Project SHINE
(SHodan INtelligence Extraction)

Findings Report

Based on intelligence gathered from the SHODAN search engine between 14 Apr 2012 through 31 Jan 2014

1 Oct
2014

This following findings report contains specific details of devices that are directly connected to the Internet that may be utilized for mission critical operations associated to one (or more) critical infrastructure sectors (and their respective industries). Information contained within this report should only be used for awareness purposes.

LEGAL DISCLAIMER

Project SHINE is a research project designed to harvest and ingest data used to correlate potential threats and risk associated with SCADA and industrial control system devices that appear to be directly connected to the Internet. The project is a collaborative effort of individuals and organizations to raise public awareness of such devices that may impact one (or more) critical infrastructure sectors (and their respective industries), while demonstrating the level of magnitude of the quantity of these devices that are publicly accessible through the Internet.

Table of Contents

Table of Contents (continued)

Table of Contents (continued)

Figures and Tables

Table of Contents (continued)

Figures and Tables (continued)

Acknowledgements

Infracritical wishes to express its sincerest gratitude by acknowledging those who have contributed and helped Project SHINE become a visible public awareness tool. As such, we wish to thank those who have participated in this effort for their contributions and support.

Contact Information

For more information about Project SHINE, please send correspondence to:

Project SHINE Inquiries
projectshine@infracritical.com

Preface

Many infrastructure security professionals have wondered as to what extent critical systems are exposed to the Internet. That is how this project got started. No one within the project team had any idea that they would encounter as much exposure as was discovered. The raw data that was accumulated is potentially dangerous information. Thus, we felt that full disclosure of any specific details about an infrastructure could be used as an attack vector against that infrastructure.

As the team has an ethical obligation to make the public aware of this situation, there were several attempts to write this document in a reasonable manner. As outlined within the tale of the *Emperor's New Clothes* by Hans Christian Andersen, there is no nice way to publicly state what needed to be said. The team is doing their best to discuss, with discretion and candor, a rampant problem of public infrastructure exposure.

Should any organization, vendor, integrator, government, or individual decide to squelch this report, the authors would like to point out that the data came strictly from publicly available sources, without any subterfuge. This is a technical problem, not a legal problem. Using the legal profession to attempt to silence this report will not address any problems identified here; it will only bring derision and ill will to those who attempt to sequester either the authors, or this report, and will provide nothing to remove any vulnerability or risk. Such efforts will merely discourage any future cooperative efforts. Additionally, this may encourage people with similar technical backgrounds to take a less open approach towards the infrastructure community.

The team was compelled to provide meaningful and relevant information from the project, outlined within this report, without compromising either the asset owners or the manufacturers of those assets.

The authors did not create this problem.

This paper is a message.

We feel a community effort should begin somewhere, starting first with public awareness.

Introduction

There have been quiet rumors over the last decade that many control systems' assets, such as remote terminal units (RTU) or programmable logic controllers (PLC) related to infrastructure, were exposed to the Internet. Information about the problem was speculative, as some have debated over whether or not these assets were lightly configured, or if their scope of exposure to the Internet existed.

Many were curious about what and where these systems were. The problem is that, although the methods of reaching out and touching these systems are not a secret, there was concern of liability. With the myriad of legal systems around World, there is no way of knowing what any given court might consider a hostile action, regardless of intentions.

Thus, a different approach became necessary.

What does "Lightly Configured" Mean?

Many control systems' implementations are usually integrated on a very tight budget. It is generally common practice that most integrators only configure SCADA/control systems to the extent necessary to get these systems working. Many SCADA/control systems today have web (HTTP), file transfer (FTP), network management (SNMP), or Telnet protocol features enabled by default – especially web servers. This means that if one of these systems were accidentally exposed to the Internet, it would probably be found, if from nothing else, a web page generated by the system's internal web server.

It is reasonable to assume that most search engines (Yahoo^{tm}, Google^{tm}, etc.) have probably discovered these devices. All anyone would have to do is to design a very selective query to find these systems.

