An Internet-Wide View of ICS Devices

Ariana Mirian[†] Zane Ma[‡] David Adrian[†] Matthew Tischer[‡] Thasphon Chuenchujit[‡] Tim Yardley[‡]
Robin Berthier[‡] Joshua Mason[‡] Zakir Durumeric[†‡] J. Alex Halderman[†] Michael Bailey[‡]

[†] University of Michigan [‡] University of Illinois at Urbana-Champaign

***Abstract*—Industrial control systems have become ubiquitous, enabling the remote, electronic control of physical equipment and sensors. Originally designed to operate on closed networks, the protocols used by these devices have no built-in security. However, despite this, an alarming number of systems are connected to the public Internet and an attacker who finds a device often can cause catastrophic damage to physical infrastructure. We consider two aspects of ICS security in this work: (1) what devices have been inadvertently exposed on the public Internet, and (2) who is searching for vulnerable systems. First, we implement five common SCADA protocols in ZMap and conduct a survey of the public IPv4 address space finding more than 60 K publicly accessible systems. Second, we use a large network telescope and high-interaction honeypots to find and profile actors searching for devices. We hope that our findings can both motivate and inform future work on securing industrial control systems.**

I. Introduction

Industrial control systems (ICS) are pervasive and control critical infrastructure ranging from power grids and chemical manufacturing plants to the environmental monitoring and fire suppression controls in commercial buildings. These systems communicate over a myriad of domain and manufacturer specific protocols that have grown organically over the past 40 years, including Modbus, BACnet, DNP3, and Siemens S7. Originally designed to operate in a closed environment, these protocols have no built-in security. However, despite this, protocols were layered on Ethernet and TCP/IP, and inevitably devices have been connected to the public Internet to support remote monitoring and management. In this paper, we investigate the devices exposed on the Internet and the malicious actors searching for them.

We first analyze the inherently vulnerable devices on the public Internet by extending ZMap [18] to support five common protocols: Modbus, DNP3, BACnet, Tridium Fox, and Siemens S7. We complete full scans of the IPv4 address space on each protocol and identify more than 65 K vulnerable control systems in 3.7 K ASes and 145 countries. Of these, over 400 use DNP3—a protocol designed to facilitate communicate between electrical substations. We similarly find upwards of 25 K Modbus and Siemens S7 hosts associated with processing control and manufacturing. We categorize devices and find that 69% of devices are generic Modbus bridges, but we also identify over 300 water flow meters and 700 solar plant data loggers. We investigate and categorize the networks, and find that devices belong to a large variety of industries—worryingly including gas and electrical companies, medical centers, and public transport providers. Lastly, we find more than 38,000 building automation devices running BACnet or Tridium Fox that control environmental systems in industries ranging from hotels and airports to water treatment plants and government buildings.

Next, we investigate who is scanning for vulnerable industrial control systems by analyzing the traffic received by a large network telescope containing roughly one million IPv4 addresses in August 2015 and launching high-interaction ICS honeypots to observe what commands actors run against systems they find. Unique from previous work on Internet-wide scanning [17], we find that only a small handful of organizations are performing regular scans—over 98% of SCADA traffic originates from ten organizations that are primarily scanning for Modbus and BACnet devices. However, we note regular traffic from bulletproof hosting providers associated with malicious actors.

The industrial control space is in disarray. A multitude of protocols have been architected with little to no thought about security, and our results show that vulnerable devices are widespread in nearly every industry and region. We hope that by bringing this ecosystem to light, our findings motivate administrators to clean up these vulnerabilities and inform future research into industrial controls.

II. Background

As industrial systems, such as electrical substations and manufacturing plants, became more complex, their components became increasingly interconnected. Historically, these devices were networked using proprietary analog control mechanisms that operated over twisted pair cable [4]. However, as microcontrollers were developed in the 1970s, these replaced analog, point-to-point wiring schemes, allowing for hundreds of components to be remotely monitored and controlled. The quick rise of digital buses led to a variety of proprietary serial communication protocols. Unfortunately, the process was uncoordinated, and each protocol was designed to address the requirements of a specific industry or manufacturer. In the 1980s, the needs for interoperability and cost-saving, vendor-agnostic solutions pushed the industry to embark on a standardization effort that went on for nearly two decades.

Standards were first developed at the national level. Germany standardized the Process Field Bus (PROFIBUS), while the Factory Instrumentation Protocol (FIP) was widely adopted in France. As time went on, companies recognized the need for an international standard and in 1999, the Instrumentation Society of America (ISA) and the Industrial Electrotechnical Commission (IEC) arrived at a compromise: IEC 61158 [23]. Unfortunately, the standard was unwieldy—it contained eight sets of protocols in a nearly 4,000 page document. Four years later, the IEC published a slimmed down IEC 61784 [24]. The

new standard remained composed of numerous protocols, but the standard selected Ethernet as the single link-layer protocol. It also pushed changes to the Ethernet standard to adapt to the real time requirements of industrial control systems. Ethernet did not, however, replace all serial field protocols, which are widely used today due to their guaranteed speed and capability of communicating longer distances. In spite of standardization efforts, approximately ten protocols are widely used today.

Common Terms There are several domain-specific terms that are used to describe components within industrial control systems. The components specific to ICS include programmable logic controllers (PLC), supervisory control and data acquisition (SCADA) systems, and distributed control systems (DCS). A PLC is a digital computer used to automatically control and monitor an electromechanical process such as a factory assembly line or a circuit breaker. A DCS is a larger system of multiple controllers organized in a hierarchy to control and monitor a complete manufacturing process or power plant. Finally, a SCADA system is used to control and monitor multiple sites that can be geographically separated in an open-loop control environment.

