US-Canada Task Force, 2004). It became increasingly difficult for Control Area operators, such as MISO, to predict energy movements. Energy marketers gained access to transmission networks without necessarily having to meet the associated infrastructure reliability costs, which fell mainly on the vertically integrated utilities. These vulnerabilities had further consequences that will be explored in subsequent diagrams.

It is also possible to trace a number of links between these public policy issues and the particular engineering events that took place during the blackout of August 14th. For example, Figure 10 shows how the changes introduced by PURPA and Order 888 arguably made it more difficult for organizations such as MISO to predict energy movements and hence ensure infrastructure reliability. This led to the drafting of Transmission Loading Relief procedures. The previous V2 diagram includes a link between the development of these procedures and the request at 15:35 from AEP to PJM to work on 350-MW Transmission Loading Relief to reduce the burdens on Star-South Canton line. This shows how our analysis can identify positive as well as negative outcomes from public policy decisions. The development of TLR procedures in response to market changes provided the transmission and reliability companies with ways of seeking relief under the uncertainties of the market. It was unfortunate, as we have seen from figure 5, that these procedures were insufficient to address the particular problems that arose from the failure of the Hanna-Juniper 345-kV line. Under other circumstances, with sufficient warning from RTCA tools, it might well have been possible to use the TLR procedures to mitigate the growing problems in the network.