Exposure to the Internet

Assuming that anyone could identify unique key words that could single out these systems, and assuming that anyone could remove duplicates from domain name server (DNS) aliases, it should be possible to find these systems on the Internet. The initial thought that the team had was to use any search engine, de-duplicate their addresses by running a reverse DNS lookup, then store everything indexed by IP addresses and search terms.

This approach has two interesting and frightening traits:

(1) The scanning has already been done. The researchers do not need to touch any system.

(2) The device owners have no way of knowing that they've been found. The only thing in their logs (assuming such logs exist) would be a search engine spider. Anyone exposed to the Internet should expect that much.

Initially, the team had no idea of the scope, or magnitude, as to how extensive this issue was.

Many things can disturb these sectors, but one very high-profiled method, is the use of computer malware to attack or destroy these cyber assets. The likelihood of someone doing this that an increasing number of these electronically-controlled devices are becoming known of being exposed to the Internet, with the possibility that someone with criminal intent or a political agenda somewhere on this planet might act against these devices is significant.

Development of a defense-in-depth strategy is crucial in securing these systems, as many systems which were once connected through private or closed-loop communications networks, are now connected through a public network. One speculative theory discussed over the years is that many private or closed-loop networks have vanished and have been replaced with newer, more modern communications equipment, in many cases, now directly connected to the Internet. Many organizations either have little to no knowledge that their once privately-connected systems may now be directly connected to the Internet, or may have little or no efforts ensuring that these systems are protected in any manner whatsoever.

What is Project SHINE?

Project SHINE ("SHINE" is an acronym meaning "SHodan INtelligence Extraction") was created to extrapolate metadata from the SHODAN search engine, which is a custom search engine designed for searching embedded devices (routers, servers, etc.) or computer systems based on a searchable term criteria set, returning service port header information. Some have also described it as a search engine of service banners in which metadata is enumerating and identified using these embedded devices or computer systems, providing additional metadata information such as contained within the following data sets:

- IP address of the device;
- Location (including latitude and longitude coordinates, if available);
- Country of origin;
- Owner of the IP address; and,
- Service port header information.

Header information may be useful in identifying information such as (but not limited to) metadata containing the following:

- Protocol type (e.g.; HTTP, SNMP, Telnet, etc.);
- Version and/or firmware release levels;
- Content expiration date and time; and,
- Other lesser known metadata values that may be directly correlated to other forms of metadata which can be utilized as aggregated data, allowing differentiation of various manufacturers of devices/software, but may use similar software and/or firmware.

Any of this metadata information may be considered of great value for a number of purposes.

Project SHINE attempted to discover if any specific organizations that were discovered by SHODAN, but to our dismay, found that majority of the IP addresses were found to be leased from Internet Service Providers (ISP), and acquiring any such additional information from those ISPs would be next to impossible, as this would violate privacy issues with the ISP's clients; we tried on several occasions in contacting one of the responsible ISPs for several dozen SCADA/control systems devices found directly connected to the Internet, inquiring the asset owner's name and were unsuccessful in our attempts.

SHODAN's information may be considered extremely valuable when attempting to discover embedded devices or computer systems that an organization may own that are directly connected to the Internet in an effort to protect their cyber assets from either tampering, manipulation of data (either input and/or output from said device), loss of integrity of its operation, disruption of service to its operation, or worse yet, total destruction of said device and/or any services or other operations relying upon said device.

On the other hand, the SHODAN search engine provides an easy one-stop shopping place to conduct a potential open source intelligence gathering capability to which either target a non-specific location (for instance, all embedded devices or computer systems operating within the United States), or a slightly more specific location (for instance, the Eastern Region of the United States), to specific organizations, or to specific devices (that may be utilized by one or more organization), provided of course, that the location data is accurate. This metadata may be utilized against any one of the parties outlined above for purposes not previously disclosed; yet, however, many may be able to ascertain that those reasons may include several purposes considered as an act of terrorism against either the United States or friendly countries aligned with the United States. Nonetheless, having such a tool available, and if sufficient knowledge may be obtained, has the potential to cause and wreck havoc against not only private infrastructure asset owners but public sectored organizations as well.

