RT-METER: A Real-Time, Multi-Layer Cyber–Power Testbed for Resiliency Analysis

Hussain M. Mustafa, Mohini Bariya, K. S. Sajan, *Member, IEEE*, Ajay Chhokra, A. Srivastava, *Senior Member, IEEE*, A. Dubey, *Senior Member, IEEE*, A. von Meier, *Member, IEEE* and G. Biswas, *Fellow, IEEE*

***Abstract*—In this work, we present a Real-Time, Multi-layer cybEr–power TestbEd for the Resiliency analysis (RT-METER) to support power grid operation and planning. Developed cyber-power testbed provides a mechanism for end-to-end validation of advanced tools for cyber-power grid monitoring, control, and planning. By integrating a host of features across three core layers—physical power system, communication network, and monitoring/ control center with advanced tools,—the testbed allows for the simulation of rich and varied cyber-power grid scenarios and the generating realistic sensor, system, and network data. Developing advanced tools to assist operators during complex and challenging scenarios is essential for the successful operation of the future grid. We detail a suite of algorithmic tools validated using the developed testbed for the realistic grid data.**

I. Introduction

THE growing complexity of electric power systems, along with more frequent natural and human-made disruptions, are challenging the operation of nation's electric grid, motivating the development of the "smart grid" [1]. The emerging smart grid is a *cyber-physical* system, incorporating large numbers of networked devices for measurement and control and encompassing many functions, participants, and components [2], [3]. This complex system must be carefully managed for resilience, defined in [4] as the system's ability to provide energy to critical loads even under adverse events [5]. Developing and demonstrating advanced tools for decision support as well as cyber-physical resilience assessment is therefore critical to the resilient power grid. However, that necessitates a cyber-physical testbed capable of capturing the complex and multi-layered smart grid for validating these tools.

The literature describes various electric grid testbeds [6]. The authors of [7] present a detailed comparison of various testbeds developed by diverse research groups, describing components, scalability, and foci. A cyber-physical testbed is used for studies of wide-area situational awareness and cyber security applications in transmission and operation [8]. The PowerCyber tested with RTDS, and various communication technologies is targeted towards voltage stability and cyber security issues [9]. None of the testbeds integrate a real-time power system simulator with a realistic control center model—including industry-standard databases and the ability to run advanced tools—connected via a communication network. While [10] and [9] present testbeds that also allow users to simulate device attacks on protection system components, they are narrow in their capabilities. Especially, the development of comprehensive, flexible testbeds for resiliency analysis is limited.

Hussain M. Mustafa, K.S. Sajan and A. Srivastava are with the Washington State University. M. Bariya and A. von Meier are with the University of California at Berkeley. Ajay Chhokra, A. Dubey and G. Biswas are with the Vanderbilt University. Authors would like to acknowledge financial support from NSF 1840192, 1840083 and 1840052 for this work.

This paper presents the development of a multilayered cyber-power testbed with three core layers: power, communication network, and control center. The control layer encompasses two primary, linked modules: algorithmic tools and data storage & interface. This testbed enables the simulation of a wide variety of cyber-power scenarios encompassing physical, communication, and control events to validate tools for grid planning, monitoring, control, and decision-support. Example of tools presented in this work target a particular cyber-power scenario and leverages data generated from the testbed layers. Our purpose here is not to detail and demonstrate novel algorithms but to illustrate how such tools' development and validation rely on an integrated cyber-power testbed. Our key contribution is a testbed with the following features:

- A power layer with a real-time physical power systems simulator and realistic emulation of grid sensors and devices developed in HYPERSIM. It also has multiple dynamic simulators to generate a massive amount of event data used for advanced tool development.
- A highly detailed, flexible communication network layer modeled in NS3 closely resembling a real communication network's behavior, which enables the modeling of various network events.
- A control layer with data storage, algorithmic tools and interface elements. The interface provides an opportunity to test the efficacy of algorithmic tools in assisting human operators under challenging and unusual system conditions.
- Example of algorithmic tools developed for inclusion in the EMS in the control layer, which have undergone preliminary validation with the testbed.

II. Multi-layered Testbed Architecture

Figure 1 shows developed hardware-in-the-loop, cyber-power testbed, with the power, network, and control layers. Some components span multiple layers or play a liminal role between layers, but this multi-layer organization is an effective mental model for understanding the testbed's various functionalities. This section describes each of the core layers' components and features.