Protocol Requirements Since ICS protocols control physical equipment—frequently in mission critical environments—these protocols have several unique constraints that have directed their design. First, protocols have strict real time constraints. For example, a relay responsible for tripping a circuit breaker in an electrical substation must respond in under 4 milliseconds. A longer response or incorrect signal can result in cascading failures and endanger human life. Further, these systems generally operate in harsh environmental conditions and must operate correctly in the face of these conditions.

These systems were never intended to be Internet connected, and given other constraints, little attention was paid to security. As a result, many legacy SCADA protocols have no built-in authentication or encryption mechanisms. Merely connecting to such services is often enough to exert complete control over the processes they oversee.

III. Common Protocols

In this section, we describe protocols commonly in use today.

A. Process Automation

Four protocols are commonly used for process automation (e.g., manufacturing facilities): Modbus, HART-IP, S7, and EtherNet/IP.

Modbus (TCP/502) Modbus was designed in 1979 to control and monitor Modicon (now Schneider Electric) PLCs. The protocol quickly became the de facto standard for industrial networks and recently Modbus has also been seen in building infrastructure, transportation, and energy management systems [5]. Modbus operates in a master/slave architecture and supports both serial and TCP/IP networks, but has two weaknesses. First, it is limited to 240 devices per network. Second, while the protocol is managed by the Modbus organization, many vendor extensions are proprietary and undocumented, resulting in interoperability issues. This protocol lacks any built-in security.

Siemens S7 (TCP/102) Siemens S7 is a proprietary, but commonly deployed, protocol used by Siemens S7 PLCs. These controllers are typically used in manufacturing, specifically the automotive and packaging industries. The protocol is command based, in which every transmission is either a command or reply. S7 is neither authenticated nor encrypted and is susceptible to spoofing, session hijacking, and denial of service attacks [11].

EtherNet/IP (TCP/44818, UDP/2222) Ethernet Industrial Protocol was developed in the 1990s by Rockwell Automation and combines standard Ethernet with the media-independent Common Industrial Protocol. It is maintained by OVDA and deployed in time-critical industrial environments. The protocol operates in a producer-consumer model, in which devices publish data to all others using an Ethernet-based multicast. The protocol lacks built-in security protections.

HART-IP (TCP/5094) HART-IP was developed to facilitate accessing HART—Highway Addressable Remote Transducer—devices over Ethernet. HART can run over 4-20mA analog wiring, making it a popular transition protocol for organizations that had previously deployed analog wiring. HART is often used in the field as a means to provide configuration and diagnostic information to remote devices. This protocol has no built-in security.

B. Building Automation

There are two leading building automation protocols: BACnet and Niagara Tridium Fox, which are typically used to control HVAC, lighting, access control, and fire detection systems.

BACnet (UDP/47808) The American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) developed BACnet (Building Automation and Control Network) in 1995 to integrate different products within a single building. While there are many standard objects and properties defined by the specification, vendors can specify proprietary objects. So while the protocol allows extensibility, it simultaneously prevents cross-manufacturer interoperability. The protocol specification includes security features, but Kaur et al. found that manufacturers do not implement these in practice [27].

Niagara Tridium Fox (TCP/1911) Tridium Fox is a proprietary protocol designed by Niagara to tunnel to remote SCADA networks. The protocol is used in building automation [2]. Unlike the other protocols, Tridium Fox does not speak directly with industrial components, but rather facilitates communication between management workstations and devices (which in turn use a lower level protocol, e.g. BACnet, to communicate with individual components). Tridium *does* have built-in authentication and basic security features.

C. Electric Power Grid

There are three protocols commonly used for power system automation: DNP3, ICCP, and IEC 61850.

ICCP (TCP/102) Inter-Control Center Protocol—also known as IEC 60870-6—is used for linking control centers (e.g., power substations). ICCP allows for both real-time commands and historical monitoring by embodying an object oriented design where devices are presented as objects with associated actions [19]. Objects can be either concrete devices (e.g., transformers and relays), or abstract data structures (e.g. transfer sets). ICCP is a cleartext protocol with no confidentiality or integrity mechanisms.

IEC 61850 (TCP/102) IEC-61850 is an international standard for networks within electrical substations. The protocol was developed to allow vendor interoperability and in turn abstracts vendor objects, enabling devices to describe their own functionality, and facilitating their communication. The protocol supports TCP/IP and switched EtherNet for long-distance communication [43]. IEC-61850 does have some basic security protections, but they vary depending on which portion of the protocol is under inspection, and many areas are deferred to other protocols, rather than being built into the protocol itself.

DNP3 (TCP/20000, UDP/20000) Distributed Network Protocol (DNP3) is a set of protocols used for electrical grid automation. When developed, IEC 61805 and 60870-5 had not yet been finalized and DNP3 was based on the partial protocols to facilitate an immediately implementable design [3]. DNP3 was developed by GE-Harris Canada (formerly known as Westronic, Inc.) in 1990 and was subsequently widely deployed by electrical and water companies. Designed for SCADA applications, the protocol optimizes the transmission of data acquisition information and control commands between master (control centers) and outstations (remote computers) using event-driven data reporting [1]. Vulnerabilities often stem from implementation errors due to the protocol's complexity [15]. A malformed frame can crash the receiving process or drive it into an infinite loop, rendering the entire device inoperable.

D. Power Meter Automation

There is one common Advanced Metering Infrastructure protocol, ANSI C12.22, which is used for communication between smart meters and utility companies.

ANSI C12.22 (TCP/1153, UDP/1153) C12.22 is prevalent in North America and leverages other industrial control system protocols such as ACSE (Association Control Service Element) to store header information, and EPSEM (Extended Protocol Specifications for Electric Metering) to carry the payloads that store metering data using the older C12.19 protocol. The data portion of the EPSEM element can be sent in clear text, authenticated clear text (EAX), or encrypted text (EAX-AES).