This method of intelligence gathering is often described as "open source intelligence" (or "OSINT")[1], and has been revitalized as a renewed method for gathering and collecting intelligence metadata. In most circumstances, open source collection methods are used to compare and aggregate against closed source collection methods; thus, utilizing the publicly available open source metadata to provide a comparative analysis against an existing state or function of a given system or environment, allows such correlation to reduce and remove any errors that may exist from the metadata collected. This provides a very powerful and useful tool when conducting such research; however, open source intelligence sources are not always accurate, and thus, is an issue when relying solely upon an open source as the only method of gathering and collecting metadata intelligence.

[1] Open-source intelligence (OSINT) is intelligence gathered and collected from publicly available sources. Within the intelligence community, the term "open" refers to overt, publicly available sources (as opposed to covert or clandestine sources); this term is not related to open-source software or public intelligence. Open-source intelligence under one name or another has been around for many years. The significance of OSINT today, both within the United States (and abroad) is the conflict between military, government, and the private sector intelligence gathering methods as to how the bulk of intelligence should be obtained. With the Internet, instant communications, and advanced media search of bulk metadata that can be easily obtained through actionable and predictive methods from public, unclassified sources. Government agencies have been slow to embrace OSINT, or believe they have suitable information feeds from the media, academia and public records.

One the other hand, if comparative analysis can provide deterministic data modeling, and thus reduce (if not remove) any extraneous metadata that was considered incorrect or erroneous in nature, this can provide a useful tool that is relatively cheap in terms of cost and resources utilized.

A search engine performs regular interrogations (scanning) of embedded devices and/or computer systems found throughout the World; it harvests metadata pertaining to them, and ingests the metadata into massive databases. Information is sorted based on a criteria set (as outlined on the previous page) and periodically interrogated by the search engine for any modification. Some devices will often show multiple scanning interrogational activities, indicating that the search engine is verifying that the device scanned is still online and available.

At no point during the activities of Project SHINE did we ever perform any scanning, or attempt to directly access any of the embedded devices and/or computer systems connected to the Internet.

Concerns with Project SHINE

Some of our concerns from our findings may include (but may not be limited to) the following:

- Lack of knowledge of devices contained within their own operations;
- Lack of knowledge of how to secure their devices within their own operations, or organizations may know, but are may not perceive the potential risk costing more than the mitigation;
- Requesting constant confirmation of asset owners' IP addresses against the Project SHINE's metadata database (indicating perhaps adherence to a compliance-driven model);
- Justification that, if no IP addresses are found within the Project SHINE metadata database, that devices are not connected to the Internet, and that the asset owners' operations are clear from any potential attack, threat or vulnerability;
- Asset owners' devices could potentially be quietly compromised by a forgotten backdoor left behind by a third-party contractor, as no one would know if it existed;
- Expectation that Project SHINE will share, either partially or in its entirety, the Project SHINE metadata extracted from the SHODAN search engine; and,
- Expectation that Project SHINE will share one (or more) of any search term(s) from the criteria set used to extract metadata from the SHODAN search engine.

Many of these concerns appear to be directly correlated to a compliance-based model (as opposed to a risk-based or consequence-based model), ensuring that an asset owners' operations is in adherence with whatever regulatory body/organization may exist for that industry, or if no regulations exist, ensure stockholders are assured of no impacts to their operations (thus, potentially devaluing the asset owners' stock prices). In either case, (observationally) this raises concerns that asset owners are not willing to secure their environments unless they are specifically told to do so (legally required by the regulatory body/organization or through their stockholders). Many asset owners will often perform security upgrades only if it is absolutely necessary, or if something severe were to occur (e.g.; such as an explosion or a leak of a toxic substance).

One observation appears to be consistent with more than one critical infrastructure (and its industries), further re-enforcing any potential liabilities that their environments are secure, or that the lack of knowledge of being capable of performing such tasks to secure their operations, and if required to secure their operations, often utilizing external resources to perform such tasks, may potential create risk of any third-party representative to introduce additional risks (e.g.; such as performing an act of sabotage, introducing hidden "backdoors"[2], etc.) into their operations.