A. Power Layer

The power layer enables real-time simulation of an electric power system. The simulation of the system utilizes HYPERSIM, a power system solver from Opal-RT technologies that enjoys extensive use in academia and industry. HYPERSIM

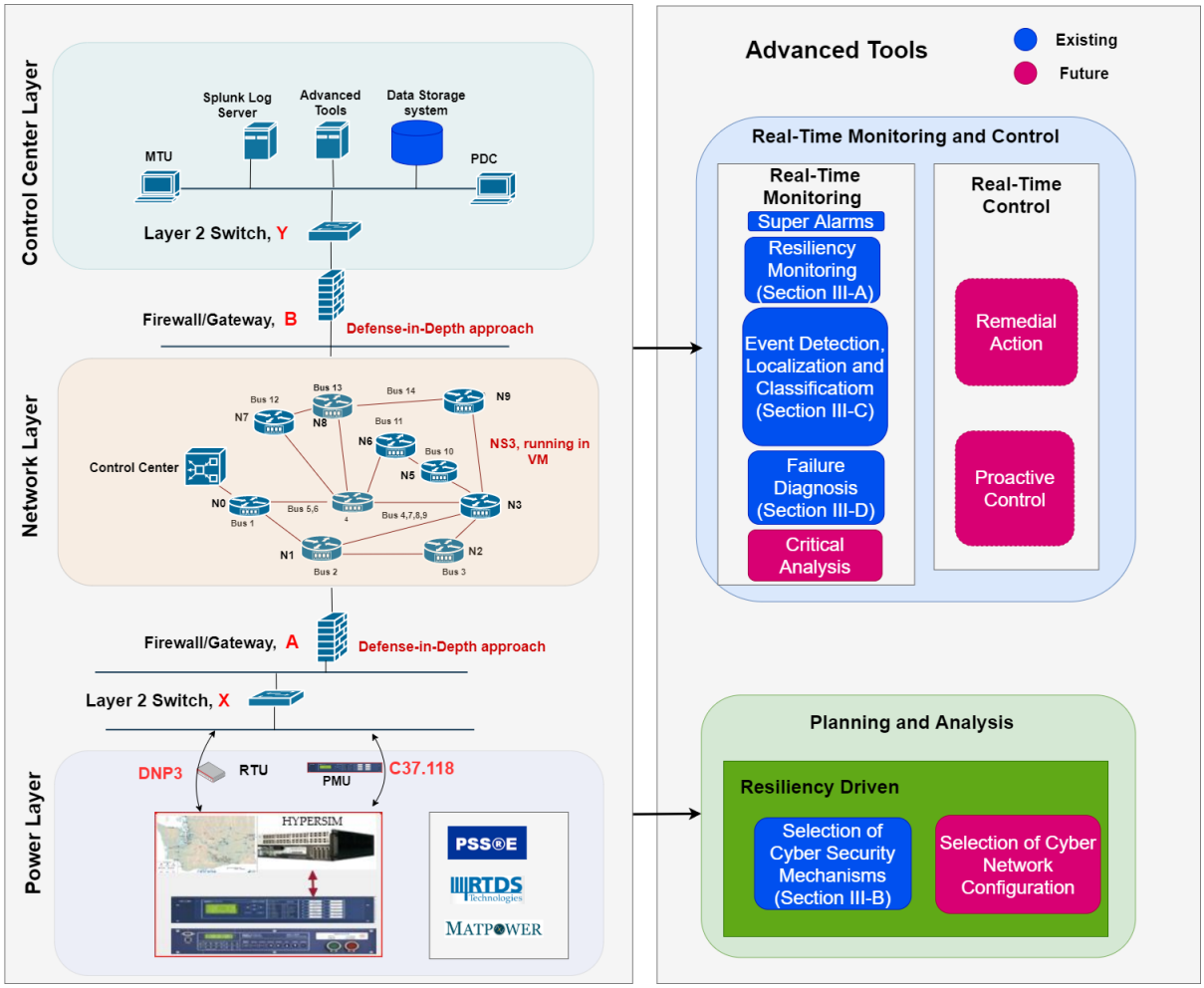

Fig. 1: RT-METER: A Real-Time, Multi-Layer Cyber–Power Testbed Architecture

provides comprehensive solvers, toolboxes, and libraries to model power grid infrastructure with a flexible, scalable, parallelizable architecture, enables high-resolution simulations on time steps of 5-200 μs [11]. In this work, the IEEE-14 bus system model is used as the physical power grid model. This relatively small, simple model is useful initial validation, but other models can also be integrated into our testbed design.

Sensors that measure physical grid quantities are also part of the power layer of the testbed. HYPERSIM supports software versions of several sensors and processors, including phasor measurement units (PMU) and remote terminal units (RTUs) that can optionally be interfaced with external hardware devices via dedicated interface cards. Measurements made by sensors in the testbed's power layer are encoded and packed according to specific communication protocols and then passed into the testbed's network layer. HYPERSIM supports several smart grid protocols such as DNP3.0, IEEE C37.118, IEC61850. Our testbed also has the capability to incorporate several different real-time power simulators, which can be used for data generation for real-time monitoring and control, along with planning and analysis tools. Currently, we can integrate PSSE, RTDS, and MATPOWER in the power layer.