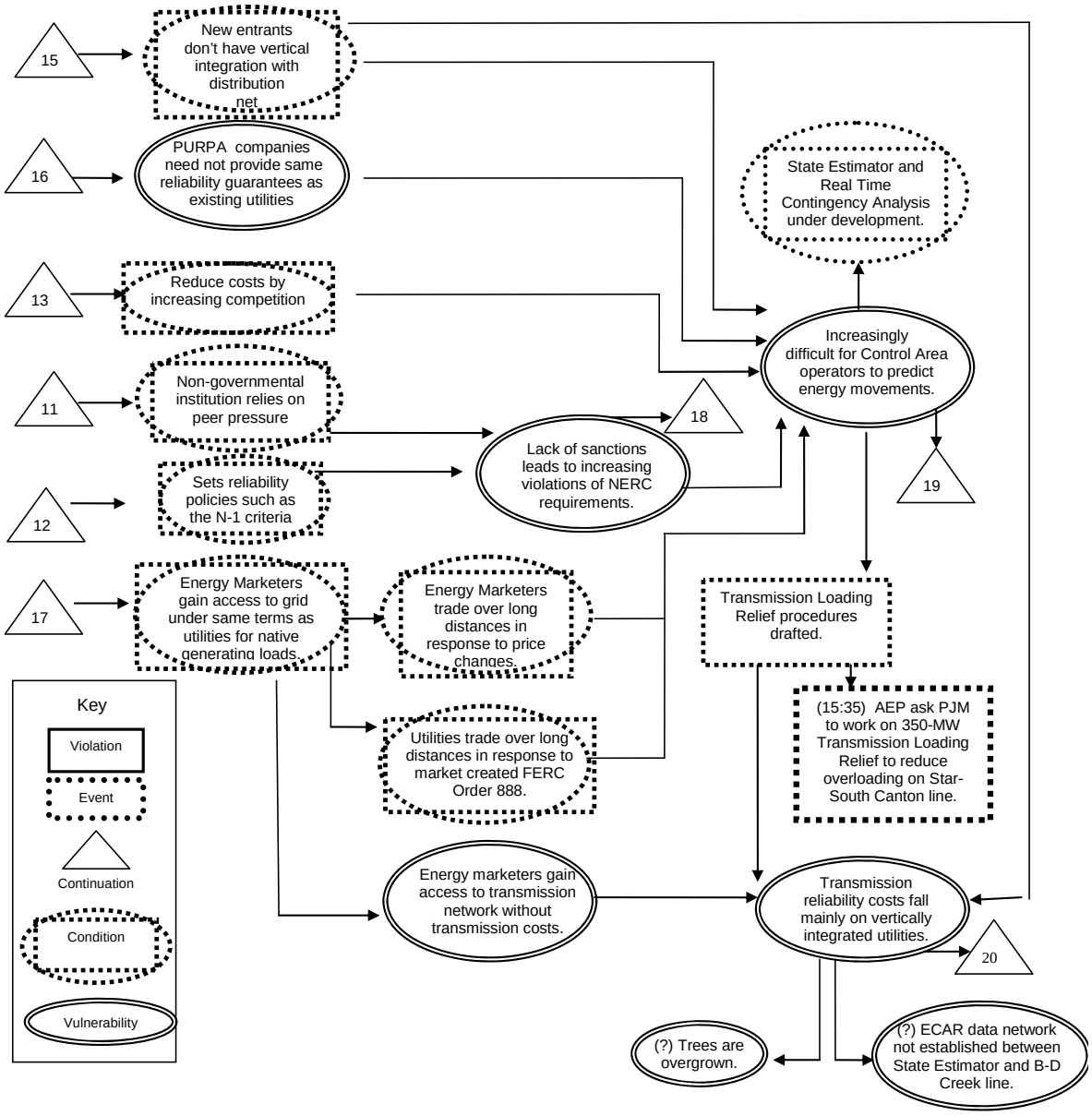

Figure 10: Instabilities in the Market Structure Prior to 2003

Figure 10 illustrates further links between public policy and the vulnerabilities that were identified in our previous analysis of engineering failures. For example, it can be argued that the way that vertically integrated utilities had to meet most of the reliability costs in a competitive market place may have contributed to the failure to prevent vegetation from growing to dangerous heights near power lines. The impact of this failure is represented in Figure 4's V2 diagram. Similarly, the uneven distribution of reliability costs may explain the lack of integration of the ECAR data network, for instance between the MISO State Estimator and Bloomington-Davis Creek 230-kV line. The V2 analysis in Figure 1 shows how this vulnerability, in turn, contributed to problems involving MISO's State Estimator and Real Time Contingency Analysis. These vulnerabilities are both annotated with a question mark in Figure 10. This indicates that additional evidence is required to support such a supposition. This is important because official reports are often weakened by considerable ambiguity in the relationship between public policy and engineering failures. Additional investigations must be performed to identify the impact that public policy

had on the managerial and organizational decision making processes, which in turn led to these vulnerabilities in infrastructure engineering.