IV. Exposed SCADA Devices

To understand how ICS devices have been publicly exposed, we extended ZMap [16], [18] to support Modbus, BACnet, Niagara Tridium Fox, and DNP3, and completed regular scans of the public IPv4 address space from December 12, 2015 to March 19, 2016. In this section, we describe the publicly accessible hosts we find. We emphasize that because most SCADA devices have no built-in authentication or security, Internet-exposed devices are inherently vulnerable to attack.

As with any active measurement, there are many ethical considerations at play. To minimize any potential harm, we thoroughly tested our scanners and follow the guidelines set forth by Durumeric et al. [18], including signaling benign nature and respecting any exclusion requests. We never alter device state, only querying broad system information. Further, we performed vulnerability notifications for the 79% of vulnerable hosts whose network's had abuse contacts in WHOIS [29].

Modbus We scanned for Modbus devices by sending the Modbus MEI Read Device Identification command, which returns an ASCII description of the host; devices that do not support the command respond with a Modbus "function not implemented" reply. In our most recent scan on March 19, 2016, we found 3.5M hosts with TCP/502 open, of which 23K responded to our Modbus query, a 7.1% increase since December 11, 2015. Of the 23K Modbus devices, 4.7K (20.3%) responded with a device information; the remainder responded with a Modbus not-implemented error.

We aggregated hosts by AS and find that devices are widespread in 1,979 ASes and 125 countries, most commonly the U.S. (19.1%), Turkey (7.9%), Spain (6.6%), France (6.5%), and Poland (4.9%). The top 10 ASes that contain the most devices all belong to ISPs and account for 30.7% of hosts. Unfortunately, this provides little indication of device owner. However, because AT&T identifies individual business subscribers in delegated WHOIS records, we were able to identity the owners of devices in AT&T ASes. We note 15 devices in medical centers and blood donation centers, and 9 that belong to energy companies including Pacific Gas and Electric Company and Dominion Energy. Similarly, we were able to obtain 102 identifiable WHOIS IP records from the French telecom Orange A.S. Unlike AT&T, 59% of the Orange hosts belong energy companies and 10% are associated with water or sanitation groups.

We are able to categorize 4,553 (97.1%) devices by their vendor name and 4,406 (94.0%) by their product code. We find that while 31.4% of devices are generic controllers, a surprising 15.8% are solar monitoring devices and 7.8% are water flow meters. The solar panels are present in 35 countries, most of which are located in western Europe, with a significant number of devices in Belgium (24.4%), Germany (23.6%), Austria (12.5%), and Italy (7.4%). The water controllers—-Nivus OCM Pro CFs—are primarily located in Japan, where there are 314 devices on the KDDI Corporation's network. Lastly, we find several devices belonging to local municipalities, hotels, and a device belonging to the U.S. Environmental Protection Agency.

Due to the large number of generic Modbus bridges and lack of fine-grained ownership data, it's difficult to discern exactly who is responsible for many of the vulnerable devices and what type of equipment they control. In most cases, these devices belong to small organizations, but we also note a handful

ICS Protocol			IPv4 Vulnerable Systems		Network Telescope	
Name	Industry	Port	Port Open	Protocol Handshake	Scan Traffic	Organizations
Modbus	ICS	TCP/502	525K	21.7K	41.7%	16
BACnet	Building Management	UDP/47808	17K	12.8K	30.6%	16
DNP3	Power Grid	TCP/20000	869K	515	5.1%	11
EtherNet/IP	ICS	TCP/44818, UDP/2222	567K	–	8.4%	12
HART-IP	ICS	TCP/5094	585K	–	2.4%	10
IEC 61850	Power Grid	TCP/102	550K	–	*8.7%	17
ANSI C12.22	Meter Reading	TCP/1153, UDP/1153	649K	–	0%	8
ICCP	Power Grid	TCP/102	550K	–	*8.7%	17
Siemens S/7	ICS	TCP/102	550K	2,292	*8.7%	17
Tridium Fox	Building Management	TCP/1911	591K	27K	3.1%	12

TABLE I: **Common SCADA Protocols**—We complete full IPv4 scans for vulnerable systems on the IPv4 address space and measure scanning activity by analyzing a large network telescope.

of devices that belong to well-known power companies, but outside of their primary network. This may suggest that groups are using other network providers to serve off-site devices, which are missed by traditional audits.

DNP3 DNP3 is a multi-layer protocol that implements custom link, transport, and application layers. In the protocol, each device is assigned a unique 2-byte address, with four addresses being reserved for broadcast. Every request must specify a source and a destination address. Unfortunately, this means that scanning for every unique address would likely overload hosts and broadcast packets have previously been noted to crash some devices [15]. We scanned for DNP3 devices by performing TCP/20000 scans against the IPv4 address space and sending a DNP link status request, which contains a status request for the first 100 address (0–99) in the destination addresses. This methodology identifies hosts that speak DNP3, but unfortunately does not provide any details about the devices themselves.

We completed a scan against the IPv4 address space on March, 19 2016 and identified 429 DNP3 devices spread across 91 networks, with the top ASes belonging to telecommunication companies or satellite communication companies (66.7%). We similarly find DNP3 devices spread out across 31 countries, with an overwhelming 68.2% of devices in the U.S. All identifiable hosts in the U.S. belonged to solar farms and small power generation companies. Due to the opaque nature of address allocation and lack of any identifying device information, it is hard to discern what the vast majority of the vulnerable devices are. However, anecdotal evidence suggests that these are being used by power companies and more than 400 devices are vulnerable.

Siemens S7 Similar to DNP3, each device in an S7 network is assigned a 2-byte ID and packets must define a source and destination address. Based on the behavior of other S7 implementations [21] and permutation tests that we ran against small samples of the IPv4 address space, we find that the set of source (0x100, 0x200) and destination (0x102, 0x200, 0x201) addresses are most commonly used and account for 87.5% of IPv4 hosts. In our scans, we issue a System Status List (SZL) request (code 0x04) that contains two sub-requests: `MODULE_IDENTIFICATION` and `COMPONENT_IDENTIFICATION`.