B. *Communication Network Layer*

Communication networks in smart grids connect remote substations with centralized and distributed control centers. Sophisticated and advanced tools running in the control center EMS need data input with acceptable latency and reliability, which largely depends on the underlying communication network. The network layer we modeled reflects the communication backbone in an operational electric system, where data flows between a substation and control center. Our communication model is not restricted to pure simulation. Instead, it is built upon a combination of actual hardware, a powerful and open-source network emulator NS3, and Linux bridging technologies.

NS3's core libraries support various communication network types, with implementations of networking protocols that closely match real-world performance. NS3 offers a TAP bridge feature that enables any physical device to be connected and communicate via the NS3 network without being installed in the NS3 network itself. We leveraged the NS3 Tap-bridge utilities to connect multiple external physical devices.

We leveraged the Ubuntu bridge-utils package to create a virtual local area network to connect multiple devices in the same network. As an added feature, our testbed employs a

defense-in-depth approach by which multiple layers of security mechanisms are incorporated into a layered system like ours. Data communication between each layer is separated by a firewall and gateway router. The firewall and gateway router functionalities are served by a single dedicated virtual machine (VM) based on the Linux kernel. The firewall is managed with a program called uncomplicated firewall (UFW) which offers a simple command-line interface to configure firewall access rules and uses iptables for configuration. We used Linux Networking Subsystem to implement routing functionalities.

As shown in Fig 1, RTU's and PMU's modeled in HYPERSIM are connected via physical ports to a real NETGEAR managed switch **(Layer 2 Switch, X)** which is also connected to the VM **(Firewall/Gateway, A)** running firewall/gateway functionalities. Our NS3 communication network, whose connectivity structure mirrors the physical IEEE 14 bus system structure, runs on a separate VM. For simplicity, we only connect a single NS3 node **(N9)**—representing a single substation—with the power layer via a firewall/gateway **(Firewall/Gateway, A)**, using the tap bridge and Linux bridge technologies. To interface with the control center in the control layer, NS3 connects to the master terminal unit (MTU) and phasor data concentrator (PDC), implementing a defense-in-depth approach by first connecting to **(Firewall/Gateway, B)** and then connect to a Linux bridge **(Layer 2 Switch, Y)**. In NS3, each node other than the control center node represents a substation. Additional substations can be integrated according to the structure we have described. Our testbed uses the DNP3 protocol for SCADA applications between RTUs and the MTU and IEEE C37.118 for synchrophasor applications between PMUs and the PDC, enabling our testbed to be used for a wide variety of power and cyber security analyses.

C. Control Layer

The MTU, PDC, and EMS all fall in the testbed's control layer as they engage in determining and executing control actions and are usually situated in a centralized control center. The real-time measurements obtained from PMUs and RTUs in the power layer are used to estimate the power system's real-time operating state which is continuously monitored through a Human Machine Interface (HMI) by the control room operator who takes corrective actions if needed. The HMI is part of the data storage & interface module within the control layer. However, currently we don't have an integrated HMI to visualize the data.

As its name suggests, the components of the data storage & interface module provide storage capabilities for generic system data and/or act as the interface between human operators and the system. This module includes many of the components that would be found in a utility's control center.

Databases are used to store sensors and communication network data. One such database, included in this layer, is the PingThings' Berkeley Tree Database (BTrDB), a highly performant database for time series data [12]. Another is a log server from software firm Splunk, which stores cyber data comprised of network traffic data and event logs. The varied data in these databases is the input to advanced algorithmic tools, which analyze the data to detect and diagnose anomalies that could indicate network faults or cyberattacks. We have included a good suite of advanced tools in the control layer, which can serve the purpose of resiliency assessment in different events. As shown in Fig 1, tools are divided into two parts, real-time monitoring and control & planning and analysis. We have also mentioned some tools as in Fig 1 to include them in the future. Detailed descriptions of the tools can be found in III.

III. Validating Advanced Tools

The tools module within the control layer includes a portfolio of algorithmic tools for grid monitoring and control. The advanced tools are categorized into two broad areas, namely (1) real-time monitoring and control and (2) Planning and analysis. Our purpose here is not to propose and validate novel tools but to illustrate more broadly how the testbed enables tool development, validation, and interface.

A. Cyber-Power Resiliency Monitoring

Resiliency to natural disasters and cyber intrusions is critical, but there is no commonly accepted metric to quantify the grid's resilience. Various cyber-physical assessment metrics are introduced to quantify the resiliency of cyber-power transmission systems. This quantization is essential to minimize the ramification of the unappealing events through appropriate decisions and appliances. It adopted to study different cyber-physical scenarios by analyzing the following components, as shown in Fig. 2 concerning the physical side and the transmission system's cyber side. This established metric considers network topology and changing real-time operating conditions by utilizing graphical analysis and measuring the network's critical parameters to capture the power grid's redundancy and vulnerabilities.

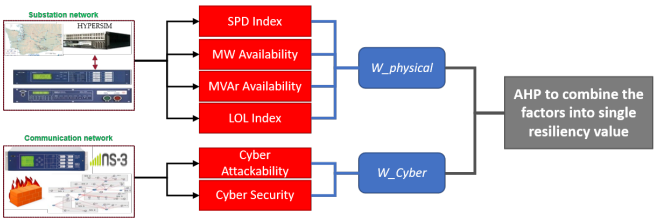

Fig. 2: AHP methods to achieve resilience metrics.