Some industries have recently considered either utilizing a consequence-based model[3], or some form of hybridized modeling methodology involving consequence-based modeling and some other form of modeling (with the exception of compliance-based modeling). The use of consequence-based modeling may be considered difficult in terms of predictability; nonetheless, several infrastructure sectors (and their respective industries) are considering such actions, as the compliance-based modeling methodology may be insufficient – more importantly – too costly, to effectively provide a secured methodology for the enterprise and its respective systems. Reasons for switching to utilizing such modeling methodologies may be due to costs, time, resources, or some form of combination thereof, as some feel that compliance-based modeling (overall) may be too expensive for long-term expenditures of an organization's capital funding, and may impact the overall net worth value of the organization. Such considerations are being evaluated today, as return-on-investments (ROI) appear to be flattening out due to increased expenditures in several aspects, which may include factors such as security based on a compliance-based model.

Role and Function of Project SHINE

Project SHINE was created with the intention of defining a searchable term criteria set metadata database that could be modified easily to establish a census baseline of all SCADA/control systems discovered through the SHODAN search engine via the Internet. The project is solely dependent upon the SHODAN search engine. Some of the problems that we have encountered through the project include (but are not limited to):

- Downtime of the SHODAN search engine (both announced and un-announced);
- Disk space availability of data within the SHODAN database; and,
- Devices (possibly) repeatedly discovered via SHODAN with different IP addresses.

[2] A "backdoor" in a SCADA, control system, computer system, et. al is a method of bypassing normal accessing methods (secured or otherwise), ensuring unauthorized access to a given system, while attempting to remain undetected for access to that system.

[3] Consequence-based modeling is used to determine consequential impacts of a given system based upon an event or incident, either intentional or accidental, and the effects resulting following the event or incident. Such a model may be used to provide predictive analysis utilizing specific criteria sets based on scenarios conducted that are used to determine such outcomes. While there may be high interest in utilizing an alternative modeling methodology, the use of consequence-based modeling may be difficult to implement, while providing useful outcomes, as many factors used for the model may be dynamically changing, thus making such predictive analysis equally as difficult.

The whole premise behind the project was to perform and demonstrate the following:

- The ability to selectively perform searches from one (or more) definable, searchable term criteria sets from (one or more) open intelligence sources;
- To correlate data into meaningful abstractive and relevant data that could be utilized to demonstrate further trending and correlation analysis based on the data given; and,
- To seek a baseline of just how many SCADA/control systems devices exist on the Internet (as of the conclusion of Project SHINE on 31-Jan-2014, we were unable to establish a baseline from which we could compare rates of growth).

It was our hope that we would be able to ascertain a baseline of devices discovered using the SHODAN search engine, and then measure any changes as the number of devices grew or shrank over a period of time. As the project was never able to accomplish this goal, we were not entirely certain why this was the case. Henceforth, this was the primary reason why Project SHINE was discontinued on 31-Jan-2014.

Search Term Selection

The crux of this exercise was choosing a search term that identified control systems devices. The team was looking for not just control systems, but also any infrastructure that supports it. For example, an uninterruptable power supply (UPS) is often a component of operations control centers. Generator controls and transfer switches were also common support elements. Serial port servers were commonplace with many control systems, as are certain brands of networking switches.

There are also certain keywords (such as MODBUS, PROFINET, etc.) that are often a search term for snagging devices that might otherwise have been missed. For example, a networking switch with a web interface that mentions IGMP snooping (e.g.; used with Ethernet/IP^{tm}) is probably a control systems' switch. However, there are many switch models and it is easy to overlook one.

Data Artifacts

Clearly, there will be other matches that are not related to an actual infrastructure. There may be demonstration units, used for marketing or testing purposes. Some units won't be counted because manufacturer's names change as firms are bought and sold all the time. Other devices may be spuriously counted because the same embedded code may be used by multiple manufacturers.

Also note that several search terms may apply to the same device. For example, an RTU with the ability to act as a MODBUS master might trigger on the search terms of the RTU make and model, as well as the search term of MODBUS.

This method depends upon the “lightly configured” assumption, that from a first-hand experience, the team knew that this assumption is more than likely. However, it is assumed that there may be a few asset owners who may actually tighten up their configurations once discovered.