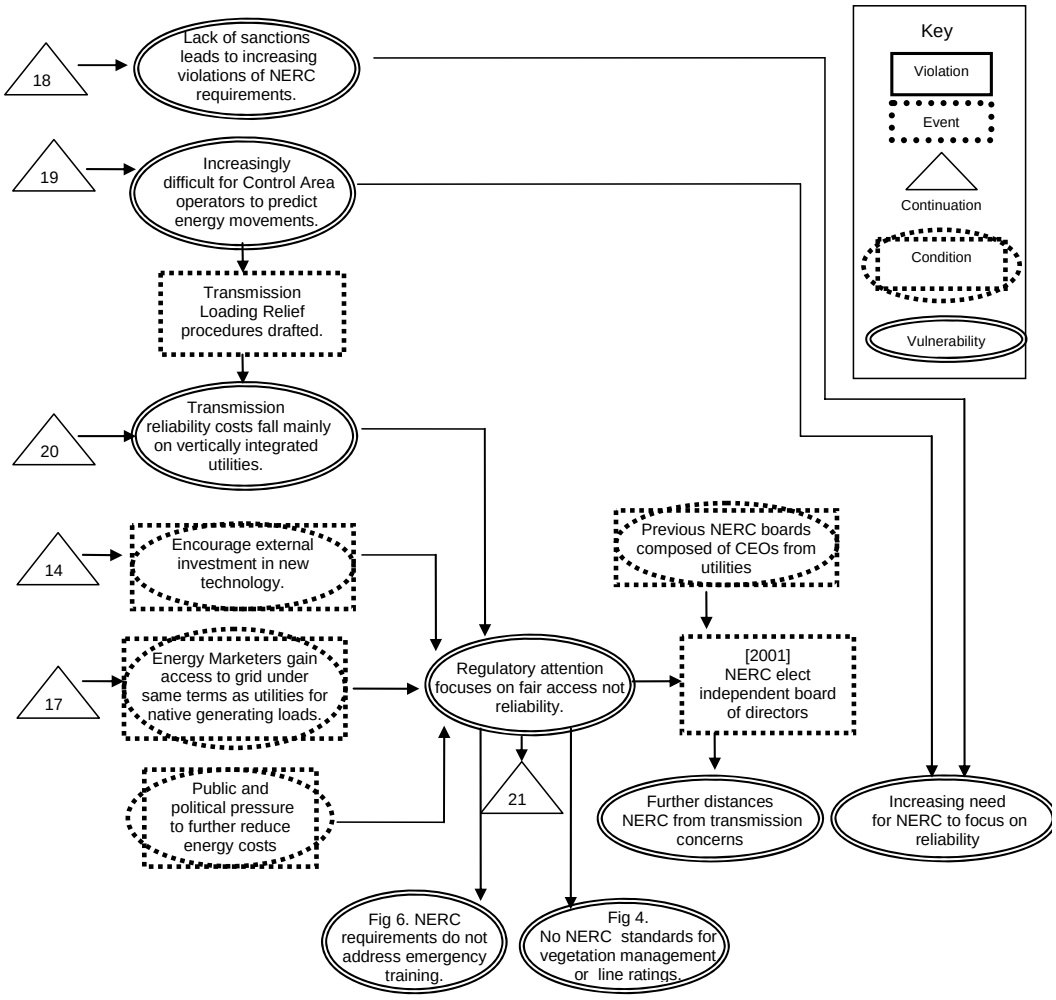

Figure 11: Market Access versus Reliability

Previous pages have shown that V2 diagrams can be used to identify the ways in which public policy creates the preconditions for infrastructure failures. Figure 11 extends this analysis by showing the conflicts that can arise within commercial and regulatory organizations. In this instance, we can see how various vulnerabilities identified in the regulatory and commercial structures made it critical that the NERC focus more on reliability. For example, the increasing number of NERC code violations combined with greater difficulty in predicting energy flows increased the need to look again at reliability issues. These vulnerabilities can be traced back to public policy decisions introduced in Figure 10. However, Figure 11 also illustrates the public policy influences that prevented the NERC from focusing more directly on reliability. These include the need to encourage external investment, public pressure to reduce costs and the role of marketers in creating new commercial opportunities. These conditions combined to focus the attention of the NERC more on market access than on infrastructure reliability. In turn, this may explain why the board became increasingly distanced from the vertically integrated utilities that helped to maintain the transmission infrastructure. As can be seen from Figure 11, this focus on market access may also explain particular engineering issues in the lead-up to the blackout. These include the lack of NERC standards for vegetation management or line emergency ratings, as shown in Figure 4 and the lack of NERC requirements for emergency training, as documented in Figure 6.