We have developed several metrics for the cyber and power layer for this analysis and monitoring of system performance. The physical resiliency metric is defined based on several primitive indexes such as the source-path-destination(SPD) index, MW Availability Index, MVAr Availability Index, and Loss of Load (LOL) Index. SPD index is calculated using graph analysis of network topology, MW Availability Index to characterize the availability of active power sources, MVAr Availability Index that indicates the system's ability to maintain voltage in the acceptable limits, LOL Index that quantify the percentage of critical loads served. These factors are combined with Analytical Hierarchical Process (AHP) method to obtain one intuitive metric for the operator to enable fast decision making. For cyber resilience metrics, some base metrics are defined. Quality of Service quantifies the communication network performance based on latency, jitter, and packet drop. Vulnerabilities in the Communication Network use graph attack to abstract the available paths for attackers to exploit vulnerabilities and provide attackability metrics. Next,

the Security metric index is developed based on deployed security mechanisms. These metrics are combined to obtain cyber resiliency metrics for fast decision-making. It is essential to note here that the developed resiliency metric integrates various factors that impact system resiliency performance. Additional factors can be added to meet specific systems requirements under study. The developed CP-TRAM methodology is validated on the IEEE 14-bus system with different physical conditions and cyber scenarios to address the cyber-physical resilience metrics. For physical system resilience, CP-TRAM metrics are applied to the following five-event cases: (1) normal operation of the system, (2) a single generator outage, (3) outage of two separate transmission lines, (4) introduced a supplementary reserve generator and (5) loss of load. For each case scenario, the system load conditions are the same and the metrics are computed based on the measurements obtained from the simulator and contingencies summarized in Table I, in which the vector of weight for physical system resilience is defined as $w_{\text{Physical}} = [0.55, 0.15, 0.10, 0.20]^T$.

TABLE I: Computed CP-TRAM Metrics for Different Operating Conditions of Transmission System

Events	SPD	MW	MVAr	LOL	RI_{Physical}	CP-TRAM (Φ_1)	CP-TRAM(Φ_2)
1	10.6867	0.4665	2.9973	1	1.000	0.3165	0.6665
2	8.7171	0.1455	2.3144	1	0.8153	0.2580	0.5434
3	7.7913	0.462	2.9191	1	0.7529	0.2383	0.5018
4	12.7552	0.8362	3.2077	1	1.1903	0.3767	0.7933
5	8.795	0.6645	3.0386	0.611	0.8332	0.2637	0.553

B. Resiliency Driven Planning and Analysis.

We can develop tools to plan/analyze different cyber security mechanisms and cyber network topologies that use the cyber-physical resiliency assessment metrics to select the best configuration by improving the system performance based on resiliency score in planning and analysis. For Example, cyber resiliency is computed by considering two combinations of cyber security strategies and finally calculating the cyber-power resiliency score under various physical systems scenarios. As shown in Table II, the combination Φ_1 adopts a lower level of security mechanisms compared to the combination Φ_2, where the security levels of different cyber mechanisms are selected based on our previous work [13]. The cyber attackability

TABLE II: Comparing Different Combinations of Cyber Security Strategies Adopted at Substation

Security Mechanisms	Combination Φ_1	Combination Φ_2
SM_1	**Low**: Basic Authentication	**High**: MFA
SM_2	**Medium**: NIDS	**High**: SIEM
SM_3	**Medium**: NIPS, Configured Firewall	**High**: NIPS, Configured Firewall, VLAN
SM_4	**Medium**: Symmetric Key	**High**: Asymmetric Key
SM_5	**Medium**: Intermediate	**High**: Advanced
SM_6	**Medium**: Educational programs	**High**: Periodic employee training and scheduled penetration tests

scores are the same for these two combinations since the network topology is not changed, i.e. $a_{\Phi_1} = a_{\Phi_2} = 0.3333$. And after normalizing the cyber security scores based on the highest security combination, i.e. Φ_2, the security scores are $s_{\Phi_1} = 0.2999$ and $s_{\Phi_2} = 1$ respectively. By applying the same weights to attackablity and security, which is $w_{\text{Cyber}} = [0.5, 0.5]^T$, we have calculated the CP-TRAM score for both combination as shaown in Table I.

