Energy and Power Engineering, 2017, 9, 723-749
http://www.scirp.org/journal/epe
ISSN Online: 1947-3818
ISSN Print: 1949-243X

Considerations for Planning a Multi-Platform Energy Utility System

Yahav Biran[1], Joel Dubow[2], Sudeep Pasricha[1,3], George Collins[1], John M. Borky[1]

[1]Department of Systems Engineering, Colorado State University, Fort Collins, CO, USA
[2]Fulcrum Co., Centreville, VA, USA
[3]Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO, USA
Email: ybiran@colostate.edu, jdubow@fulcrumit.com, sudeep@colostate.edu,
gcollins@colostate.edu, mborky@colostate.edu

How to cite this paper: Biran, Y., Dubow, J., Pasricha, S., Collins, G. and Borky, J.M. (2017) Considerations for Planning a Multi-Platform Energy Utility System. *Energy and Power Engineering*, **9**, 723-749.
https://doi.org/10.4236/epe.2017.912046

Received: August 30, 2017
Accepted: November 5, 2017
Published: November 9, 2017

Abstract

A federated cloud-based multi-platform Power System is presented to meet the growing challenges confronting power system operators. It uses a federated architecture to provide a group sourced increase in cyber security, in reducing the need for computing resource overcapacity, for sharing computing and power resources during emergencies, for minimizing energy costs, and for sharing information on threats and incident responses. In the face of nation-state and organized crime complex, multi-technology, coordinated attacks, a single organization stands an ever reducing chance of remaining safe. The proposed federated cloud preserves the economic efficiency advantages of marketplace of non-monopolistic organizations innovating to obtain competitive advantage with shared preparation, resources, information, and resiliency enabled by individual Power System cloud-based computing creating a federated System. The paper applies earlier advances. This paper combines the results previously published in different publications and applies them to a single paradigmatic example of a power system consisting of a number of individual asset owners. It includes the architecture, model of energy, and resource sharing as well as a novel, self-learning, semantic-less breach detection system for detecting anomalous behavior in resource usage across the power system participants. The paper extends previous work published about federated cloud. The simulations results provided to demonstrate the usefulness of the proposed system.

Keywords

Outlier Sensing, Metamorphic Malicious Software,
System Protection

1. Introduction

Introduction Energy sector organizations, or market segments of the energy sector, that serve and process a large number of simultaneous users and large quantities of data require "cloud computing services" as part of their information technology process. These services enable convenient, on-demand network access to a shared pool of configurable computing resources. Cloud computing provides a rapidly growing share of industry-wide IT resources. IT related spending toward workload processing increased 32.8 percent in the year 2015, 29 percent in 2016 and 29 percent during 2017[1]. The Energy sector is no exception to this trend. Energy-based companies are using more IT to gather and process data for their business processes. The energy sector faces dynamic load patterns and generation and distribution processes that are complex and changeable at multiple rates. This, in turn, requires broadband, decentralized processing, digital storage and data management.

As an increasing fraction of computing services move to the cloud, there will be a proliferation of software characteristics, service models, and deployment options. Many organizations are moving towards hybrid cloud/hosted computing models. Energy sector organizations are most interested in availability, adaptability, and security. Availability refers to reliable service conditions that make energy related services available to the users it serves. Adaptability relates to the Vendor lock-in risk [1]. Security covers the risk energy sector systems exposed while hosted in the cloud or the organization's premise. The security risk applies to the likelihood of service interruption caused by malicious intent. The single Cloud Service Provider (CSP) model provides a sub-optimal solution for the electric utility or electric power holding companies from adaptability, availability, and cyber security perspectives. This paper applies to Power Systems the non-specific results developed in earlier publications [2] [3].

In addition, upsets in generation and distribution can more likely exhibit instabilities if multiple clouds are interconnected via the internet but not coordinated to share critical information and resources to compensate for outages or unplanned loads. To exchange information, computing resources and enable load sharing, multi-cloud architectures become an attractive solution. However, a multi-cloud solution implemented by a single electric utility or electric power holding company won't provide the resilience and adaptability that multi-organization cloud services provide. Also, a single organization managing multiple cloud instances requires expensive adaptations to each CSP's tools and service constructs that may vary among different CSPs. A federated cloud allows multiple CSPs to scale and optimize cross-datacenter deployments and cross-regional deployments by suggesting a cross-cloud provider's resource sharing collaboration via a cloud aggregator. Furthermore, it removes the lock-in risk and minimizes the security risk by spreading the total risk among

[1] Gartner Says Worldwide Cloud Infrastructure-as-a-Service Spending to Grow $246.8 billion in 2017.

the CSPs' platforms.

A basic goal of an energy sector CSP organization is to maximize its market share among the energy service provider (ESP). Accommodating variable demands for computing resources requires an immense capacity, as it calls for providing for the maximum demand. In some cases, this drives CSP's to underutilize massive datacenter deployments. In other situations, the CSPs suffer over-utilization because of a misestimate of the market share, load, and reliability projections. These cases lead to suboptimal utilization and suboptimal revenue projection.

Cyber security is essential for successfully operating a business. ESPs, from government agencies to private companies, rely on information technology to perform critical and essential functions. All organizations face cyber risks today. Varied threat actors can exploit vulnerabilities in software, hardware, or processes, leading to significant and occasionally catastrophic consequences for an enterprise, industry, region, or even an entire country. Cyber risks are increasing in number and scope and are constantly evolving. Every day, malicious actors attack, disrupt, or steal sensitive information assets or processes from public and private sector computer networks, causing significant damage to business and government. Though quantifying this problem precisely is impossible, experts have estimated that the aggregate economic impact from cyber-attacks may range from $300 billion to $1 trillion globally when totaling direct and indirect costs from actual money stolen, the value of intellectual property and business secrets stolen, the cost of downtime caused by malicious disruption, and recovery expenses for repairing or replacing damaged networks and equipment [4].

What has become increasingly clear is that preventing cyber-attacks is not possible. What needs to be done, after defending information resources, is managing risks and the impacts of successful cyber-attacks. This, in turn, requires organizations to mitigate as many vulnerabilities as possible, observe indicators of attack, deploy resources to mitigate impacts and pursue attackers, and adapt the system to deter future attacks. To make matters more critical and complex, attacks themselves have become more complex in multi-platform environments. Whereas older cyber-attacks originated and terminated on information systems, newer attacks originate from information systems, cyber physical systems, or physical security systems and terminate in various permutations and combinations of each type of inter-networked computing environment. In this paper we describe a specific use case of the polymorphic malware detection system described earlier [2] [3].

In the face of coordinated attacks from nation-states and organized crime complexes, a single organization stands little chance of remaining safe. Yet economically, the best performance is with market places of non-monopolistic organizations innovating to obtain competitive advantage. Thus the key challenge of the coming decade is to maintain competitive environments while pooling cyber defense resources and developing a capability to be more agile than attackers. That is the purpose of this paper and of earlier papers in our body of work [2] [3] [5] [6] [7] [8]. A proposed solution to help meet these needs is the federation of cloud computing resources amongst organizations with common interests or which operate in a common market segment. The federation provides shared services for smoothing energy load variations, computing load variations, and security-based resource optimization.

2. Federation of Multiple Cloud Services

Cloud federation is an emerging new paradigm that interconnects the cloud computing environments of two or more CSPs for the purpose of load balancing traffic and accommodating spikes in demand. The approach allows numerous cloud service providers to use computing resources optimally [3]. Also, it allows ESP to avoid the CSP lock-in risk and deliver service availability that can not be provided by a single CSP. No matter what the architecture, there is a need to ensure security and information assurance to users, to manage energy costs, and to share computing capacity to smooth loads and provide mutual backup. [3] specifies the selection principals ESPs needs. Below we name few: 1) Availability: Reliable service conditions that make its services available to the users it serves. Reliability is defined by SLA and is measured by the allowed unavailability, aka, downtime. Such measurements are done with the help of independent third party service .e.g., Gartner's CloudHarmony[2].

2) Latency: Some of the SPs workloads are sensitive to network latency, which is defined by the time the service takes to respond to a user or other sub-system request. In the case of a service that runs in a geo-location, which is different from that of the end-user or other sub-systems, the network latency can impact the overall service performance. Therefore, SPs who run latency-sensitive workloads prefer to provide their service by maintaining optimal proximity to its end-users or sub-systems in which SPs' workloads interoperate with.