One consideration is that many devices are hidden behind network address translation (NAT) firewalls. It should not be unexpected that multiple search terms indicating different devices might be radiating from a single address. This is why the team made no effort to de-duplicate addresses with different search terms. Thus, a “hit” is an address that has a search term discovered. Multiple hits on a single address are not only possible, but actually somewhat likely.

Because of these limitations, this study is only intended as a gross first order estimate of the scope of the problem. The reader of this report should not infer too much detail into the data presented here as it is based upon assumptions that are probably not as solidified as one might like.

Found Discoveries

The team had to spot-check some of the discoveries, mostly to verify that the search terms were picking up what was sought. These spot-checks consisted of plugging a sample IP address in to the location bar of a web browser. In most cases, something responded back. Occasionally, such interrogation would yield graphics that would identify who the owner of the unit was. Unfortunately, due to the sheer volume of responses, checking every IP address was impractical.

Another discovery should have been obvious from the start: the team attempted quite a few reverse address lookups. Many of the hits led to dynamically assigned ISP addresses. In other words, it was almost impossible to know from the address alone what something belonged to.

Nevertheless, the Project SHINE stumbled across some interesting things: mining equipment, a surprising number of wind farms, a crematorium, a few water utilities, several substations, and many innocuous accessories that are commonly used for control systems, such as serial port servers, UPS equipment, HVAC systems, etc.

What is SHODAN?

Unlike most traditional search engines (such as Google[tm], Yahoo[tm] or Bing[tm]), the SHODAN search engine is a very custom and specialized search engine that lets users find specific types of embedded devices and computer systems (routers, servers, etc.) using a variety of filters and searchable terms (criteria sets) that are directly connected to the Internet. Several individuals have defined this search engine as a specialized search engine that harvests and ingests metadata from service banners. Service banners can provide some relevant information about the type of device or software being utilized, what version of software or firmware exists, and if there are any conditions or flags that currently exist, often indicating limitations of the embedded devices or computer systems scanned as to its functionality, identify any possible options that the service supports, perhaps a welcome message or anything else that the client would like to know before interacting with that embedded device or computer system. More importantly, what makes SHODAN important is, based on the software or firmware levels identified can help determine any potential vulnerability which may exist for these embedded devices or computer systems; this is (probably) the most significant reason of concern for the SHODAN search engine. This was recently demonstrated when the "Heartbleed"[4] vulnerability was identified in early April 2014, utilizing SHODAN as one method of identifying systems through a openly public venue that were vulnerable to the Heartbleed vulnerability.

The name "SHODAN" is a reference to the *"Sentient Hyper-Optimized Data Access Network"* ("SHODAN")[5]and is a fictional artificial intelligence with the main antagonist of the cyberpunk-horror themed action role-playing video game series called System Shock and System Shock 2.

The search engine began as a side project for John Matherly who determined that it would be interesting in determining how many embedded devices or computer systems were connected to the Internet. SHODAN users are able to find systems of all types including (but not limited to) devices such as traffic lights, security cameras, home heating systems, control systems for water parks, gas stations, water plants, power substations, and much more.

This is both a positive and negative aspect of the SHODAN search engine in that – depending on its users' intent – may be construed as either a tool (positive) or as a weapon (negative). The fact is, SHODAN may not necessary be the problem for devices discovered and harvested by the search engine; however, it makes its availability of metadata that (otherwise) was not publicly or openly available as now a freely available tool for adversaries with a one-stop location to perform surveillance and intelligence gathering methods against (perhaps) a specific target or organization.

[4] Heartbleed is a security bug within the open-source OpenSSL cryptographic library, which is widely used to implement the Internet's Transport Layer Security (TLS) protocol. This vulnerability, classified as a buffer over-read, results from a missing bounds check in the handling of the Transport Layer Security (TLS) heartbeat extension, the heartbeat being behind the bug's name.
[5] http://en.wikipedia.org/wiki/SHODAN

What Service Ports does SHODAN Look For?