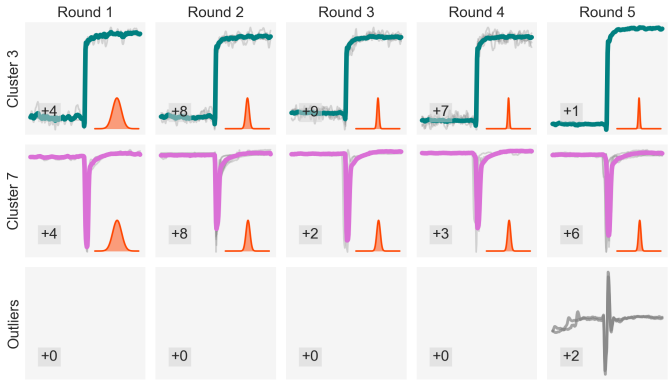

Fig. 3: Two clusters and outliers visualized over five rounds of k-ShapeStream clustering. Gray and colored lines indicate individual events and cluster centers respectively. Inset numbers show number of events added to the cluster in each round. The inset distribution visualizes the probability metric associated with each cluster: notice the narrowing distribution representing increasing certainty in the cluster.

C. Event detection, Classification, and localization

The detection, classification, and localization of physical events is an essential algorithmic tool to enable situational awareness in the grid. In its broadest meaning, *events* can include any change in the system. In this work, we highlight a statistical approach for event detection, an unsupervised clustering approach for event classification, and a straightforward, physics-based approach for event localization, all using high-resolution *voltage* phasor data from PMU devices.

Event Detection. Our event detection tool uses a statistical approach—well-suited for generic anomaly detection [14]. Such an approach presumes that events result in a change in the measurements' statistics and, therefore, doesn't assume a specific event type or measurement signature. The tool can flag both physical events and communication layer issues that manifest in the measurement streams: bad or missing data points. The approach associates a probability measure with each detected event, capturing our certainty that it is a real event. The algorithm, termed Bayesian event detection, tracks and updates moment estimates and then used a Bayesian score to identify events. Mathematically, it assumes that voltage measurements arise from a Gaussian distribution, whose mean and variance parameters follow a Normal-inverse-chi-squared (NIX) prior. If a new measurement point is sufficiently unlikely given current parameter values, it is classified as an event. The Gaussian parameters are reset after an event is detected, reflecting the renewed uncertainty in the measurement statistics following an event. This algorithm is suited for event detection in PMU stream.

Event Classification. Our event classification tool—termed k-ShapeStream—uses a streaming clustering approach to identify recurring event signatures [15]. A window of measurements capturing the event signature can be passed to k-ShapeStream once an event is detected, which will either

be able to match the event to previous events with similar signatures or flag the event as an outlier. There is no explicit event labeling with k-ShapeStream. However, the evidence of prior, similar events can enable classification and diagnosis. k-ShapeStream not only identifies clusters of recurring events but attaches a probability metric to the cluster membership, enabling an intuitive understanding of the algorithm's confidence in the cluster membership.

This algorithm's application to grid data is visualized in Fig. 3. This data is a distribution PMU data from an operational feeder and demonstrates the algorithm's efficacy in clustering recurrent event signatures and separating outliers. The visualization of clusters and their attached probability metrics is intuitive and can be presented to a human operator through the testbed HMI interface.

Event Localization Events are localized to the network node, which witnessed the largest *change* in measured voltage magnitude at the event time [14]. The physical justification for this approach is based on two simplifying assumptions. First, that an event consists of a change in the power injection at a single bus. Second, that the following power flow linearization holds: $V \approx RP$ where $V \in \mathbb{R}^N$ is the set of bus voltage magnitudes, $P \in \mathbb{R}^N$ is the set of real power injections at each bus, and $R \in \mathbb{R}^{N \times N}$ is the $N \times N$ system resistance matrix. Under these assumptions, and using the fact that the diagonal elements of R have the largest magnitude [16], our algorithm will correctly localize the event to the source bus.

The tools described in this section rely on the testbed's ability to simulate physical and cyber system events, generate realistic, synthetic PMU data, and store this data in the high-performance time-series BTrDB. BTrDB's structure enables rapid access to large volumes of very high-resolution PMU data [17], which is critical for these tools' efficiency. The results of these tools will eventually be communicated back to the data storage & Interface Layer, where they can be visualized for a human user. In the near future, we will generate an extensive library of physical grid events in the testbed, on which these algorithms can be extensively validated from both a technical and human user perspective.

D. Model based diagnosis

We present a discrete model-based methodology, Temporal Causal Diagrams (TCDs) [18], as part of this testbed to infer missed and spurious detection problems of protection system while correctly diagnosing faults in the physical system. The two core components of any model-based diagnosis system are 1) *Fault model* i.e., the representation of faults and their effects on the system in terms of observable discrepancies or alarms, 2) *Reasoning algorithm* that utilizes the fault model to produce hypotheses that explain the observed alarms. For TCD based diagnosis system, these components are defined as follows:

1) Fault Model: A TCD fault model uses Temporal Fault Propagation Graph (TFPG) [19] to relate the effects of faults in physical equipment with the response of protection relays. A TFPG is a directed graph where a node can be a fault or discrepancy, and an edge between them implies fault affect propagation. In the power transmission system, a fault node can represent a grounding fault in a segment of a transmission line, whereas the connected discrepancies denote the response of different zone elements associated with the distance relays in the vicinity. Table III shows a subset of the TFPG model for IEEE 14 bus system that illustrates the relationship between faults injected in the transmission line, $\text{TL}^{11\text{-}10}$ (between buses 11,10) and three forward-looking zone elements (z1, z2, z3) of primary ($\text{DR}^{11\text{-}10}$, $\text{DR}^{10\text{-}11}$) and backup ($\text{DR}^{6\text{-}11}$, $\text{DR}^{9\text{-}10}$) distance relays. The transmission line, $\text{TL}^{11\text{-}10}$, is divided into four segments such that for each segment, the response of all zone elements in every distance relay is consistent as indicated in Table III. For instance, if a fault is injected anywhere in the segment, $\text{TL}_1^{11\text{-}10}$, then all three zone elements of distance relay, $\text{DR}^{11\text{-}10}$, should detect the decrease in impedance. Consequently, time-stamped events, labeled as z1_$\text{DR}^{11\text{-}10}$, z2_$\text{DR}^{11\text{-}10}$, z3_$\text{DR}^{11\text{-}10}$ are logged to signify detection. Similarly, detection events related to the remaining relays' zone elements are also logged as described in Table III. A TFPG edge consists of an attribute to signify the minimum or maximum time taken for fault effect to propagate from source to destination node. The last column of the table captures this range which is equal to fault detection time for distance relays, i.e., $\leq$ 0.032 secs [20].

TABLE III: TFPG sub-model for IEEE 14 bus system

Segment ID	Fault Node	Discrepancies	Time
$\text{TL}_1^{11\text{-}10}$ (0-20%)	F_1	z1_$\text{DR}^{11\text{-}10}$, z2_$\text{DR}^{11\text{-}10}$, z3_$\text{DR}^{11\text{-}10}$, z2_$\text{DR}^{10\text{-}11}$, z3_$\text{DR}^{10\text{-}11}$, z2_$\text{DR}^{6\text{-}11}$, z3_$\text{DR}^{6\text{-}11}$, z3_$\text{DR}^{9\text{-}10}$	0-0.032 secs
$\text{TL}_2^{11\text{-}10}$ (20-50%)	F_2	z1_$\text{DR}^{11\text{-}10}$, z2_$\text{DR}^{11\text{-}10}$, z3_$\text{DR}^{11\text{-}10}$, z1_$\text{DR}^{10\text{-}11}$, z2_$\text{DR}^{10\text{-}11}$, z3_$\text{DR}^{10\text{-}11}$, z2_$\text{DR}^{6\text{-}11}$, z3_$\text{DR}^{6\text{-}11}$, z3_$\text{DR}^{9\text{-}10}$	0-0.032 secs
$\text{TL}_3^{11\text{-}10}$ (50-80%)	F_3	z1_$\text{DR}^{11\text{-}10}$, z2_$\text{DR}^{11\text{-}10}$, z3_$\text{DR}^{11\text{-}10}$, z1_$\text{DR}^{10\text{-}11}$, z2_$\text{DR}^{10\text{-}11}$, z3_$\text{DR}^{10\text{-}11}$, z2_$\text{DR}^{9\text{-}10}$, z3_$\text{DR}^{9\text{-}10}$, z3_$\text{DR}^{6\text{-}11}$	0-0.032 secs
$\text{TL}_4^{11\text{-}10}$ (80-100%)	F_4	z1_$\text{DR}^{10\text{-}11}$, z2_$\text{DR}^{10\text{-}11}$, z3_$\text{DR}^{10\text{-}11}$, z2_$\text{DR}^{11\text{-}10}$, z3_$\text{DR}^{11\text{-}10}$, z2_$\text{DR}^{9\text{-}10}$, z3_$\text{DR}^{9\text{-}10}$, z3_$\text{DR}^{6\text{-}11}$	0-0.032 secs

2) Reasoning Methodology: The reasoning algorithm is based on the parent-child relationship between nodes in the TFPG fault model. After receiving the first alarm, the reasoner identifies the corresponding discrepancy in TFPG and traverses backward to create one hypothesis for each reachable fault node. A hypothesis, h is a tuple, $< \mathcal{F}_p, \mathcal{F}_c, \mathcal{C}, \mathcal{E}, \mathcal{M}, \mathcal{I} >$, where $\mathcal{F}_p$ is a set of physical faults that corresponds to fault nodes in TFPG. $\mathcal{F}_c$ is a set of cyber faults (missed and spurious detection) related to zone elements of distance relays. $\mathcal{C}$ is the set of alarms consistent with the hypothesis. $\mathcal{E}$ is a collection of alarms expected to be active in the future. An alarm can remain in the expected set for a fixed amount of time, dictated by the time attribute of the fault propagation edge between the corresponding discrepancy and its parent node. If an alarm is received within the appropriate time period, then it is moved to a consistent set; otherwise, it is added to the missing set, $\mathcal{M}$. For every missing alarm, $a \in \mathcal{M}$, a missed detection fault, f_{md}^a, related to the zone element is added to $\mathcal{F}_c$. Finally, $\mathcal{I}$ contains a collection of alarms that cannot be explained by the hypothesis h. For every alarm, $a \in \mathcal{I}$, a spurious detection fault, f_{sd}^a associated with the zone element is added to $\mathcal{F}_c$. The reasoning algorithm ranks the hypothesis with minimum number of fault nodes, $|\mathcal{F}_p \cup \mathcal{F}_c|$ first, following the law of parsimony [21].

