



# CloudZero Phase 2 Technical Report

Anthony Wallace,<sup>1</sup> Anna Liao,<sup>1</sup> David Rager,<sup>1</sup> Adarsh Hasandka,<sup>1</sup> Abhijeet Sahu,<sup>1</sup> Nicholas Ryan,<sup>1</sup> Steven Drake,<sup>1</sup> Josh Rivera,<sup>1</sup> Paul Snyder,<sup>1</sup> Bryan Richardson,<sup>2</sup> and Keith Schwalm<sup>2</sup>

*1 National Renewable Energy Laboratory  
2 Patria Security*

**NREL is a national laboratory of the U.S. Department of Energy  
Office of Energy Efficiency & Renewable Energy  
Operated by the Alliance for Sustainable Energy, LLC**

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at [www.nrel.gov/publications](http://www.nrel.gov/publications).

Contract No. DE-AC36-08GO28308

**Technical Report**  
NREL/TP- 5R00-88566  
May 2024



# CloudZero Phase 2 Technical Report

Anthony Wallace,<sup>1</sup> Anna Liao,<sup>1</sup> David Rager,<sup>1</sup> Adarsh Hasandka,<sup>1</sup> Abhijeet Sahu,<sup>1</sup> Nicholas Ryan,<sup>1</sup> Steven Drake,<sup>1</sup> Josh Rivera,<sup>1</sup> Paul Snyder,<sup>1</sup> Bryan Richardson,<sup>2</sup> and Keith Schwalm<sup>2</sup>

*1 National Renewable Energy Laboratory  
2 Patria Security*

## **Suggested Citation**

Wallace, Anthony, Anna Liao, David Rager, Adarsh Hasandka, Abhijeet Sahu, Nicholas Ryan, Steven Drake, Josh Rivera, Paul Snyder, Bryan Richardson, and Keith Schwalm. 2024. *CloudZero Phase 2 Technical Report*. Golden, CO: National Renewable Energy Laboratory. NREL/TP-5R00-88566. <https://www.nrel.gov/docs/fy24osti/88566.pdf>.

**NREL is a national laboratory of the U.S. Department of Energy  
Office of Energy Efficiency & Renewable Energy  
Operated by the Alliance for Sustainable Energy, LLC**

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at [www.nrel.gov/publications](http://www.nrel.gov/publications).

Contract No. DE-AC36-08GO28308

**Technical Report**  
NREL/TP-5R00-88566  
May 2024

National Renewable Energy Laboratory  
15013 Denver West Parkway  
Golden, CO 80401  
303-275-3000 • [www.nrel.gov](http://www.nrel.gov)

## NOTICE

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by the Department of Energy Office of the Chief Information Officer (OCIO). The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government.

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at [www.nrel.gov/publications](http://www.nrel.gov/publications).

U.S. Department of Energy (DOE) reports produced after 1991 and a growing number of pre-1991 documents are available free via [www.OSTI.gov](http://www.OSTI.gov).

*Cover Photos by Dennis Schroeder: (clockwise, left to right) NREL 51934, NREL 45897, NREL 42160, NREL 45891, NREL 48097, NREL 46526.*

NREL prints on paper that contains recycled content.

## Acknowledgments

For their support, input, feedback, and guidance throughout the first phase of this effort, we extend a special thanks to the project sponsors, Dan LaGraffe and Shannon Hughes; and to Luis Valencia, Dylan Sundy, and the rest of the Office of Chief Information Officer Control System Working Group. The cloud services and architecture would not have been possible without the generous guidance and support of Amazon Web Services.

## Executive Summary

The electric sector's rapid adoption of cloud services for data analysis and reliability operations has forced industry and regulators to consider the risks of the cloud environment at an untenable pace. The U.S. Department of Energy and the North American Electric Reliability Corporation have accelerated research in identifying threats and mitigation strategies in these new network architectures that involve commercial cloud services for bulk electric system reliability operations and forecast data analysis.

Researchers from the National Renewable Energy Laboratory (NREL) created the CloudZero project to study the threats and potential benefits to reliability provided by these new services. The project was accomplished in 2 phases designed to research the feasibility of cloud-based services to replace or supplement existing utility on-premises solutions.

Phase 1 of the CloudZero project provided a review for previous work and literature and laid foundation for securely moving power system data to commercial cloud infrastructure with live connections. These results indicated cloud services have matured to a state that brings improved reliability to our current infrastructure by decreasing recovery time in failure scenarios and improving the ability to scale services.

The primary goal of CloudZero Phase 2 was to increase the scale of testing with higher-fidelity modeling, adding rigor to the U.S. Department of Energy's ongoing efforts to address the issue of electric utility migration to the cloud. In this phase, researchers used the Advanced Research on Integrated Energy System (ARIES) Cyber Range at the National Renewable Energy Laboratory to establish a large-scale co-simulation environment of an urban portion of Austin, Texas, to evaluate utility system forecasting and control along with bulk electric system reliability operations using cloud native services. Researchers created a short-term load forecasting pipeline and a virtual, cloud-based remote terminal unit for the experiment. They then consulted with stakeholders in the electric sector and Amazon Web Services to create a short list of prioritized use cases for evaluation in the Advanced Research on Integrated Energy System Cyber Range, such as electric load forecast, system redundancy, and bulk electric system control operations.

Prioritized takeaways from this research include:

- Utilities can quickly leverage "out-of-the-box" cloud services including real-time machine learning forecast capabilities and data visualization that produces accurate models without Machine Learning expertise.
- The demonstrated system architecture was able to successfully show the capacity for failover using cloud-based services at a speed sufficient for most basic Supervisory Control and Data Acquisition (SCADA) services in switching servers within 2 minutes.
- A co-simulation framework within the cloud was able to enable the study of inter-dependencies between large-scale transmission and distribution systems. For example, how malicious control signals in the transmission system impact the voltages of the distribution feeder and how faults in the distribution feeder affect the transmission system are analyzed.

### Load Forecasting

Short-term load forecasting is a notable use case for the adoption of cloud across the sector; predicting load patterns supports planning and operations, including dispatching generation and switching operations. Researchers implemented a load forecast pipeline in the cloud, creating a high-fidelity representation of a real-world forecast system. By partnering with cloud subject matter experts, National Renewable Energy Laboratory researchers fed live power data from the on-premises distribution and transmission co-simulation to a live cloud sandbox. The setup further substantiated outcomes from the CloudZero Phase 1 report about connection stability. Researchers created a series of data transformation (extract, transform, and load) operations that produced a "live" forecast using cloud-native machine learning services. These services create a "point-and-click" interface that allows operation teams to create accurate forecast models using their own data with minimal training and knowledge. By shifting load forecasting processes to cloud native services, utilities can save orders of magnitude on labor and on-site equipment required for these artificial intelligence/machine learning capabilities.

Cloud services generated predictors for the time series data fed live to the cloud. These services performed within acceptable bounds even without the need for a machine learning expert to perform specialized hyper-parameter

tuning. Additionally, weather data included in the cloud-native forecast service were a valuable feature that increased the accuracy of some models.

A major find in our deployment discovered a cost reduction of several thousand dollars using Elastic File System storage instead of Simple Storage Service Buckets. Although this may appear counter intuitive, this discovery is due to high costs associated with sending or receiving data to and from Amazon Web Services, respectively. Amazon Web Services send and receive operations take place when a user executes either a PUT process function to send data, or a GET process function to receive data. These transactions become progressively more expensive if they trigger follow on functions or processes that consume resources every time one of these transactions occur. Given on-premises hardware and development costs to create a unique model for predictor creation, the switch to Elastic File System storage represents significant savings for a smaller utility.

### *Service Availability*

A loss of supervisory and control system availability can threaten grid reliability. Cloud service provider disaster recovery and failover capabilities are attractive for these types of systems. Phase 2 served as a proving ground for testing critical service availability use cases for utilities and grid operators. The environment was configured to test multiple failover or redundant architectures of varying costs for different objectives and needs of utilities. Our experiments in CloudZero Phase 1 proved out foundational concepts.

The Phase 2 architecture of the co-simulation was more complex and higher fidelity than Phase 1. For example, the Phase 2 environment contained approximately 78,000 loads on 89 distribution feeders compared with a small military installation microgrid in Phase 1. The testing resulted in failover times of approximately two minutes, expressed as the recovery time objective. The setup was a hot-standby failover, which does not implement an "always-on" secondary but still provides adequate service levels for most data processing circumstances. This setup also reported zero data loss, which could improve confidence in data integrity in a degraded cloud environment.

### *Bulk Electric System Reliability Operations*

CloudZero Phase 1 proved that latency on commodity cloud connections were well within the tolerable limits for certain bulk electric system reliability operations (Henry et al. 2022). Researchers deployed operational technology controller devices in the cloud and were able to test automatic generation control and other operations that have a higher tolerance for delayed action. They also deployed and tested remedial action schemes in the environment, increasing the model's small scale fidelity from Phase 1 to a much larger scale with more than 78,000 advanced metering infrastructure devices simulated within a city level environment. Our modeling team simulated a dynamic and interactive cyber-physical power system model that represents a major urban area of the United States with both transmission and distribution layers interacting together. The environment enabled researchers to begin answering questions at much higher scale. Cascading effects from attacks were simulated to the scale of a city, and recovery plans were validated.

### *Threat Scenario Execution*

Prior work in CloudZero Phase 1 focused primarily on system failure and recovery, so did not emulate threat actors. In Phase 2, threats were executed from the cloud assets to create realistic scenarios identified by utilities and cloud service providers as areas of concern. These threats were executed in a realistic manner using well-known attack frameworks and software. Additional cyber mitigations, such as replacing Telnet with Secure Shell Protocol, were explored to protect against these future threats and outcomes. These mitigations are pointed out in section 6 of this report. In addition to the preexisting on-premises solutions, cloud services provide supplemental features listed below that were evaluated in our research that encourage and promote secure configuration:

- **Encrypted data storage and network traffic:** Encrypt data stored in each process step, in managed databases, or in simple object storage. Encryption can be based on keys managed by either the organization or the cloud service provider.
- **Secure networking groups and subnet access:** Network access control lists should be leveraged to block unwarranted traffic within and external to specific areas of the network deployment.
- **Comprehensive logging:** Logging at the network and host level should be used liberally throughout critical points in the cloud deployment and the links from any on-premises systems.

- **Multifactor authentication:** This point should go without saying, but cloud service providers have vastly improved the security of their services by enhancing their service offerings around multifactor authentication. This feature should be leveraged in all elevated or critical context within infrastructure both in control and data analysis processes.

This report details these findings, highlights the benefit of cloud services to the energy sector, and demonstrates ways to address future security configurations of the electric grid.

# Table of Contents

<b>Acknowledgments</b>	<b>iv</b>
<b>Executive Summary</b>	<b>v</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Prior Work in CloudZero Phase 1	1
1.2 Objective of CloudZero Phase 2	1
<b>2 Enterprise Emulation Environment</b>	<b>3</b>
2.1 Experiment Enterprise Emulation	3
2.1.1 Enterprise Threat Scenarios	3
2.2 Experiment Malleability	4
2.2.1 Infrastructure as Code	4
2.2.2 Configuration as Code	4
2.2.3 Minimeta	4
<b>3 Load Forecast</b>	<b>5</b>
3.1 Overview	5
3.2 Deployed Load Forecast Cloud Architecture	5
3.3 Cloud Architecture	5
3.3.1 AWS IoT Core	5
3.3.2 Amazon Kinesis Data Firehose	6
3.3.3 Amazon Web Services Lambda	6
3.3.4 Amazon EventBridge	6
3.3.5 Amazon Web Services Step Functions	6
3.3.6 Amazon Forecast	6
3.3.7 Amazon OpenSearch	6
3.4 Cloud Usage Cost Estimate	6
3.4.1 Single-Feeder Cost Estimate	7
3.4.2 Ninety-Feeder Cost Estimate	7
3.5 Data Analysis	7
3.5.1 Feeder Data	7
3.5.2 Comparison of Forecast Versus Actual	10
3.5.3 Metric of Evaluation	10
3.5.4 Impact of Weather Data on Forecast Accuracy	11
3.5.5 Impact of Training Sample Size on Forecast Accuracy	11
3.5.6 Comparison of Forecasts on Two Different Feeders With Unique Load Profiles	11
3.5.7 Comparison of Amazon Forecast Versus Open-Source Long Short-Term Memory Approach	11
3.6 Load Forecasting Summary	12
<b>4 High Availability</b>	<b>13</b>
4.1 Historical Availability of Cloud Services	13
4.2 Failover Architecture	13
4.3 Highly Available Infrastructure and Services in Amazon Web Services	13
4.4 High-Availability Scenario Description	13
4.4.1 Scenario Details	14
4.5 Results	14
4.6 Analysis	15
<b>5 Bulk Electric System Reliability Operations</b>	<b>17</b>
5.1 Cosimulation Studies	17
5.1.1 Mapping Transmission to Distribution	18



5.1.2	Cosimulation Implementation . . . . .	18
5.1.3	Transient Modeling of the Travis County Generator . . . . .	18
5.2	Automatic Generation Control . . . . .	19
5.2.1	Test of Local Automatic Generation Control . . . . .	20
5.2.2	Test of Global Automatic Generation Control . . . . .	20
5.2.3	Implementation of Automatic Generation Control in Cyber Range: Operational Technology Simulator Logic Module . . . . .	21
5.3	Remedial Action Schemes Using Contingency Analysis and Situational Awareness . . . . .	23
5.3.1	Implementation of Remedial Action Scheme in Cyber Range: Operational Technology Simulator Logic Module . . . . .	23
5.4	Voltage Control . . . . .	24
5.4.1	Transmission-Side Voltage Control Using Exciter Control . . . . .	24
5.4.2	Implementation of Voltage Control in Cyber Range: Operational Technology Simulator Logic Module . . . . .	26
5.4.3	Distribution-Side Voltage Control Using Capacitor Banks and Transformer . . . . .	26
<b>6</b>	<b>Threat Scenario Integration . . . . .</b>	<b>27</b>
6.1	Data Pipeline Pollution Attack . . . . .	28
6.2	Command and Control Attack . . . . .	29
6.3	Cloud Provider Tools and Best Practices for Threat Mitigation . . . . .	30
6.3.1	Summary . . . . .	32
<b>7</b>	<b>Conclusion . . . . .</b>	<b>33</b>
7.1	Key Takeaways . . . . .	33
7.1.1	Load Forecasting Key Takeaways . . . . .	33
7.1.2	High Availability Key Takeaways . . . . .	33
7.1.3	Bulk Electric System Reliability Operations Key Takeaways . . . . .	33
7.2	Impacts on Utilities . . . . .	34
7.3	Future Phase Opportunities . . . . .	34
	<b>References . . . . .</b>	<b>37</b>
	<b>Appendix A Tables, Figures, and Code Snippets . . . . .</b>	<b>38</b>
A.1	Security Operations Center . . . . .	38
A.1.1	Enterprise Data Shippers . . . . .	38

## List of Figures

Figure 1.	Summary of Phase 1 Results . . . . .	1
Figure 2.	Load forecast architecture. . . . .	5
Figure 3.	Raw data showing aggregate residential load profiles from Feeder 9 <sub>0</sub> . Top plot shows load variation across several days. Lower plot shows the same data, but over a four-month period. . . . .	8
Figure 4.	Raw data showing aggregate commercial load profiles from Feeder 9 <sub>1</sub> . Top plot shows load variation across several days. Lower plot shows the same data, but over a two-month period. . . . .	9
Figure 5.	Raw data showing aggregate mixed load profiles from Feeder 26 <sub>0</sub> . Top plot shows load variation across several days. Lower plot shows the same data, but over a two-month period. . . . .	10
Figure 6.	Forecast vs. actual load data for Feeder 9 <sub>0</sub> . . . . .	11
Figure 7.	AWS system architecture . . . . .	14
Figure 8.	U.S. east failover charts . . . . .	15
Figure 9.	U.S. west failover charts . . . . .	16
Figure 10.	Chart highlighting failover . . . . .	16
Figure 11.	Power System Use Case Considered for the Cosimulation . . . . .	17
Figure 12.	Impact of generator outage on the ACE of the test area, followed by the new generation set points for the other two generators (Generator 2 MW and Generator 3 MW) in the test area. The negative ACE caused the system frequency (Generator 2 frequency) in the area to decrease. . . . .	21
Figure 13.	Certain loads in Travis County at substations DS 31, 102, 107, 109, and 110 are tripped. . . . .	22
Figure 14.	The ACE is positive, which causes the generator megawatt set point to decrease, causing the ACE to dip even further into negative values. . . . .	22
Figure 15.	The ACE is positive, which causes the generator megawatt set point to reduce to 150 MW, causing the ACE to stabilize. . . . .	23
Figure 16.	One line fault and two solid-phase faults at two buses . . . . .	24
Figure 17.	Voltage collapse without any exciter in any of the three generators in the Sam Gideon Power Plant in the Test area . . . . .	25
Figure 18.	Impact of number of generators when the exciter control is enabled . . . . .	25
Figure 19.	Impact of exciter voltage set point . . . . .	25
Figure 20.	The application monitoring the per-unit voltage at the terminals of the Sam Gideon Power Plant when there is a fault on the load Bus 17. . . . .	26
Figure 21.	Geographic representation of distribution and transmission . . . . .	27
Figure 22.	Command and control attack architecture . . . . .	28
Figure 23.	Snapshot from Internet of Things (IoT) Core service showing one feeder’s aggregated active power polluted to zero . . . . .	28
Figure 24.	Attacker pivoting from virtual Remote Terminal Unit (vRTU) in the cloud to Remote Terminal Unit in Substation . . . . .	29
Figure 25.	Load bus is opened by Telnet attack . . . . .	30
Figure 26.	Geographic representation of blackout caused by breaker opening . . . . .	31

## List of Tables

Table 1.	Single-Feeder Cost Estimate . . . . .	7
Table 2.	Ninety-Feeder Cost Estimate . . . . .	7
Table 3.	Seventy-Thousand-Device Cost Estimate . . . . .	7
Table 4.	Feeder Mix of Residential Meters and Commercial Meters . . . . .	8
Table 5.	Residential Feeder 9 <sub>0</sub> Predictor Errors . . . . .	9
Table 6.	Commercial and Mixed Feeder Predictor Errors . . . . .	10
Table 7.	Root Mean Square Error for the Long Short-Term Memory (LSTM) Trained Predictors for Feeder 9 <sub>0</sub>	12
Table A.1.	Point of Common Coupling Between Transmission and Distribution Systems . . . . .	39
Table A.2.	Generator Transient Modeling . . . . .	40
Table A.3.	Impact of Line Outage Contingencies. The impact factor represents the percentage by which a transmission line is overflowing. An impact factor of 104.0 indicates that a transmission line is 4% above its maximum power carrying capacity. . . . .	41
Table A.4.	Impact of Line Outage Contingencies After One of the RAS is Implemented. . . . .	41

## ACRONYMS

AAF	Active Active Failover
ACE	Area Correction Error
AiTM	Adversary-in-the-Middle
AGC	Automatic Generation Control
API	Application Programming Interface
ARIES	Advanced Research on Integrated Energy Systems
ARP	Address Resolution Protocol
AWS	Amazon Web Services
BES	Bulk Electric System
BRO	BES Reliability Operations
CaC	Configuration as Code
CDR	Connected Data Rate
DLT	Data Loss Tolerance
DNP3	Distributed Network Protocol 3
EC2	Elastic Cloud Compute
EFS	Elastic File System
ERCOT	Electric Reliability Council of Texas
FedRAMP	Federal Risk and Authorization Management Program
FTP	File Transfer Protocol
HA	High Availability
HELICS	Hierarchical Engine for Large-Scale Infrastructure Co-Simulation
HSF	Hot Standby Failover
HTTP	Hypertext Transfer Protocol
ICCP	Inter-Control Center Communications Protocol
ICS	Industrial Control Systems
IEEE	Institute of Electrical and Electronics Engineers
IoT	Internet of Things
LSTM	Long Short-Term Memory
MAPE	Mean Absolute Percentage Error
MQTT	Message Queuing Telemetry Transport
NERC	North America Electric Reliability Corporation
NREL	National Renewable Energy Laboratory
OpenDSS	Open Distribution System Simulator
OT	Operational Technology
OT-Sim	Operational Technology Simulator

MITRE	MIT Research Establishment of Adversarial Tactics, Techniques and Common Knowledge
PLC	Programmable Logic Controller
PLF	Pilot Light Failover
PSS	Power System Stabilizer
p.u.	per unit
PWDS	PowerWorld Dynamic Studio
RAS	Remedial Action Schemes
RDS	Relational Database Service
RMSE	Root Mean Square Error
RTO	Recovery Time Objective
RTU	Remote Terminal Unit
vRTU	Software Based Virtual Remote Terminal Unit
S3	Simple Storage Service
SANS	SysAdmin, Audit, Network, and Security
TELNET	Teletype Network - Insecure Communication Protocol Transmitting Plain Text
VPC	Virtual Private Cloud
VPN	Virtual Private Network
VVC	Volt-Volt ampere reactive Control
WSUS	Windows Server Update Services

# 1 Introduction

On Tuesday, September 6, 2022, California reached a historic and nearly catastrophic milestone. The load demand for the state electric power coordinator peaked above 52,000 MW, and controlled blackouts were staged for execution (California State Assembly: Committee on Utilities and Energy 2022). The governor made an urgent plea to the public for assistance, and air conditioning control (mainly through residential thermostats) contributed significantly to maintaining grid stability. This demand response action via cloud controlled devices is just one example of the many electric sector applications where cloud infrastructure may have a more meaningful impact on grid reliability at scale. Continued adoption of electric vehicles, distributed renewable generation, and behind-the-meter battery systems will further drive further cloud adoption for electric sector applications.