3) Adaptability: The Vendor lock-in risk is one of the core business risks that every enterprise, who wishes to offload its workloads to the cloud, faces. Current IT practices rely on common standards and protocols that allow organizations to switch components or elements in their IT operations. However, cloud computing disrupts most of these practices. Onboarding into a single CSP introduces a risk vector that locks the SP to use the CSP's platform, API's and tools. Adopting a single CSP requires an operational adaptation to CSP's methods. New needs on the SP side or changes in the CSP service terms might sub-optimize the operations of the SP. A Federated Cloud removes that risk vector by creating a CSP agnostic apparatus that allows the SP to adapt when the vendor lock-in plays a critical role in the migration decision.

Cloud Federation is an advantageous structure for aggregating cloud based services under a single umbrella. It is designed to share resources and responsibilities for the benefit of member cloud service providers. [3] [6] discusses the

[2] Research and compare cloud providers and services https://cloudharmony.com/status

required framework of managing multiple cloud providers. Federation is useful not only for sharing resources amongst cloud service providers but also for providing enclaves for performing domain-specific missions such as management of electrical grids and supply chains [2].

Some responsibilities of an effective federation include assuring that data transfers amongst the federation's CSPs are secure. The Federation will, above all, need to detect any anomalous behavior occurring in transactions and resource sharing. In addition to the growing number of security tools, there is a need to log and identify security issues requiring attention early on in the process. In particular, breach detection in inter-cloud data transfer and communications is a particularly serious security issue because of the possibility of an attacker potentially gaining access to more than one CSP federation member [2].

As an example of the benefits of federated cloud, we propose an energy-sector cloud federation that enables energy organizations' IT workloads to operate across numerous CSPs. The federation will enable an energy-based common dataset that can be used for breach detection and provide other benefits described below.

The federation is comprised of a multi-cloud sector-based enclave within a CSP network that interconnects with other enclaves hosted in other CSPs. The core idea is not bounded to any sector and is equally valid for any industry sector. This paper focuses on the energy sector and its computing needs. The cloud computing performance patterns analyzed by [3] [6]. The analysis includes the federation behavior that might emerge out of the performance patterns. e.g., pricing, datacenter carbon footprint.

The reminder of the paper is organized as follows, Section 3 (Terminology) discusses the terminology used in this paper; Section 4 (Energy Sector Threat Landscape) summarizes the main threats the energy sector is currently facing; Section 5 (Why Energy-based Federated Cloud?) describes the solution federated cloud provides to the energy sector threats; Section 6 (Cyber Security Challenges in Federated Cloud) discusses the operational and Infrastructure security challenges of federated cloud; Section 7 (Semantic-less Breach Detection) discusses the tool we suggest for detecting the anomalous behavior of the systems that run and connect within the Cloud Federation; Section 8 (Evaluation) discusses the breach-detection tool prototyped and presents the prototype results; finally, Section 9 (Conclusions) summarizes the main results and the original research question discussed herein.

3. Terminology

The following section briefly defines the important terms used in this paper.

ESP (Energy-Service-Provider): A utility, grid operator or power plant, usually an organization with end-consumers who require processing of IT workloads. We will use the terms Service Provider (SP) and Energy Service Provider (ESP) interchangeably throughout the paper.

CSP (Cloud-Service-Provider): One of the System of Systems constituent that offers computing resources, digital storage and network bandwidth to its customers and the cloud federation to process its workloads. Also, it provides the software that provisions and manages cloud services.

Public cloud. It offers computing resources, e.g. broadband network, computing, storage, and infrastructure applications over the public Internet. The organization, that chooses to run their workloads on the public cloud is considered as a cloud tenant. Public cloud providers adhere to generic service level agreements for service availability and security.

Private cloud: It can include public cloud offerings, excluding the multi-tenancy property. However, multi-tenancy can be implemented within the enterprise that operates the private cloud. Therefore, the implementation might be customized to adhere to specific enterprise needs.

Hybrid cloud: The hybrid cloud aggregates several public and private clouds to run heterogeneous workloads that might span across different geographical locations and enterprises.

IT Workload. These are the organization's IT needs to serve internal, external, and users' IT services and data. Cloud workloads are broadly of two types: online and offline. The former provides low-latency, read/write access to data. For example, a web user requests a web page to load online and serve within a fraction of a second. The latter provides batch-like computing for tasks that process the data offline, which is reported later to users by the systems servers; for example, the search results based on a pre-calculated index. Production offline workloads usually comprise mainly unstructured data sets, such as click stream, web graph, and sensors data. The service level objectives (SLO) for online jobs span a fraction of a second, and those for offline job goals span hours, days and, sometimes weeks.

OT (Operational Technology): These are the workloads generated by the cyber physical systems such as SCADA systems in the control rooms of power companies. They are systems that run on specific functionality that does not include an operating system even though they can communicate with the network. Unlike IT workloads, OT workloads runs on the organization premises and not in the cloud or other remote facility. Also, OT workloads omit operational logs that can be used by other OT workloads or pushed to other systems and turn into IT workloads.

ICS (Industrial control system): general term for OT systems used in manufacturing for system that control energy, water, and other critical commodities.

Control Plane: This is the software that automatically controls the operations of software-based systems. It is a rule-based system that accepts signals from various system components and acts, based on a pre-defined policy.

SLA (Service Level Agreement): This is an agreement between the Cloud-Service Provider and its customers, the Service-Providers. It often includes guaranteed levels of availability, network latency, and numerous other provisions.

4. Energy Sector Threat Landscape

The IT workloads and Operational Technology (OT) systems in the energy sector are different enough from those of other infrastructure sectors that a separate approach to forensics is justified. Cyber forensic technologies are advancing rapidly, especially for IT systems. Section 5 explores various challenges and solutions the cloud federation offers.

OT forensics, and specifically, energy sector cyber technologies, are not advancing as rapidly as IT systems and networks. A cyber-security threat-aware system that keeps tracks of the control system configuration and behavior and that shares security information with similar systems will reduce reaction time and increase the agility if resource sharing in crisis will enhance systematic resiliency against cyber threats. Section 7 introduces a method for enabling high level of security for combined IT and OT workloads.

The following sections will describe various threat actors and survey recent major cyber-attacks against the energy sector. The survey will be used to build a system model that implements an energy centric Cloud Federation model and proposes an approach for managing multi-threat cyber-attacks.

4.1. Threat Actors

Threat actors are motivated to launch attacks for a variety of reasons: 1) Espionage, including governments stealing information to gain geopolitical advantage and governments or government-sanctioned actors stealing information to give businesses competitive advantage. 2) War, in which government actors cause severe destruction or disruption as part of an international conflict. 3) Disruption, including attacks by non-government actors or governments pursuing hostile actions that cause serious damage but that do not rise to the level of war. 4) Hacktivism, in which individuals or groups cause disruption or repetitional damage for ideological or political reasons. 5) Crime, in which individuals, small groups, or large criminal organizations commit financial theft or steal information to sell in the dark web.

Cyber-attacks against companies can happen in three primary ways. The first is through inside access, including insiders such as Edward Snowden or physical network access. A second way is remote penetration, such as through spear-phishing emails. The third method is by exploiting the IT/OT supply chain, as in the case of the malware dubbed “Zombie Zero”, which was deployed against the shipping and logistics industry across the globe.

In the “Zombie Zero” episode, researchers discovered that hand-held terminal scanners used for inventory at ports were being shipped from their Chinese manufacturer with malware embedded alongside their OT operating system. The malware looked for financial information being processed by the hand-held scanner and exfiltrated that data. The IT supply chain path can also be used via software, without involvement by a hardware manufacturer[3].

[3] “Windows OS Hacked, Supply Chain Poisoned”, Tech News World, http://www.technewsworld.com/story/80742.html

Within these three methods, attackers use many techniques: stealing passwords and identifying information to compromise legitimate users' access; exploiting vulnerabilities in Internet-facing websites or applications; utilizing insiders; using backdoors in hardware; intercepting Internet communications through "man-in-the-middle" attacks; or entering a company's corporate or industrial control system networks through software patches and vendor updates. In many campaigns and attacks, attackers combine techniques to gain entry undetected and to maintain persistent access and control within a target's networks.

As companies assess dynamic cyber risks today and for years to come, it is critical to note that cyber capabilities proliferate, giving previously less-sophisticated actors such as non-state groups access to more advanced attack skills and techniques: "Whatever a nation state uses today, organized crime will use tomorrow and hacktivists will use the day after that." In fact, the emerging threat space is characterized by sharing and cooperation between different cyber-attackers. Alexander Klimburg, a security researcher with Harvard's Kennedy School of Government and The Hague Centre for Strategic Studies, writes that "Cybercrime, cyberterrorism, and cyberwarfare share a common technological basis, tools, logistics and operational methods." This is particularly true of tools and techniques for delivering malware. For example, hackers who research "zero days", meaning vulnerabilities in code that have not yet been discovered or disclosed and therefore can be exploited before they are "patched", can sell their zero days on a black market for as much as $200,000 or, as in the case of hackers, "lending their zero-day hacks to the government for espionage purposes, then using them for crime later."