Table IV shows the best hypothesis generated by the

TABLE IV: TCD Reasoner Output

Scenario	Alarms	Reasoner Hypothesis
Missed detection in z1, z2, z3 elements DR^{10_11} at t = 0 sec, Fault injected at 40% from bus 11 at t = 1 sec	z1_DR^{11_10}, z2_DR^{11_10}, z3_DR^{11_10}, z2_DR^{6_11}, z3_DR^{6_11}, z3_DR^{9_10} logged at t = 1.016 secs	ID = H1, Rank = **2** $\mathcal{F}_p$ = **F_2**, at t = [0.84-1.048 secs], $\mathcal{F}_c$ = $\{f_{md}^{z1\_DR^{10\_11}}, f_{md}^{z2\_DR^{10\_11}}, f_{md}^{z3\_DR^{10\_11}}\}$, $\mathcal{C}$ = {z1_DR^{11_10}, z2_DR^{11_10}, z3_DR^{11_10}, z2_DR^{6_11}, z3_DR^{6_11}, z3_DR^{9_10}}, $\mathcal{E}$ = $\varnothing$, $\mathcal{M}$ = {z1_DR^{10_11}, z2_DR^{10_11}, z3_DR^{10_11}}, $\mathcal{I}$ = $\varnothing$ **(Ground truth)** ID = H2, Rank = **1** $\mathcal{F}_p$ = **F_1**, at t = [0.84-1.048 secs], $\mathcal{F}_c$ = $\{f_{md}^{z2\_DR^{10\_11}}, f_{md}^{z3\_DR^{10\_11}}\}$, $\mathcal{C}$ = {z1_DR^{11_10}, z2_DR^{11_10}, z3_DR^{11_10}, z2_DR^{6_11}, z3_DR^{6_11}, z3_DR^{9_10}}, $\mathcal{E}$ = $\varnothing$, $\mathcal{M}$ = {z2_DR^{10_11}, z3_DR^{10_11}}, $\mathcal{I}$ = $\varnothing$
Spurious detection in zone 3 element of DR^{11_10} at t = 1 sec	z3_DR^{11_10} logged at t = 1.016 secs	ID = H0, Rank = **1**, $\mathcal{F}_p$ = $\varnothing$, $\mathcal{F}_c$ = $\{f_{z3\_DR^{11\_10}}^{sd}\}$, $\mathcal{C}$ = $\varnothing$, $\mathcal{E}$ = $\varnothing$, $\mathcal{M}$ = $\varnothing$, $\mathcal{I}$ = z3_DR^{11_10} **(Ground truth)**

reasoner for two scenarios, wherein first scenario a fault is injected in line TL^{11_10} at a distance 40% from the bus 11 with a missed detection in all zone elements of relay DR^{10_11}. The second scenario shows a reasoner response when zone 3 element of relay DR^{11_10} incorrectly detects a fault condition. As illustrated in the last column of Table IV, the highest-ranked hypothesis of the TCD reasoner correctly identifies the root cause(s) of the observed alarms in the second scenario. However, the second-best hypothesis accurately diagnoses physical fault and missed detection in all zone elements of DR^{10_11} in the first scenario.

IV. Summary

A flexible and integrated cyber-power testbed is needed for the validation of advanced tools to enable resiliency. A cyber-power testbed is also essential for training grid engineers and researchers to adapt to new conditions and effectively use new technologies. This work describes the design and implementation of a multi-layer cyber-power testbed incorporating real-time simulation and realistic models of various measurements, devices, and networks. Our testbed is capable of storing the large volumes of data required by advanced tools at the control center layer. The incorporation of firewalls according to a defense-in-depth approach creates a realistic cyber context, providing an opportunity to explore different security mechanisms with rule-based access policies for untrusted and trusted traffic. This range of features also allows the testbed to be used to understand and quantify cyber-physical resilience.

In the future, we will continue to use this testbed to (i) model cyberattack scenarios in different layers, capturing network data under normal and attack scenarios for use in attack detection tools, (ii) integrate a human-machine interface to visualize, (iii) enable remote access to this testbed to be used by a wide range of communities for learning and research purposes, (iv) Develop and validate additional number of advanced tools.

References

[1] M. H. Cintuglu, O. A. Mohammed, K. Akkaya, and A. S. Uluagac, "A survey on smart grid cyber-physical system testbeds," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 446–464, 2016.