We identify 2.8 K S7 devices in 75 countries and 501 ASes. Of these, 75.3% are located in ten countries and unlike the other protocols we scanned, we find a higher prevalence of devices in Europe, particularly Poland, which accounted for 37.6% of hosts. Most notably, three S7 devices provided ownership metadata that indicated they belong a European rail provider. Oddly, we find one S7 device with system parameter `PG[random.randint(0,1) f`. This host also serves an HTTPS certificate with an issuer's common name "Nepenthes Development Team", which suggests that the device is a misconfigured instance of the Nepenthes honeypot [8]. We found ten of these honeypots in five ASes.

BACnet We scanned for BACnet devices on March 18, 2016, finding 16.8K devices, of which 13,162 devices (78.3%) provide vendor name, most commonly Reliable Controls (12.7%) and Tridium (10.6%). 13K hosts (78.8%) also volunteer details about the type of device they control: generic controllers (29.6%), SCADA servers (16.7%), HVACs (11.4%), SCADA routers (7.2%), SCADA webcontrollers (4.2%), and power monitors (1.2%). The devices are located across 86 countries and 1330 ASes, most commonly the United States (64%) and France (16%). 44 (3.0%) belong to government entities, 33 (1.9%) belong to medical facilities, and 32 (1.8%) belong to financial institutions.

Tridium Fox We scanned for Tridium Fox hosts by sending a Fox Hello message on March 14, 2016, and identified 26,535 public Tridium Fox hosts, of which 98% are variants on the JACE and NPM line of controllers running the QNX real-time OS. These controllers are manufactured by Tridium and integrated into 3rd party products. Of the devices we can identify, 22K are generic SCADA controllers (97.3%); we note 522 HVACs, 45 solar panels, 40 cinema controllers, and 22 light controllers. Ten countries account for 94.3% of devices; 72.8% are in the U.S. We found 2.6K devices in AT&T ASes, which belong to a plethora of organizations, ranging from airports and defense contractors to utility companies and water treatment facilities. However, we note that the protocol is primarily used for building automation and likely does not

Protocol	December 2015	March 2016	Percent Increase
BACnet	16,752	16,813	0.4%
DNP3	419	429	2.3%
Modbus	21,596	23,120	7.1%
Fox	26,299	26,535	0.9%
S7	2,357	2,798	18.7%

TABLE II: **Change in Vulnerability**—We find an increase in vulnerable hosts between December 2015 and March 2016.

indicate vulnerabilities in processing facilities, whereas the presence of Modbus, S7, and DNP3 hosts might.

Estimating Other Protocol Use We chose to implement one protocol from each category but note that there are several common protocols that we did not develop application scanners for: EtherNet/IP, ICCP, and HART-IP. To estimate a rough upper bound of devices, we completed TCP port scans of the IANA designated port for each protocol. We note a very high number of devices that responded on every port. We filtered out these hosts by performing a secondary scan on unused ephemeral port (TCP/58372) and only counted hosts that responded on the ICS port, but not the ephemeral port. We present the numbers in Table I. There are a few caveats to this methodology. First, Siemens S7, IEC 61850, and ICCP all use port 102, and our methodology cannot distinguish between the protocols on that port. Second, other protocols may use the ports for non-standard protocols. In the end, we see roughly 500–900K responses per port.

Honeypots Some of the hosts we find in our scans are undoubtedly honeypots—similar to the ones we deploy in Section V. As later described, the de facto ICS honeypot—Conpot—deploys a HTTP server on TCP/80 along with Modbus and Siemens S7. This page includes text unique to Conpot, which is easily fingerprintable. To measure the number of honeypots, we completed a full scan of TCP/502 and TCP/102 and then performed a follow-up `GET /` request on port 80. In our scan, 69 S7 devices responded with this page, and 68 Modbus devices responded with the page. A search for the HTTP body on Censys [16] returned one additional host with the Conpot page, but did not host S7 or Modbus, for a total 70 hosts. However, we also find that most conpot instances use the default S7 values, including plant ID "Mouser Factory" and a system parameter "Technodrome". Including these hosts with default values, but no HTTP page, we find a total 103 honeypots. This is nearly 5% of the S7 hosts we found, but less than 0.5% of Modbus hosts. These devices are primarily located in known cloud providers, including Amazon EC2, Digital Ocean, and SingleHop—further indicating that these are user deployed honeypots, not legitimate devices. Approximately 30 of these honeypots belong to our team.

V. Who is Scanning?

While it is well known that Internet-wide scanning is pervasive [17], there is little known about who is scanning for industrial control systems nor what attackers do when they find connected systems. To understand who is searching for ICS devices, we analyze both horizontal scans detected by a passive network telescope as well as observe how scanners interact with 20 high interaction honeypots we deploy.

A. Network Telescope

We tracked who scanned for Modbus, BACnet, DNP3, EtherNet/IP, HART-IP, IEC 61850, ANSI C12.22, ICCP, Siemens S7, and Niagara Tridium Fox by analyzing the packets received by a network telescope composed of nearly one million addresses during August 2015. None of the addresses in our telescope respond to connection attempts and observed packets are typically the result of scanning, misconfiguration, or denial of service backscatter (e.g., [9], [37]). We apply the same methodology as Durumeric et al. [17], in which we assume that scanners pick addresses uniformly at random. During this time, we recorded 1.6 billion packets, from which we identified 22.1 M scans, 2.1 K of which targeted an ICS protocol. The scans targeting ICS protocols account for 0.9% of all scan traffic. We defer to the referenced work [17] for a complete description of the scan detection methodology and its weaknesses. We summarize our results in Table I.