An increasingly distributed grid requires platforms for testing demand response strategies and other important reliability operations for a secure electric infrastructure. We created the CloudZero project to develop a proving ground for these unique problems and buy down the risk to utilities looking to migrate workloads and control to the cloud.

## 1.1 Prior Work in CloudZero Phase 1

Phase 1 of the project focused on background research of the current state of cloud applications and explored uncertainty regarding the performance of infrastructure such as latency and bandwidth constraints. Those results were compared to on-site deployments of private cloud infrastructure, and data integrity was verified at multiple points in the cloud and on site in the Advanced Research on Integrated Energy Systems (ARIES) Cyber Range. Additionally, it established industry partnerships leveraged to build a test bed to lay the foundation for more complex phases of analysis and demonstration. The power system deployed for these tests was a simple microgrid based on a previous National Renewable Energy Laboratory (NREL) project at a small military installation. Figure 1 includes some abbreviated results of Phase 1.





Scenario	Description	Results
<b>On-premise latency/bandwidth test</b>	Bandwidth and latency testing using ping and iPerf3 between two test machines in two virtual local area networks within the experiment. This scenario was used as a control for comparison with the measurements from the cloud tests.	 Avg. Mbps = 100 Mbps Avg. latency = 2.87 ms
<b>Cloud tunnel latency/bandwidth test</b>	Bandwidth and latency testing using ping and iPerf3 between a test machine in the on-prem environment and a test machine in the AWS virtual private cloud.	 Avg. Mbps = 11.6 Mbps Avg. latency = 38.48 ms
<b>Data integrity verification</b>	A comparison of a data sample from both the on-prem historian and the cloud historian within AWS.	 No difference (Consistency was verified)
<b>Disaster recovery</b>	Simulated cloud outage and recovery of the data historian connection and replication.	 Time to recovery was 119 seconds without reconfiguration or data loss

Figure 1. Summary of Phase 1 Results

## 1.2 Objective of CloudZero Phase 2

Phase 2 of CloudZero took advantage of the foundations laid by Phase 1 and demonstrated a much more robust deployment and set of scenarios. Using a large data lake of synthetic power models, researchers created test scenarios consisting of a distribution and transmission simulation of an urban area of Austin, Texas. "The [Synthetic Models for Advanced, Realistic Testing: Distribution Systems] (SMART-DS) data sets provide standardized distribution network models that have been validated against thousands of real utility systems" ("SMART-DS: Synthetic Models for

Advanced, Realistic Testing: Distribution Systems and Scenarios” 2023). The distribution layer contained more than seventy thousand loads (commercial and residential), and the transmission layer represented all of Travis County. This co-simulation was run with the Hierarchical Engine for Large-scale Industrial Control Systems (HELICS), and the system was coupled to a corporate control network via an NREL virtual control system device simulator called Operational Technology Simulator (OT-Sim).

The notional corporate networks created to represent live virtual systems in two distribution companies and a transmission company contain faithful representations of utility systems, such as domain controllers, email services, and file sharing. These networks lay the groundwork for future work in creating a balancing authority to manage relationships (e.g., energy markets, dispatch, demand response, etc.) between the three entities.

*Note* - Throughout this report, Amazon Web Services (AWS) is referenced as the cloud solution provider used by NREL researchers to conduct the research and obtain use case results. It is important to note, however, that although we selected AWS for this project, other cloud service providers offer similar services.

## 2 Enterprise Emulation Environment

The convergence between information technology, operational technology, and cloud services is critical in our efforts to understand how information technology and operational technology systems interconnect with cloud services across a wide-area network. Although the interaction between information technology, operational technology, and cloud base services might seem straightforward at first glance, the optimal way in which these systems interconnect is typically not understood well enough to achieve the most secure and resilient state possible. One way to bolster the understanding of optimizing the convergence items is to evaluate them through lens of the Purdue model (Williams 1993). The Purdue model is a structural model for industrial control systems (ICS) that describes segmentation of physical processes, sensors, supervisory controls, operations, and logistics. Using principles from this model, the team created a way to rapidly deploy an emulated enterprise environment and related services to enable analysis of threat vectors across the enterprise, including integration points between cloud and edge-level energy services. During this research, our team built automated scripts to ease several challenges encountered when deploying and evaluating corporate services. The methods described in this section automated the integration of interconnections within a generalized corporate environment and external cloud service providers such as AWS.

### 2.1 Experiment Enterprise Emulation

We scoped NREL's enterprise emulation scenario to include key corporate services that are standard systems and applications services that integrate and interconnect across utilities and cloud scenarios. For our emulation, we deployed the following systems and services as virtual machines and integrated them through software-defined networks within NREL's ARIES Cyber Range:

- Domain controller: Windows Server 2019, active directory users and groups
- Email server: Windows Server 2019, Windows workstation access to Microsoft Exchange email service
- FTP server: Windows Server 2019, Windows workstation file share service access to users
- WSUS server: Windows Server 2019, Windows workstation and server updates and patching service
- HTTP server: Windows Server 2019, Windows workstation access to web services
- Windows Workstation: Windows 10/11, workstation authorized to access services across the domain.

#### 2.1.1 Enterprise Threat Scenarios

The following list highlights the MITRE ATT&CK framework tactics and techniques that NREL considered for the threat scenarios. These threat scenarios detail how a threat actor would likely attempt to access the enterprise network. The focus was primarily on vulnerabilities and exploits on physical systems in the operation network. For the threat emulations and research in this report, NREL assumed that the threat actor gained access to the enterprise network through common and well-known phishing attacks:

- Initial access: External remote services ID: T1133, phishing ID: T1566
- Execution: Exploitation for client execution ID: T1203, command and scripting interpreter ID: T1059
- Persistence: Create or modify system process ID: T1543
- Privilege escalation: Valid accounts ID: T1078, exploitation for privilege escalation ID: T1068
- Defense evasion: Impair defenses ID: T1562, masquerading ID: T1036
- Credential access: Exploitation for credential access ID: T1212, steal or forge authentication certificates ID: T1649.
- Discover: Account discovery ID: T1087.

For more information on the threat scenarios compromising the operational technology network system scenarios, see Section 6, Threat Scenario Integration.



## 2.2 Experiment Malleability

To establish a foundational model for the emulated enterprise, researchers developed and integrated an enterprise system leveraging development, security, and operations (DevSecOps) principles. Using infrastructure as code and configuration as code methods, the researchers ensured that, from the foundation up, all systems within the emulated enterprise were developed using a secure and resilient design. The team created a rapid development workflow and build process using tools such as Ansible, Packer, minimega, and Phēnix. Although it is not directly within the scope of the analysis, note that the planning, design, and development of "as code methods" allows system security engineers to lead development with system security principles at its core. In addition, we included a module for scenario development that will continue to be developed, scaled, cloned, and optimized based on the needs of the enterprise scenario. The key points of the environment orchestration include:

- Translation into how modern security managers should be looking at security operations, management, and integration solutions
- Methods of emulation in the NREL Cyber Range related to the optimization and evaluation of enterprise integration/interconnection scenarios between edge-level operational technology and cloud-based systems
- Introduction of information technology/operational technology and cloud scenarios enabling intrinsic design principles that extend through the edge and across enterprise services to cloud services and across an emulated wide-area network.

### 2.2.1 Infrastructure as Code

Researchers developed and orchestrated infrastructure as code modules to produce the baseline enterprise emulation services. Infrastructure as code methods allowed the development team to iterate on different approaches and enable system services and configurations based on the needs of defined scenarios. The team used Packer as the standard tool set to build the evaluation's virtual disk images. The evaluation also used plug-ins for tools like PowerShell and Ansible to build the emulation service as a versioned module. We built Linux systems using Phēnix, a custom solution through our experiment services. The Phēnix disk image functionality allowed developers to build Linux virtual disk images via streamline automation. The virtual images of both Windows and Linux were staged as modules and deployed in the scope of the experiment. We deployed additional configurations in the system based on the defined scenario via alternative configuration as code and scenario orchestration methods.

### 2.2.2 Configuration as Code

We developed configuration as code to enable specific kinds of system scenarios and allow the development team to iterate on different approaches, enabling system services and configuration based on the needs of specific scenarios. We used Phēnix to inject and orchestrate systems and components in active experiment scenarios and Ansible to automate configuring and running the build process of Packer. For example, we employed file injections through Phēnix to start and/or stop services for data collection and transport based on scenario system state requirements.

### 2.2.3 Minimeta

Minimeta is a minimega extension for hosting multiple containers as smart meters within each virtual machine in the minimega space. The purpose of developing Minimeta is to scale thousands of smart meters in the advanced metering infrastructure network. For instance, in this project, NREL researchers considered one urban region of the city of Austin, TX, comprising 75,000 smart meters (Figure 11b).

## 3 Load Forecast

### 3.1 Overview

Amidst the rapid adoption of intermittent inverter-based resources at the bulk electric system (BES) and distribution system levels, it is increasingly important to accurately forecast load and distributed energy resource generation to maintain reliable energy delivery. With the increased frequency and volume of distribution data and by leveraging artificial intelligence/machine learning analysis, NREL researchers demonstrated how this capability supplemented with cloud native services will improve load forecasting and increase overall grid stability.

Traditional industry solutions are on-site systems that gather 15- to 30-minute time series data. Generally, these data are aggregated at the distribution feeder in an on-site historian. The CloudZero team created a high-fidelity system similar to this setup, but the data were directly streamed to the AWS Internet of Things Core for storage into Simple Storage Service (S3) and analysis by the Amazon Forecast service to produce predictors that perform short-term load forecasting. This is a scalable, reliable system that can handle 1-minute intervals of "live" data simulated from virtual systems. Although the CloudZero team selected AWS and its associated services as the cloud service provider for this research, it is important to note that other cloud service providers exist that we could have selected to provide the services and conduct the research performed and described throughout this document.

The proposed reference architecture aims to deploy a hybrid on-site/cloud system that can reliably handle high volume, high frequency data streams. To achieve this, NREL researchers also created a reference hybrid data pipeline to stream data from the on-site cyber range to data storage and machine learning forecast capabilities in the cloud.

### 3.2 Deployed Load Forecast Cloud Architecture

Message Queuing Telemetry Transport (MQTT) messages from the on-site cyber range are received at IoT Core. The data are routed to Amazon Kinesis via IoT Topic and IoT Rule. There is a Lambda function integrated with Amazon Kinesis to format the data into time stream comma-separated values required for import to Amazon Forecast. Time series data for the past hour are imported to Amazon Forecast, and the next forecasted load profile for the next 6 hours is generated. The 6-hour forecast data are then visualized in OpenSearch.

### 3.3 Cloud Architecture

We implemented a hybrid streaming data pipeline architecture with data from the on-site cyber range to the AWS cloud services (Figure 2).

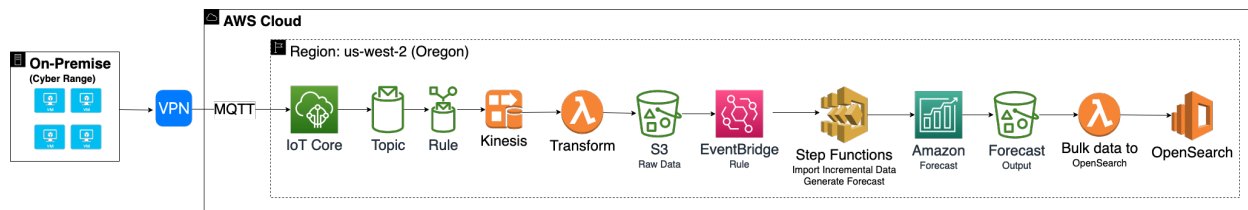


Figure 2. Load forecast architecture.

#### 3.3.1 AWS IoT Core

AWS IoT Core serves as a secure, bidirectional internet communication gateway, providing a link between various remote devices and the AWS cloud. In this use case, IoT Core is used to receive streaming data from the on-site cyber range as MQTT messages. IoT Core has a message broker that provides a mechanism for secure communications, which is used to publish and receive messages via direct MQTT.

MQTT topics identify AWS IoT messages. AWS IoT clients identify the messages they publish by giving the messages topic names. The message broker uses topic names and topic filters to route messages from publishing clients to subscribing clients.

### 3.3.2 Amazon Kinesis Data Firehose

An IoT Core rule is defined to route messages to Amazon Kinesis Data Firehose, which receives real-time streaming data to be processed and formatted for input to the machine learning forecast service. Amazon Kinesis Data Firehose can be integrated with an AWS Lambda function to implement extract, transform, and load.

### 3.3.3 Amazon Web Services Lambda

AWS Lambda is an event-driven, serverless compute service that executes code in response to events. AWS Lambda is used to transform and consolidate the MQTT message contents to a format compatible as input to the Amazon Forecast service. We used another AWS Lambda function for the bulk data upload of the generated forecast data to OpenSearch.

### 3.3.4 Amazon EventBridge

Amazon EventBridge is a serverless event bus service. An Amazon EventBridge rule triggers the AWS Step Functions state machine when time stream data are available to import to Amazon Forecast.

### 3.3.5 Amazon Web Services Step Functions

AWS Step Functions can orchestrate multiple AWS services into serverless workflows to quickly build and update applications. In the AWS console for Step Functions, users can view the architecture diagram and visualize the workflow as it progresses through each stage. AWS Step Functions is leveraged to orchestrate all the application programming interface calls to import incremental time stream data to Amazon Forecast and generate forecast data.

### 3.3.6 Amazon Forecast

Amazon Forecast is a fully managed time series forecasting service that uses the same machine learning technology as Amazon.com. A Forecast predictor is a model that is trained using a target time series. By default, Amazon Forecast creates an AutoPredictor, where the Amazon Forecast applies the optimal combination of algorithms to each time series in the users' data sets. NREL researchers applied two built-in features for Weather Index and U.S. holiday. Weather Index incorporates historical and projected weather information into the model. The holidays feature also incorporates a feature engineered data set of national holiday information.

We trained predictors with 2 months of historical data, from July 2, 2018, to Aug. 31, 2018. We chose July 2, 2018, as the start date because that is the earliest date available in the weather index feature for the United States. By default, Amazon Forecast computes the weighted quantile loss at P10, P50, and P90. 'PXX' specifies the threshold such that the true value is expected to be lower than the predicted value XX% of the time. For example, P50 specifies the threshold such that the true value is expected to be lower than the predicted value 50% of the time. Ideally, the actual feeder load profile will be between P10 and P90 because the actual is within the bounds of overestimation and underestimation.

### 3.3.7 Amazon OpenSearch

Amazon OpenSearch is the AWS-managed service equivalent to Elasticsearch for visualizing the forecast time series plot. There is an event-triggered AWS Lambda function to bulk import the forecast data into OpenSearch and to visualize the forecast time series plot.

## 3.4 Cloud Usage Cost Estimate

For our experiments, we simulated feeder-level aggregation in the on-site cyber range and sent MQTT messages to IoT Core.

We used S3 as the primary storage service for our experiments, but when the data are scaled to millions of meters, the transaction costs to transfer files are significant. To reduce these transaction costs, our recommendation to grid operators leveraging AWS cloud services is to use an Elastic File System (EFS) storage service mounted to both transform Lambda functions. EFS storage costs are three times greater than S3 storage costs, but they carry no transactional costs. EFS is comparable to the common on-site Network File System (NFS) service. Assuming data storage of 8 GB, which is equivalent to 1 day of data, using EFS would reduce storage costs by \$4,000 per month.

### 3.4.1 Single-Feeder Cost Estimate

The cost estimate shown in Table 1 assumes data for a single aggregated feeder. The data are aggregated in the on-site cyber range and sent to one MQTT device in AWS IoT Core.

**Table 1. Single-Feeder Cost Estimate**

Service	Assumption	Data Frequency	Monthly Cost
IoT Core	Single MQTT device	720 messages/month	\$0
Forecast	2 KB of 1-minute data produced/hour	1.44 MB/month	\$1,000
Kinesis	2 KB of 1-minute data produced/hour	2 KB/hour	\$10
VPC	Always on, with active VPN and transit gateway	1.44 MB/month in, 9.46 MB/month out	\$350
EFS	Retains data for last 24 hours	48 KB/day	\$0

### 3.4.2 Ninety-Feeder Cost Estimate

The cost estimate shown in Table 2 assumes data for several aggregated feeders. Data are aggregated in the on-site cyber range and sent to one MQTT device in AWS IoT Core per aggregated Feeder.

**Table 2. Ninety-Feeder Cost Estimate**

Service	Assumption	Data Frequency	Monthly Cost
IoT Core	90 MQTT devices	32,400 messages/month	\$5
Forecast	2 KB of 1-minute data produced/hour/aggregate	139 MB/month	\$1,000
Kinesis	45 records ingested/hour	180 KB/hour	\$10
VPC	Always on, with active VPN and transit gateway	1.44 MB/month in, 9.46 MB/month out	\$350
EFS	Retains data for last 24 hours	4.3 MB/day	\$0

Table 3 shows the monthly cost estimates for a load forecast pipeline hosted in AWS with the capacity to ingest all 70,000 MQTT devices worth of data, given certain assumptions, as shown:

**Table 3. Seventy-Thousand-Device Cost Estimate**

Service	Assumption	Data Frequency	Monthly Cost
IoT Core	70,000 MQTT devices	10,000 messages/month/device	\$1,000
Forecast	Hourly forecast, 9 hours/month training	4 GB/month	\$1,000
Kinesis	300 records ingested/second	5.4 GB/hour	\$200
VPC	Always on, with active VPN and transit gateway	240 GB/month out	\$350
EFS	Retains data for last 24 hours	4.3 MB/day	\$0

## 3.5 Data Analysis

### 3.5.1 Feeder Data

The feeder data shown in Table 4 are from the Open Energy Data Initiative data lake (<https://data.openei.org/>) based on the Synthetic Models for Advanced, Realistic Testing: Distribution Systems and Scenarios (SMART-DS) project, Austin city model for the year 2018.