4.2. Canonical Examples of Cyber-Attacks against the Energy Sector

The energy sector has been an important target for cyber-attacks around the world. Intelligence from news media, various threat reports, U.S. Government organizations such as Industrial Controls Systems Computer Emergency Response Team (ICS-CERT), and other sources confirm the rising sophistication of malware and a growing interest not only against corporate networks but also against industrial control systems and energy critical infrastructure in particular.

Based on incidents reported to the Department of Homeland Security (DHS), the energy sector led all other sectors in 2014 in the number of reported cyber-attacks. Attacks reported outside the energy sector, for example in other critical infrastructure, are also of interest as some were targeted at ICS equipment manufacturers, showing increasing malicious activity in this space. The ICS vendor community may be a target for sophisticated threat actors for a variety of reasons, including espionage and reconnaissance to prepare for possible sabotage. Of the attacks reported, roughly 55 percent involved Advanced Persistent Threat (APT) indicating sophisticated actors sources. Other actor types included hacktivists, insider threats, and criminals[4].

Among the most widespread and renowned espionage campaigns targeted against the energy sector is the "Dragonfly" campaign (also called the "Energetic Bear" campaign), in which data from thousands of energy companies in the United States and Europe has been compromised in an ongoing cyber espionage campaign. The target list includes various electricity generation companies, petroleum suppliers, and industrial energy equipment providers across the United States, France, Italy, Germany, Spain, Poland, and Turkey. The campaign is being carried out allegedly by an Eastern European hacker group called Dragonfly. The cyber security company Symantec noted that the primary goal of the campaign was espionage and that it "bears the hallmarks of a state-sponsored operation, displaying a high degree of technical capability", signaling possible cooperation between governments and sanctioned non-government attackers. While the Energetic Bear campaign is among the most widespread, it is only one of many campaigns observed in recent years.

One alarming trend in cyber threats against the energy sector is that they are moving toward disruption or even destruction, in addition to espionage and theft. This trend was recently confirmed in the cyber-attack against the Ukrainian electric grid [9]. In the "Shamoon" attack, Saudi Arabian state-run oil giant Saudi Aramco came under a targeted and advanced attack from a hacktivist group known as the Cutting Sword of Justice, allegedly because of the company's role as a financial hub for the Saudi regime. The attackers used "wiper" malware which was also used in the Ukrainian attack, which renders computers useless, to disable 30,000 OT workstations and disrupt the internal operations and workstations of Saudi Aramco for days [10]. Even more alarming, the ultimate target of the attack was not the corporate network but the ICS systems that control production and distribution of oil and gas. Abdullah al-Saadan, Aramco's vice president for corporate planning, stated that the attackers sought to "stop the flow of oil and gas to local and international markets." Fortunately, the malware contained an error and was unable to spread from the corporate network to the production ICS network.

"Wiper" malware, like that used in the Shamoon attack, is one form of disruptive attack. Another is ransomware, in which malware uses the attacker's secret encryption algorithm to encrypt data on the target network and render it inaccessible by the target company, unless or until the company pays a ransom. The value of the ransom demanded can vary widely depending on the attacker and the target. Ransomware increased in frequency, complexity, and geographic spread in 2014 and 2015, with two major campaigns, CBT-Locker and TorrentLocker, currently affecting thousands of users around the world. A third kind of disruptive attack is the Distributed Denial of Service (DDOS) attack in which attackers disrupt a target's Internet connectivity or public profile. DDOS attacks were for some time considered a solved problem that was adequately mitigated

[4] https://www.dhs.gov/topic/protecting-critical-infrastructure

by existing defensive technologies. However, a new generation of high-volume DDOS attacks have made earlier mitigation measures inadequate to prevent disruption, prompting further development of defensive technologies.

In addition, the energy sector has experienced several "cyber-physical" attacks in which cyber methods are used to cause "real-world" physical damage to ICS systems. Bloomberg News recently reported an attack on a major oil pipeline in Turkey in which alleged cyber-attackers caused an explosion. Also recently, the German Federal Office for Information Security (BSI) published information on an attack against a steel facility that caused "massive damage", though more details have not been forthcoming[5]. These incidents, as well as the alleged U.S.-Israel Stuxnet attack that disrupted Iranian nuclear operations, highlight the significant risk for the energy sector of cyber-attacks causing physical damage [11].

The threat of operational disruption is difficult to assess and prevent because the line between espionage and disruption is blurred. Although the majority of cyber intrusions in the energy sector to date can be characterized as espionage, the information being exfiltrated could easily be used to enable disruption or even sabotage at some time in the future. Espionage attacks also allow attackers to maintain a presence inside a corporate or production network that could be used to cause disruption or sabotage at a later time, for example in the event of an international conflict, as suggested by Admiral Michael Rogers, Director of the National Security Agency (NSA). In highlighting the risk of disruption or destruction, Eugene Kaspersky, the Moscow-born founder of security research company Kaspersky Labs, noted that there has been a dramatic surge in targeted attacks against power grids, banks, and transportation networks around the world and warned that groups targeting crucial infrastructure have "the capacity to inflict very visible damage. The worst terrorist attacks are not expected".

5. Why an Energy-Based Federated Cloud?

One of the main organizational failures in the enumerated attacks is the lack of cooperation amongst energy providers. The attacks' precursors and potential impact failed to provide timely warning and options for incident response throughout the existing tools, techniques and staff that were available to the energy organizations. Moreover, even if these data were available it lacked the attack context, and preparation, and thus the cyber attack detection was delayed, often resulting in significant damage to the organization. This paper describes how the federated cloud will provide sharing of security data, energy supply loads, information processing and forensic response amongst the Federation member organizations while enabling them to maintain their independence and corporate critical business information. This latter consideration is currently a main barrier for organizational cyber security collaboration.

Federated clouds represent a new paradigm that allows multiple cloud pro-

[5] www.securityweek.com/cyberattack-german-steel-plant-causes-significant-damage-report

viders (energy corporate private, public, or hybrid) to optimally utilize computing, energy, cyber security, incident response and forensic resources. It will also increase productivity by reducing operating costs. For example, in terms of energy use, which represents over 20% of cloud provider expenses, it does this by, 1) lowering the Federation members' data center's deployments per provider ratio, by sharing energy, and 2) scheduling available energy via aggregators and 3) employing, where appropriate, more renewable and carbon-free energies. The earlier paper provides a more in-depth model of combining renewable and baseline energy to save costs and reduce non-renewable energy consumption in cloud provider data centers [3].

The federated cloud utilizes a software containerized software instead of virtualized software. The goal of containerized software is to obtain resource density and and allocation elasticity. It uses Linux-based Containers orchestrated by a Kubernetes[6] resource management system. The Kubernetes system governance the job scheduling and resource allocation [12]. Furthermore, the proposed solution scales out and optimizes cross-datacenters and cross-regional deployments by computing and suggesting useful cross-cloud provider's collaboration via the Federation cloud aggregator. Furthermore, it suggests operating data centers employing maximum intermittent green energy sources [3] [5] [7]. Yet all of these advantages will be for naught if federation services cannot be supplied securely in the face of growing sophistication and quantities of cyber security threats. The baseline metric is that the losses arising from cyber breaches are substantially less than the value of services provided. This issue is detailed in an earlier paper from our group [2].

Prototype Federated Cloud Architecture

The Cloud Federation architecture is comprised of multiple CSPs, a Cloud-Coordinator, and a Cloud-Broker system. These are defined below.

Cloud-Coordinator. Acts as an information registry that stores the CSPs dynamic pricing offers and demand patterns. Clouds Coordinators periodically update the CSPs availability and offering prices. Also, a Cloud-Coordinator will help integrate, where appropriate, more renewable and carbon-free energies [7].

Cloud-Broker. Manages the membership of the CSP constituents. Both CSPs and ESPs will use the Cloud-Broker to add new cloud federation members. Also, the Cloud-Broker will act on behalf of the individual ESP for resource allocation and provisioning requests. Cloud-Broker also provides a continuous ability to deploy SP software, configuration, and data to one or more member CSPs. In this manner it will provide a multi-utility power system with agility, resilience, capacity smoothing and financial efficiencies.

6. Cyber Security Challenges in Federated Cloud

The Cloud Federation has a global scale software and hardware infrastructure.