Normally, SHODAN harvests and ingests metadata information from several service ports:

- File Transfer Protocol (FTP) – port 21
- Telnet (TELNET) – port 23
- Unsecured Web (HTTP) – port 80
- Secured Shell (SSH) – port 22
- Simple Network Mgmt Protocol – port 161
- Secured Web (HTTPS) – port 443

On a concerning note, there have been several discussions from several media sources, blogs and public forums indicating whether or not the SHODAN search engine is or will locate other service ports or banners from lesser known protocols. This, too, can be considered useful for both positive and negative reasons.

Recently, there have been five (5) SCADA/control systems ports discovered that are scanned and metadata collected by the SHODAN search engine, which include:

- Siemens SIMATIC / ICCP – port 102
- DNP3 – port 20000
- BACNet – port 47808
- MODBUS/TCP – port 502
- Ethernet/IP – port 44818

For example, if additional service port information were included that provided commented header information relevant to a specific company name (e.g.; "CoGen, Inc."), location of the plant or base of operation (e.g.; "Western Region Pump Location A-11, City of Metdata, Metadata County, US"), function of the operation (e.g.; "Waste Extraction Unit #3, Pump #102"), this could provide additional information that adversaries could now have available at their fingertips. Through this method, they wouldn't need the IP addresses (or address range) to quickly determine what or who was their target or organization, and with some further analysis, could ascertain functions, roles, or perhaps, even the individuals responsible for the now recently identified operations. All of this is openly, freely and publicly available via SHODAN.

One note that we would like to point out, are that the service ports of the SCADA/control systems scanned by SHODAN, are not used solely by SCADA/control systems, as those services ports are utilized by other non-SCADA/control systems specific functionalities. These service ports have been defined for other service-specific roles and functions pertaining to network operations within an enterprise or Internet-connected systems.

Results from Project SHINE: Manufacturers

Some statistics of SCADA/control systems, as well as embedded device manufacturers are shown below. Breakdown is based on manufacturer's name, and count of number of search term criteria sets for each manufacturer, along with a globalized count, an adjusted globalized count based on redundancies, an adjusted globalized count based on manufacturers that do not fall under traditional SCADA/control systems manufacturers, and their redundancies.

Redundancies would be search term criteria sets in which there are globalized search terms, along with more specific search terms, based on model number, or some other relevant data that would make the search term more unique. Information is provided in tabular format, and shown on subsequent pages.

Manufacturer Results

As of the end of the project on 31-Jan-2014, we have discovered (roughly) 927 search term criteria sets that the project has monitored since the project's inception on 12-Apr-2012.

Approximately 886 search term criteria sets were considered as unique search terms ("<manufacturer>" [non-unique] versus "<manufacturer>+<specific device>" [unique]), and removing what some might consider non-traditional SCADA/control systems, reduced the count to (roughly) 578 total.

There are a total number of 207 manufacturers monitored, and with the reduction of non-traditional SCADA/control systems, reduces the count to (roughly) 182 manufacturers total; a difference of 25 manufacturers who might not be considered traditional SCADA/control systems manufacturers.

927 total number of search terms	**886** unique number of search terms; traditional and non-traditional	**578** unique number of search terms; non-traditional removed	**207** total number of manufacturers	**182** unique number of traditional manufacturers

Figure 1.

41 total search term difference; traditional and non-traditional	**349** total search term difference; non-traditional removed	**25** total manufacturer difference

Figure 2.

Although the total number of manufacturers monitored may appear to be paltry in its count in comparison to the total of manufacturers in the information technology industry, the number of manufacturers is by no means complete, as there are thousands of manufacturers who specialize in SCADA/control systems in some fashion thereof.

Those manufacturers, who were discovered, were found through several sources, such as trade magazines, blogging web sites, etc. The premise behind the project was to gauge, through sampling, any trending (if any) of the total number of devices discovered through the SHODAN search engine.

The project was started with the notion that eventually the team would notice that the harvesting process was not occurring at a rate seen initially. When that point was reached, one could presume most of the assets had been discovered.

Unfortunately, that didn't happen. There were two reasons for this. First, it was presumed that whatever was found wouldn't be removed from service any time soon and new exposures wouldn't happen quickly. This would be wrong. Second, search terms were being added frequently all throughout the project. Even when examining individual search terms, there appeared to be no leveling off of the rate of ingestion. Some of this problem may have been an artifact of SHODAN's data gathering methods being quite irregular, not having been designed for this purpose. Nonetheless, our efforts proved unsuccessful, as the project was unable to obtain a baseline.