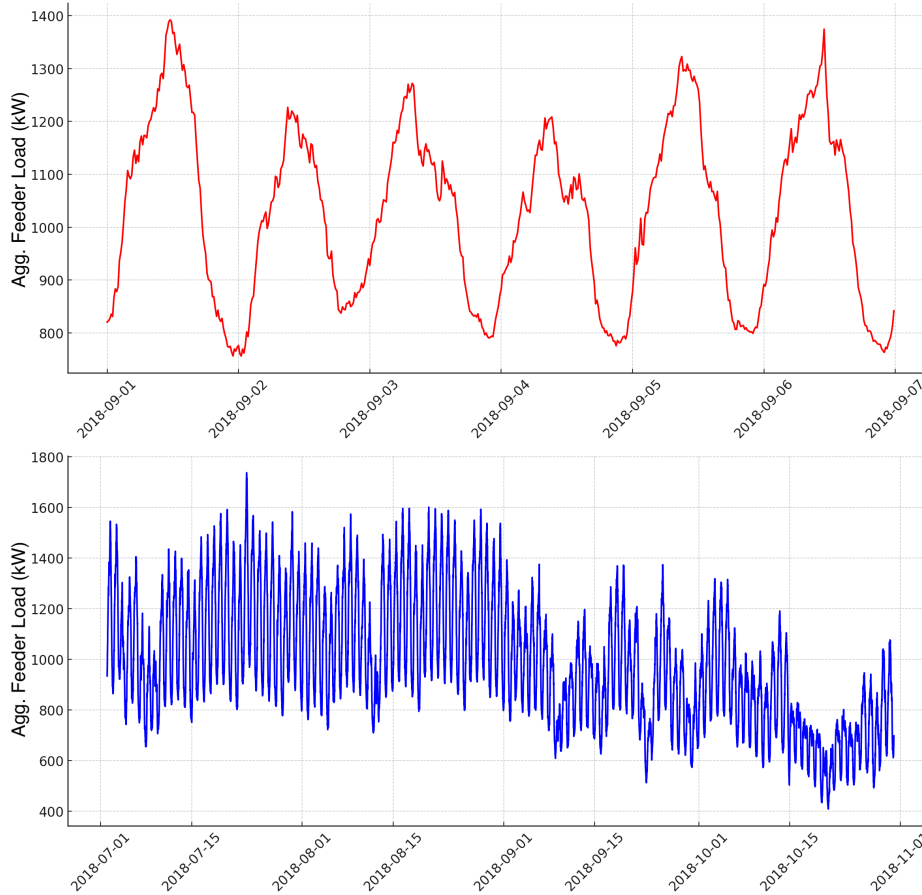
We designated feeders using a specific notation, for example, residential Feeder 9<sub>0</sub>, where 9 represents the substation number, and the subscript 0 refers to the first feeder of that substation.

**Table 4. Feeder Mix of Residential Meters and Commercial Meters**

Feeder	Residential Meters	Commercial Meters
9 <sub>0</sub>	139	56
9 <sub>1</sub>	0	22
26 <sub>0</sub>	311	760

*Residential Feeder*

Within Feeder 9<sub>0</sub>, 71% of the total meter counts is residential while 29% are commercial customers. We expect that variability in load on this feeder will be strongly driven by residential buildings, compared with the other feeders in this study. The aggregate raw electricity load profile is shown in Figure 3. The top plot shows the load profile for six days. The the bottom plot shows the load profile for 4 months. There is a noticeable seasonal change from August to September, correlating to the higher load in summer and lower load in autumn. There is less variability in September and October, as evidenced by reduced daily load fluctuation going from August into September. Table 5 shows a comparison of the forecast predictor errors trained on the raw data set. We observed during the research that as we increase weather data features such as temperature, humidity, and precipitation, it requires additional training data for improvement in prediction. Hence, when trained with 4 months of data versus 2 months of data, there was improvement reflected in forecast accuracy with the inclusion of weather data.



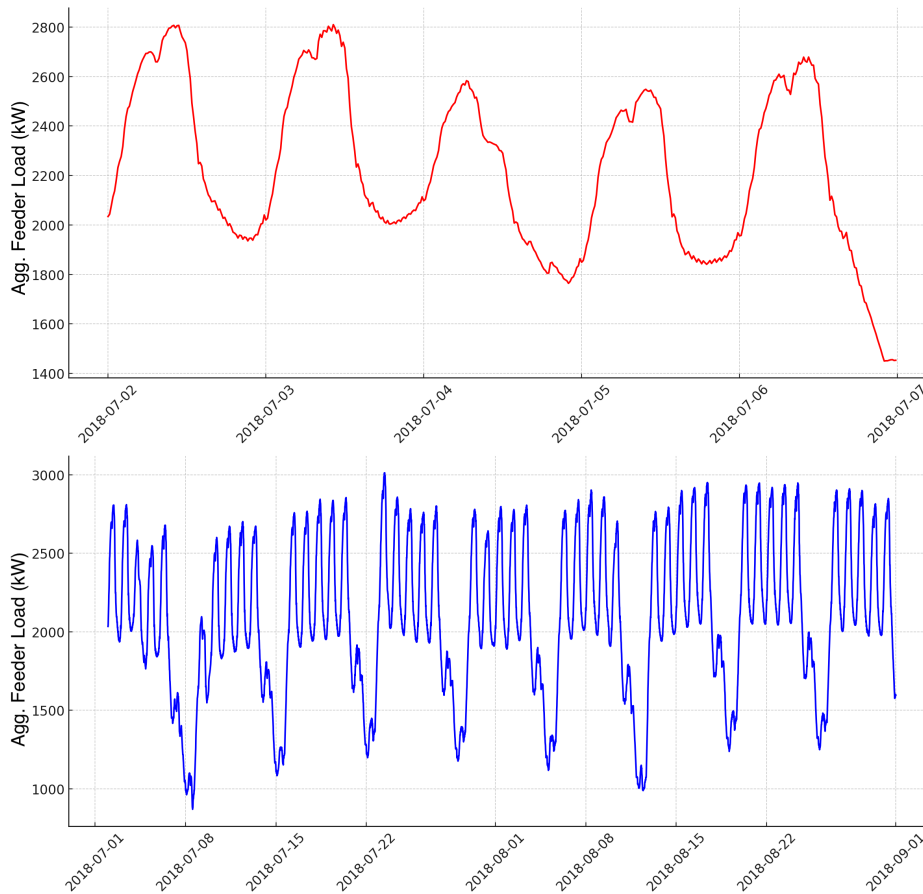
**Figure 3. Raw data showing aggregate residential load profiles from Feeder 9<sub>0</sub>. Top plot shows load variation across several days. Lower plot shows the same data, but over a four-month period.**

**Table 5. Residential Feeder 9<sub>0</sub> Predictor Errors**

Training Data Time Period	Number of Months	Weather Index N		Weather Index Y	
		RMSE	MAPE	RMSE	MAPE
Jul–Aug	2	39.6	0.0408	52.5	0.0551
Sep–Oct	2	23.6	0.0313	25.8	0.0357
Jul–Oct	4	18.0	0.0233	16.1	0.0215

*Commercial and Mixed Feeders*

Feeder 9<sub>1</sub> is all commercial load. The raw electricity load profile is shown in Figure 4. There is a distinctive pattern in weekday versus weekend load demand. The demand significantly drops on the weekends.



**Figure 4. Raw data showing aggregate commercial load profiles from Feeder 9<sub>1</sub>. Top plot shows load variation across several days. Lower plot shows the same data, but over a two-month period.**

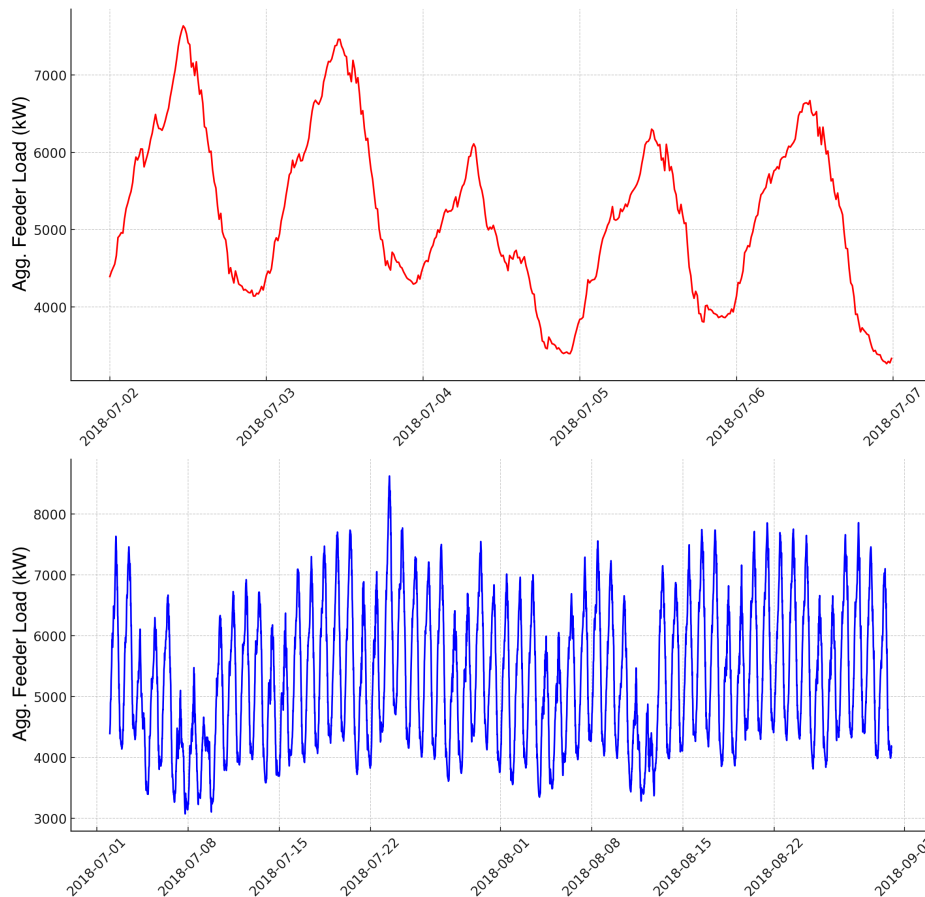
Feeder 26<sub>0</sub> is 71% commercial load and 29% residential load. The raw electricity load profile is shown in Figure 5. Although the load decreases on the weekends, the residential load demand is still evident in the daily load demand cycles on the weekends. The decrease in load demand around July 4 is likely due to the holiday. The Amazon Forecast holiday feature would factor this date into the machine learning algorithm for this scenario. We trained all these forecast predictors with data from July 2, 2018, to Aug. 31, 2018.

Feeder 9<sub>1</sub> is all commercial loads so the demand is very uniform, probably following a typical commercial building schedule, and we expect this feeder load profile to have fewer errors. Because feeder 26<sub>0</sub> has a mix of residential and commercial loads, we hypothesize that the demand is variable and would have more errors.

Comparison of commercial feeder predictor errors is shown in Table 6. Feeder 9<sub>1</sub> only has commercial load. The weather index significantly decreases the error in the predictors.

**Table 6. Commercial and Mixed Feeder Predictor Errors**

Feeder	Weather Index N		Weather Index Y	
	RMSE	MAPE	RMSE	MAPE
9 <sub>1</sub>	208.5	0.1201	15.1	0.0069
26 <sub>0</sub>	222.0	0.0527	360.8	0.0857



**Figure 5. Raw data showing aggregate mixed load profiles from Feeder 26<sub>0</sub>. Top plot shows load variation across several days. Lower plot shows the same data, but over a two-month period.**

### 3.5.2 Comparison of Forecast Versus Actual

Figure 6 provides an example of the forecast error bounds generated by Amazon Forecast, compared to the actual load data of feeder 9<sub>0</sub>. The actual data are within the p10 and p90 bounds. The actual load data have a sudden jump at 10 a.m. on Sept. 1, 2018. The forecast assumes a smooth aggregated trend line and it makes sense that it may not be able to predict a sudden jump in load, unless that sudden change occurs regularly at 10 a.m..

### 3.5.3 Metric of Evaluation

Amazon Forecast produces accuracy metrics to evaluate predictors, including root mean square error (RMSE), weighted quantile loss, mean absolute percentage error (MAPE), mean absolute scaled error, and weighted absolute percentage error. This portion of the evaluation considers MAPE and RMSE.

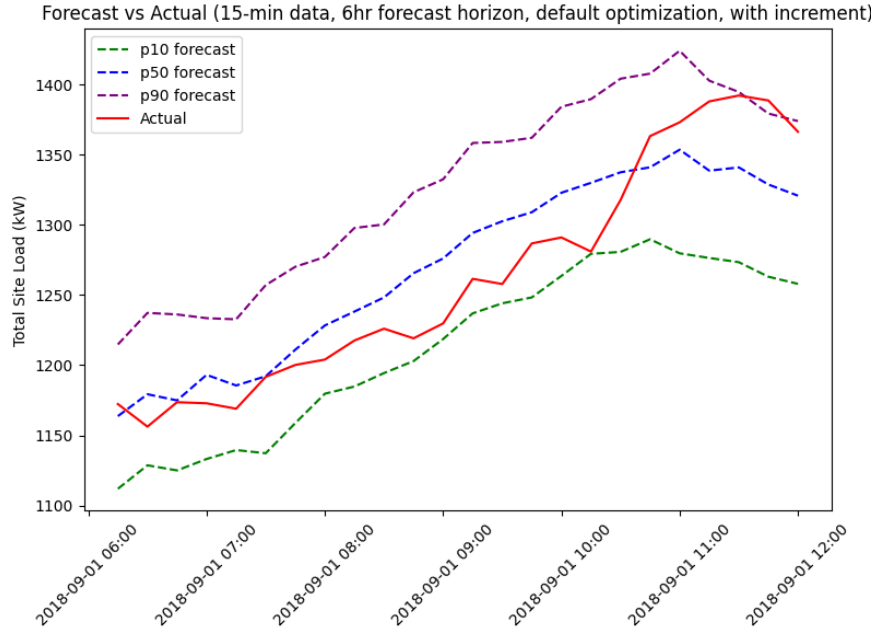


Figure 6. Forecast vs. actual load data for Feeder 9<sub>0</sub>

### 3.5.4 Impact of Weather Data on Forecast Accuracy

The Amazon Forecast Weather Index is a built-in feature that incorporates historical and projected weather information in the learned model; hence, we trained our predictors with weather data to understand its sensitiveness to prediction accuracy. The weather data are integrated based on the Austin city weather information from July 2018 to synchronize with the load data built in the SMART-DS project for Austin, Texas, for 2018 (Palmitier et al. 2020).

Weather Index could have a significant impact on feeder load data because of possible changes in heating, ventilating, and air-conditioning use in Austin for summer versus winter. Holidays could also impact feeder load data because electricity usage patterns change on holidays compared to regular business days.

### 3.5.5 Impact of Training Sample Size on Forecast Accuracy

As shown in Table 5, when we train the predictors including the weather index data with 2 months of data for Feeder 9<sub>0</sub>, the predictor error increases. Conversely, when we train the predictor with 4 months of data, the inclusion of the weather index data decreases the predictor error. This indicates that for data spanning longer time frames, time-synchronized weather data are a more significant factor in improving load forecasts. As more detailed weather information is added to the data set—such as temperature, humidity, or precipitation—the forecast model operates with more dependent variables, making the model more complex, and generally improving performance and reducing error.

### 3.5.6 Comparison of Forecasts on Two Different Feeders With Unique Load Profiles

A comparison of a feeder with a high ratio of residential to commercial load, e.g., Feeder 9<sub>0</sub>, to one with a low ratio, e.g., Feeder 9<sub>1</sub> (Table 5 and Table 6), shows that more residential load causes higher variance in power usage than feeders with commercial load. The inclusion of the weather data results in low forecast errors for Feeder 9<sub>1</sub>.

### 3.5.7 Comparison of Amazon Forecast Versus Open-Source Long Short-Term Memory Approach

A variant of a recurrent neural network, long short-term memory (LSTM) is very well suited for problems with temporal and sequential data; thus, we performed LSTM-based forecasting using an open-source PyTorch package with an aggregated data set from Feeder 9<sub>0</sub> (Table 7) to compare its performance with Amazon Forecast. The Open Energy Data Initiative data set comprises load data for 15 minutes using load interpolation to generate higher-resolution 1-minute data for solving the OpenDSS feeder and for generating active and reactive power data for an



interval of 1 minute. A comparison of the results in our evaluation (Table 7 and Table 5) shows that Amazon Forecast performed better than the open-source LSTM implementation based on the RMSE values. This is due to the hyperparameter tuning.

**Table 7. Root Mean Square Error for the Long Short-Term Memory (LSTM) Trained Predictors for Feeder 9<sub>0</sub>**

LSTM sequence	1 min	15 min
10	28.66	145.53
20	612.4	169.70
30	170.03	475.09
40	817.189	976.86

### 3.6 Load Forecasting Summary

This load forecast use case evaluates how cloud can be leveraged to ingest and handle 1-minute data and correspondingly higher volumes of data. It validated that cloud services can immediately scale as needed to process data at higher rates and volume. This is compared to on-prem systems, where dedicated engineers would be required to scale the hardware up to handle more data. One can quickly leverage out-of-the-box cloud services including real-time data ingest, on-the-fly extract transform load services, data storage, machine learning forecast capabilities, and data visualization. Some experiments also validate the efficacy of using cloud services rather than third party or open-source services.

## 4 High Availability

### 4.1 Historical Availability of Cloud Services

Based on historical data (Li et al. 2013), despite the high service-level agreement advertised by cloud service providers, there is always an inherent risk of a service outage. As an example, the service-level agreements provided by AWS Amazon Web Services 2023b show most services used in the selected architecture are guaranteed to provide at least 99.5% uptime or better. This inherent risk can come from simple human error, hardware failure, or even natural disasters. From the data set analyzed in (Li et al. 2013), collected using surveys from major cloud service providers in 2013, natural disasters were responsible for only 6.4% of outages. In comparison, power outages are much more frequently responsible, at 38.5% of outage incidents.

### 4.2 Failover Architecture

Various failover architectures are available within AWS depending on the criticality, sensitivity to data loss, cost tolerance, and availability requirements of the cloud service provided. The failover types considered in this report are pilot-light, hot-standby, and active-active failover architectures. Generally, the faster time for recovery and lower data loss risk provided by the high-availability architecture increases operational costs. The choice of failover architecture is evaluated by considering the trade-offs between the criticality of service, tolerance to outage, and cost sensitivity. For this project, we selected an active-active failover architecture involving parallel redundant operations due to the advantages and benefits it provides (described in the next section) compared to the other commonly used methodologies. Due to the highly available nature of this architecture, this is what AWS recommends, especially when implementing redundancy, for any systems controlling critical infrastructure.

### 4.3 Highly Available Infrastructure and Services in Amazon Web Services

AWS has several services and features within services that might offer high availability and resilience. Some features are architectural, such as Regions and Availability Zones, whereas others are service-specific features, such as S3 replication. An AWS region is a physical geographic location hosting a cluster of data centers. Each region is independent and has its own set of services, infrastructure, and availability zones. AWS provides the option to deploy applications and services in one or multiple regions, as needed. By having multiple regions across the world, AWS purports to offer low-latency connectivity, fault tolerance, and disaster recovery options for their customers that comply with applicable regulatory standards (Amazon Web Services 2023d).

AWS Availability Zones are logical data centers within an AWS region. Each availability zone is isolated and physically separate from other availability zones within the same region, with its own power source, networking, and connectivity to the internet. It can be an economic solution to provide resilience by deploying applications across multiple availability zones within the same region to ensure that services remain available even if there is a disruption in one availability zone. This does not protect from an AWS region-wide outage, however, which can occur when there are catastrophic or extreme weather events in a specific geographic region.