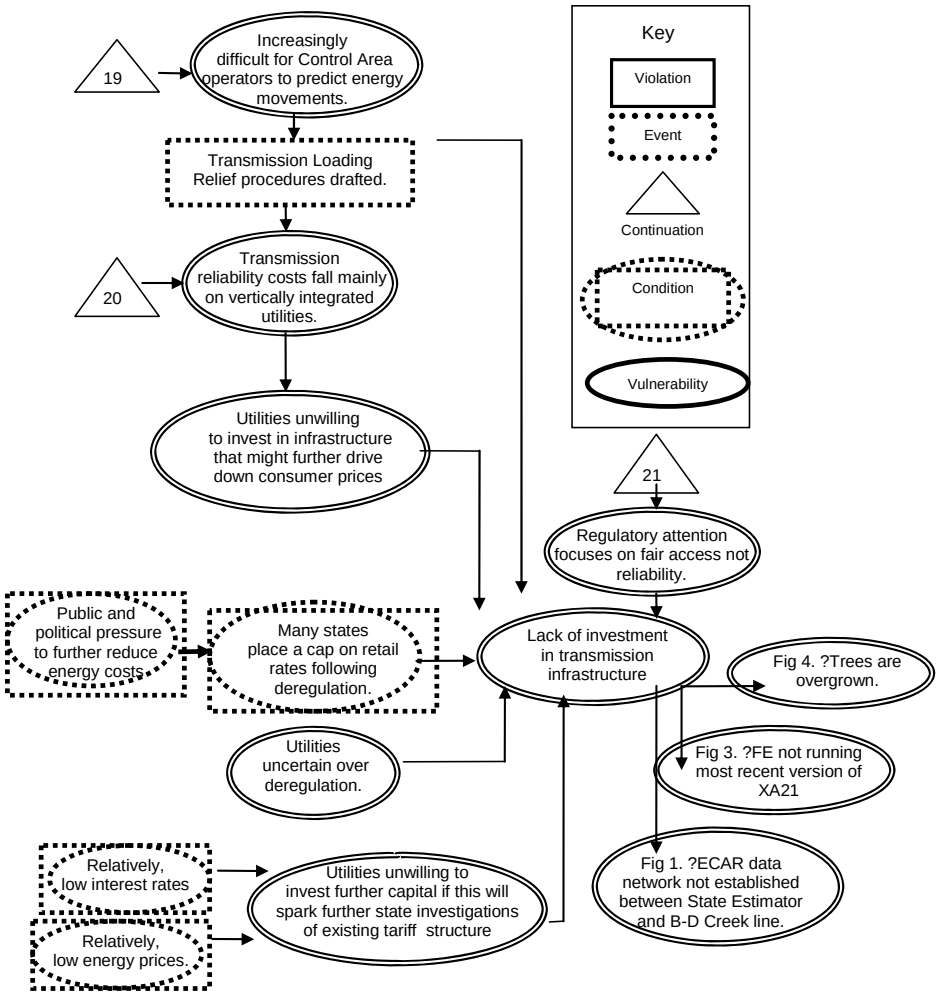

Figure 12: Deregulation as a Causal Factor in the August Blackout

Figure 10 illustrated the way in which V2 diagrams help to identify the need for additional evidence to support subjective claims about the role that public policy, such as market deregulation, can have upon the reliability of national infrastructures. Figure 11 extended this analysis to show how the same technique can identify conflicting public policy requirements, for example between ensuring open access in deregulated markets and ensuring infrastructure reliability. These are significant benefits because the opening paragraphs have described the considerable disagreements that have emerged over the causes of the blackout. For example, Hughes (2005) argues that the US-Canadian task force failed to address the underlying causes of the failure. In his view "deregulated companies are averse to building new generation that will drive down consumer prices and, therefore, their profits". Utilities were dissuaded from commissioning infrastructure improvements because they might have been forced into a more general review of their rate structure in order to justify any additional funding. They were reluctant to trigger these reviews in a 'partially deregulated' market given the relatively low interest rates and oil/gas prices. Further barriers to investment were created by the cap that many states placed on retail rates following deregulation. This limited the utility's ability to recover investments in new transmission through prince increases to retail customers (Wood 2005).

Figure 12 captures some of the previous arguments that link deregulation to the immediate engineering failures that led to the 14th August blackout. It is possible to trace the conditions that were created by FERC Order 888 and the PURPA legislation through the previous V2 diagrams to the vulnerabilities identified in the previous diagram. The utilities' concern about driving down consumer prices, the cap on retail rates and potential state reviews of existing tariff arrangements all dissuaded the utilities from further

infrastructure investment. The regulatory attention on ensuring open access to deregulated markets also removed some of the pressure to focus on reliability issues also acted to dissuade further investments. All of these issues, in part, explain the particular engineering failures introduced in the first half of this paper. The lack of investment that is linked to aspects of deregulation is linked in Figure 12 to the limitations of the ECAR network, introduced in figure 1 and to the lack of vegetation control, from Figure 4. The previous V2 diagram also suggests a relationship between the difficulties of planning for infrastructure investment in the deregulated market and delays in the planned upgrades to FE's XA21 monitoring servers.

As mentioned, this analysis is controversial. We cannot simply conclude that market deregulation caused the blackout simply because one occurred some time before the other. Causal inferences require that we present the mechanisms which connect an event to its outcome. Figure 12 sketches such relationships at a relatively high-level of abstract. Further analysis and evidence is required to support arguments that reduced infrastructure investment across the utilities led to delays in the XA21 project. Similarly, evidence must be provided to support arguments that infrastructure investment was falling in the manner described. This is necessary because we can also use the same V2 diagrams to sketch out causal arguments that identify very different relationships between public policy and the engineering failures that occurred on the afternoon of the 14th August.

The argument that deregulation created the latent causes for the blackout is not universally accepted, especially by those who benefited from the newly created energy market structure. In the aftermath of the blackout, PJM argued that "electric competition enhances, rather than compromises, grid reliability. Competition, supported by regional grid managers such as Regional Transmission Organizations (RTOs), brings stronger information, grid management tools and locational prices that make all market participants partners in reliability protection and reinforce and improve grid reliability" (Harris, 2005). In this view, the changes of the 1980s and 1990s helped utilities to lower costs and increase efficiency. The reduction in capital outlay by the utilities in the years immediately before the blackout can be explained in terms of a reduction in *over capacity* that had built up as a result of earlier investments that were based on over estimates of demand growth.

There is some evidence to suggest that long-distance energy transfers were increasing before the full impacts of deregulation affected the utilities in North America. It can, therefore, be argued that the increased flows were not caused by increased competition within the industry but by the rate caps imposed by states at a time when the costs of fossil fuels were rising. Existing utilities, therefore, had to look for cheaper energy sources from outside their immediate region. The motivation for utilities to import energy came not simply from intervention in the market through artificial price setting but also through the regulatory hurdles and local opposition that often frustrated attempts to build new power generation capacity close to the point of need. This analysis argues that the transmission grid is now being used in ways that were never intended when it was initially developed reflecting the influence of local environmental and business pressure groups as well as the supply and demand of energy.