These 2.1 K scans originated from 101 hosts in 34 ASes. We find that most traffic targeted TCP/502—Modbus (41.7%) and BACnet (30.6%). For each of the protocols, less than ten organizations were responsible for 90% of the scan traffic and were part of regular, scheduled scan campaigns. We identify the scanners using the fingerprints developed by Durumeric et al. and note that 79.8% of traffic is from ZMap, 2.4% is generated by Masscan, and 17.8% is from other scanners.

Most of the ICS scan traffic (as determined by packet count) was from Kudelski Security, who regularly scans for Modbus and BACnet devices and was responsible for 52% of all ICS scan traffic. Notably, Kudelski only scans four ports: 123 (NTP, 58.7M packets), 47808 (BACnet, 38.4M packets), 502 (Modbus, 37.2M packets), and port 80 (HTTP, 1.3M packets). A second organization, Shodan, scanned for Modbus, BACnet, TCP/102, DNP3, EtherNet/IP, Niagara Fox, and Hart-IP and account for 19% of traffic. Shodan scanned for an additional 207 ports outside of the ICS space. We find two academic institutions scanning: the University of Michigan, who scanned for Modbus on a scheduled basis, and Reseau National de telecommunications pour la Technologie, who scanned for devices on TCP/102. Hosts in ChinaNet scanned for Modbus, DNP3, Niagara Tridium Fox, EtherNet/IP, and TCP/102, and accounted for 9.2% of traffic. The remaining organizations that regularly scan for ICS are all cloud hosting providers and none of the hosts provide any identifying information (Table III).

B. Low Interaction Honeypots

To understand what attackers do when they find exposed systems, we launched 20 instances of Conpot [44]—an open source SCADA honeypot—on Amazon EC2. Conpot emulates a Siemens SIMATIC S7-200 programmable logic controller and will log requests and in some cases appropriately respond to Modbus (TCP/502), Siemens S7 (TCP/102), BACnet (UDP/47808), HTTP (TCP/80), IPMI (TCP/623), and SNMP

	Modbus	BACnet	TCP/102	DNP3	Ethernet	Fox	Hart	All Protocols
All ICS Traffic	41.7	30.6	8.7	5.1	8.4	3.1	2.4	
Shodan Search Engine	5.1	7.2	24.5	65.5	51.8	71.2	90	18.5
Kudelski Security	61.1	86.2						51.8
Chinanet	4.2		20.3	29.3	19.3	21.2		9.1
University of Michigan	16.2							6.7
SoftLayer Technologies*	3.5				23			3.5
ECATEL/Quasi Networks*	3.8		9.3	2.7	2.8		4.0	2.4
FDC Servers*				1.8	2.2	3.0	3.8	2.5
Amazon EC2*			13					1.1
PlusServer AG*	1.8		8.7					1.6
Reseau National de telecommunications pour la Technologie			5.7					0.5
Ukrainian Data Center*			5.3					0.5
Other	4.3	6.6	13.2	0.7	0.9	4.6	2.2	1.8

TABLE III: **Top Scanners**—We analyze who is scanning for industrial control systems by analyzing the traffic received by a large network telescope in August 2015. * denotes shared hosting provider.

	Modbus	BACnet	Siemens S7	All
All ICS Traffic (total)	1954	520	2778	5252
All ICS Traffic (%)	37.2%	9.9%	52.9%	100%
University of Michigan	18.1%	58.5%	29.2%	27.9%
Shodan Search Engine	23.5%	9.4%	24.1%	22.4%
PlusServer AG*	13.4%	0.2%	6.5%	8.4%
ChinaNet	3.8%	0.0%	12.0%	7.8%
Kudelski Security	13.5%	16.7%	0.0%	6.7%
ECATEL: PLCScan*	10.3%	0.0%	5.0%	6.5%
China169	2.1%	0.0%	8.4%	5.2%
ZNet*	3.1%	2.9%	3.6%	3.3%
ECATEL: Other*	4.0%	3.3%	2.6%	3.2%
Amazon EC2*	1.5%	1.9%	0.0%	1.0%
Rapid7	0.0%	6.5%	0.0%	0.6%
Other	6.7%	0.4%	8.6%	7.0%

TABLE IV: **Top Conpot Scanners**—We analyze what commands scanners operate by running 20 high interaction Conpot honeypots. * denotes shared hosting provider.

(TCP/161). We then ran 20 instances from nonconsecutive addresses in the Amazon EC2 Eastern Region for 10 weeks from December 4, 2015 to February 14, 2016.

We received 2,778 S7 connections, 1,954 Modbus connections, and 520 BACnet connections (Table IV). These connections originated from 338 unique hosts, located in 24 ASes belonging to different organizations. Notably, 55.9% of hosts belong to known security scanning organizations, 35.8% of hosts belong to nondescript ISP organizations in China, and 7.7% of hosts belong to bulletproof hosting providers. For the most part, scanning is temporally consistent. However, we note one outlier: on December 19, 2015, we observed a large spike in S7 connections, originating from 87 hosts in Chinese ISPs. Despite the influx of connections, we only detect relatively benign module information requests. The longitudinal data also reveals a decrease in ICS requests during January 2016, which we primarily attribute to reduced Modbus scanning from Michigan and reduced S7/Modbus scanning from Shodan. It is also possible that non-research organizations are capable of fingerprinting conpot devices to avoid further scanning.

For Modbus, 70% are requests to read device identification and the remaining 30% are *report slave ID* requests, which provides the type of controller and status of a particular attached unit. Half of these targeted unit 0; the other half targeted unit 255. Unit 0 is the first address and unit 255 is typically used to address a gateway itself. In other words, requests attempt to extract information about device models. For Siemens S7, we observe a consistent, but more complex series of commands from scanners. In all cases, scanners attempt to start an outer COTP connection to 0x102 and, if they fail, connect to device 0x200. Next, scanners execute two system status list request commands for additional information: request component information and request module information. We also detected seemingly-benign failed COTP connection requests. While we did not discover any explicit control-oriented commands that alter device state, we suspect that if malicious agents are scanning, they may be first using the diagnostic information-gathering commands that we observed.