AWS provides solutions designed to achieve optimal performance, reliability, and redundancy for applications and systems hosted on the AWS platform. These services are built on a distributed infrastructure across multiple availability zones within a specific region. These services employ automatic scaling, load balancing, and fault-tolerant mechanisms to provide reliability. Some examples of well-known services that support high availability include Amazon Elastic Cloud Compute (EC2), S3, and Relational Database Service, among others. This allows for the deployment of these services independent of any specific virtual private cloud (VPC), with the cost of the reliability services amortized over the individual service cost.

### 4.4 High-Availability Scenario Description

To better understand the impacts of implementing high availability and failover in a cloud native implementation of the BES reliability operations use case as well as the load forecasting use case, we developed an active-active setup for the implementation in a different region.

#### 4.4.1 Scenario Details

##### Step 1: Initialize Experiment

We initialized the experiment with data at an arbitrarily selected point in time from the previously used load forecasting data set. We then deployed the experiment with this initialization to represent an outage occurring at a random point in time during the normal operation of the system. We allowed the system to run for a few minutes before proceeding to the next step.

##### Step 2: Execute Failover During Experiment

After we verified that the experiment was correctly operating as before, we simulated the natural disaster scenario by applying a policy in the primary region to break all active connections to the primary region from any meters. These IoT meters were configured to automatically failover to the secondary region when they lose the primary connection. By shifting this failover function into the devices themselves, the cloud architecture's ability to provide high availability was simplified in both complexity and cost. Once the data started to flow into the pipeline for the secondary region, we monitored the secondary S3 bucket and the Amazon Forecast service for impacts.

##### Step 3: Evaluate Results

When we had executed the failover scenario, we immediately verified the capacity of the devices to switch to the secondary region. After we concluded the experiment, however, we needed to analyze the metrics that identify the recovery time objective and the recovery point objective provided by this system from the metrics collected from the AWS CloudWatch.

The overall architecture diagram is shown in Figure 7.

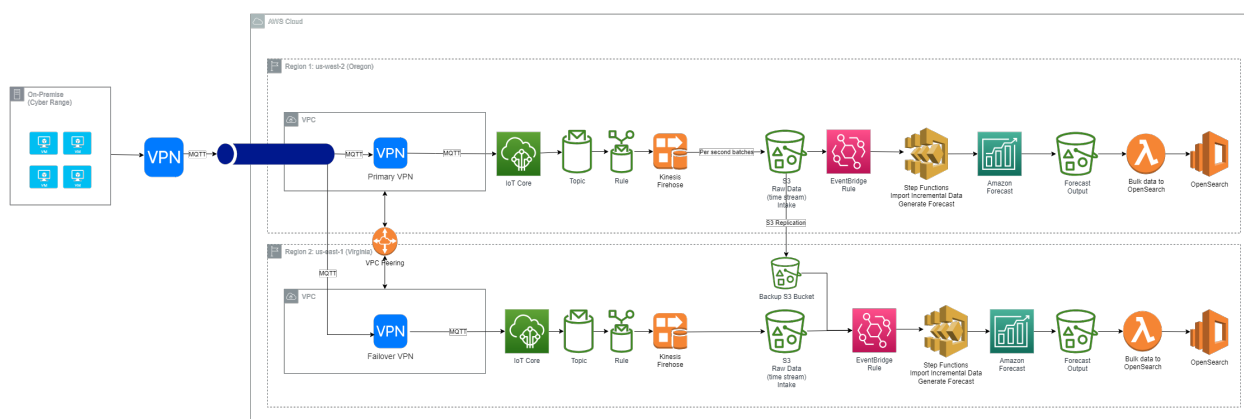


Figure 7. AWS system architecture

## 4.5 Results

We performed the failover test 2 minutes after we initialized the experiment and the data flow into the primary region started. This short time period between the initialization and the testing reduced the density of the extraneous data and allowed for simpler analysis. During this period, only connections related to this scenario were allowed into AWS to use the more precise measurements and metric collection capabilities available in AWS CloudWatch.

### Speed of Failover

As shown in figures 8 and 9, each showing data flow into the region in the local time zone, all 89 aggregated meter connections were active and connected after an approximate period of 2 minutes. This indicates that the established cloud control and monitoring architecture was capable of achieving a minimum time before restoration of data flow, referred to as the recovery time objective (RTO), of 2 minutes. This is also shown in Figure 10 in the highlighted segment of time showing the data flow into both regions. The data is normalized to Coordinated Universal Time for easy comparison.

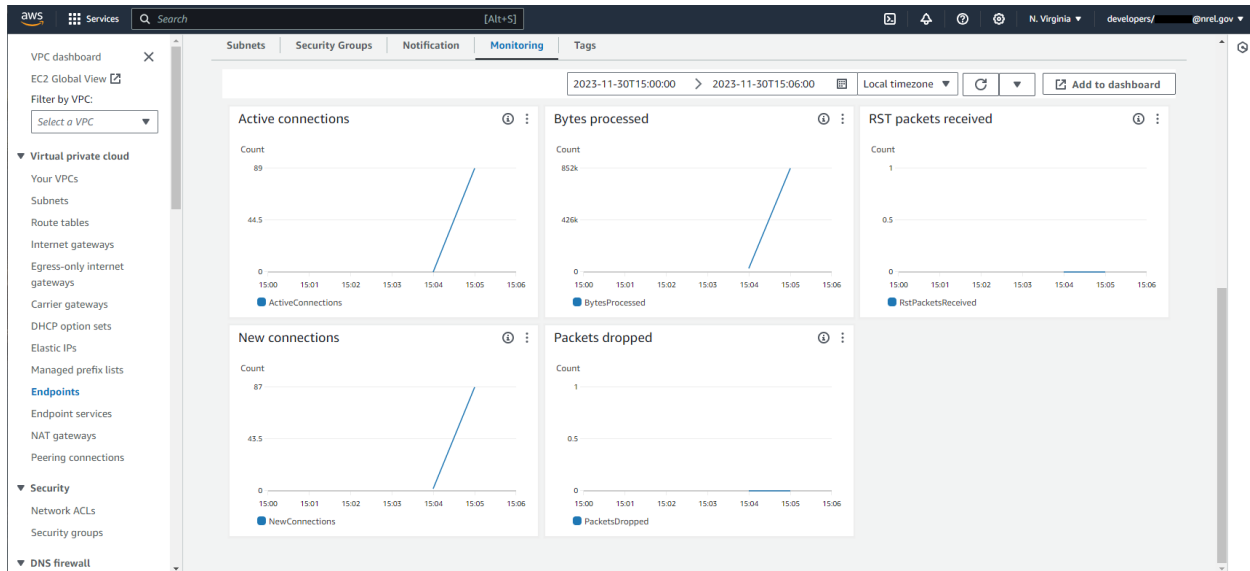


Figure 8. U.S. east failover charts

### Data Loss

Due to the design of the MQTT clients in this scenario, where an emulated attacker attempted to resend data on the alternate route when a connection was lost, and due to the speed of the failover to the alternative site occurring within a single aggregated data time step, in this use case we observed no data loss. This was verified by comparing the time stamps of the data available in the primary and secondary S3 buckets. The achieved result indicates that the deployed cloud architecture provided a very low recovery point objective of 1 minute due to the inherent backup and restoration features available in S3.

## 4.6 Analysis

From the observed results of this scenario, we obtained several metrics of performance related to the resilience of the deployed system.

### Data Loss Tolerance

From the measured time to failover, or RTO, as well as the average connected data rate (CDR), the maximum data loss tolerance (DLT) the system must accept can be calculated using the formula  $DLT = RTO * CDR$ . Provided that the assumption that the MQTT client implementation does not trigger any internal data loss protection mechanisms, the average expected data loss would be calculated by multiplying the normal connected data rate with the RTO. With the average data rate observed to be 18.5 kB/minute and an RTO of 2 minutes, this would indicate a maximum data loss potential of 37 kB, within the bounds of our experimental setup and assumptions.

### Architectural Resilience

The deployed architecture used multiple availability zones as well as multiple regions. AWS provides the highest level of resilience available using AWS cloud native services. This shows resilience to various forms of outages that might impact an individual availability zone or service in a VPC. These outages might impact a single availability zone, which can happen with some frequency, or, more uncommonly, a larger-scale region-wide service disruption could occur. Historical lists of region-wide outages within the last 5 years of the current date are available as post-event summaries on the AWS support website (Amazon Web Services 2023a).

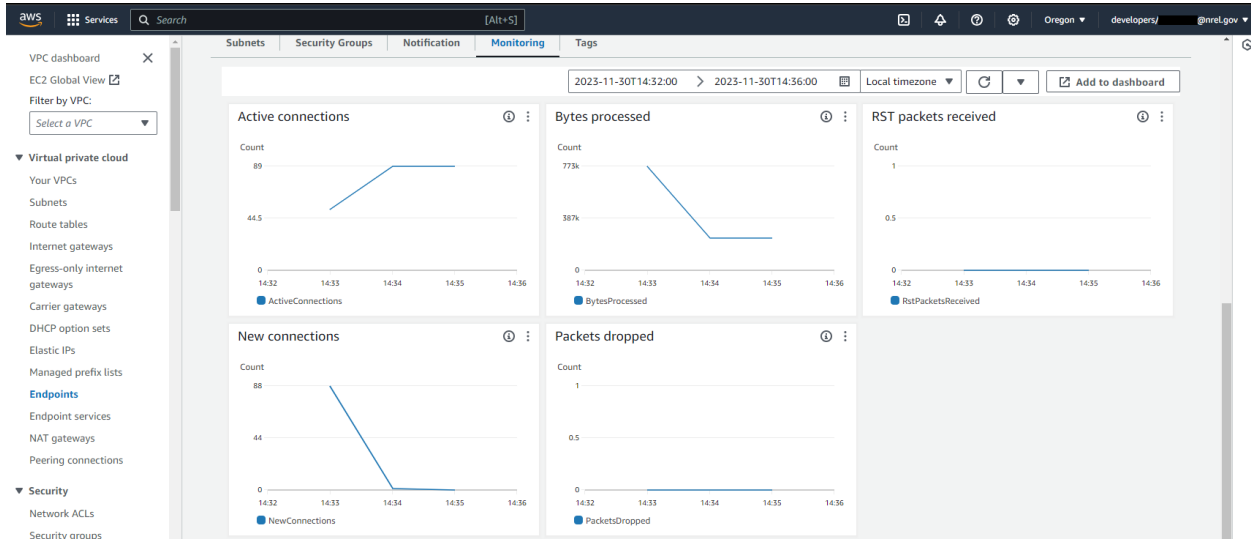


Figure 9. U.S. west failover charts

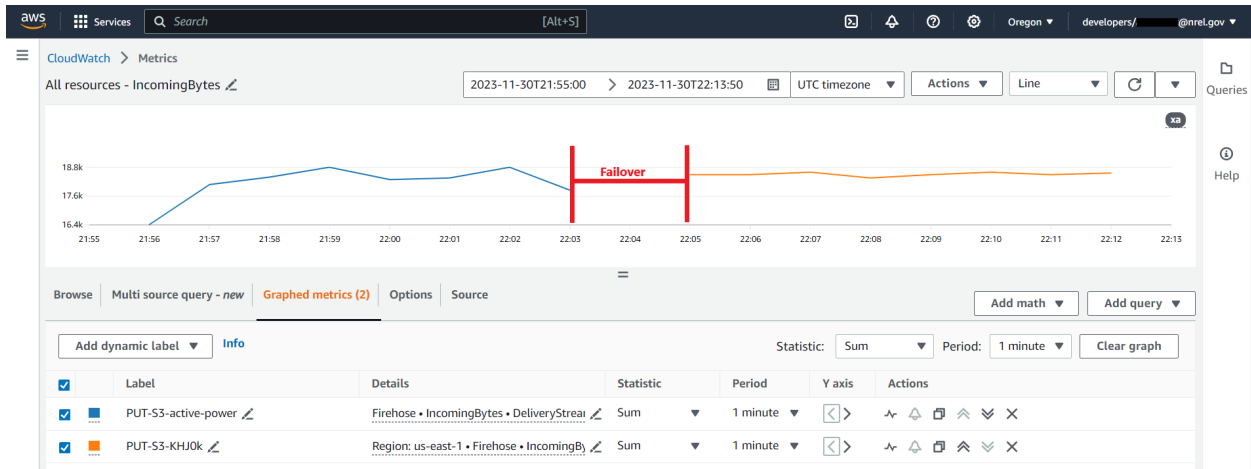


Figure 10. Chart highlighting failover

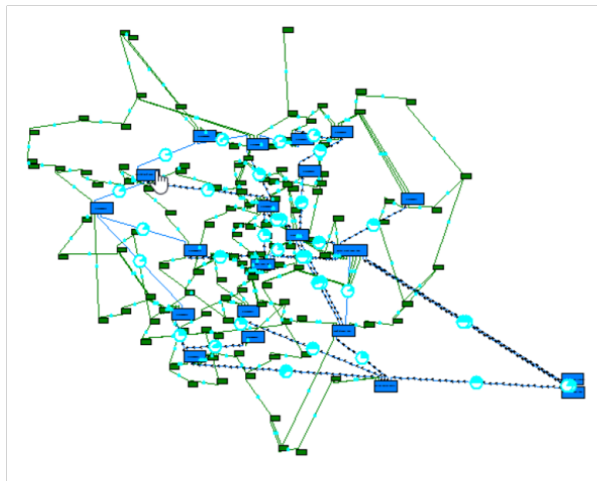
## 5 Bulk Electric System Reliability Operations

### 5.1 Cosimulation Studies

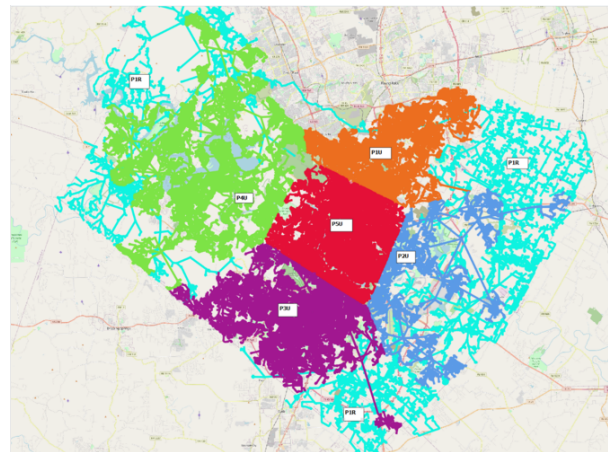
For the cosimulation setup in this experiment, we used PowerWorld Dynamic Studio (PWDS) as a transmission system simulator. It continuously solves the power flow equations for the system in real time using the latest values of the net megawatts and megavolts ampere reactive power injected at the load buses. It runs the built power transmission systems using the PowerWorld simulator, which facilitates large-scale power system modeling in the steady-state and transient stability time frames. The load profiles come from the SMART-DS project (Palmitier et al. 2020) for the P3U region of the city of Austin. Open Distribution System Simulator (OpenDSS) then solves the power flow by obtaining the voltage magnitude and angle obtained from the load bus from the transmission system running in PWDS.

For scalable distribution system models, we reviewed the literature from the SMART-DS project (Palmitier et al. 2020). The scenarios in SMART-DS include multiyear time-series load and weather patterns, solar power and battery adoption patterns, demand-responsive customers, and outage scenarios for several U.S. regions. Because the SMART-DS models include Austin, Texas, and a regional transmission system model is also available for Travis County where Austin is located, we selected Austin as the use case for our cosimulation (Panossian et al. 2021). In the current demonstration, we considered one of the six distribution regions (five urban, one rural) of the Austin area for the distribution system modeling (Figure 11b), although we also considered the whole of Travis County (Figure 11a) for transmission system modeling. The primary reason for considering one of the six regions is to limit the number of smart meters for each residential or commercial building to host as a container within the cyber range.

The controllers considered for the BES reliability operations are developed using OT-sim (Patria Security, LLC 2023) within the cyber range. OT-sim consists of a set of modules that run simulated OT devices in virtual machines or containers. The OT-sim modules facilitate Modbus and Distributed Network Protocol 3 (DNP3) communications, and their input/output interfaces act as Hierarchical Engine for Large-scale Infrastructure Co-Simulation (HELICS) federates. The logic module within OT-sim is used to define the control logic for the BES reliability operations use cases.



(a) Travis County Power Transmission Model in PowerWorld



(b) Austin City Power Distribution Model's P3U Region (magenta) in OpenDSS from the SMART-DS Project

Figure 11. Power System Use Case Considered for the Cosimulation

### 5.1.1 Mapping Transmission to Distribution

There are five areas in the Travis County transmission model:

1. Travis
2. Area2
3. Area3
4. Area4
5. TestAGCArea.

The P3U Austin area from the SMART-DS distribution model has a point of common coupling with the respective transmission system, as shown in Table A.1 in Appendix A.1.1.

### 5.1.2 Cosimulation Implementation

The transmission and distribution cosimulation is implemented using HELICS (Hardy et al. 2024), a flexible and scalable open-source cosimulation framework. The major components of HELICS constitute both federates and a broker. The federate is a specific instance of a simulation executable. For instance, in our cosimulation, we modeled each distribution substation as one federate, defined as a power federate. Alternatively, the complete transmission system is modeled through a single transmission federate. The publication and subscription of measurements, along with a command execution, takes place through the broker.

#### Transmission Federate

The transmission federate is an instance of the transmission simulation executable that models a group of components—such as generators, buses, loads, branches, and transformers—in PowerWorld. This federate interacts with the PWDS server to monitor attributes such as the bus status, frequency, and voltage. The federate is also used to control the set points and statuses of the transmission model in PWDS based on the updates from the distribution/power federate (real and reactive power from the distribution substation) or commands from the controller (OT-sim device) from the remote terminal units (RTUs) modeled in the cyber range and/or the OT-sim instance in the cloud. These federates are a *combination federate*, which acts as a *value federate* while publishing values to the kafka or distribution federates, and which acts as a *message federate* while receiving commands. This transmission federate is modeled to monitor the area for automatic generation control (AGC) and to test remedial action schemes (RAS) and the control of the generator terminal voltage (i.e., exciter-based control).

#### Power Federate

The power/distribution federate is an instance of a distribution simulation OpenDSS model that models a group of components, such as capacitor banks, buses, regulators, branches, and transformers. They are modeled in OpenDSS using the SMART-DS Austin city model. The federate solves the load flow in OpenDSS and monitors attributes such as the bus status, frequency, and voltage from the OpenDSS instance. It also controls components such as when the federate receives cosimulation updates from the transmission federate (substation's voltage magnitude and angle) or commands from the controller (OT-sim device) from the RTUs modeled in the ARIES Cyber Range and/or the OT-sim instance in the cloud. The power federate also acts as a combination federate. For the cosimulation, 29 instances of power federates are operational, based on the 29 distribution substations emulated for the P3U Austin region.

### 5.1.3 Transient Modeling of the Travis County Generator

The following sections outline the four major components of a generator that play a major role in regulating voltage, frequency, and grid synchronization: the machine model, the exciter model, the governor model, and the stabilizer model.

#### Machine Model

The machine model plays the major role in grid synchronization. All the nuclear plants in the Travis County model uses a GENROU machine model from PowerWorld (PowerWorld 2024c), which models the round rotor machines. A

representative model of the electromagnetic dynamic behavior of a synchronous machine can be constructed entirely on the basis of the observed behavior of its stator voltage and current.

#### Exciter Model

In the Travis County transmission system, the *ESST4B* exciter model is used for all the hydropower and natural gas power plants—*ESST4B*: Excitation System *ST4B* is based on the Institute of Electrical and Electronics Engineers (IEEE) *ST4B* standard (IEEE-PES 2016). All the *ST* systems are static excitation systems. In static excitation systems, the power for providing field excitations is derived from the generator output terminals. A transformer known as an excitation transformer is connected to the output terminals of the generator to step down the voltage to the required voltage level. Because DC supply is needed, the transformer output is connected to a thyristor full bridge rectifier. The firing angle of the thyristor full bridge rectifier is controlled by a regulator (voltage regulator) to provide the required field excitation. The secondary terminal of the current transformer and the potential transformer is connected to the generator output terminals and is fed to the regulator. On the basis of the generator terminal voltage, the regulator adjusts its firing angle. The advantage of using static excitation systems is that there are no separate rotating-type exciters, and the system is free from friction, windage, and commutator loss occurring in the exciter. All the static excitation system models from *ST1A* to *ST10C* have common features except for the over- and under excitation limit and additional proportional-integral block, as shown in Table 3 of IEEE Std 421.5 (IEEE-PES 2016). Parameters of this model are given in (PowerWorld 2023b).