Figure 13 shows how the same V2 diagrams that were used to associate the effects of deregulation with particular engineering failures can also be used by the proponents of deregulation to clarify their arguments in favor of market forces as a mechanism for ensuring infrastructure reliability. As can be seen, this interpretation draws upon many of the same events and conditions that are referred to in a critical account of deregulation. However, an additional vulnerability is introduced by the overcapacity of the 1980s and 1990s that were themselves the result of an over regulated market. This led to public pressure to reduce costs and is one reason for local and environmental opposition to building additional generating capacity. Public pressure to reduce costs encouraged states to intervene further in the market by introducing the price caps, mentioned above. This market intervention acted as a direct restraint on investment. It also created additional structural vulnerabilities because utilities were forced to look further a field for lower cost generating sources. The increased transmission of power from those sources contributed to network instability. Further barriers to investment came not from deregulation itself but from the manner in which that deregulation was implemented; utilities were uncertain about the long term viability of their position in the market as they bore the transmission costs for new entrants.

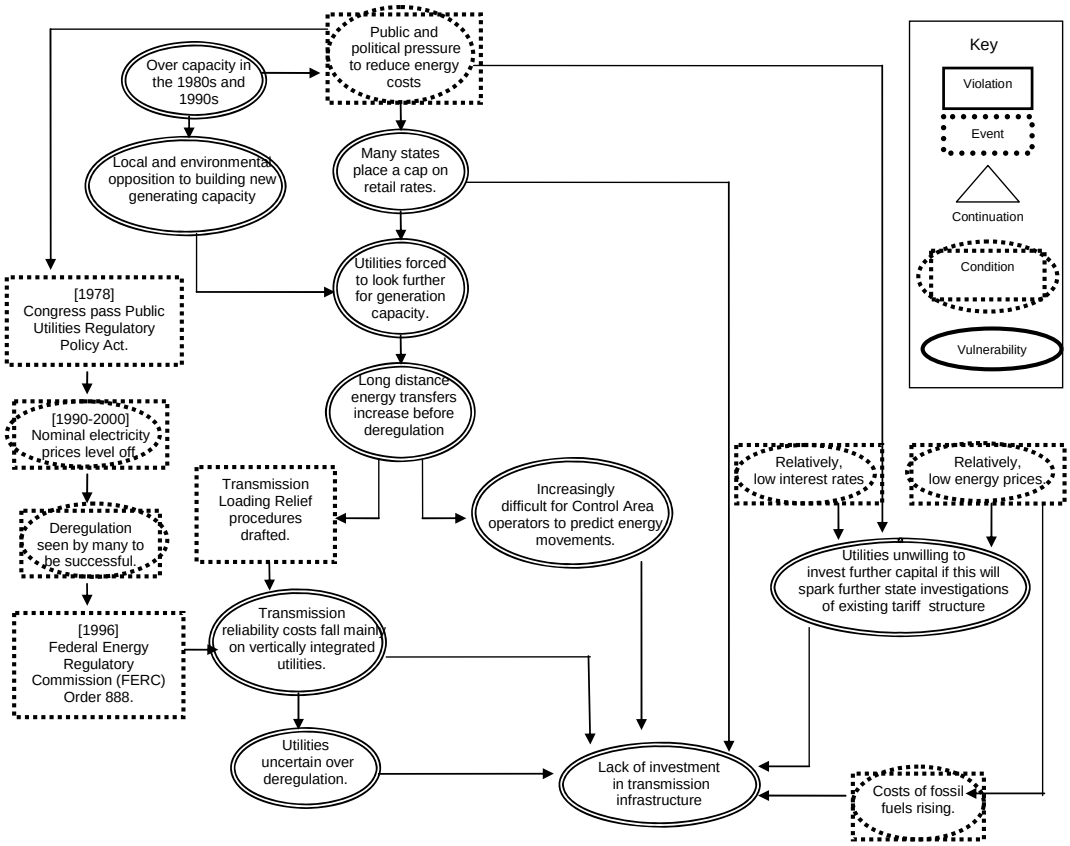

Figure 13: Counter-Arguments to Deregulation as a Causal Factor in the August Blackout

Engineering Recommendations and Public Policy Responses

The V2 diagrams introduced in this paper can chart the changes in public policy following adverse events. For example, many different organizations identified lessons from the August 2003 "blackout". For example, the Federal Energy Regulatory Commission (FERC) created an energy reliability division. This helps to form policy and develop standards as the generation and distribution industries respond to changing market conditions. Figure 14 illustrates how the creation of the new division can be seen as a response to some of the vulnerabilities that were identified from the blackout. This response is denoted by a pentagon. We do not distinguish between events and conditions, as we do between violations and vulnerabilities in the lead-up to an infrastructure failure. Future work must determine whether there are significant benefits from introducing additional symbols to distinguish between different types of response. Future work must also consider the distinction between completed and planned interventions in the aftermath of an adverse event. Many reliability standards are still under revision following the "blackout". It may be useful to provide a different graphical representation those responses that have been implemented and those that are still under consideration.