Most requests originated from the University of Michigan and Shodan Engine [34], who scan for Modbus, BACnet, and Siemens S7. Unlike Michigan, Shodan issued "report slave ID" commands instead of "read device identification". We similarly observe identification connections to all hosts from Kudelski Security and Rapid7, both security consulting companies. We observe 26 hosts scanning from cloud providers: Amazon EC2, Quasi Networks (ECATEL), SoftLayer Technologies, Zenlayer, Dedicated Panel, PlusServer AG, Leaseweb, Hosting Solutions, and Digital Ocean. Unfortunately, none of the hosts expose any identifying information in reverse DNS or WHOIS records, nor host informative HTTP pages on port 80. Our observations are consistent with the coarse grained network telescope findings.

ICS scanning differs from broad Internet-wide scanning behavior. In the ICS space, there are only a handful of organizations regularly scanning—many of which have web presences. There are a handful of bulletproof hosting providers that are used for scanning, but they are a minority. In comparison, Durumeric et al [17] found Internet-wide scanning consisted of nearly 18K scans targeting more than 1% of the IPv4 address and that scans targeting >10% of the IPv4 address space originated from 350 ASes. More than half of the scans they found could not be identified. This is likely due to the specialized nature of the protocols and differing incentive for attack. While attacking devices can result in catastrophic

	BACnet	Modbus	Fox	S7	DNP3
Our Scans	16,813	23,419	26,535	2,798	429
Shodan	11,911	16,922	20,312	3,828	257
Difference	(41.2%)	(38.4%)	(30.6%)	(-26.9%)	(66.9%)

TABLE V: **Shodan Comparison**—We find significantly different results than previously reported by Shodan.

damage for a company, they have little monetary value to a naïve attacker looking for potential botnet workers or to steal financial credentials.

VI. Related Work

Security concerns around the electric power grid have been present for years [7], [56] and our work illustrates that despite a large body of work surveying these challenges [22], [31], [35], [42], [45], [46] and providing best practices [28], tens of thousands of devices remain vulnerable and the number continues to grow.

There have been three main efforts to address this gap. First, several solutions have been proposed to improve the security of control systems themselves [10], [14], [26], [38], [39]. Second, researchers have sought to build new attack detection capabilities. For example, specialized IDS have been introduced for smart meters [12], [55], home-area networks [25], and process control systems [30], [51]. Finally, researchers have sought to model the security of power grids [32] and SCADA devices [20], [50].

The techniques used in our study draw from a long tradition in other security domains. Network telescopes [37] have been used to detect global scanning [17] and high interaction honeypots [40] have been used to gain deeper insight into attack behavior. Recently, there has been considerable development of ICS-specific honeypots [52]–[54], which we leverage. The most similar honeypot related work to ours [47] describes the authors' experiences with a honeynet deployed on Amazon EC2 that exposed various combinations of the Modbus, DNP3, ICCP, IEC-104, SNMP (v1/2/3), TFTP, and XMPP protocols and collected data over 28 days. In contrast, our work combines both high and low interaction honeypot data and uses scanning to highlight the exposed hosts on the Internet.

There is a long history of using scanning to measure vulnerabilities [6], [18], [33] and in the ICS domain, authors have explained the simplicity of discovering and exploiting ICS devices [36], [48], [49]. Most similar to our work, Shodan [34] has completed regular IPv4 scans for several ICS protocols and several groups have completed further analysis from their dataset [13], [41].

We compared our protocol scans from the week of March 14, 2016 to Shodan, and found drastically different results. Because Shodan does not present a single snapshot in time, it's difficult to compare results, but we find that any one of our single scans generally finds more than the sum of the results that Shodan found from March–April, 2016 (Table V). We do, however, note that we find fewer hosts for Siemens S7. This is likely because many S7 scanners attempt the Cartesian product of two source address and three destination addresses when sending the SZL command, whereas to limit the load on remote S7 devices, we only attempt two of the six possible address combinations. As discussed in Section IV, our methodology only detects 87% of S7 devices compared to attempting all six combinations, which may account for the discrepancy between our results and Shodan. Given these discrepancies, we strongly encourage other research groups to perform their own scans rather than rely on Shodan's results. We have worked with the Censys team [16] to continue running scans for the protocols discussed. The data from weekly scans will be available at https://censys.io/data.

VII. Conclusion

The SCADA protocols in use today were designed over twenty years ago and were originally intended for closed, serial systems. However, despite the lack of built-in security, these protocols have been layered on top of Ethernet and TCP/IP to support long distance communication. In this paper, we first analyzed devices exposed on the public Internet by implementing five popular protocols in ZMap and completing regular scans of the IPv4 address space. We find over 60,000 vulnerable SCADA devices. Unfortunately, most are hosted by large ISPs, making attribution difficult. However, in the cases where we can identify the owner, we observe a worrying glimpse of systems ranging from electrical substations to HVAC installations in government buildings. Second, through a combination of low-interaction and high-interaction honeypots, we characterized who is searching for these vulnerable devices. Unlike the bulk of Internet scan traffic, ICS scanning is dominated by a small handful of known actors ranging including academic institutions and security firms, who scan on a regular basis. By providing an aggregated view into both the current vulnerability landscape and nascent adversarial behavior, we hope to raise the issue of SCADA security and inform future protocol development.