#### Governor Model

The governor is a device that controls the speed and output power of the generator by, for example, changing the water flow of the turbine in the case of the hydropower generators. The prime mover provides the mechanism for controlling the synchronous machine speed and hence the terminal voltage frequency. To control the speed, a device must sense the speed in such a way that a comparison with a desired value can be used to create an error signal to take a corrective action. Using this error signal, the control valve, which controls the flow of the fuel (water, steam, or air) to the turbine, can be modified. A positive error signal would make the control valve open more, allowing more fuel to enter the turbine and increase the speed, resulting in an increased frequency and a negative error signal that enforces the control valve to close more, thereby decreasing the speed.

There are two types of generators in the transmission model that use the governor model. The *GAST2A* model is used for natural gas plants (PowerWorld 2024a), and the *HYGOV* model is used for hydropower plants (PowerWorld 2024b). Because the type of turbine depends on the fuel that moves the turbine, the governor models are unique to the specific types of generators. The governor model governs the primary-type frequency control, also called the droop speed control, whereby the power output of a generator reduces as the line frequency increases. This is commonly used as the speed control mode of the governor of a prime mover. Further, the secondary-level control of the frequency is responsive to an automatic generation control (AGC) signal, which is based on multi-area transactions declared across each area.

#### Stabilizer Model

A power system stabilizer (PSS) is a control system within generators that monitors the current, voltage, and turbine/generator shaft speed. When necessary, it then sends the appropriate control signals to the voltage regulation control system block to damp the system oscillations so that the frequency does not stray beyond tolerances. If there were no stabilizers, the frequency fluctuations resulting from a certain event would not dampen. The stabilizers that were considered for the Travis transmission model are made of the *PSS4B* model (PowerWorld 2024d). Because the stabilizer output goes to the voltage regulator, which is part of the exciter model, we used the *PSS4B* model, which complies with the *ESST4B* exciters.

Table A.2 in Appendix A.1.1 illustrates the machine, exciter, governor, and stabilizer model used for the generators in the P3U region of Austin city.

## 5.2 Automatic Generation Control

An AGC system receives power flow and frequency measurements from sensors at substations and then outputs control commands with the primary goal of ensuring frequency stability. AGC is a real-time control application operating on the order of a few minutes, and it is very sensitive to the sensor measurements it receives. When it



comes to power system applications in modeling and operation, one main purpose of an AGC system is to minimize the area control error (ACE). The ACE can be defined as the difference between the actual power flow and the scheduled power flow between two different areas, as in Eq. 5.1. The ACE value (PowerWorld 2023a) accounts for the produced electricity frequency, measured and nominal; the frequency bias factor in MW/0.1Hz; and the sum and initial values of the power flow, as in Eq. 5.2.

$$ACE = P_{Actual} - P_{Scheduled} \quad (5.1)$$

where,  $P_{Actual}$  is the actual power flow between two areas, and  $P_{Scheduled}$  is the scheduled power flow, defined through a simulation attribute  $MWTransactions$  in PWDS, which is the transaction defined presumably under a contract between the two areas.

$$ACE = (f_{meas,bus} - f_{nom,bus}) * 10 * Bias + (tieflows_{sum} - tieflows_{initials}) \quad (5.2)$$

where,  $f_{meas,bus}$  is the frequency measured at bus, and  $f_{nom,bus}$  is the nominal frequency which is 60 Hz for the North American electric grid.  $tieflows_{initials}$  is the initial tie line flow (tie line is the transmission line connecting two areas), and  $tieflows_{sum}$  is the sum of actual tie line flows.  $Bias$  is the frequency bias, usually expressed in megawatts per 0.1 Hertz (MW/0.1 Hz), associated with a balancing authority area that approximates the balancing authority area's response to Interconnection frequency error (FERC 2014).

$$G_{setpoint} = G - 2 * ACE * P_f \quad (5.3)$$

Not only is it important to minimize the ACE, but it is also ideal to obtain an ACE value of zero; i.e., the actual power flow needs to match the scheduled power flow, a core function of AGC. Another important variable to consider for AGC system implementation is the participation factor,  $P_f$ , of the generators (Eq. 5.3), which is defined as the amount of real power that a generator contributes relative to the amount of change in the load consumption in the system.

### 5.2.1 Test of Local Automatic Generation Control

In the local AGC, we monitored how the new AGC set point is altered in the generators to address contingencies such as generator outage and load surge events. The AGC controller used in this case is based on the AGC algorithm running inside PowerWorld. The algorithm for obtaining the AGC set point for the generators is based on Eq. 5.1 and Eq. 5.2.

### 5.2.2 Test of Global Automatic Generation Control

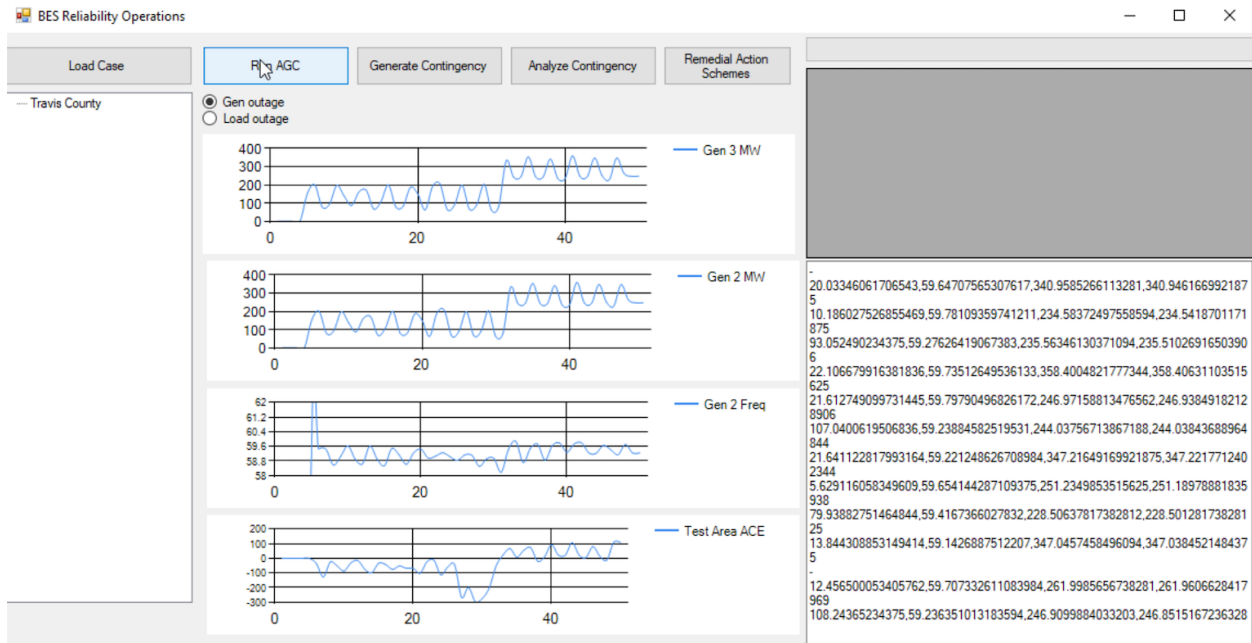
The primary purpose of a balancing authority is to maintain grid stability in terms of voltage level, frequency control, and the flow of electricity to customers. The balancing authority collects periodic measurements from all utilities within their jurisdiction and can send commands to the utilities via Inter-Control Center Communications Protocol (ICCP) (ERCOT 2018). For example, the Electric Reliability Council of Texas (ERCOT) is the balancing authority for Texas. ERCOT provides utilities with network routers so that they have a dedicated connection between the ICCP node in the utility and ERCOT. Their communications happen over a private wide-area network link. ERCOT is also responsible for encrypting the ICCP data between the utility's router and its own routers. Instead of using ICCP as a proof of concept, we used DNP3 to monitor the ACE and to change the generator's megawatt set point.

In the global AGC, the balancing authority monitors the frequency and the ACE of different areas modeled in Travis County and controls the generator of Area 5 (i.e., the test area) by giving a new generator set point from the balancing authority. With the use of the ACE and  $P_f$ , the AGC set point of each generator can then be calculated as in Eq. 5.3.

### Generator Outage Event

The project team focused on the fifth area, *TestAGCArea*, to study the impact of a generator outage on the ACE and to determine how this is used to update a new megawatt set point for the rest of the generators in the test area. There

are three generating units in the Sam Gideon Power Plant in the test area, connected to Bus 172 circuits 1, 2, and 3, respectively. In the generator outage event (Sam Gideon Power Plant Unit 1, i.e., Bus 172, Circuit 1), the ACE value as well as the frequency is expected to reduce, because the net generation decreases to less than the net load in that area. Figure 12 refers to the increase in ACE in the *TestAGCArea*.



**Figure 12. Impact of generator outage on the ACE of the test area, followed by the new generation set points for the other two generators (Generator 2 MW and Generator 3 MW) in the test area. The negative ACE caused the system frequency (Generator 2 frequency) in the area to decrease.**

### Load Fluctuation Event

In contrast to the previously defined generator outage event, which modeled load outages assuming no protection systems and causing the frequency to rise, these load fluctuation events are reflected in the system as observed from the PWDS logs (Figure 13). Due to these outages, our team observed that the ACE is positive, but the set point given to Generator 2 and Generator 3 is as low as 86 MW, which causes the ACE to dip even further into negative values (Figure 14).

The previous control of the AGC set point caused the ACE to reduce, but a corrective set point resulted in better control, as shown in Figure 15. We demonstrated these figures through a stand-alone application running in the virtual machine in the cloud. For control, the OT-sim is also deployed within the cloud. In the same way that this application monitors the ACE, the OT-sim module monitors this altering ACE to compute the corrective set point and send the command to the substation devices running on-site in the cyber range.

### 5.2.3 Implementation of Automatic Generation Control in Cyber Range: Operational Technology Simulator Logic Module

OT-sim has a logic module that has input and output interfaces for implementing logic, similar to a programmable logic controller. This monitors the current generator's output and the ACE and determines a new set point for the generator. For instance, in the following work, the ACE of Area 5 i.e. *TestAGCArea* is considered for updating the MW set point for *gen2* and *gen3*. The detailed OT-sim logical program code snippet for this use case is as shown in the `<logic></logic>` snippet in the next page:

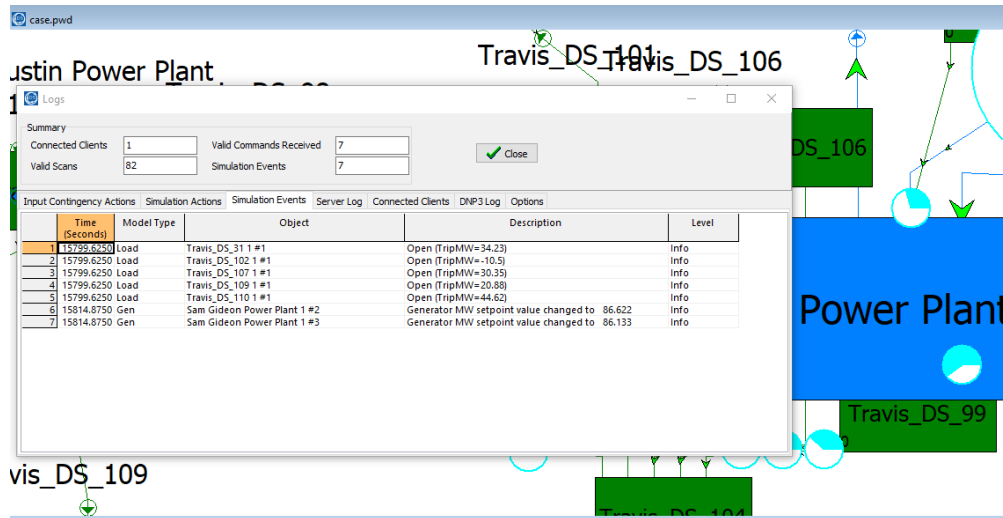


Figure 13. Certain loads in Travis County at substations DS 31, 102, 107, 109, and 110 are tripped.

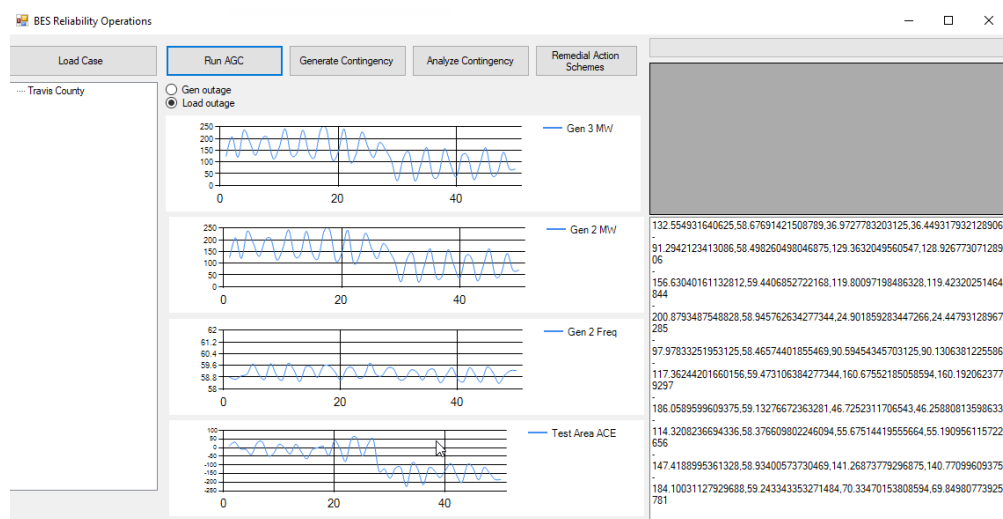


Figure 14. The ACE is positive, which causes the generator megawatt set point to decrease, causing the ACE to dip even further into negative values.

```

<logic>
  <period>1s</period>
  <program>
    <![CDATA[
      gen2_ace = gen2_mw - area5_ace * 2.0
      gen3_ace = gen3_mw - area5_ace * 2.0
      gen2_mwsp = min(gen2_ace + 50.0, 280.0)
      gen3_mwsp = min(gen3_ace + 50.0, 280.0)
    ]]>
  </program>
  <variables>
    <area5_ace tag="area-5.ace">0</area5_ace>
    <gen2_mw tag="generator-2_bus-172.mw">0</gen2_mw>
    <gen2_mwsp tag="generator-2_bus-172.mw_setpoint"></gen2_mwsp>
    <gen3_mw tag="generator-3_bus-172.mw">0</gen3_mw>
    <gen3_mwsp tag="generator-3_bus-172.mw_setpoint"></gen3_mwsp>
  </variables>
</logic>

```

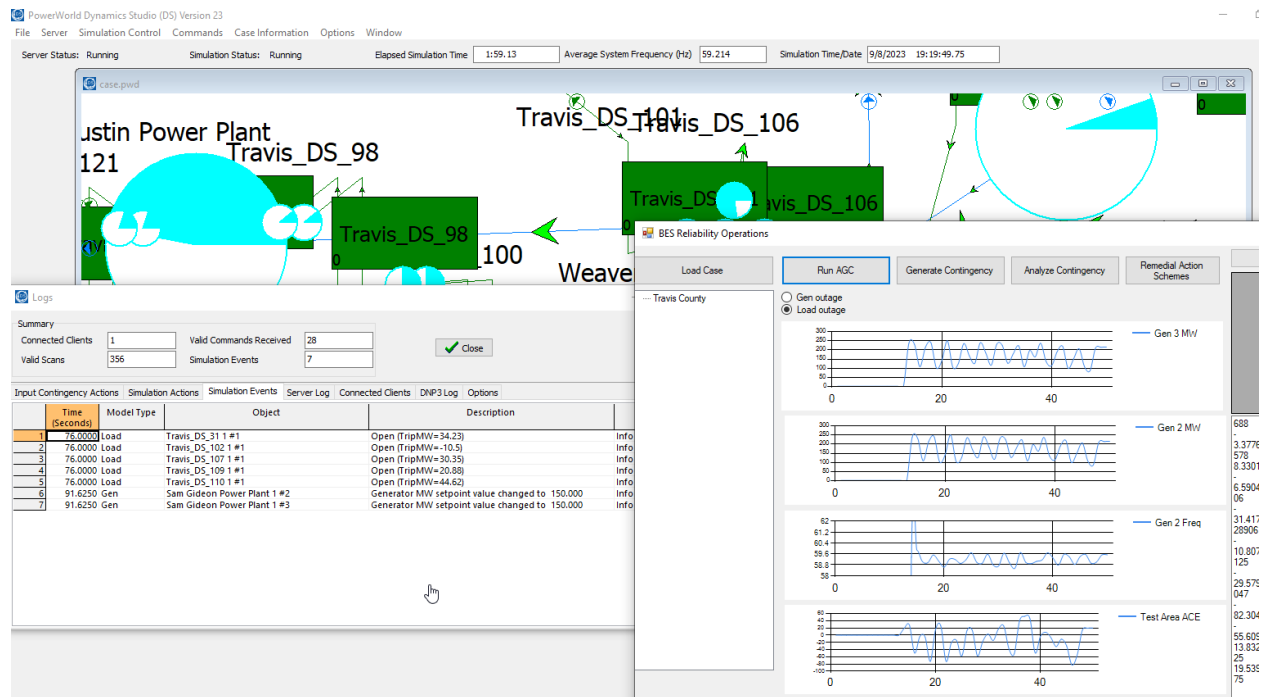


Figure 15. The ACE is positive, which causes the generator megawatt set point to reduce to 150 MW, causing the ACE to stabilize.

### 5.3 Remedial Action Schemes Using Contingency Analysis and Situational Awareness

In this use case, a contingency analysis tool ranks the contingencies in Travis County based on an N-1 contingency. Based on these contingencies, we devised RASs.

We performed an N-1 contingency analysis on the Travis County model that generated a total of 337 contingencies, of which 9 contingencies cause violations based on the line overload. The list of line outage contingencies ranked based on their impact is shown in Table A.3 in Appendix A.1.1. Every contingency can have a set of RAS. For demonstration purposes, we considered only one contingency, C1, as enumerated in the contingency list (Table A.3 in Appendix A.1.1).

First, we implemented the RAS for the first contingency, C1: **RAS Condition:** If the line between Bus 101, Circuit 1 to Bus 158, Circuit 2 or between Bus 101, Circuit 1 to Bus 158, Circuit 1 has an overflow, i.e., 102% of the upper limit, then the associated response is given by: **RAS Action:** Open Line connecting Bus 124, Circuit 1 (Travis\_DS\_106) and Bus 119, Circuit 1 (Travis\_DS\_101). C1 causes the transmission line between Bus 101, Circuit 1 to Bus 158, Circuit 2 and between Bus 101, Circuit 1 to Bus 158, Circuit 1 to overflow to compensate for one unit of the Weaver Power Plant generation; hence, this RAS will assist in reducing the violations. The ranked line outage contingencies that cause violations after implementation of one RAS are listed in Table A.4 in Appendix A.1.1.