It is important not to underestimate the value of the simple annotations illustrated in Figure 14. They denote the relationship between the causes of an adverse event and the recommendations that are intended to avoid future recurrences. These links must be drawn if we are to prevent organizations from using previous incidents as an excuse or 'smoke screen' to justify recommendations that have little relationship to the incident or accident. In this example, Figure 14 shows that the introduction of FERC's reliability division can be related to the problems of maintaining a focus on reliability at a time when undue attention was focused on 'fair access' to deregulated markets. The creation of new organizational structures within FERC is intended to: "Allow prompt recovery of prudent expenses to safeguard reliability, security and safety; oversee the development and enforcement of grid-reliability standards; work with other agencies to improve infrastructure security; work with the states to support robust programs for customer demand-side participation" (FERC, 2004).

In August 2005, President Bush approved the Domenici-Barton Energy Policy Act. This created the Electric Reliability Organization (ERO) to establish and enforce standards throughout North America, including Canada and Mexico (Hughes, 2005). The Act is also intended to ensure that generation and distribution companies comply with these reliability standards. Figure 14 represents these initiatives. Federal enforcement actions address the lack of sanctions that led to increasing violations of NERC reliability requirements in the months before the blackout. By encouraging compliance with NERC standards, enforcement actions can also help regulatory organizations to predict energy movements. A key issue here is that V2 diagrams can also be used to identify subsequent audit requirements by explicitly linking recommendations to particular vulnerabilities. In this example, it is important to identify metrics to determine whether or not the Domenici-Barton Energy Policy Act increases compliance with NERC requirements. Similarly, Figure 14 illustrates the need to determine whether, in turn, these compliance actions can help reliability organizations to more accurately predict energy transfers.