[6] http://kubernetes.io

We describe a progressive layers security model starting from the physical security of datacenters (part of the networked triad of physical, ICS and IT systems), progressing to the hardware and software that underlies the infrastructure. The model also incorporates the constraints and processes to support the cloud federation operational security. The following section describes the cloud federation cyber security design throughout the data processing life cycle of the cloud federation to enable secure communication with tenants (SP) and their customers or control plane communication including CSP, Cloud-Brokers, and Cloud-Coordinator.

Figure 1 describes the cyber security layers offered by the cloud federation. The following paragraph briefly describes the security elements corresponding to each layer[7]. [2] extended the cyber security model and emphasizes the operational security employing our proposed novel breach detection methodology. This was done since the operational security corresponds to the perimeter security of an enterprise system and the interface to the Federation members. It is readily configured to handle multi-technology and digital domain threats such as those to IT, IoT, and physical technology system. Also, it will suggest a systems for encryption of inter cloud provider micro-services communication, with emphasis on cross-CSPs (Power System Federation members) for tenants' workloads.

6.1. Infrastructure Security

The required baseline security level needed for cloud federation constituent's systems is referenced in **Figure 2**. It includes deployed facilities and computer systems managed by the individual CSPs or the Federation. The larger CSP's often exceed these baselines.

Datacenter Premises. CSPs design and build their datacenters based on their expected computing capacities and service reliability defined by their SLA and the resulting redundancy levels of sub systems needed to ensure reliability [13] [14]. The datacenter incorporates various components for physical security protection. These, in turn rely on networked alarm systems and indicators of power system malfunctions that are serious enough to require attention. Access to such facilities is governed by the CSP security operations. It uses technologies such as biometric identification, metal detection, metal detectors, and CCTV solutions [15].

Hardware Design. CSPs data centers run computing server machines fed by local power distribution units and connected to a local network that is connected to the edge of the wider network. The computing, digital storage, and networking equipment need to comply with standards that ensures the required audit and validation of the regulatory and industry security requirements [16] [17], e.g., hardware security chip [18].

[7] We extrapolate Google Cloud security model from https://cloud.google.com/security/security-design/

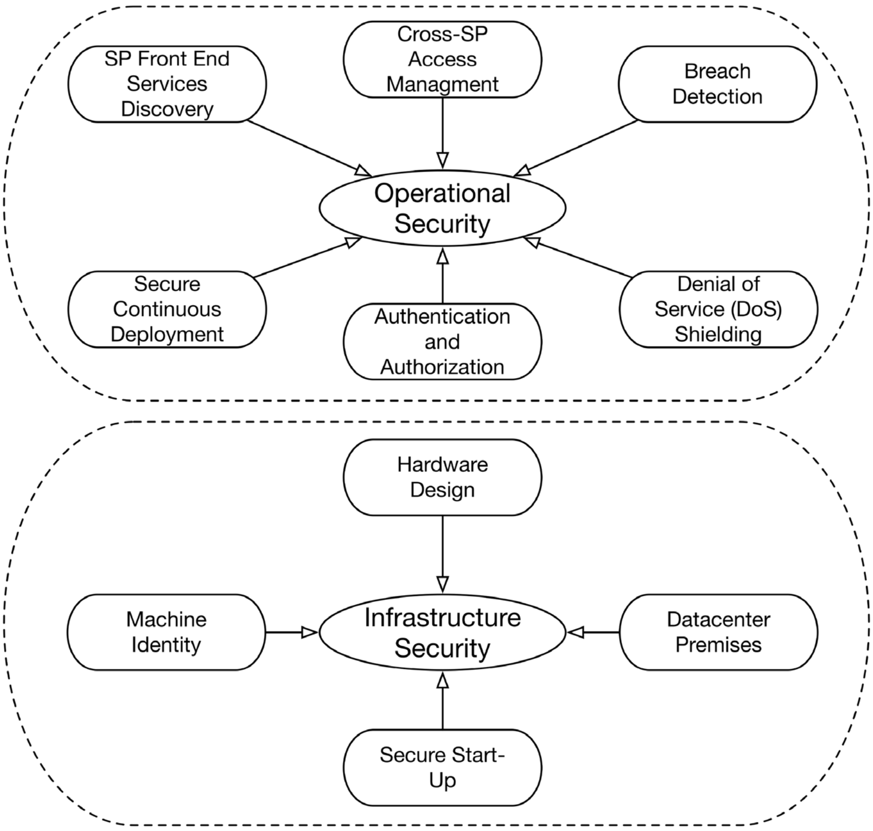
SP Front End Services Discovery
Cross-SP Access Managment
Breach Detection
Operational Security
Secure Continuous Deployment
Authentication and Authorization
Denial of Service (DoS) Shielding
Hardware Design
Machine Identity
Infrastructure Security
Datacenter Premises
Secure Start-Up

Figure 1. Federation of Power System CSP's security design. Comprises of infrastructure security and operational security layers.

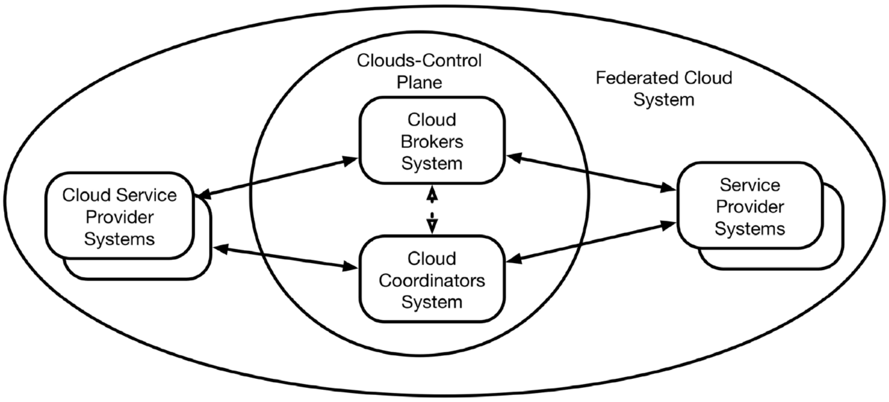

Figure 2. Proposed Federation of Power System CSP's, dedicated as s SoS includes cloud service providers and energy service providers. The Clouds-Control plane Clouds-Broker and Clouds-Coordinator.

Machine Identity. This critical configuration repository confirms that any participating computing server in the cloud federation can be authenticated to its CSP machine pool through low-level management services [18]

Secure Start-Up. Ensures that CSPs servers are booting the correct software stack. Securing underlining components such as Linux boot loaders, OS system images and BIOS by cryptographic signatures can prevent an already compromised server from being continuously compromised by an ephemeral malware or other non-malware attack vector.

6.2. Operational Security

Operational security comprises the business flows between the SP with the cloud federation and the CSP it uses for processing workloads. The following section briefly discusses the required cybersecurity measures needed for SP and CSP business scenarios in a cloud federation.

Cross-SP Access Management: SP workloads are manifested in two workload types, 1) short-lived workload. *i.e.*, jobs that are terminated upon completion, and 2) long-lived workload. *i.e.* services. The former workload might require connectivity to external services during its processing. The latter might expose serving endpoints to other services, e.g., short-lived jobs might require persistent storage to write their job results hence connecting to BigTable[8] storage server provisioned by other CSPs, which, in turn, require access management that uses credentials and certificates stored centrally within the Cloud Federation.

SP Front End Service Discovery: Long-lived workloads might expose public facing endpoints for serving other workloads or end-user requests. SP front-end services require publishing endpoints to allow other workloads within or external to the cloud Federation to discover their public facing entry point and this requires service discovery capabilities. Service discovery endpoints, and the actual service endpoints, are prone to risks such as Denial of Service attacks or intrusions originated by an attacker. We argue that current solutions offered by individual CSP's are sub-optimal because the target scope of the intrusion doesn't limit propagation or allow an active incident detection and response strategy. For example, assuming an attack probability for a given CSP, running several CSPs reduces the attack impact by a factor of the number of CSPs. Section 7 formulates the risk function and shows how a cloud-federation minimizes those attack impacts by using the semantic-less breach detection system and show how the most serious impacts originate by crossing machine boundaries.

Secure Continuous Deployment: Continuous Deployment (CD) is the function that allows cloud-native applications to get updated through an automated pipeline that is initiated by a new or updated code submission that is compiled, tested through various quality gates until it is certified for deployment of the production systems and deployed seamlessly. Continuous deployment enables cloud applications to innovate faster and safer no matter what number of machines are in the service pool. A secure continuous deployment service requires secured SP code and a configuration repository that authenticates to the target computing resource regardless of the CSP network segmentation. Traditional network segmentation, or fire walling, is a secondary security mechanism that is used for ingress and egress filtering at various points in the local network segment to prevent IP spoofing [19] [20].