Acknowledgments

The authors thank Frank Li, Vern Paxson, and Jonathan Pevarnek for their advice and help, as well as the exceptional sysadmins at the University of Michigan for their ongoing support. This work was supported in part by the National Science Foundation under awards CNS-1505790, CNS-1530915, CNS-1518741, CNS-1345254, CNS-1409505, CNS-1518888, and CNS-1530915, by the Department of Energy under award DE-OE0000780, by a Google Ph.D. Fellowship, and by an Alfred P. Sloan Foundation Research Fellowship.

References

[1] A DNP3 protocol primer. http://www.dnp.org/AboutUs/DNP3PrimerRevA.pdf.

[2] Fox tunneling and HTTP tunneling. http://www.hvacc.net/pdf/tridium/docs_3.5.25/EngNotes/tunneling/docEn_Tunneling.pdf.

[3] IEEE standard for electric power systems communications — distributed network protocol (DNP3). http://ieeexplore.ieee.org.proxy2.library.illinois.edu/stamp/stamp.jsp?tp=&arnumber=6327578.

[4] Industrial control system survey. http://www.rfidblog.org.uk/Preprint-GallowayHancke-IndustrialControlSurvey.pdf.

[5] Modbus application protocol specification v1.1b3. http://www.modbus.org/docs/Modbus_Application_Protocol_V1_1b3.pdf.

[6] D. Adrian, Z. Durumeric, G. Singh, and J. A. Halderman. Zippier ZMap: Internet-wide scanning at 10 gbps. In *USENIX Workshop on Offensive Technologies*, 2014.

[7] M. Amin. Security challenges for the electricity infrastructure. *Computer*, 35(4):8–10, Apr. 2002.

[8] P. Baecher, M. Koetter, T. Holz, M. Dornseif, and F. Freiling. The Nepenthes platform: An efficient approach to collect malware. In *Recent Advances in Intrusion Detection*. 2006.

[9] M. Bailey, E. Cooke, F. Jahanian, and J. Nazario. The Internet motion sensor: A distributed blackhole monitoring system. In *Network & Distributed System Security Symposium*, 2005.

[10] H. Balasubramanian. Incremental design migration support in industrial control systems development. Master's thesis, Virginia Tech, 2014. https://vtechworks.lib.vt.edu/handle/10919/50990.

[11] D. Beresford. Exploiting Siemens Simatic S7 PLCs, July 2011. http://www.cse.psu.edu/~sem284/cse598e-f11/papers/beresford.pdf.

[12] R. Berthier and W. H. Sanders. Specification-based intrusion detection for advanced metering infrastructures. In *Pacific Rim Intl. Symposium on Dependable Computing*. IEEE, 2011.

[13] R. C. Bodenheim. Impact of the Shodan computer search engine on internet-facing industrial control system devices. Technical report, 2014. http://oai.dtic.mil/oai/oai%3Fverb%3DgetRecord%26metadataPrefix%3Dhtml%26identifier%3DADA601219.

[14] A. Costin, J. Zaddach, A. Francillon, and D. Balzarotti. A large-scale analysis of the security of embedded firmwares. In *USENIX Security Symposium*, 2014.

[15] J. A. Crain and S. Bratus. Bolt-On Security Extensions for Industrial Control System Protocols: A Case Study of DNP3 SAv5. *IEEE Security & Privacy*, (3):74–79, May 2015. http://dx.doi.org/10.1109/msp.2015.47.

[16] Z. Durumeric, D. Adrian, A. Mirian, M. Bailey, and J. A. Halderman. A search engine backed by Internet-wide scanning. In *ACM Conference on Computer and Communications Security*, Oct. 2015.

[17] Z. Durumeric, M. Bailey, and J. A. Halderman. An Internet-wide view of Internet-wide scanning. In *USENIX Security Symposium*, 2014.

[18] Z. Durumeric, E. Wustrow, and J. A. Halderman. ZMap: Fast Internet-wide scanning and its security applications. In *22nd USENIX Security Symposium*, Aug. 2013.

[19] G. Ericsson and A. Johnsson. Examination of ELCOM-90, TASE. 1, and ICCP/TASE. 2 for inter-control center communication. *Power Delivery, IEEE Transactions on*, 12(2):607–615, 1997.

[20] A. Giani, G. Karsai, T. Roosta, A. Shah, B. Sinopoli, and J. Wiley. A testbed for secure and robust SCADA systems. *SIGBED*, 2008.

[21] S. Hilt. s7-info NSE script. https://nmap.org/nsedoc/scripts/s7-info.html.

[22] V. M. Igure, S. A. Laughter, and R. D. Williams. Security issues in SCADA networks. *Computers & Security*, 25(7):498–506, Oct. 2006.

[23] International Electrotechnical Commission and others. IEC 61158: Digital data communications for measurement and control—Fieldbus for use in industrial control systems, 2003.

[24] International Electrotechnical Commission and others. IEC 61784-1: Digital data communications for measurement and control - part 1: Profile sets for continuous and discrete manufacturing relative to fieldbus use in industrial control systems. 1, 2003.

[25] P. Jokar, H. Nicanfar, and V. Leung. Specification-based intrusion detection for home area networks in smart grids. In *Smart Grid Communications (SmartGridComm), 2011 IEEE International Conference on*, pages 208–213. IEEE, 2011.

[26] D.-H. Kang, B.-K. Kim, and J.-C. Na. Cyber threats and defence approaches in SCADA systems. In *Conference on Advanced Communication Technology*. IEEE, 2014.

[27] J. Kaur, J. Tonejc, S. Wendzel, and M. Meier. Securing BACnet's pitfalls, June 2015. http://www.researchgate.net/publication/273773837_Securing_BACnets_pitfalls.

[28] W. Knowles, J. M. Such, A. Gouglidis, G. Misra, and A. Rashid. Assurance Techniques for Industrial Control Systems (ICS). In *Workshop on Cyber-Physical Systems-Security and/or PrivaCy*. ACM, 2015.