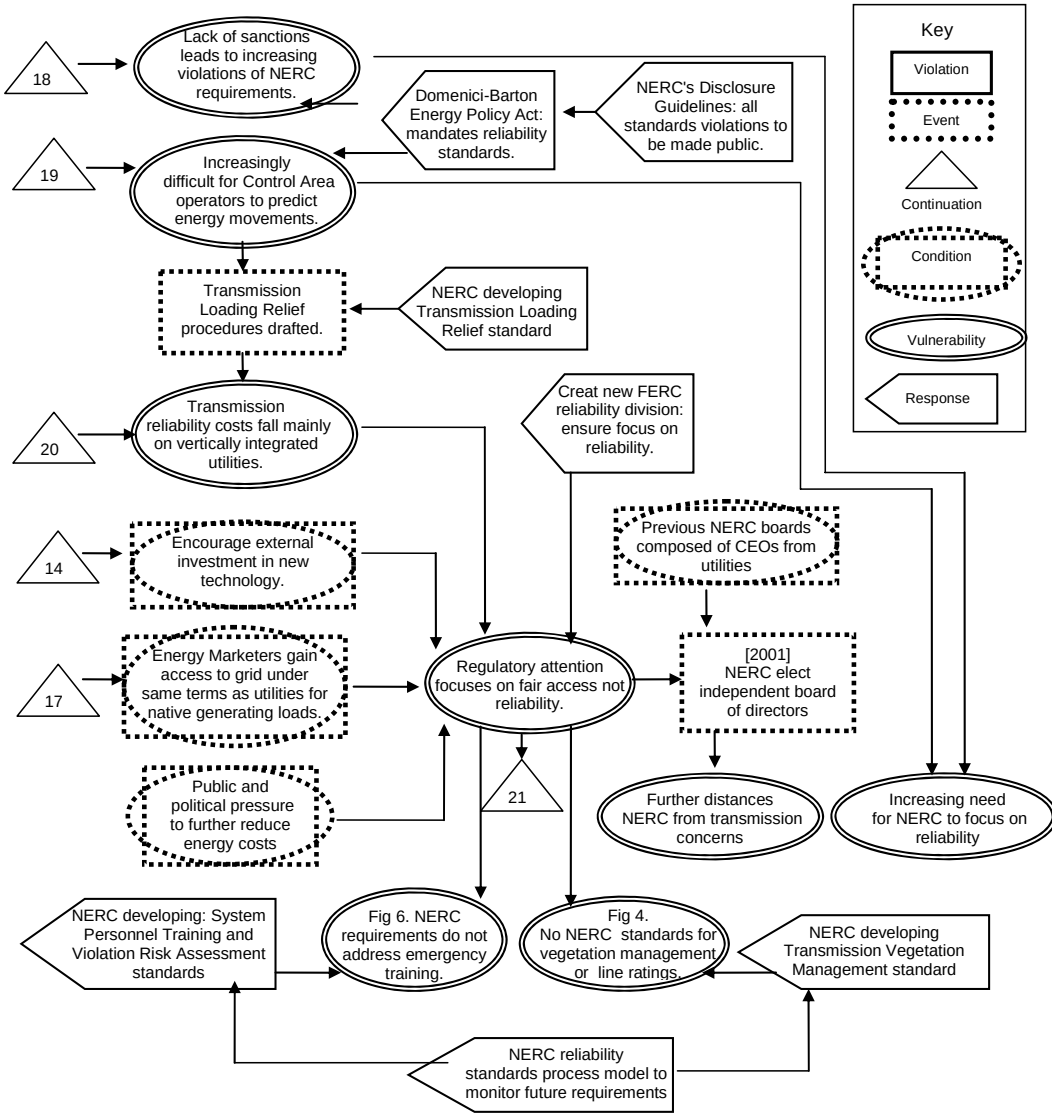

Figure 14: Responses to the August 2003 Blackout

The August 2003 "blackout" provided important insights into the limitations of existing reliability standards. Federal initiatives to encourage compliance would have few benefits unless actions were taken

to revise those standards in the aftermath of the infrastructure failure. For instance, the NERC is drafting new standards on: the management of tree growth in transmission line rights-of-way; operator emergency training requirements; real-time diagnostic and analytic tools for managing power flows on the power grids etc. Figure 14 illustrates the relevance of these draft standards by linking them to the vulnerabilities that the joint report identified as causes of the August 2003 blackout. It is important to note that additional work is required to determine whether the detail provisions of these new standards address specific concerns raised by the blackout. For example, any new standard on Transmission Loading Relief must clarify those situations in which this procedure can be used. As we have seen, PJM and AEP spent valuable time trying to negotiate a TLR that would only have had a very limited impact upon the developing failure. These specific requirements for a new TLR standard might be integrated into future analysis by extending the V2 diagram in Figure 5 to include detailed provisions from the NERC response.

The NERC has developed a reliability standards process model to establish a framework for these revisions. This model helps to identify the process of consultation and approval that is intended to ensure the coverage of, and compliance with, the new standards being developed following the blackout. Figure 14 illustrates the meta-level importance of this process model. In particular, this framework is intended to prevent future infrastructure failures being caused by a lack of applicable standards. It can be argued that the specific revisions to Transmission Loading Relief procedures or to vegetation management requirements only address the symptoms of the blackout. In contrast, the new process model addresses the underlying problems of reliability standards provision in a deregulated market.

The V2 diagram in Figure 14 sketches the interactions between different initiatives from commercial, regulatory and governmental organizations. This helps to ensure the 'joined-up' thinking that is often lacking with piecemeal public policy reforms to highly technical, infrastructure provision. In this instance, there is a danger that the provisions of the Domenici-Barton Energy Policy Act would be ineffective if organizations were reluctant to disclose NERC violations. Figure 14, therefore, also illustrates the importance of the NERC's new Guidelines for Reporting and Disclosure. This is intended to ensure that all confirmed violations of NERC standards are to be made public. The Federal sanctions rely upon effective reporting of NERC requirement violations. These reporting requirements rely upon Federal sanctions if reports are not filed.

Conclusions and Further Work

The relevance of this work should not be underestimated. The scope of this work also extends well beyond North America's energy infrastructure. For example, the liberalization of European energy markets have created conditions that are similar to those in the United States before 2003. Recent fluctuations in gas prices have made some countries reluctant to pass supplies across national borders without first ensuring the security of their own supply. This makes it difficult for transmission companies, utilities and regulators to make accurate predictions about future supplies. Similarly, plans to allow for the symmetric distribution of electricity by plants that consume power at some times but then generate electricity at others, for instance using renewable sources, will only work if we have a reliable and stable information technology infrastructure. This IT infrastructure must balance the supply and demand of base and reactive power. It must also provide for transparent and equitable systems of payment for both generators and infrastructure providers.

The relevance of this work is also illustrated by the continuing interaction between public policy and infrastructure engineering in the North American energy markets. The blackout continues to have an impact on public and political opinion. There is a renewed interest to ensure that industry bodies take direct action to avoid any repetition. For instance, the NERC has revised operator training requirements to include at least 32 hours of emergency drills including 'realistic' simulations for staff monitoring the reliability of bulk transfers. Not only has there been considerable pressure to revise the self-regulatory framework that supports the energy infrastructure, there is also growing political and public interest to ensure the effective policing of these requirements. It seems likely that any further reliability problems will trigger greater market intervention and regulation.