[8] https://cloud.google.com/bigtable/

Authentication and Authorization. In a federated cloud architecture, deployed workloads might require, and benefit from, access to other services deployed by the federation. The canonical example will be an end user request service deployed in the Federation that triggers another micro-service within the SP architecture. Such cascading requests require multilayered authentication and authorization processes. *i.e.*, a micro-service calls another micro-service and authenticates on behalf of the end user for audit trails supported by the end-user authentication token and the cascading micro-service tokens generated throughout the end user request. Figure 3 depicts the data flow during a call initiated by an ESP Micro-Service that runs in one of the federation's CSP (power system) members denoted by CSP_i and CSP_j. A call initiated from CSP_j that was provisioned in the federation as ms_n. The call destination runs on a different and sometimes the same SP. Let SP_m denote the destination SP. The call payload is encrypted by the SP_n private key. The call arrived at an SP_m endpoint and checked for authenticated admission. SP_m admission control decrypts the call payload using the SP_n public key that was submitted throughout the on-boarding process to the cloud federation. It is verified for authenticity and authorization of allowed call-sets. If admitted, SP_m calls and processes the get_data() call and sends back the response to the originating SP, SP_n.

Breach Detection: The Cloud federation comprises various workload types that are owned by different autonomous organizations. Breach detection includes a complex data processing pipeline that integrates system signals originating from specific users of a CSP service as well as the potential cloud federation tenants. System signals are comprised of network devices as well as signals from infrastructure services. Only in recent years, after the growing numbers of data breaches and liabilities arising from losses have organizations started to incorporate business related metrics for breach detection [21] [22] [23] [24]. Both data pipelines need to generate operational security warnings of potential incidents. The output of such warnings usually alerts security operations staff to

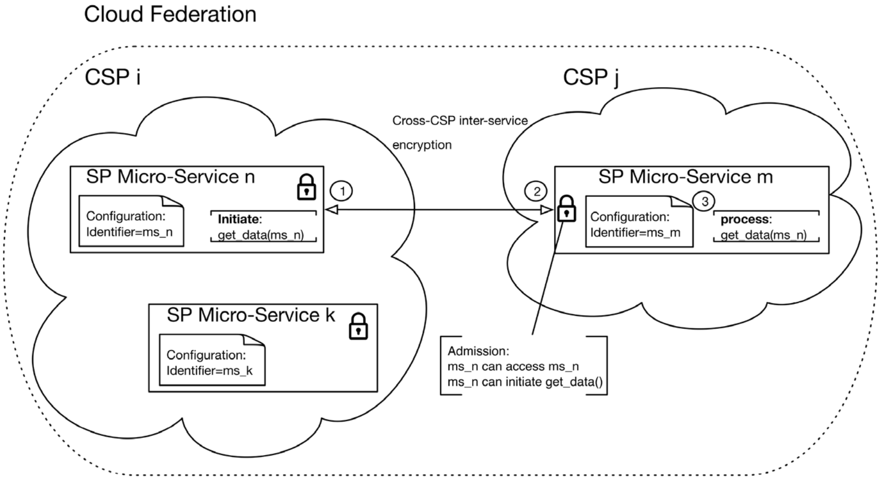

Figure 3. Authentication and authorization in Cloud Federation Cross-SP model.

potential incidents that require the relevant team's triage and response as appropriate. An estimation of the most significant attack patterns and prepared response patterns will go far to prevent catastrophic damage.

Such methods are sub-optimal in a federated cloud for two main reasons: 1) different data sets are owned by different organization departments that are not integrated physically, schematically, or semantically, 2) Lack of unification of both data sets as accomplished by fusion requires a complex transformation of both data sets semantics into a single data set. The above situation exacerbated when migrating the workload to the cloud as it introduces another orthogonal data set that contributes to complexity. The following sections propose a method for breach detection that collapses the three individual CSP silos into a cohesive semantic-less data set that will enhance the ability of the Cloud Federation services to detect breaches to an extent limited by available data and their investment in detection technology tools, *i.e.* allowing methods to the tenants to incorporate more data about their workload for more automatic detection.

7. Semantic-Less Breach Detection

As indicated in previous sections, this new cyber-security model proposed for the power system federated cloud assumes the network is breached. Thus a breach-detection system is needed to continuously attempt determining whether a workload is infected as well as the exploited risk type as enumerated above. Since the workload presented by power systems is substantially different than that of on-demand data streaming we extend that work to show how it applies to the subject of this paper [2]. Breach detection system effectiveness is influenced by a number of factors. For the sake of clarity, we focused on the human social factor and the emergent public cloud offering. The following section describes the important factors required for optimized breach detection. This mode of breach detection has to span the heterogeneous schema employed by various federation members CSPs.

7.1. The Control Systems Factor

Energy-based organizations such as power plants or utility services use supervisory control and data acquisition (SCADA) systems for managing the facility standard operations, e.g., power generation or power distribution for energy organizations. Those OT workloads are typically comprised of networked computers and data instrumented in graphical user interfaces. Its goal is a real-time control logic of the operational modules through field sensors and actuators. An attack on a facility OT Control Systems will induce unexpected commands to or behavior of the control systems that can result in unexpected system behavior. In cases like Stuxnet, the malware manipulated the feedback loop exacerbating the system [25]. Generating independent system feedback through a separate set of field sensors and actuators into a shared repository with statistical learning algorithms would enable anomaly detection and provide the needed cohesive context into the status of other individual, interconnected, and correlated facilities.

7.2. Cloud Federation Benefits to Power Systems

Public Cloud services exacerbate the organization's human factor risk by introducing an additional organizational entity that is often separated from the organization it serves. Public cloud operations are agnostic to its tenan's workload semantics by definition. CSPs configure their multi-tenancy to allow even businesses with conflict of interest to run their workload on the same platform. Such practices and policies do nothing to implement the cohesive view required for optimized malware detection and other Power System functions that provide value added through collaboration.

Workloads deployed in public cloud services are not limited to known machine boundaries as are traditional on-premise models offer. Although CSPs feature cyber security mechanisms that attempt mimicking the traditional computing workload hosting, workloads artifacts are under the CSP control. As such, the cloud client workloads might be compromised. Thus, there is a need for another cyber security dimension for the individual Power System ESP workload that overcomes the lack of control when running in the cloud.

A novel self-learning methodology is described that removes the need for Power System tenant information that streamlines semantic-less information flow from the various software stacks of the Cloud Federation. These include both independent Operational Technology (OT) field signals, tenant metrics, and control-plane metrics. Also, it streamlines the aggregation of useful training data for security incidents shared in collaborative platforms both inside and outside the Cloud Federation. The results obtained indicate that a Cloud Federation optimizes such collaboration and self-learning process. A system that implements such self-learning was prototyped and resulted in an initial up to 95% True-Positive rate with 96% True-Negative.

Workload data and usage patterns are critical process underpinning the ESP business success. Yet the leakage of some of the workload data and usage patterns impose a threat to the ESP business. This challenge represents a new threat of organizational espionage as well as attacks on the ESP service that can degrade ESP business continuity. Therefore, sharing semantics breaks the isolation between the two systems and might hold the hosting system accountable for security attacks in CSP or Cloud Federation platforms. Also, transforming every workload semantics into a coherent model that aggregates numerous ESP workloads requires a significant amount of investment. SPs will be reluctant to make such an investment, especially since it doesn't produce income. This method thus has a low likelihood of being implemented. Therefore, enabling a method that eliminates SP investment and business risks is a key for a modern breach detection system success. That is what is described here. Finally, a Cloud-Federation provides a centralized view of cross-CSP operations. Such centralized views allows SP workload deployment to different CSPs to gather a rich data set that will be available for malware identification, for predictive analytics and for improved performance of the datacenters and IT systems of individual power systems. We suggest a method that captures computing resources usage and intra federation traffic and infers potential breach or disruption to proactively alerts CSP security stakeholders about suspicious cyber instances.

From Workload Semantic to Semantic-Less

According to [3] [7], the organization IT workloads composed of online and offline systems. The semantic-less detection will address the polymorphic malware case as its data stream are abstracted from computing activity. The semantic-less detection can be extended independently by generating additional critical signals that can later be inspected for anomalies intrinsically to the signal sets supports by the supply-chain or manufacturing equipment.

In our case, a tenant's workload in a federated cloud manifested by sandboxed software containers that are limited to not more than 1) namespace per tenant for isolation and 2) limited to a resources control groups (aka cgroup)[9] Control groups are the mechanism for limiting computing server host CPU, Memory, Disk I/O and Network I/O usage per namespace. That is the foundation of Linux Containers, which alludes to the existing methods of measurements of the metrics set, CPU, memory and I/O usage. We call this set the behavioral attributes set. Access to cgroup and namespace configuration and control is available on the host level *i.e.* the host OS that runs the multi-tenant workloads *i.e.* a control-plane component [2].