#### 5.3.1 Implementation of Remedial Action Scheme in Cyber Range: Operational Technology Simulator Logic Module

This OT-sim logic module monitors the measurements defined in the RAS condition and implements a control action based on the suggested RAS action. For instance, in this scenario, OT-sim monitors the power flow in the two transmission lines mentioned in the RAS condition, and if either one exceeds 102% of the upper limit of the transmission line power carrying capacity, the breaker in another transmission line is opened. The detailed OT-sim logical program code snippet for this use case is as shown in the following <logic></logic> snippet:

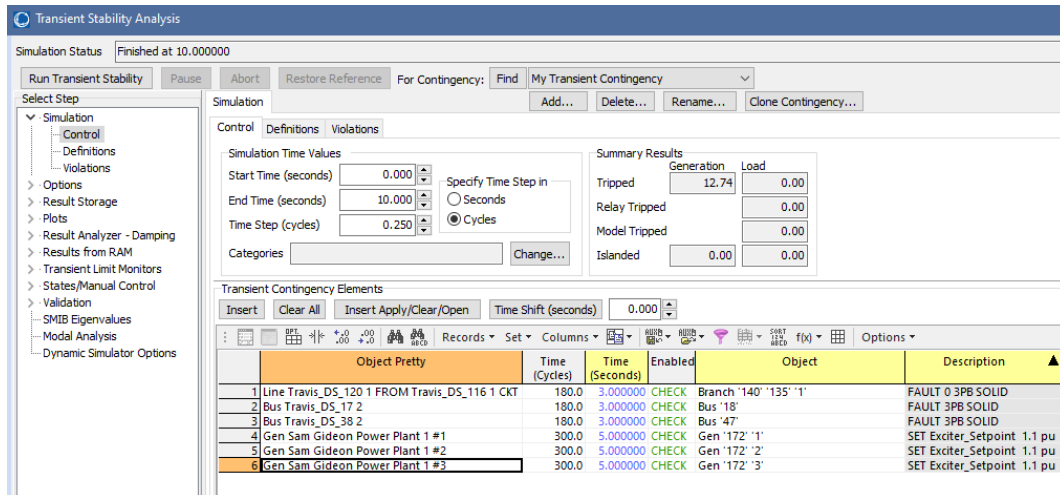


Figure 16. One line fault and two solid-phase faults at two buses

```

<logic>
  <period>1s</period>
  <program>
    <![CDATA[
      max_flow = 500.0
      closed = closed && (abs(mva_to) / max_flow) < 1.02]]>
  </program>
  <variables>
    <mva_to tag="branch-1_bus-101-58.mva_to">0</mva_to>
    <closed tag="branch-1_bus-124-119.status">1</closed>
  </variables>
</logic>

```

## 5.4 Voltage Control

For some generators in the voltage control scenario, the default exciter control in the simulator was disabled, which showed how a set of line and bus faults in the test area would cause a voltage collapse. Then by enabling the exciter control, we observed how the voltage collapse was prevented. We will evaluate this further by changing the voltage set points through the exciter control using the OT-sim logic module.

### 5.4.1 Transmission-Side Voltage Control Using Exciter Control

A solid 3 phase fault event occur at 3 seconds of the dynamic simulation, as shown in Figure 16. Due to this fault, the voltage suddenly increases and undergoes a voltage collapse, as shown in Figure 17 when the exciter controls are disabled in all the generators. When the exciter is added with the set point of 1.04 per unit (p.u.), there is an improvement in the terminal voltage of the generators, as shown in Figure 18a and Figure 18b with one and two generator exciters enabled, respectively. The figures show that using one excitor the voltage is restored to an average voltage of 0.87p.u.; with two exciters, the voltage recovered to 0.91 p.u. Increasing the voltage set point to approximately 1.1 p.u. compared to 1.04 results in an improvement under multiple fault scenarios, as shown in Figure 19a and Figure 19b. The graphical user interface showing the impact of a fault followed by restoration using exciter control is shown in Figure 20.

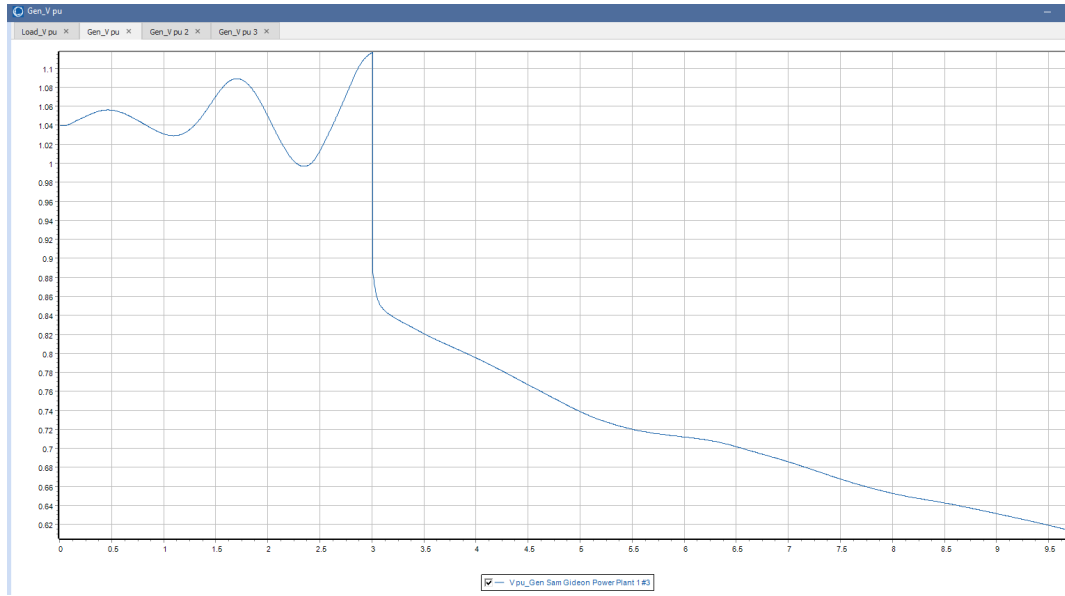
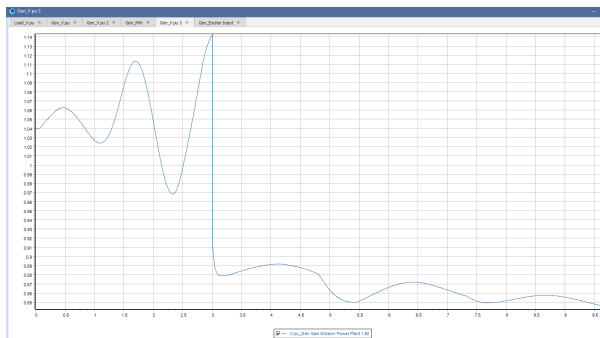
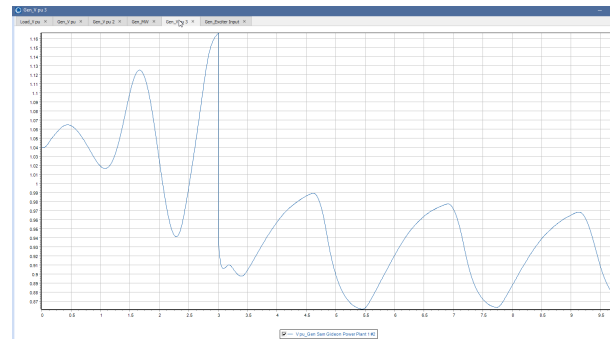


Figure 17. Voltage collapse without any exciter in any of the three generators in the Sam Gideon Power Plant in the Test area

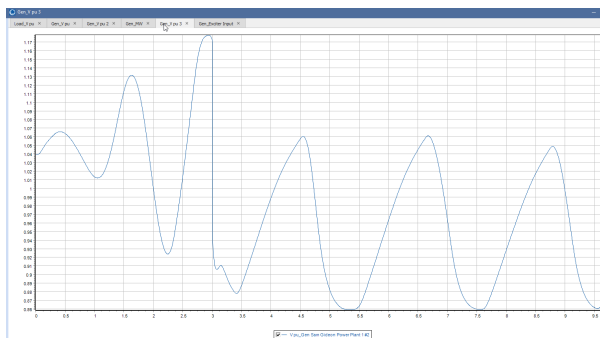


(a) Impact of exciter set point: The exciter is enabled in Unit 1 of the Sam Gideon Power Plant with a voltage set point of 1.04 p.u.

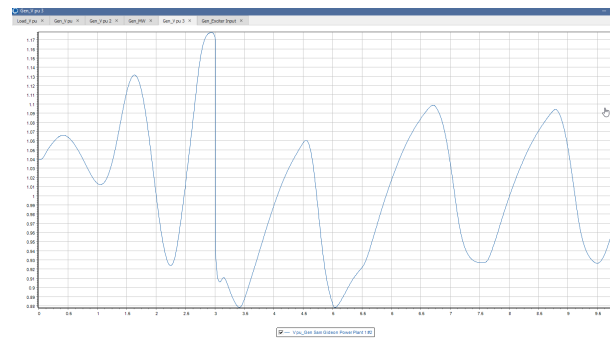


(b) Impact of exciter: The exciter is enabled in Unit 1 and Unit 2 of the Sam Gideon Power Plant.

Figure 18. Impact of number of generators when the exciter control is enabled

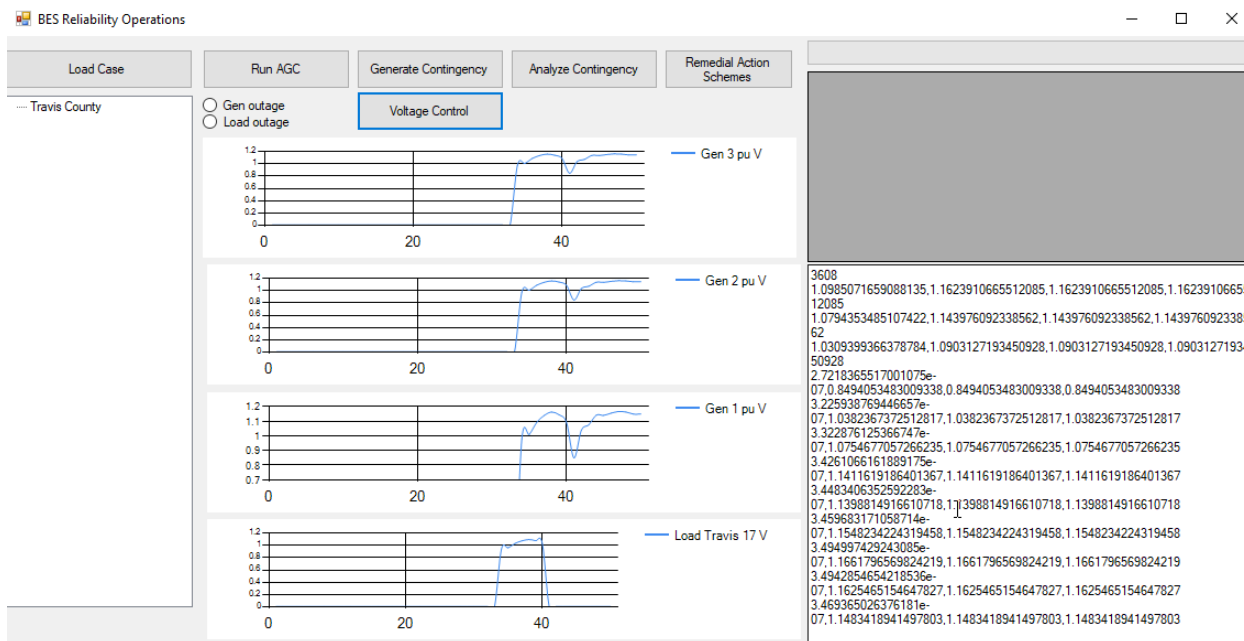


(a) Impact of exciter set point: The exciter is enabled in all units (1, 2, and 3) of the Sam Gideon Power Plant with a voltage set point of 1.04 p.u.



(b) Impact of exciter set point: The exciter is enabled in all units (1, 2, and 3) of the Sam Gideon Power Plant with a voltage set point of 1.1 p.u.

Figure 19. Impact of exciter voltage set point



**Figure 20. The application monitoring the per-unit voltage at the terminals of the Sam Gideon Power Plant when there is a fault on the load Bus 17.**

#### 5.4.2 Implementation of Voltage Control in Cyber Range: Operational Technology Simulator Logic Module

This OT-sim logic module monitors the terminal voltages of the generator. To address the voltage collapse due to the fault, OT-sim sends a command to increase the set point of the voltage regulator through the exciter of the generator within a given time. The detailed OT-sim logical program code snippet for this use case is shown below:

```
<logic>
  <period>1s</period>
  <program>
    <![CDATA[
      gen1_vctrl = 1.0 / gen1_pu
      gen1_pusp = min(gen1_vctrl, 1.15)
      gen1_pusp = max(gen1_pu, 0.85)
    ]]>
  </program>
  <variables>
    <gen1_pu tag="generator-2_bus-172.voltage_pu"></gen1_pu>
    <gen1_pusp tag="generator-2_bus-172.voltage_pu_setpoint"></gen1_pusp>
  </variables>
</logic>
```

#### 5.4.3 Distribution-Side Voltage Control Using Capacitor Banks and Transformer

For the voltage control problem, NREL researchers planned to leverage volt-VAR control using an artificial intelligence-based approach with AWS Sagemaker based on similar work implemented by Siemens (Fan, Lee, and Wang 2021). This approach uses a reinforcement learning framework for performing volt-VAR control through controlling the batteries, transformer, and capacitor banks. They implemented the technique for a 8,500-node feeder. In this phase, they implemented volt-VAR control in one substation of the 29 substations in the P3U region of the Austin distribution model, and NREL researchers saw an improvement by considering a proximal policy optimization reinforcement learning agent. In the future, NREL researchers may want to migrate this into AWS Sagemaker to efficiently train it with better reinforcement learning algorithms available from the cloud services.





## 6.1 Data Pipeline Pollution Attack

The following figures lay out the implementation of a data pollution attack against aggregated distribution feeder data transmitted to AWS from the on-premises power modeled in the Cyber Range. MITRE’s newly added support for operational technology in Caldera released in September 2023, and scripted automation built by NREL made to work with Caldera’s new red team test tool, allowed the research team to emulate how malicious script could be successfully executed on a compromised VPN endpoint in the cloud. See Figures 22 and 23

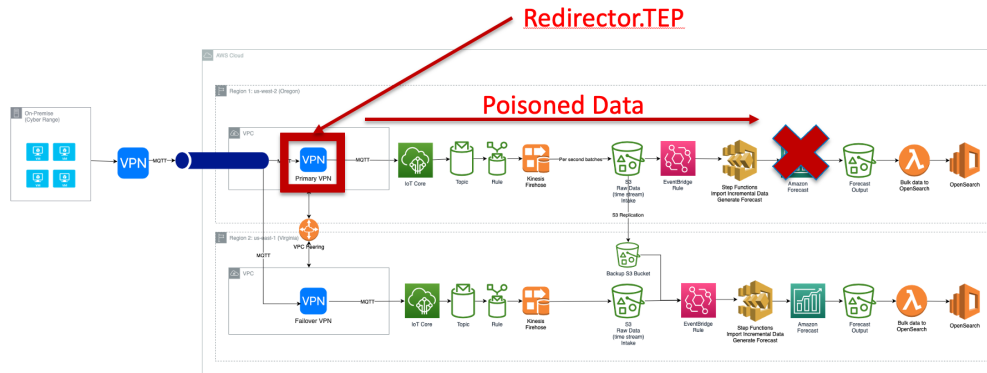


Figure 22. Command and control attack architecture

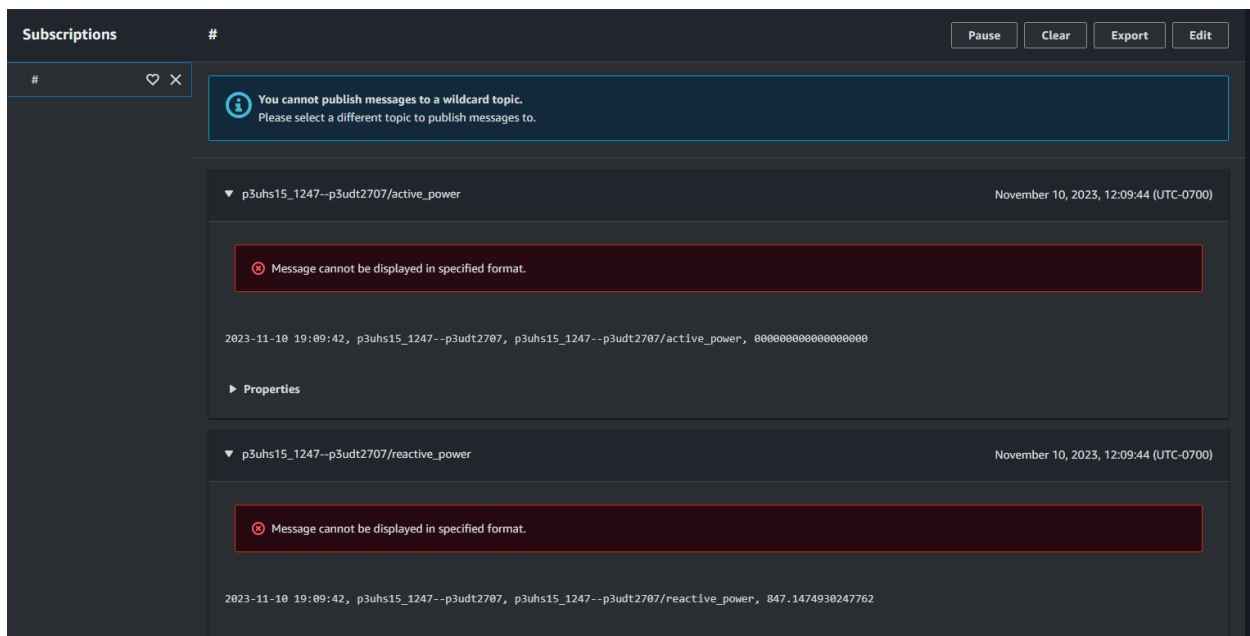


Figure 23. Snapshot from Internet of Things (IoT) Core service showing one feeder’s aggregated active power polluted to zero

This script modifies the active power values as they are received from the distribution feeder instruments and forwards the false values down the forecast pipeline deployed in the cloud. The zero value renders the resulting forecast useless for load operations. The attack could be modified to only slightly adjust the values producing an errant forecast and potentially triggering other automated system responses like load shedding.

*Note that the cloud-hosted components in this scenario are executing the instructions they receive from the ARIES Cyber Range and are not directly compromised. Because these workflows are built via multiple AWS services, configuration of error checks on the inbound data would ensure they do not violate a boundary of expected parameters, preventing this attack. Additional defensive measures are addressed later in this section.*

As represented in Figure 22, a poorly maintained and monitored VPN could allow a threat actor access to MQTT traffic containing important predictive data that are en route to the IoT cloud broker on the Amazon cloud. In this threat scenario, the attacker successfully intercepts the traffic, mimicking an intelligent electronic device. Instead of arriving at the IoT broker, the man-in-the-middle attack redirects the data to a local process where actual data are swapped with false values, in this case zeros. In this threat emulation, the raw data, or real power data, are modified, not the reactive power. This exploit shows that some data are still going through, but others are not.

## 6.2 Command and Control Attack

The control attack scenario considers how a threat actor could manipulate a breaker at the point of common coupling between the transmission network and the distribution via a compromised controller based in the cloud. "The nation's electric grid is becoming more vulnerable to cyberattacks, particularly those involving Industrial Control Systems (ICS) that support grid operations" (U.S. Government Accountability Office 2019). Increasingly, threat actors seek to control the power transmission in utility systems. A variety of threat scenarios could gain a foothold in the operation network through network resources on the enterprise side, and these attacks are very similar in the cloud based resources. MITRE's description of external remote services, T1133, or phishing, and T1566, are just two frequently achieved initial access methods (MITRE 2023). After successful network penetration, the threat actor attempts to pivot to accessing the controller. Successful access to the on-premises controller is achieved through compromising a virtual remote terminal unit (vRTU) based in the cloud, as shown in Figure 22. The virtual RTU and the local breaker controller are connected through the cloud VPN using a well-established and common Distribution Network Protocol version 3 (DNP3) communication protocol. Although DNP3 is not inherently vulnerable, the communication between key devices could allow invalidated control messages to reach adjacently connected ICS devices. These controllers play a key role in energy delivery at electric grid substations.