As we have seen, the blackout also continues to inform and motivate Federal intervention, including FERC reliability requirements for network analysis, transactions scheduling, grid forecasting etc. Many of these

regulations focus on the provision of information technology infrastructure. Initial studies again illustrate important differences in the impact of such public policy initiatives across the energy market; "a few very large utilities have invested in development and installation of the sophisticated, complex software tools identified as best practices needed for reliable grid operations" (Harris, 2004). In contrast, many smaller utilities retain "old, patched EMS, state estimator and contingency analysis software that does not allow precise, near-real-time evaluation of grid conditions and threats". Such technological disparities create the preconditions for future failures. You do not need to look far within the current market structure to realize that it contains the seeds of tomorrow's failures.

Biography:
Chris Johnson is Professor of Computing Science at the University of Glasgow in Scotland. He heads the Glasgow Accident Analysis Group, this small team of multi-disciplinary researchers is interested in understanding the role of computer applications in the failure of complex systems. He has held a NASA fellowship for his work on incident investigation techniques and helped to author incident reporting guidelines for European Air Traffic Management. He has published over 100 peer reviewed papers, including a Handbook of Accident and Incident Reporting. He coordinated the EC ADVISES Research Training Network supporting human factors approaches to the design of safety-critical systems.

References
J.P. Hughes, Reliability Risks during the Transition to Competitive Electricity Markets, Issue Papers on Reliability and Competition, Technical Workshop on Competition and Reliability in North American Energy Markets, US Department of Energy and Natural Resources Canada, Washington DC, August 2005. Available on http://www.energetics.com/meetings/reliability/pdfs/hughes.pdf, last accessed 21st January 2006.

J. Delgado, The Blackout of 2003 and its Connection to Open Access, Issue Papers on Reliability and Competition, Technical Workshop on Competition and Reliability in North American Energy Markets, US Department of Energy and Natural Resources Canada, Washington DC, August 2005.
Available on http://www.energetics.com/meetings/reliability/pdfs/hughes.pdf, last accessed 21st January 2006.

FERC, Division of Reliability, Presentation at Federal Energy Regulatory Commission Open Meeting, Washington, DC, October 6, 2004 by Joseph H. McClelland, Director, Division of Reliability, Office of Markets, Tariffs, and Rates.
Available on http://www.ferc.gov/EventCalendar/Files/20041013132551-A-2.ppt#258,2, FERC's Strategic Plan, last accessed 21st January 2006.

W.W. Hogan, Shedding Light, Wall Street Journal, Commentary, Page A20, 19th April 2004.

P.G. Harris, Relationship between Competitive Power Markets and Grid Reliability: The PJM RTO Experience, Issue Papers on Reliability and Competition, Technical Workshop on Competition and Reliability in North American Energy Markets, US Department of Energy and Natural Resources Canada, Washington DC, August 2005. Available on
http://www.energetics.com/meetings/reliability/pdfs/hughes.pdf, last accessed 21st January 2006.

J.P. Hughes, Reliability Risks during the Transition to Competitive Electricity Markets, Issue Papers on Reliability and Competition, Technical Workshop on Competition and Reliability in North American Energy Markets, US Department of Energy and Natural Resources Canada, Washington DC, August 2005.
Available on http://www.energetics.com/meetings/reliability/pdfs/hughes.pdf, last accessed 21st January 2006.

C.W. Johnson, A Handbook of Incident and Accident Reporting, Glasgow University Press, Glasgow, UK, 2003. http://www.dcs.gla.ac.uk/~johnson/book

Commonwealth of Massachusetts, Report of the Governor's Task Force on Electric Reliability and Outage Preparedness, March 2004, Available on http://www.mass.gov/dte/225repgtf.htm, last accessed 21st January 2006.

U.S.-Canada Power System Outage Task Force, Final Report on the August 14, 2003 Blackout in the United States and Canada: Causes and Recommendations, April 2004.
Available on https://reports.energy.gov/BlackoutFinal-Web.pdf, last accessed 21st January 2006.

United States Government Accountability Office, Electricity Restructuring: Key Challenges Remain, Report to The Chairman, Subcommittee On Energy And Resources, Committee On Government Reform, House Of Representatives, GAO-06-237, November 2005.
Available on http://www.gao.gov/new.items/d06237.pdf, last accessed 21st January 2006.