7.3. Data Collection

Both cyber security leaders and national agencies agree that addressing emerging cyber risks require sharing cyber attacks retrospects and their historical behavior, and discovered vulnerability reports as a foundation for collaboration, predictive time series analysis, risk quantification and risk allocations all leading to safer cyber services [26] [27]. Incidents are often documented in unstructured reports that require a manual analysis to identify trends [28].

To assess whether or not a system was breached, it is required to establish malicious system behavior patterns and then decompose those patterns into generic computing system metrics that can later be classified as harmful or safe. The following section discusses the source datasets we chose to assess the initial malicious patterns and their detectability by our method. We continued by decomposing the data and removing the tenant semantics. That allows a generic pattern of malicious activity dataset that can be used as a training data for the supervised model.

7.3.1. Source Datasets

We extended the data sets used by [2] that used the National Vulnerability Database (NVD) [29] and the Vocabulary for Event Recording and Incident Sharing (VERIS) [30]. Both datasets included thousands of reported incidents span-

[9] https://www.kernel.org/doc/Documentation/cgroup-v1/cgroups.txt

ning across various categories. In addition, the datasets used in this paper included online malware analysis sandboxes services such as VirusTotal[10]. Specially, with malware that is deployed in stages, Stage I is usually called the "dropper". The dropper is a relatively small piece of code that breaches the existing defenses and gains a foothold in the target system or network. The dropper would normally contact the command and control (C & C) server to download Stage II, which could: contain the malware that further infects the target, starts to exfiltrate data, or otherwise executes a malicious payload, which may or may not take immediate action depending on the environment. Some Stage I droppers are sandbox or virtual environment aware to the point that if they sense they are not in the actual target environment (*i.e.*, some virtual environment), they may not take any further action. This can thwart any further analysis. Nonetheless, some information may possibly be gleaned through a reverse engineering of the Stage I dropper to determine the IP address of the C & C server. Asset owners could then be warned about this particular piece of information and block that specific IP address.

Our model focuses on 1) Unauthorized access attempts, 2) Suspicious Denial of Service, and 3) Data Stealing Malicious Code, including ransomware instances. We filtered the incidents that conform to the categories and performed a qualitative assessment of the identified breach impacts. Lastly, for simplicity, we applied an additional category that distinguishes the target component reported, service-based or client-based. We included only the service-based incidents. *i.e.*, reported incidents that clearly targeted desktops and workstations were not included in defining tenant semantic structures.

We applied filters for training data accuracy. Filters for VERIS dataset included server workloads as indicated in Section 4.2.1, *i.e.*, Authentication Server, Backup Server, Database Server, DHCP Server, Directory Server(LDAP, AD), Distributed control system, Domain Name Server, File Server, Mail Server, Mainframe Server, Web Application Server, and Virtual Machine Server [30]. Assets operating systems were filtered to Linux and Unix as such operating systems are more prevalent in servers than Windows, MacOSX, and mobile device operating systems.

The VERIS dataset includes incident actions. We filtered the action types that fit the paper focus workloads. *i.e.* Brute Force, Cache Poisoning, Cryptanalysis, Fuzzing, and HTTP Request Smuggling attacks. We excluded Buffer overflow cases as such attacks can be prevented in deterministic methods and common in Windows-based operating systems [31]. The dataset size following the refinement is 5015 incidents from an original set of over 10,000 incidents. **Table 1** summarizes the dataset we used for the training data.

7.3.2. Removing the Tenant Semantics

Our approach attempts to detect anomalies in both the control-plane and tenant activities that conform to suspicious patterns. In Section 7.4.1, we defined a

[10] https://www.virustotal.com

Table 1. Summary of datasets used.

Malware Category	modus operandi	Number of Incidents
Brute Force	Exhaustive effort of data encryption	946
Cache Poisoning	Corrupt data is inserted into the cache database e.g., DNS	894
Cryptanalysis	Exhaustive effort of data encryption	750
Fuzzing	Injects random bad data into an application to break it	946
HTTP Request Smuggling	Exhausting a proxy cache by sending HTTP requests	639
Data stealing malware	Data transmitting across unencrypted network	840

categorical dataset that adheres to real incident data. This data applies to potential breaches for server-based workloads. We stipulate, for the purposes of this paper that such server-based workloads will obey similar suspicious patterns when deployed in the cloud.

We transformed the categorical dataset into a multivariate time series data that can be used for supervised anomaly detection. The multivariate set is comprised of general operating system observations that do not include any workload semantics but could be used for contextual anomaly detection. The contextual attributes are used to determine the position of an instance on the entire series. We found that, based on collected incident data, the conversion of behavior patterns to multivariate time-series satisfies effective breach detection of any malware, conventional or polymorphic.

We gathered the operations reported in the incident reports (Table 1) and inferred about the operating system resources consumed during the malware lifespan. Table 2 depicts the relationship between the malware characteristics and operating system usage. Figure 4 describes a workload sample comprising of tele-signalization, tele-metering, tele-control and tele-regulation [32]. It shows the common pattern of the operating system resource usage that will be used as multivariate time series data sequences. The Evaluation section describes in more details the nature of the data and how it translates into meaningful time series data.

7.3.3. Prediction Methodology

We used the data gathered in Table 2 for formulating the anomaly detection problem of polymorphic malware [33]. The detection approach includes three distinct methods: 1) Detecting anomalous sequences in OS usages time series events, 2) Detecting anomalous subsequences within OS usages time series, and 3) Detecting anomalous OS usage events based on frequency. Let T denote a set of S training sequences based on OS usage generated by CSPs, SPs, and the Federation control plane. Also, S denotes a set of m test sequences generated based on Table 2. We find the anomaly score $A(S_q)$ for each test sequence $S_q \in S$, with respect to T. T mostly includes normal OS usage sequences, while S includes anomalous sequences.

Table 2. Dataset classification.

Malware Category	OS Resources Patterns
Brute Force	Extensive CPU, Memory, I/O to disk or network
Cache Poisoning	Extensive I/O to disk or network
Cryptanalysis	Extensive I/O to disk or network
Fuzzing	Network I/O Ingress
HTTP Request Smuggling	Network I/O Egress
Data stealing malware	Network I/O Egress

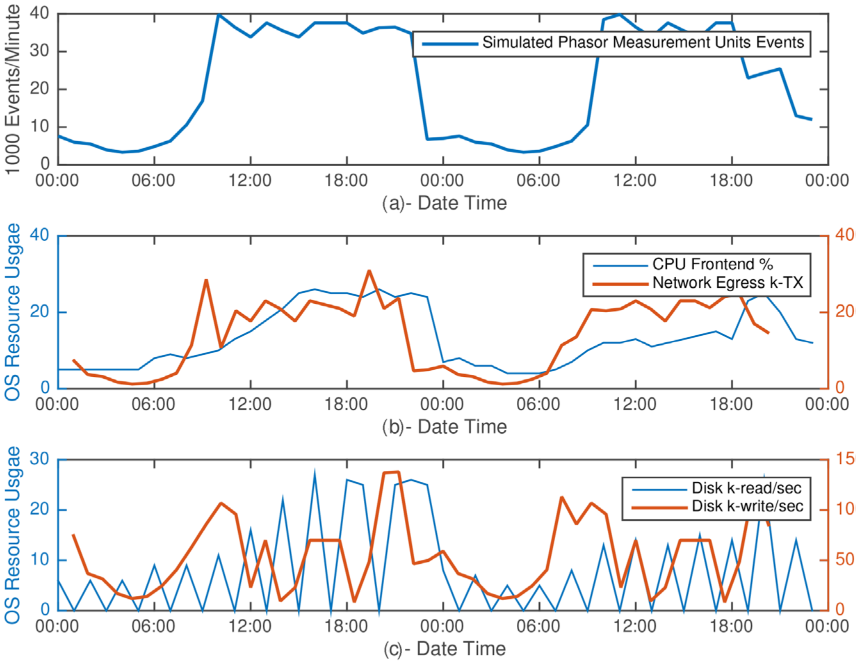

Figure 4. (a) shows a simulation data of a Phasor Measurement Units (PMU) recording device workload. (b) and (c) shows a normal system usage patterns, CPU, network and Disk usage respectively.

The semantic-less tool output produces a score for a scanned training sequence T using Regression, *i.e.*, a forecast of the next observation in the time series, using the statistical model and the time series observed so far, and compares the forecasted observation with the actual observation to determine if an anomaly has occurred [34]. For simplicity, our model uses Tensor Flow [33] for regression calculation. Our tool is not limited to that tool or the regression type.