[29] F. Li, Z. Durumeric, J. Czyz, M. Karami, M. Bailey, D. McCoy, S. Savage, and V. Paxson. You've got vulnerability: Exploring effective vulnerability notifications. In *USENIX Security Symposium*, 2016.

[30] H. Lin, A. Slagell, C. Di Martino, Z. Kalbarczyk, and R. K. Iyer. Adapting Bro into SCADA: Building a specification-based intrusion detection system for the DNP3 protocol. In *Cyber Security and Information Intelligence Research Workshop*. ACM, 2013.

[31] J. Liu, Y. Xiao, S. Li, W. Liang, and C. L. P. Chen. Cyber security and privacy issues in smart grids. *Communications Surveys & Tutorials*, 2012.

[32] R. Liu, C. Vellaithurai, S. S. Biswas, T. T. Gamage, and A. K. Srivastava. Analyzing the Cyber-Physical Impact of Cyber Events on the Power Grid. *Smart Grid, IEEE Transactions on*, 6(5):2444–2453, 2015.

[33] G. F. Lyon. *Nmap Network Scanning: The Official Nmap Project Guide to Network Discovery and Security Scanning*. Insecure.org, USA, 2009.

[34] J. C. Matherly. Shodan search engine. https://www.shodan.io.

[35] D. Maynor and R. Graham. SCADA security and terrorism: We're not crying wolf. 2006. https://www.blackhat.com/presentations/bh-federal-06/BH-Fed-06-Maynor-Graham-up.pdf.

[36] B. Meixell and E. Forner. Out of control: Demonstrating SCADA exploitation. 2013. https://media.blackhat.com/us-13/US-13-Forner-Out-of-Control-Demonstrating-SCADA-Slides.pdf.

[37] D. Moore, G. M. Voelker, and S. Savage. Inferring Internet denial-of-service activity. In *Proceedings of the Tenth USENIX Security Symposium*, pages 9–22, Washington, D.C., Aug. 2001.

[38] D. M. Nicol, W. H. Sanders, S. Singh, and M. Seri. Usable Global Network Access Policy for Process Control Systems. *Security & Privacy, IEEE*, 6(6):30–36, Nov. 2008.

[39] S. C. Patel, G. D. Bhatt, and J. H. Graham. Improving the Cyber Security of SCADA Communication Networks. *Commun. ACM*, 52(7):139–142, July 2009.

[40] N. Provos and T. Holz. *Virtual Honeypots: From Botnet Tracking to Intrusion Detection*. Addison-Wesley Professional, first edition, 2007.

[41] B. Radvanovsky. Project SHINE findings report. 2014. http://www.slideshare.net/BobRadvanovsky/project-shine-findings-report-dated-1oct2014.

[42] P. A. S. Ralston, J. H. Graham, and J. L. Hieb. Cyber security risk assessment for SCADA and DCS networks. *ISA Transactions*, 2007.

[43] I. R.E. Mackiewicz, Member. Overview of IEC 61850 and benefits. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1709546&tag=1.

[44] L. Rift, J. Vastergaard, D. Haslinger, A. Pasquale, and J. Smith. CONPOT ICS/SCADA honeypot. http://conpot.org.

[45] SANS Institute. The state of security in control systems today. 2015. https://www.sans.org/reading-room/whitepapers/analyst/state-security-control-systems-today-36042.

[46] A. Sarwate. SCADA security: Why is it so hard? 2011. https://media.blackhat.com/bh-ad-11/Sarwate/bh-ad-11-Amol_SCADA_WP.pdf.

[47] A. V. Serbanescu, S. Obermeier, and D.-Y. Yu. ICS threat analysis using a large-scale honeynet. In H. Janicke and K. Jones, editors, *ICS-CSR*, Workshops in Computing. BCS, 2015.

[48] A. Soullié. Industrial control systems: Pentesting PLCs 101. 2014. https://www.blackhat.com/docs/eu-14/materials/eu-14-Soullie-Industrial-Control-Systems-Pentesting-PLCs-101.pdf.

[49] R. Speers, T. Goodspeed, I. R. Jenkins, R. Shapiro, and S. Bratus. Fingerprinting IEEE 802.15. 4 devices with commodity radios. Technical report, Dartmouth Computer Science, 2014. http://www.cs.dartmouth.edu/reports/TR2014-746-rev2.pdf.

[50] C.-W. Ten, C.-C. Liu, and M. Govindarasu. Vulnerability assessment of cybersecurity for SCADA systems using attack trees. In *Power Engineering Society General Meeting*. IEEE, 2007.

[51] A. Valdes and S. Cheung. Intrusion monitoring in process control systems. In *System Sciences, 2009. HICSS'09. 42nd Hawaii International Conference on*, pages 1–7. IEEE, 2009.

[52] E. Vasilomanolakis, S. Srinivasa, and M. Mühlhäuser. Did you really hack a nuclear power plant? An industrial control mobile honeypot. In *IEEE Conference on Communications and Network Security*, 2015.

[53] K. Wilhoit. Who's really attacking your ICS equipment? *Trend Micro*, 2013. http://www.trendmicro.com/cloud-content/us/pdfs/security-intelligence/white-papers/wp-whos-really-attacking-your-ics-equipment.pdf.

[54] M. M. Winn. Constructing cost-effective and targetable ICS honeypots suited for production networks. Master's thesis, Air Force Institute of Technology, 2015. http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA615223.

[55] Y. Zhang, L. Wang, W. Sun, R. C. Green, M. Alam, et al. Distributed intrusion detection system in a multi-layer network architecture of smart grids. *Smart Grid, IEEE Transactions on*, 2(4):796–808, 2011.

[56] B. Zhu, A. Joseph, and S. Sastry. A taxonomy of cyber attacks on SCADA systems. In *International Conference on Internet of Things*. IEEE, 2011.