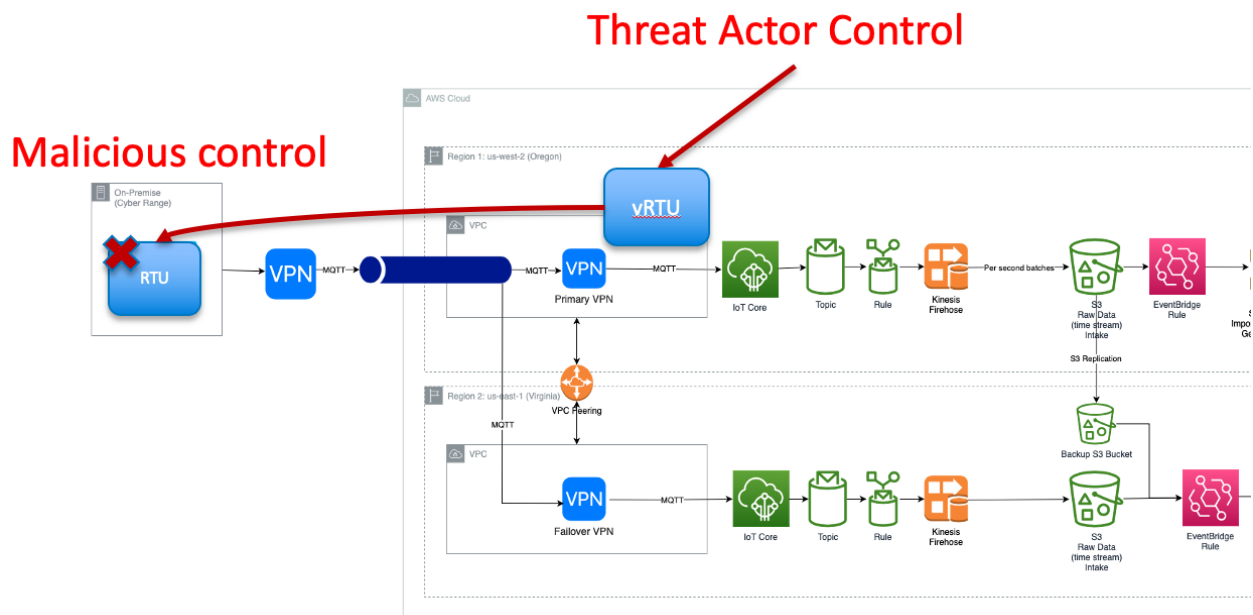


Figure 24. Attacker pivoting from virtual Remote Terminal Unit (vRTU) in the cloud to Remote Terminal Unit in Substation

As part of the threat actor’s arsenal to successfully access vital OT network systems and assets, attackers look for and continue to find legacy application tools, such as Telnet. It is well-known that Telnet transmits everything in plain text. Occasionally, organizations are not be aware that Telnet protocols are still enabled. The North American Electric Reliability Corporation (NERC) recognizes that legacy protocols such as Telnet and other legacy solutions are still in use because most ICS devices still widely support remote connectivity using Telnet and plain text transmission (NERC 2023). A SANS Institute report provides additional examples of this ongoing use of Telnet: “We [not only] see this in legacy systems, but in ICS that control factories, water or hydro utilities we see this all the time in production” (Vandenbrink 2017). Further, the report states that this vulnerable practice might be due to the continued use of equipment that does not support Secure Shell Protocol (SSH) and in some cases does not support credentials.

Once a threat actor successfully compromises the OT network using vulnerable and insecure applications like Telnet, there remain very few ICS within the environment that the attacker cannot access. Figure 24 illustrates how critical systems are linked by communication protocols. Enterprising threat actors will take advantage of any means to pivot from a foothold in the OT network to key critical systems in the generation, transmission, and distribution.

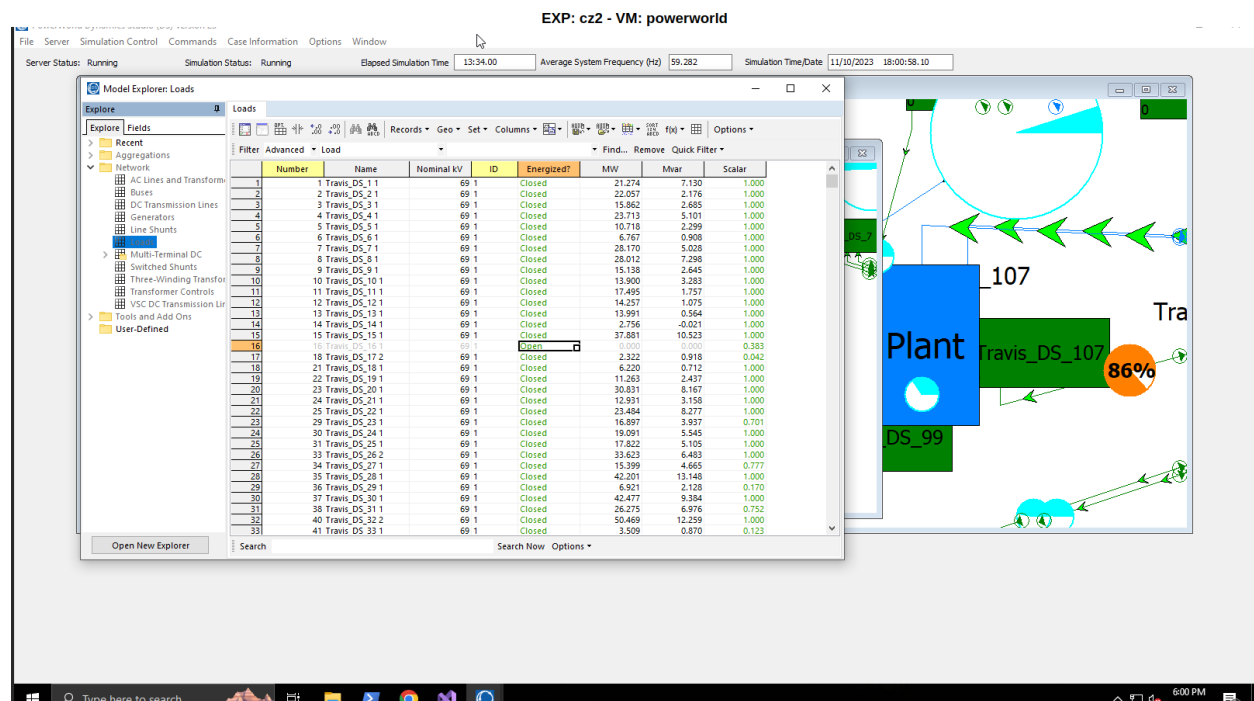


Figure 25. Load bus is opened by Telnet attack

In this scenario, the emulated attacker uses Telnet commands while in the cloud OT network. A successful compromise of the RTU allows the attacker to send malformed communications from the vRTU, the Level 2 control center, and the breaker controller on the common point of coupling between the transmission and the distribution. From this malformed communication, the emulated attack successfully opens the substation breaker on Bus 16, as represented in Figure 25. An attack of this magnitude and depth on an energy producer’s operation network against a breaker controller demonstrates how critical power generation could be adversely manipulated. The emulated attack opens the breaker disconnecting power generation placed on the grid, which in turn triggers the system to draw power from other regions to cover the shortage. A well-orchestrated attack that could open several breaker controllers at the point of power generation at targeted substations could ultimately achieve load shedding events and jeopardize grid stability. The result may achieve power outages across a wide geographic region, as shown in Figure 26.

### 6.3 Cloud Provider Tools and Best Practices for Threat Mitigation

Understanding the shared responsibility model is crucial when using cloud providers. This includes a comprehensive understanding of how cloud providers secure workloads in their data centers, which forms the baseline for securing

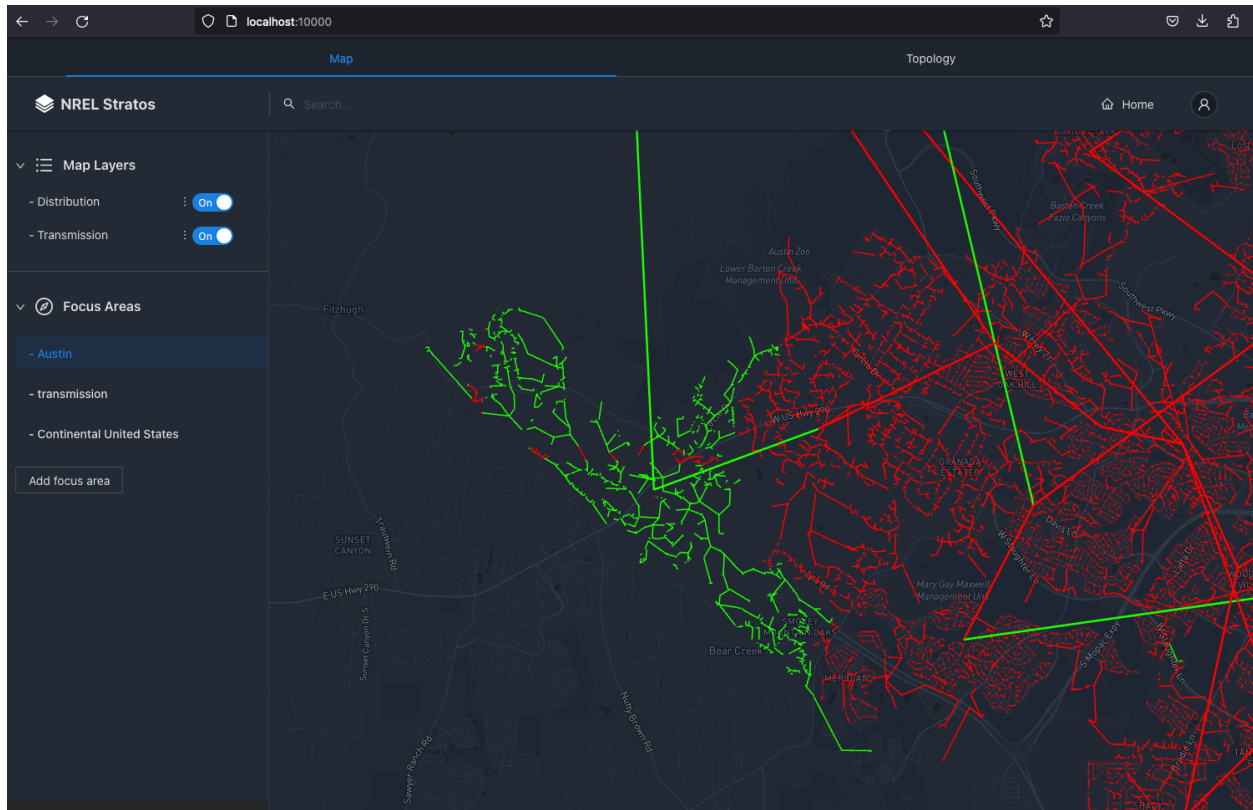


Figure 26. Geographic representation of blackout caused by breaker opening

customer workloads and defines the responsibility of the cloud service provider’s customer. AWS, for instance (amongst other cloud solution providers), provides artifacts on compliance and agreements that cover items such as specific compliance reporting and Federal Risk and Authorization Management Program certification status (Amazon Web Services, n.d.(a)).

By building upon this foundation of secured infrastructure, the cloud services generally offer enhanced security and transparency by exposing API, network, and applications logs as well as an event bus for related service events, all of which help mitigate against potential threats (Amazon Web Services, n.d.(b)).

### Cloud Security Infrastructure

Cloud service providers like AWS provide tools and infrastructure designed to counteract security threats for systems hosted on their platforms. For example, key features in the AWS services used in this report include:

- **Encrypted data storage and network traffic:** AWS encrypts data stored in each process step, in managed databases, or in S3 object storage. Encryption can be based on keys managed by either the organization or Amazon.
- **VPC:** AWS VPC ensures encrypted end points for customer applications. Although AWS secures networking within the VPC, customers are responsible for encrypting traffic in customer-hosted applications. Additionally, VPC-hosted workloads are safeguarded against network sniffing by other systems and Address Resolution Protocol cache poisoning (Amazon Web Services 2011).
- **Internet Protocol Spoofing Prevention:** AWS prevents EC2 in each process step in the AWS cloud from sending spoofed network traffic (Amazon Web Services 2011).

- **Security Groups:** These groups control the Internet Protocol ranges that can interact with EC2 compute nodes or network-based AWS services. They support ephemeral configurations for unpredictable Internet Protocol addresses and can restrict traffic to specific networking devices.
- **VPC configurations:** Consisting of both public and private subnets, VPCs also include access control lists for managing traffic between subnets.
- **Network firewalls and flow logs:** AWS offers a network firewall for VPC resources and provides near-real-time flow logs to monitor network requests.

#### *Secured Data Transmission and IoT Workloads*

- **MQTT and IoT core:** AWS issues certificates for MQTT workloads, ensuring encrypted data transmission over Transport Layer Security (TLS) connections.
- **Federal Risk and Authorization Management Program certification:** Both the VPC and IoT Core services have achieved Federal Risk and Authorization Management Program certification, ensuring high-level security standards.

#### *API Security and Monitoring*

Despite robust security measures, the risk of attackers gaining user-level access remains, particularly if multifactor authentication is not used. AWS addresses this through:

- **AWS CloudTrail:** This service monitors all API calls to the cloud services, including authorization details and source locations, to help identify unauthorized or abnormal activities.
- **Amazon GuardDuty:** Working alongside AWS CloudTrail, Amazon GuardDuty applies anomaly detection to API activities and can alert when anomalous behavior is identified.
- **Multifactor authentication and role-based access:** To prevent unauthorized access, AWS recommends not issuing default root credentials and enforcing multifactor authentication and time-constrained role-based access for configuring core services.

#### **6.3.1 Summary**

The depth and effectiveness of cloud security, as exemplified by AWS's comprehensive suite of tools and best practices, are significantly enhanced when applied in alignment with the shared responsibility model. This model, which delineates the security obligations of both the cloud provider and the user, is pivotal in creating a robust security posture. By leveraging AWS's advanced encryption, network protection, and monitoring services, combined with diligent customer practices such as implementing multifactor authentication and role-based access, organizations can achieve a highly secure cloud environment. It is important that users of cloud services understand and actively engage in their part of the shared responsibility model because this collaboration is the cornerstone of safeguarding data and applications in the cloud. Through this synergy, cloud providers may offer a compelling resilient and secure infrastructure for diverse computing needs.

## 7 Conclusion

In Phase 2 of the CloudZero project, NREL researchers successfully designed and deployed a test environment using the ARIES Cyber Range as a base with connections to AWS cloud-based services in order to evaluate feasibility, limitations, risks, and costs of using cloud-based services as a viable method of replacement for or supplementation to utility on-premises solutions. Following exercises to ensure the developed test system was stable and accurate, NREL researchers executed three use cases to explore this feasibility: Load Forecasting, High Availability, and Bulk Electric System Reliability Operations. The following summarizes the key takeaways, impacts on utilities, and possible future phase work resulting from that research.

### 7.1 Key Takeaways

NREL researchers evaluated the three use cases described above, which resulted in key takeaways that may be useful to utilities, especially as the industry continues to shift towards the cloud to take advantage of its benefits.

#### 7.1.1 Load Forecasting Key Takeaways

- Cloud-based services can be successfully leveraged to ingest and handle 1-minute data as well as correspondingly higher volumes of data as needed.
- Cloud-based services can immediately scale as needed to process data at higher rates and volume. This is compared to on-premises systems where dedicated engineers would be required to scale the hardware up to handle more data.
- Utilities can quickly leverage out-of-the-box cloud services including real-time data ingest, on-the-fly extract transform load services, data storage, machine learning forecast capabilities, and data visualization. Additionally, these services are well integrated into a single platform versus having to work with several third-party services.

#### 7.1.2 High Availability Key Takeaways

- Reliability, like many cloud services and functions continues to be largely a cost-benefit analysis. NREL's demonstrated system architecture was able to successfully show the capacity for failover using cloud-based services at a speed sufficient for most basic Supervisory Control and Data Acquisition (SCADA) services in switching servers within 2 minutes. However, this RTO is unacceptable for managing critical services or high frequency grid operations. With the demonstrated topology and the level of cost it requires, this lends credence to the hypothesis that many smaller utilities or those serving largely noncritical loads can feasibly shift the monitoring and operation of those services to the cloud with current levels of service.
- Given industry trends of increasing implementation and availability of cloud-native reliability features offered by cloud service providers and the development of highly available reference architectures employing these new features, we expect the cost to enable reliability to go down over time for a specific level of service.

#### 7.1.3 Bulk Electric System Reliability Operations Key Takeaways

- A co-simulation framework within the cloud was able to enable the study of inter-dependencies between large-scale transmission and distribution systems. For example, how generator outages in the transmission system impact the voltages of the distribution feeder and how faults in the distribution feeder affect the transmission system are analyzed.
- As a proof-of-concept, a set of BRO applications including AGC and RAS were able to address the outages and faults that occurred in the test system.
- Cloud-based services such as machine learning based prediction and control, security features for authenticating the IEDs, and other services such as failover services for high availability are flexible and could allow for expansion to other critical and time-sensitive BRO operations applications.

## 7.2 Impacts on Utilities

Using the findings from NREL's research, utilities can benefit from the use of the cloud in the following ways.

- Utilities can leverage cloud-based services to enhance their existing on-premises systems to process data and generate load forecasts at costs that are potentially lower than on-premises counterparts.
- Utilities can use these findings as a resource to perform cost-benefit analyses with more accurate information and better-informed estimates of cost to deliver reliable grid services with a cloud-based deployment of their grid monitoring and control services. These findings may also prove valuable to cloud service providers to help them produce better configurations and more accurate pricing of their offerings to customers.
- Utilities can better identify which BES reliability operations and on-premises applications can be deployed in the cloud, as a result of the response time of the AGC control actions, remedial actions to contingencies, and voltage control under system faults.

## 7.3 Future Phase Opportunities

The electric sector is already leveraging cloud services for some aspects of grid operations. Creating frameworks in which to test new deployments of data analysis and control operations will bolster the efforts by organizations to secure gains in reliability achieved with the cloud. Creating more experiments in these test beds to demonstrate effective use of cloud-native services to secure these new business processes is paramount. As an example, this work shows that the ARIES Cyber Range is poised to answer more complex questions for both utility and non-utility stakeholders about the security and reliability of virtual power plants and electricity aggregators. The Cyber Range will support modeling these systems and simulating their effects on greater grid stability at the distribution and transmission layers. This research will unlock the next generation of secure networks and reliable grid operation of the future.

The risks are fully characterized and understood, the industry can properly prepare for securely transitioning to the use of cloud resources for certain operational needs. Testing capabilities such as those developed in CloudZero are necessary to demonstrate the value and evaluate risks of this transition. CloudZero has built and validated an environment that is effective for testing solutions to these issues. Future testing may also include natural hazard threats that are statistically much more likely to disrupt grid operations.