8. Evaluation

We prototyped cloud federation system that mimics the properties analyzed in Section 2. The prototyped system includes the components that are depicted in Figure 2. For the scope of the prototype, we enabled semantic-less metrics from both ESPs and CSPs to improve correlation efficiency. CSP data sharing limits the effectiveness of any cyber analytical technique and, in practice, will represent a compromise between improved cyber security and CSP privacy and confidentiality. With that proviso, in the following section, we evaluate a computer load coordination system component that manages typical power system, such as tele-signalization, tele-metering, tele-control and tele-regulation. It generates T, a set of n training sequences based on OS usage generated by CSPs, ESPs, and the federation control plane.

We chose Wide-Area Measurement System (WAMS) workloads for the simulation as they are the common digital recording devices being installed at different points in the North American grid, especially under the smart grid initiatives of the US Department of Energy [35]. WAMS creates a predictable processing, networking and digital storage usage [36]. We showed that WAMS workloads follows a pattern of usage that can be monitored for breach detection that can help ESPs to seamlessly improve their service availability with minimal ESP investments.

8.1. Experiment Planning

We conducted a simulation of a cross-regional platform that is comprised of a control-plane, workload-plane, and coordinating components. This is embodied in a resource allocation system (Kubernetes). This system provisions resources to have a priority for being near users. The control-plane enables an effective compute resource provisioning system that spans across different public cloud providers and regions. Also, it collects operating systems usages for both the ESP workload and control-plane. The coordinating components record and communicate GPS-synchronized, high sampling rate (6 - 60 samples/sec), dynamic power system data. The workload-plane is comprised of edge servers that process the tele-signals. It is built on standard Apache HTTP[11] servers that run on the edge location.

The control-plane software infrastructure is based on Kubernetes[12], it facilitates internal discovery between Apache HTTP server instances so instances can connect across different cloud boundaries and regions. This architecture provides an open architecture that enables continuous monitoring. In a real world federation, the data load may require several big data nodes and substantial compute capacity. This paper is a demonstration an proof of concept on a finite scale to permit model and parameter tracking and adjustment.

8.2. Execution

8.2.1. The System Preparation

The prototype experiment was based on energy consumption projections in the U.S. The experiment included the setup of three virtual datacenters deployed in different regions: 1) Central US, 2) West US and 3) East US. The clusters were sized based on US population distribution[13] by regions *i.e.* 20% for West US,

[11] Apache Web Server reference retrieved from https://httpd.apache.org

[12] Kubernetes reference retrieved from http://kubernetes.io

[13] US Population Distribution retrieved from https://www.census.gov/popclock/data tables.php

40% for East US and 40% Central US. The cluster sizes for West US, Central US, and East US are 3, 7 and 7 machines respectively. Each machine is standard 2-CPU cores with 7.5 GB of memory.

The control-plane comprised of a Kubernetes API server and controller-manager. The controller coordinator component will need to allocate resources across several geographic regions to different cloud providers. The API server will run a new federation namespace dedicated for the experiment in a manner that such resources are provisioned under a single system. Since the single system may expose external IPs, it needs to be protected by an appropriate level of asynchronous encryption[14].

For simplicity, we use a single cloud provider, Google Container Engine, as it provides a multi-zone production-grade compute orchestration system. The compute instances that process the user workloads are deployed as Docker containers that run Ubuntu 15 loaded with the Apache HTTP server. For simplicity, we avoided content distribution by embedding the tele-signaling simulator in the Docker image. We ran 52 Docker containers that span across the three regions and acted as WAMS edges.

8.2.2. Baseline and Execution

The baseline execution included data populations for power system tele-signalization. The data population was achieved by using Kubernetes Jmeter batch jobs. The loader jobs goal is to generate traffic that obeys the observed empirical patterns depicted in **Figure 4**. The system usage for both control-plane and ESP was measured through cAdvisor, a Kubernetes resource usage analyzer agent. The agent, from every node in a cluster, populates system usage data to Heapster, a cluster-wide aggregator of monitoring and event data[15].

We labeled the system usage with semantic-less dimensions, as shown in **Figure 4(c)** graph. The Network egress was measured by thousands of transmitted packets (k-TX), Disk writes per second (k-write/sec) and CPU usage per container (%). The Heapster aggregated the data based on the labels that are later pushed to a centralized database, influxDB. The anomalous sequences $S_q \in S$ was injected as synthetic randomized system usage data to the influxDB using HTTP API. The date was tagged as well according to the three labels, CPU, network and disk usage. We used data in **Figure 4(b)**, **Figure 4(c)** as a baseline sequence that was randomized using NumPy[16]. The randomization followed the malicious usage patterns described in **Table 2** and resulted in data shown in **Figure 5(c)**.

The execution required a TensorFlow session that looped through the dataset 2000 times, updated the model parameters and obtained the anomaly score $A(S_q)$ for each test sequence $S_q \in S$, on T. The breach and anomaly

[14] Simulation code and data retrieved from https://github.com/yahavb/green-content-delivery-network

[15] https://kubernetes.io/docs/user-guide/monitoring/

[16] Package for scientific computing with Python, numpy.org

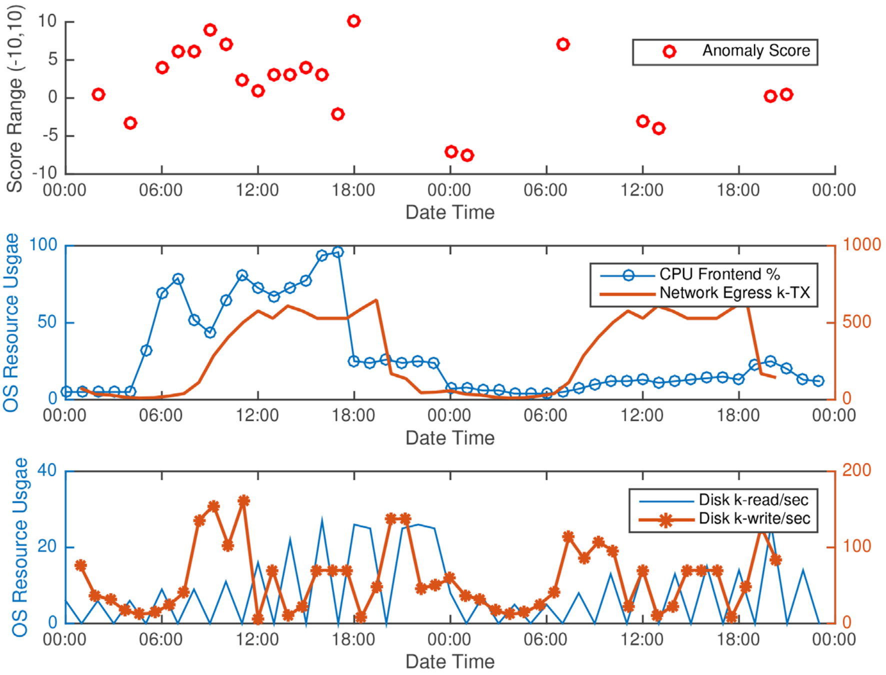

Figure 5. Simulation of anomalies found in the PMU recording devices obtained after 2000 training epochs of the TensorFlow tool. It shows anomaly scores that exceed a threshold value. The threshold value depends on data quality, on interference, noise levels, and disturbances. Thus the threshold value is a dynamic system property. For the purpose of this demonstration various singular values of the threshold were chosen.

detection was performed using the data streams depicts in Figure 4 as training sequences and Figure 5 as observed anomalous sequence using the TensorFlow logistic regressions supervised learning algorithm[17].

8.3. Analysis

The prototype included two core datasets, normal (Figure 4) and malicious (Figure 5(c)). The CPU usage in the normal dataset fit the normal tele-signaling patterns at the first half of the run. The second half required less CPU due to the caching mechanism applied in the Apache HTTP server that alleviates the need for the CPU when served through a cache. The Disk write pattern manifested similar content caching schema. The network egress ratio was not impacted by the caching schema.

The malicious dataset used the malware classification table (Table 2). Figure 5, shows a semantic-less behavior for ransomware malware that attempts to encrypt data while serving workload. Suspicious signals denoted by a star and o markers for CPU and disk write respectively. Based on the dataset classification, ransomware requires no network egress but requires the CPU for data encryption and writing back the encrypted payload to disk. Our prototype included similar patterns depicted in Table 2 with a similar approach as done for ransomware.