[2] Y. Yan, Y. Qian, H. Sharif, and D. Tipper, "A survey on cyber security for smart grid communications," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 4, pp. 998–1010, 2012.

[3] J. Liu, Y. Xiao, S. Li, W. Liang, and C. P. Chen, "Cyber security and privacy issues in smart grids," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 4, pp. 981–997, 2012.

[4] Tushar, V. Venkataramanan, A. Srivastava, and A. Hahn, "CP-TRAM: Cyber-Physical Transmission Resiliency Assessment Metric," *IEEE Transactions on Smart Grid*, pp. 1–1, 2020.

[5] P. S. Sarker, A. S. Saini, K. Sajan, and A. K. Srivastava, "Cp-sam: Cyber-power security assessment and resiliency analysis tool for distribution system," in *2020 Resilience Week (RWS)*. IEEE, 2020, pp. 188–193.

[6] "Measurement challenges and opportunities for developing smart grid testbeds," Available at https://www.nist.gov/system/files/documents/smartgrid/SG-Testbed-Workshop-Report-FINAL-12-8-2014.pdf, December, 2014, summary report by NIST.

[7] V. Venkataramanan, P. S. Sarker, K. S. Sajan, A. Srivastava, and A. Hahn, "Real-time federated cyber-transmission-distribution testbed architecture for the resiliency analysis," *IEEE Transactions on Industry Applications*, vol. 56, no. 6, pp. 7121–7131, 2020.

[8] T. Morris, A. Srivastava, B. Reaves, W. Gao, K. Pavurapu, and R. Reddi, "A control system testbed to validate critical infrastructure protection concepts," *International Journal of Critical Infrastructure Protection*, vol. 4, no. 2, pp. 88–103, 2011.

[9] A. Hahn, A. Ashok, S. Sridhar, and M. Govindarasu, "Cyber-physical security testbeds: Architecture, application, and evaluation for smart grid," *IEEE Transactions on Smart Grid*, pp. 847–855, 2013.

[10] J. Hong, S. Wu, A. Stefanov, A. Fshosha, C. Liu, P. Gladyshev, and M. Govindarasu, "An intrusion and defense testbed in a cyber-power system environment," in *2011 IEEE Power and Energy Society General Meeting*, 2011, pp. 1–5.

[11] "Real-Time simulation | Real-Time Solutions | OPAL-RT," Opal-RT Technologies, Montreal, QC, Canada, H3K 1G6. [Online]. Available: https://www.opal-rt.com/

[12] M. P. Andersen, S. Kumar, C. Brooks, A. von Meier, and D. E. Culler, "Distil: Design and implementation of a scalable synchrophasor data processing system," in *2015 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. IEEE, 2015, pp. 271–277.

[13] Anshuman, K. S. Sajan, and A. K. Srivastava, "ML-based data anomaly mitigation and cyber-power transmission resiliency analysis," in *2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, 2020.

[14] K. Sajan, M. Bariya, S. Basak, A. Srivastava, A. Dubey, A. von Meier, and G. Biswas, "Realistic synchrophasor data generation for anomaly detection and event classification," in *2020 8th Workshop on Modeling and Simulation of Cyber-Physical Energy Systems*. IEEE, 2020.

[15] M. Bariya, J. Paparrizos, A. von Meier, and M. J. Franklin, "k-shapestream: Probabilistic streaming clustering for electric grid events *(forthcoming)*," in *2021 POWERTECH Conference*. IEEE.

[16] P. Van Mieghem, K. Devriendt, and H. Cetinay, "Pseudoinverse of the laplacian and best spreader node in a network," *Physical Review E*, vol. 96, no. 3, p. 032311, 2017.

[17] M. P. Andersen and D. E. Culler, "Btrdb: Optimizing storage system design for timeseries processing," in *14th {USENIX} Conference on File and Storage Technologies ({FAST} 16)*, 2016, pp. 39–52.

[18] A. Chhokra, N. Mahadevan, A. Dubey, and G. Karsai, "Qualitative fault modeling in safety critical cyber physical systems," in *Proceedings of the 12th System Analysis and Modelling Conference*, ser. SAM '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 128–137. [Online]. Available: https://doi.org/10.1145/3419804.3420273

[19] S. Abdelwahed, G. Karsai, N. Mahadevan, and S. C. Ofsthun, "Practical implementation of diagnosis systems using timed failure propagation graph models," *IEEE Transactions on instrumentation and measurement*, vol. 58, no. 2, pp. 240–247, 2008.

[20] "SEL-421 Protection, Automation, and Control System." [Online]. Available: https://selinc.com/products/421/

[21] A. Chhokra, A. Dubey, N. Mahadevan, S. Hasan, and G. Karsai, *Diagnosis in Cyber-Physical Systems with Fault Protection Assemblies*. Cham: Springer International Publishing, 2018, pp. 201–225. [Online]. Available: https://doi.org/10.1007/978-3-319-74962-4_8