## References

- Amazon Web Services. 2011. *Amazon Web Services: Overview of Security Processes*. Seattle, WA: AWS. [https://personal.utdallas.edu/~muratk/courses/cloud11f\\_files/AWS\\_Security\\_Whitepaper.pdf](https://personal.utdallas.edu/~muratk/courses/cloud11f_files/AWS_Security_Whitepaper.pdf).
- . 2023a. “AWS Post-Event Summaries.” Amazon Web Services. (accessed: 12.12.2023). <https://aws.amazon.com/premiumsupport/technology/pes/>.
- . 2023b. “AWS Service Level Agreements (SLAs).” Amazon Web Services. (accessed: 12.12.2023). <https://aws.amazon.com/legal/service-level-agreements/>.
- . 2023c. “Disaster recovery options in the cloud.” Amazon Web Services. (accessed: 12.12.2023). <https://docs.aws.amazon.com/whitepapers/latest/disaster-recovery-workloads-on-aws/disaster-recovery-options-in-the-cloud.html>.
- . 2023d. “New Compliance Guide: NERC CIP standards for BES Cyber System Information on AWS.” Amazon Web Services. (accessed: 12.12.2023). <https://aws.amazon.com/blogs/industries/new-compliance-guide-nerc-cip-standards-for-bes-cyber-system-information-on-aws/>.
- . n.d.(a). “AWS Services in Scope by Compliance Program.” Amazon Web Services. <https://aws.amazon.com/compliance/services-in-scope/>.
- . n.d.(b). “Shared Responsibility Model.” Amazon Web Services. <https://aws.amazon.com/compliance/shared-responsibility-model/>.
- Bergengruen, V. 2023. “‘Is There Something More Sinister Going On?’ Authorities Fear Extremists Are Targeting U.S. Power Grid.” *Time*, <https://time.com/6244977/us-power-grid-attacks-extremism>.
- California State Assembly: Committee on Utilities and Energy. 2022. “2021-2028 California Summer Grid Reliability Planning & Procurement Actions.” <https://autl.assembly.ca.gov/sites/autl.assembly.ca.gov/files/CPUC%20-%20Energy%20Resource%20Planning%20%26%20Procurement%20Actions%20Nov%202022.pdf>.
- Chen, Y., Y. Tan, and B. Zhang. 2019. “Exploiting Vulnerabilities of Load Forecasting Through Adversarial Attacks.” *arXiv*, <https://doi.org/10.48550/arXiv.1904.06606>.
- Colonial Pipeline Company. 2023. “Colonial Pipeline System Disruption.” *Colonial Pipeline Company*, <https://www.colpipe.com/news/press-releases/media-statement-colonial-pipeline-system-disruption>.
- CrowdStrike. 2023. *CrowdStrike 2023 Global Threat Report*. Sunnyvale, CA: CrowdStrike. <https://go.crowdstrike.com/2023-global-threat-report.html>.
- Cybersecurity and Infrastructure Security Agency. 2023. “APT28 Exploits Known Vulnerability to Carry Out Reconnaissance and Deploy Malware on Cisco Routers.” *Cybersecurity and Infrastructure Security Agency*, <https://www.cisa.gov/news-events/cybersecurity-advisories/aa23-108>.
- ERCOT. 2018. *ERCOT Nodal ICCP Communication Handbook*. Austin, TX: ERCOT. [https://www.ercot.com/files/docs/2019/02/08/07.ERCOT\\_Nodal\\_ICCP\\_Communications\\_Handbook\\_v3\\_09\\_10\\_15\\_2018.doc](https://www.ercot.com/files/docs/2019/02/08/07.ERCOT_Nodal_ICCP_Communications_Handbook_v3_09_10_15_2018.doc).
- Fan, T.-H., X. Y. Lee, and Y. Wang. 2021. “PowerGym: A Reinforcement Learning Environment for Volt-Var Control in Power Distribution Systems.” *arXiv*, <https://doi.org/10.48550/arXiv.2109.03970>.
- FERC. 2014. *Frequency Response and Frequency Bias Setting Reliability Standard*. Federal Energy Regulatory Commission. <https://www.federalregister.gov/documents/2014/01/23/2014-01218/frequency-response-and-frequency-bias-setting-reliability-standard>.
- Hardy, T. D., B. Palmintier, P. L. Top, D. Krishnamurthy, and J. C. Fuller. 2024. “HELICS: A Co-Simulation Framework for Scalable Multi-Domain Modeling and Analysis.” *IEEE Access* 12:24325–24347. <https://doi.org/10.1109/ACCESS.2024.3363615>.



Henry, J., W. Folger, A. Hasandka, A. Liao, A. Wallace, B. Buchanan, B. Richardson, et al. 2022. *CloudZero Phase 1: A Preliminary Analysis of Cloud Performance for Energy Systems (Citation Only)*. <https://research-hub.nrel.gov/en/publications/cloudzero-phase-1-a-preliminary-analysis-of-cloud-performance-for>.

IEEE-PES. 2016. *IEEE Recommended Practice for Excitation System Models for Power System Stability Studies*. Piscataway, NJ: IEEE Power and Energy Society. <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7553421>.

Li, Z., M. Liang, L. O'Brien, and H. Zhang. 2013. "The Cloud's Cloudy Moment: A Systematic Survey of Public Cloud Service Outage." *International Journal of Cloud Computing and Services Science (IJ-CLOSER)* 2 (5). <http://dx.doi.org/10.11591/closer.v2i5.5125>.

MITRE. 2023. "MITRE ATT&CK Framework: Authoritative Reference to Map Cyber Attack." MITRE. <https://attack.mitre.org>.

———. 2024. *Caldera: A Scalable, Automated Adversary Emulation Platform*. <https://caldera.mitre.org/>.

NERC. 2023. *Consideration of Comments*. <https://shorturl.at/AEIS8>.

Palmintier, B., M. D. Carlos, P. Marcos, F. Emilio, G. S. Roman, F. Tomas de Cuadra, G. Nicolas, T. Elgindy, and P. Duenas. 2020. "SMART-DS Synthetic Electrical Network Data OpenDSS Models for SFO, GSO, and AUS." Open Energy Data Initiative. <https://data.openei.org/submissions/2981>.

Palo Alto Networks. 2023. *Unit 42 Cloud Threat Report, Volume 7*. Santa Clara, CA: Palo Alto Networks. <https://start.paloaltonetworks.com/unit-42-cloud-threat-report-volume-7>.

Panossian, N., T. Elgindy, B. Palmintier, and D. Wallison. 2021. "Synthetic, Realistic Transmission and Distribution Co-Simulation for Voltage Control Benchmarking." In *2021 IEEE Texas Power and Energy Conference (TPEC)*, 1–5. <https://doi.org/10.1109/TPEC51183.2021.9384935>.

Patria Security, LLC. 2023. "Operational Technology (OT) Simulator." <https://ot-sim.patsec.dev/>.

PowerWorld. 2023a. "Available Generation Control (AGC) Modeling." [https://www.powerworld.com/WebHelp/Content/MainDocumentation\\_HTML/Available\\_Generation\\_Control\\_Modeling.htm](https://www.powerworld.com/WebHelp/Content/MainDocumentation_HTML/Available_Generation_Control_Modeling.htm).

———. 2023b. "Exciter Model: ESST4B." [https://www.powerworld.com/WebHelp/Content/TransientModels\\_HTML/Exciter%20ESST4B.htm](https://www.powerworld.com/WebHelp/Content/TransientModels_HTML/Exciter%20ESST4B.htm).

———. 2024a. *Governor Model: GAST2A*. [https://www.powerworld.com/WebHelp/Content/TransientModels\\_HTML/Governor%20GAST2A.htm](https://www.powerworld.com/WebHelp/Content/TransientModels_HTML/Governor%20GAST2A.htm).

———. 2024b. *Governor Model: HYG0V and HYG0VD*. [https://www.powerworld.com/WebHelp/Content/TransientModels\\_HTML/Governor%20HYGOV%20and%20HYGOVD.htm](https://www.powerworld.com/WebHelp/Content/TransientModels_HTML/Governor%20HYGOV%20and%20HYGOVD.htm).

———. 2024c. *Machine Model: GENROU*. [https://www.powerworld.com/WebHelp/Content/TransientModels\\_HTML/Machine%20Model%20GENROU.htm](https://www.powerworld.com/WebHelp/Content/TransientModels_HTML/Machine%20Model%20GENROU.htm).

———. 2024d. *Stabilizer Model: PSS4B*. [https://www.powerworld.com/WebHelp/Content/TransientModels\\_HTML/Stabilizer%20PSS4B.htm](https://www.powerworld.com/WebHelp/Content/TransientModels_HTML/Stabilizer%20PSS4B.htm).

"SMART-DS: Synthetic Models for Advanced, Realistic Testing: Distribution Systems and Scenarios." 2023. <https://www.nrel.gov/grid/smart-ds.html>.

U.S. Government Accountability Office. 2019. "Critical Infrastructure Protection: Actions Needed to Address Significant Cybersecurity Risks Facing the Electric Grid." U.S. Government Accountability Office. <https://www.gao.gov/products/gao-19-332>.

Vandenbrink, R. 2017. "Migrating Telnet to SSH without Migrating." Internet Storm Center. <https://isc.sans.edu/diary/Migrating+Telnet+to+SSH+without+Migrating/22376>.

Williams, T. 1993. "The Purdue Enterprise Reference Architecture." 12th Triennial World Congress of the International Federation of Automatic control. Volume 4 Applications II, Sydney, Australia, 18-23 July, *IFAC Proceedings*

*Volumes* 26 (2, Part 4): 559–564. ISSN: 1474-6670. [https://doi.org/https://doi.org/10.1016/S1474-6670\(17\)48532-6](https://doi.org/https://doi.org/10.1016/S1474-6670(17)48532-6).  
<https://www.sciencedirect.com/science/article/pii/S1474667017485326>.

## Appendix A. Tables, Figures, and Code Snippets

### A.1 Security Operations Center

To create health metrics for the virtual machines deployed in the emulated enterprise, the researchers developed and integrated a set of security operation services regarding each system, network, and service. These checks included external threat scenarios that enabled the researchers to generate network and system level data that could be used to validate events that have occurred across the system, such as a cyber threat.

The development of a baseline security operation center provided a platform to gather, store, and analyze data.

- Security information and event manager: Elasticsearch and Kibana were deployed in the security operation scenario to enable active data gathering from systems across the enterprise.
- Intrusion detection system: Snort was leveraged as a simple signature-based solution to enable active logging of events across the enterprise network.
- Network-level packet capture: Packet captures via virtual network taps were used to gain exposure to network-level data.
- System-level log data: The beat suite was used as a method to log host data across the enterprise system.

#### A.1.1 Enterprise Data Shippers

As a baseline method to enable visibility into the emulated enterprise, researchers enabled the collection of data across the system. The following logging tools were deployed:

- Winlogbeat: Windows logs for all systems within the enterprise emulation model (domain controller, email server, etc.)
- Packetbeat: Packet capture and log generation from sniffer on mirror ports across the enterprise. This provides network-level data, allowing us to discover when an event is taking place across the system.
- Metricbeat: State of health information that is gathered to help understand if and when an event occurred
- Filebeat: Logging data from both Linux and Windows systems to enable the log analysis.

These data tools were leveraged to better understand entry and existing points between information technology and operational technology systems and the challenges with domain-level system integration with operational technology systems and supporting cloud services, such as AWS or Azure (or other private on-site cloud services).

**Table A.1. Point of Common Coupling Between Transmission and Distribution Systems**

Index	Area	Transmission Substation Name	Distribution Substation Name
1	Travis	Travis_DS_16	P3uhs0_69
2	Travis	Travis_DS_36	P3uhs1_69
3	Area4	Travis_DS_38	P3uhs2_69
4	Travis	Travis_DS_26	P3uhs3_69
5	Area3	Travis_DS_29	P3uhs4_69
6	Area3	Travis_DS_15	P3uhs5_69
7	TestAGCArea	Travis_DS_34	P3uhs6_69
8	Area3	Travis_DS_23	P3uhs7_69
9	Area4	Travis_DS_17	P3uhs8_69
10	Area4	Travis_DS_41	P3uhs9_69
11	TestAGCArea	Travis_DS_31	P3uhs10_69
12	Area3	Travis_DS_27	P3uhs11_69
13	TestAGCArea	Travis_DS_35	P3uhs12_69
14	Area3	Travis_DS_28	P3uhs13_69
15	TestAGCArea	Travis_DS_30	P3uhs14_69
16	Travis	Travis_DS_32	P3uhs15_69
17	Area3	Travis_DS_21	P3uhs16_69
18	Area3	Travis_DS_19	P3uhs17_69
19	TestAGCArea	Travis_DS_18	P3uhs18_69
20	Area3	Travis_DS_24	P3uhs19_69
21	TestAGCArea	Travis_DS_20	P3uhs20_69
22	Area3	Travis_DS_22	P3uhs21_69
23	Area3	Travis_DS_25	P3uhs22_69
24	Travis	Travis_DS_33	P3uhs23_69
25	TestAGCArea	Travis_DS_37	P3uhs24_69
26	Area4	Travis_DS_39	P3uhs25_69
27	Area4	Travis_DS_40	P3uhs26_69
28	Area4	Travis_DS_42	P3uhs27_69
29	Area4	Travis_DS_43	P3uhs28_69

**Table A.2. Generator Transient Modeling**

Generator	Governor	Excitor	Machine Model	Stabilizer
Marshall Ford Power Plant Unit 2	HYGOV	ESST4B	GENROU	PSS4B
Marshall Ford Power Plant Unit 3	HYGOV	ESST4B	GENROU	PSS4B
Austin Power Plant Unit 1	HYGOV	ESST4B	GENROU	PSS4B
Austin Power Plant Unit 2	HYGOV	ESST4B	GENROU	PSS4B
Weaver Power Plant Unit 3	GAST2A	ESST4B	GENROU	PSS4B
Weaver Power Plant Unit 4	GAST2A	ESST4B	GENROU	PSS4B
Weaver Power Plant Unit 5	GAST2A	ESST4B	GENROU	PSS4B
Weaver Power Plant Unit 6	GAST2A	ESST4B	GENROU	PSS4B
Mueller Energy Center Unit 1	GAST2A	ESST4B	GENROU	PSS4B
Sand Hill Power Plant Unit 1	GAST2A	ESST4B	GENROU	PSS4B
Sand Hill Power Plant Unit 2	GAST2A	ESST4B	GENROU	PSS4B
Sand Hill Power Plant Unit 3	GAST2A	ESST4B	GENROU	PSS4B
Sand Hill Power Plant Unit 4	GAST2A	ESST4B	GENROU	PSS4B
Sand Hill Power Plant Unit 5	GAST2A	ESST4B	GENROU	PSS4B
Sand Hill Power Plant Unit 6	GAST2A	ESST4B	GENROU	PSS4B
Decker Creek Power Plant Unit 1	GAST2A	ESST4B	GENROU	PSS4B
Decker Creek Power Plant Unit 2	GAST2A	ESST4B	GENROU	PSS4B
Decker Creek Power Plant Unit 3	GAST2A	ESST4B	GENROU	PSS4B
Decker Creek Power Plant Unit 5	GAST2A	ESST4B	GENROU	PSS4B
Decker Creek Power Plant Unit 6	GAST2A	ESST4B	GENROU	PSS4B
Bastrop Energy Center Unit 1	GAST2A	ESST4B	GENROU	PSS4B
Bastrop Energy Center Unit 2	GAST2A	ESST4B	GENROU	PSS4B
Bastrop Energy Center Unit 3	GAST2A	ESST4B	GENROU	PSS4B
Lost Pines 1 Power Project Unit 1	GAST2A	ESST4B	GENROU	PSS4B
Lost Pines 1 Power Project Unit 2	GAST2A	ESST4B	GENROU	PSS4B
Lost Pines 1 Power Project Unit 3	GAST2A	ESST4B	GENROU	PSS4B
Sam Gideon Power Plant Unit 1	No	ESST4B	GENROU	PSS4B
Sam Gideon Power Plant Unit 2	No	ESST4B	GENROU	PSS4B
Sam Gideon Power Plant Unit 3	No	ESST4B	GENROU	PSS4B
Webberville Solar Project	No	No	No	No
IKEA Solar Array	No	No	No	No

**Table A.3. Impact of Line Outage Contingencies. The impact factor represents the percentage by which a transmission line is overflowing. An impact factor of 104.0 indicates that a transmission line is 4% above its maximum power carrying capacity.**

Contingency Name	Description (From and To Bus)	Impacted Lines with impact factor
C1	Bus 119, Circuit 1 to Weaver Power Plant 2 Circuit 3	a) Travis_DS_101 1 (119) TO Weaver Power Plant 2 (158) CKT 1 <b>104.0</b> b) Travis_DS_101 1 (119) TO Weaver Power Plant 2 (158) CKT 2. <b>104.0</b>
C2	Bus 119, Circuit 1 to Weaver Power Plant 2 Circuit 2	a) Travis_DS_101 1 (119) TO Weaver Power Plant 2 (158) CKT 1. <b>104.0</b> b) Travis_DS_101 1 (119) TO Weaver Power Plant 2 (158) CKT 3. <b>104.0</b> .
C3	Bus 108, Circuit 1 to Marshall Ford Power Plant 2 Circuit 1	a) Travis_DS_90 1 (108) TO Marshall Ford Power Plant 2 (151) CKT 2. <b>109.5</b>
C4	Bus 140, Circuit 1 to Bus 138, Circuit 2	a) Travis_DS_118 2 (138) TO Travis_DS_-123 1 (143) CKT 1. <b>101.6</b>
C5	Bus 120, Circuit 1 to Bus 118, Circuit 2	a) Travis_DS_102 1 (120) TO Travis_DS_-100 1 (118) CKT 1. <b>101.0</b>
C6	Bus 108, Circuit 1 to Bus 107, Circuit 1	a) Travis_DS_88 1 (106) TO Travis_DS_-95 1 (113) CKT 1. <b>100.3</b>
C7	Bus 108, Circuit 1 to Marshall Ford Power Plant 2 Circuit 2	a) Travis_DS_90 1 (108) TO Marshall Ford Power Plant 2 (151) CKT 1. <b>109.5</b>
C8	Bus 120, Circuit 1 to Bus 118, Circuit 1	a) Travis_DS_102 1 (120) TO Travis_DS_-100 1 (118) CKT 2. <b>101.0</b>
C9	Bus 119, Circuit 1 to Weaver Power Plant 2 Circuit 1	a) Travis_DS_101 1 (119) TO Weaver Power Plant 2 (158) CKT 2. <b>104.0</b> b) Travis_DS_101 1 (119) TO Weaver Power Plant 2 (158) CKT 3. <b>104.0</b>

**Table A.4. Impact of Line Outage Contingencies After One of the RAS is Implemented.**

Contingency Name	Description (From and To Bus)	Impacted Lines with impact factor
C2	Bus 119, Circuit 1 to Weaver Power Plant 2 Circuit 2	a) Travis_DS_101 1 (119) TO Weaver Power Plant 2 (158) CKT 1. <b>104.0</b> b) Travis_DS_101 1 (119) TO Weaver Power Plant 2 (158) CKT 3. <b>103.6</b> .
C3	Bus 108, Circuit 1 to Marshall Ford Power Plant 2 Circuit 1	a) Travis_DS_90 1 (108) TO Marshall Ford Power Plant 2 (151) CKT 2. <b>110.2</b>
C4	Bus 140, Circuit 1 to Bus 138, Circuit 2	a) Travis_DS_118 2 (138) TO Travis_DS_-123 1 (143) CKT 1. <b>101.2</b>
C5	Bus 120, Circuit 1 to Bus 118, Circuit 2	a) Travis_DS_102 1 (120) TO Travis_DS_-100 1 (118) CKT 1. <b>101.2</b>
C6	Bus 108, Circuit 1 to Bus 107, Circuit 1	a) Travis_DS_88 1 (106) TO Travis_DS_-95 1 (113) CKT 1. <b>101.4</b>
C7	Bus 108, Circuit 1 to Marshall Ford Power Plant 2 Circuit 2	a) Travis_DS_90 1 (108) TO Marshall Ford Power Plant 2 (151) CKT 1. <b>110.2</b>
C8	Bus 120, Circuit 1 to Bus 118, Circuit 1	a) Travis_DS_102 1 (120) TO Travis_DS_-100 1 (118) CKT 2. <b>101.2</b>
C9	Bus 119, Circuit 1 to Weaver Power Plant 2 Circuit 1	a) Travis_DS_101 1 (119) TO Weaver Power Plant 2 (158) CKT 2. <b>104.0</b> b) Travis_DS_101 1 (119) TO Weaver Power Plant 2 (158) CKT 3. <b>103.6</b>