[17] https://www.tensorflow.org/tutorials/wide

Our model yielded a series of anomaly scores $A\left(S_q\right)$ for $S_q \in S$ anomalous patterns (**Figure 5**). The model attempted to detect anomalous subsequences within S_q. e.g., S_{net} for network, S_{cpu} for CPU, S_{read}, and S_{write} for disk read and write transactions. We used a scoring based techniques for each of the observed sequences. The score range is between $\left[-10,10\right]$ and executed 2000 training epochs. Scores that are closed to 10 indicate a potential breach. Negative scores indicate normal behavior. The scores shown in **Figure 5** are the $\max\left(S_{net}, S_{cpu}, S_{write}\right)$. The maximum-based aggregation was chosen because of the suspected anomaly nature. Specifically, extensive write and CPU operations plays an equal role in the potential for breach. Other anomalies might use different type of aggregations. According to the anomaly scores shown in the experiment, up to 95% were True-Positive of detected breaches and 96% of the detections were True-Negative. The advantage of more data and more specific data labeling is shown be the improved detection scores from earlier paper by 10% [2]. This result shows the potential for continual improvement of results in the field and less dependence on initial score than is needed with a non-learning, less adaptive system.

9. Conclusion

Power systems are facing cyber threats that are growing in quantity, complexity, and sophistication. Collaborative security in the form of a federated cloud will add coordination for security, for load sharing/redundancy and for improved asset utilization while maintaining organizational independence. This publication integrates and applies earlier results on cyber security, equipment design, energy efficiency, and load management to create a paradigm for next generation power systems. It improves security by employing well known effective ways that minimize the potentially crippling costs associated with providing security, resilience and capacity redundancy. Large scale complex technology such as that of power systems is already driving the emergence of disparate, distributed, large scale, multi-tenant environments such as the proposed cloud federation. These environments are redefining traditional perimeter boundaries along with traditional security practices. Defining and securing architectural configurations that maintain, secure and adapt Power system functionality asset boundaries is more challenging. This paper presents a federated cloud-based multi-platform power system utility that facilitates communication, load resilience under stress, energy cost management, and a proactive approach for detecting breaches, utilizing general system usage patterns that help to predict potential breach proactively from prior history and regular power system computing service provider workloads. This approach shares and minimizes upfront investments and provides savings that generate a return on investment within a year of deployment. Our proposed system of systems ultimately provides a way to meet power system regulatory service level requirements and maximizes system owner control.

References

[1] Leavitt, N. (2009) Is Cloud Computing Really Ready for Prime Time. *Growth*, **27**, 15-20. https://doi.org/10.1109/MC.2009.20

[2] Biran, Y., Collins, G., Borky, J.M. and Dubow, J. (2017) Semantic-Less Breach Detection of Polymorphic Malware in Federated Cloud. Advances in Science. *Technology and Engineering Systems Journal*, **2**, 553-561.

[3] Biran, Y. Collins, G. and Dubow, J. (2017) Cloud Computing Costand Energy Optimization through Federated Cloud SOS. *Systems Engineering*, **20**, 280-293.

[4] Lewis, J. and Baker, S. (2013) The Economic Impact of Cybercrime and Cyber Espionage. McAfee.

[5] Biran, Y., Sudeep, P., George, C. and Joel, D. (2016) Enabling Green Content Distribution Network by Cloud Orchestration. 2016 3*rd Smart Cloud Networks & Systems* (*SCNS*), IEEE, Dubai, 3-12. https://doi.org/10.1109/SCNS.2016.7870553

[6] Biran, Y., Collins, G., Azam, S. and Dubow, J. (2017) Federated Cloud Computing as System of Systems. *International Conference Computing*, *Networking and Communications*, 711-718.

[7] Biran, Y., Collins, G. and Liberatore, J. (2016) Coordinating Green Clouds as Data-Intensive Computing. *IEEE Green Technologies Conference*, 130-135.

[8] Biran, Y., Pasricha, S., Collins, G. and Dubow, J. (2017) Clean Energy Use for Cloud Computing Federation Workloads. *Advances in Science*, *Technology and Engineering Systems Journal*, **2**, 1-12.

[9] North America Electric Reliability Corp (2016) Defense Use Case. Analysis of the Cyber Attack on the Ukrainian Power Grid.

[10] Bronk, C. and Tikk-Ringas, E. (2013) The Cyber Attack on Saudi Aramco. *Survival*, **55**, 81-96. https://doi.org/10.1080/00396338.2013.784468

[11] Lindsay, J.R. (2013) Stuxnet and the Limits of Cyber Warfare. *Security Studies*, **22**, 365-404. https://doi.org/10.1080/09636412.2013.816122

[12] Collins, G. and Biran, Y. (2015) Multi-Tenant Utility Computing with Compute Containers. 5*th International Conference on Consumer Electronics-Berlin*, 213-217.

[13] Nelson, R. (2016) IOT Spans Edge to Data Center. *Evaluation Engineering*, **55**, 18-21.

[14] Kim, D., Chung, H.R. and Thompson, P.R. (2009) Cloud-Based Automation of Resources.

[15] Orr, R.J. and Abowd, G.D. (2000) The Smart Oor: A Mechanism for Natural User Identification and Tracking. *Extended Abstraction Human Factors in Computing Systems*, 275-276.

[16] Biran, Y. and Collins, G. (2016) Open Compute-Equipment Design Specification as a Standard for Cloud Computing. *Zooming Innovation in Consumer Electronics International Conference*, 70-75.

[17] Krutz, R.L. and Vines, R.D. (2010) Cloud Security: A Comprehensive Guide to Secure Cloud Computing. Wiley Publishing.

[18] Skorobogatov, S.P. (2005) Semi-Invasive Attacks: A New Approach to Hardware Security Analysis. PhD Thesis.

[19] Andersen, D.G., *et al.* (2003) Mayday Distributed Filtering for Internet Services. *USENIX Symposium on Internet Technologies and Systems*, 20-30.

[20] Peng, T., Leckie, C. and Ramamohanarao, K. (2007) Survey of Network-Based De-

fense Mechanisms Countering the Dos and Ddos Problems. *ACM Computing Surveys*, **39**, 3. https://doi.org/10.1145/1216370.1216373

[21] Liu, Y., Sarabi, A., Zhang, J., Naghizadeh, P., Karir, M., Bailey, M. and Liu, M. (2015) Cloudy with a Chance of Breach: Forecasting Cyber Security Incidents. *USENIX Security*, 1009-1024.

[22] Agrawal, T., Henry, D. and Finkle, J. (2014) JP Morgan Hack Exposed Data of 83 Million, among Biggest Breaches in History.

[23] Krebs, B. (2014) The Target Breach, by the Numbers. Krebs on Security, 6.

[24] Sidel, R. (2014) Home Depot's 56 Million Card Breach Bigger than Target's. *Wall Street Journal*.

[25] Langner, R. (2011) Stuxnet: Dissecting a Cyber Warfare Weapon. *IEEE Security & Privacy*, **9**, 49-51. https://doi.org/10.1109/MSP.2011.67

[26] Jaquith, A. (2007) Security Metrics. Pearson Education.

[27] Chew, E., Swanson, M., Stine, K., Bartol, N., Brown, A. and Robinson, W. (2008) Performance Measurement Guide for Information Security. National Institute of Standards and Technology Special Publication 800-55 Revision 1.

[28] Tufte, S.E. (2015) Documenting Cyber Security Incidents. Department of Management Science and Engineering, Stanford University, School of Information, UC Berkeley.

[29] CERT Coordination Center (2004) Vulnerability Notes Database. CERT Coordination Center.

[30] Verizon RISK Team (2015) 2015 Data Breach Investigations Report. Verizon Data Breach Investigations Report (DBIR).

[31] Wartell, R., Mohan, V., Hamlen, K.W. and Lin, Z. (2012) Binary Stirring: Self-Randomizing Instruction Addresses of Legacy x86 Binary Code. *Proceedings of the* 2012 *ACM Conference on Computer and Communications Security*, 157-168.

[32] Corsi, S. (2015) Voltage Control and Protection in Electrical Power Systems: From System Components to Wide-Area Control. Springer. https://doi.org/10.1007/978-1-4471-6636-8

[33] Chandola, V. (2009) Anomaly Detection for Symbolic Sequences and Time Series Data. PhD Thesis, University of Minnesota.

[34] Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., *et al.* (2016) Tensorow: Large-Scale Machine Learning on Heterogeneous Distributed Systems.

[35] Chakrabortty, A. and Khargonekar, P.P. (2013) Introduction to Wide-Area Control of Power Systems. *American Control Conference*, 6758-6770.

[36] Karlsson, D., Hemmingsson, M. and Lindahl, S. (2004) Wide Area System Monitoring and Control-Terminology, Phenomena, and Solution Implementation Strategies. *IEEE Power and Energy Magazine*, **2**, 68-76. https://doi.org/10.1109/MPAE.2004.1338124