Manufacturers Discovered – Top 11 Manufacturers

As of 31-Jan-2014, there are eleven (11) top manufacturers identified that are used (in some form or another) for SCADA/control systems environments, which include:

- ALLIED-TELESYS
- DIGI INTERNATIONAL
- EMBEDTHIS-WEB
- ENERGYICT
- GOAHEAD-WEB
- INTOTO
- LANTRONIX
- MOXA
- NIAGARA
- SIEMENS
- VXWORKS

The devices discovered are estimated at 586,997, roughly 26.84% of the 2,186,971 devices sampled. Some of the devices discovered are widely used; therefore, the total count is estimated for this reason.

Manufacturer	Count	% Out of 100%
ENERGYICT	106235	18.10%
SIEMENS	84328	14.37%
MOXA	78309	13.34%
LANTRONIX	56239	9.58%
NIAGARA	54437	9.27%
GOAHEAD-WEB	42473	7.24%
VXWORKS	34759	5.92%
INTOTO	34686	5.91%
ALLIED-TELESYS	34573	5.89%
DIGI INTERNATIONAL	30557	5.21%
EMBEDTHIS-WEB	30381	5.17%

Figure 3.

Results from Project SHINE: HVAC/Building Automation

Some statistics of SCADA/control systems are broken down based on HVAC and building automation control environments. Data acquired from the SHODAN search engine confirms that SHODAN is periodically scanning and collecting metadata for these devices.

HVAC/Building Automation Manufacturer Results

As of 31-Jan-2014, there are sixteen (16) HVAC/BACNet manufacturers currently being scanned and metadata collected for SCADA/control systems specific devices, which include:

- BACNET INTERNATIONAL
- BOSCH AUTOMATION
- CARRIER
- CENTRALINE
- CONTROL4
- CUMMINGS
- HEATMISER
- HONEYWELL
- JOHNSON CONTROLS
- LENNOX
- LG ELECTRONICS
- LIEBERT
- STULZ
- TEMPERATURE GUARD
- TRANE
- YORK

The total number of HVAC/BACNet devices discovered is estimated at approximately 13,475, roughly 0.62% of the 2,186,971 devices sampled. This does not include other software that may control HVAC and/or building automation control system environments (such as Niagara/AX, Windweb, EmbedThis, etc.); however, as these software manufacturers are utilized for functions in addition to HVAC and building automation, those figures would pollute the overall tally counted. Therefore, the total count is estimated for this reason.

Manufacturer	Count	% Out of 100%
HEATMISER	6487	48.13%
HONEYWELL	3588	26.63%
YORK	921	6.83%
BACNET INTERNATIONAL	560	4.16%
TRANE	506	3.76%
JOHNSON CONTROLS	460	3.41%
CARRIER	234	1.74%
TEMPERATURE GUARD	180	1.34%
LG ELECTRONICS	145	1.08%
LIEBERT	126	0.94%
CENTRALINE	81	0.60%
STULZ	77	0.57%
CONTROL4	38	0.28%
BOSCH AUTOMATION	37	0.27%
LENNOX	24	0.18%
CUMMINGS	11	0.08%

Figure 4.

One reason for listing HVAC/building automation control systems is this is a potential attack vector for accessing other assets. This is an emerging attack vector[6], as many asset owners often incorporate their building automation control system environments with the remainder of the enterprise as: (1) a convenience of utilizing the enterprise to access these devices; and (2) as asset owners have not conceived this as a potential threat vector:

- Target Corporation (February 2014)[7] – target was the point-of-sale environments, with an estimated number of approximately 70 million credit cards potentially compromised;
- Google Australian Division (May 2013)[8] – target was disruption of the enterprise environments by researchers; there was no damage, as the devices controlled HVAC environments only; and,
- Carrell Clinic (July 2009)[9] – target was disruption of the enterprise environments, and obtain patient record information by a former employee; no patient records were compromised.

HVAC/Building Automation Manufacturers - Top 5 Manufacturers

[Sampling data collected as of 31-Jan-2014]

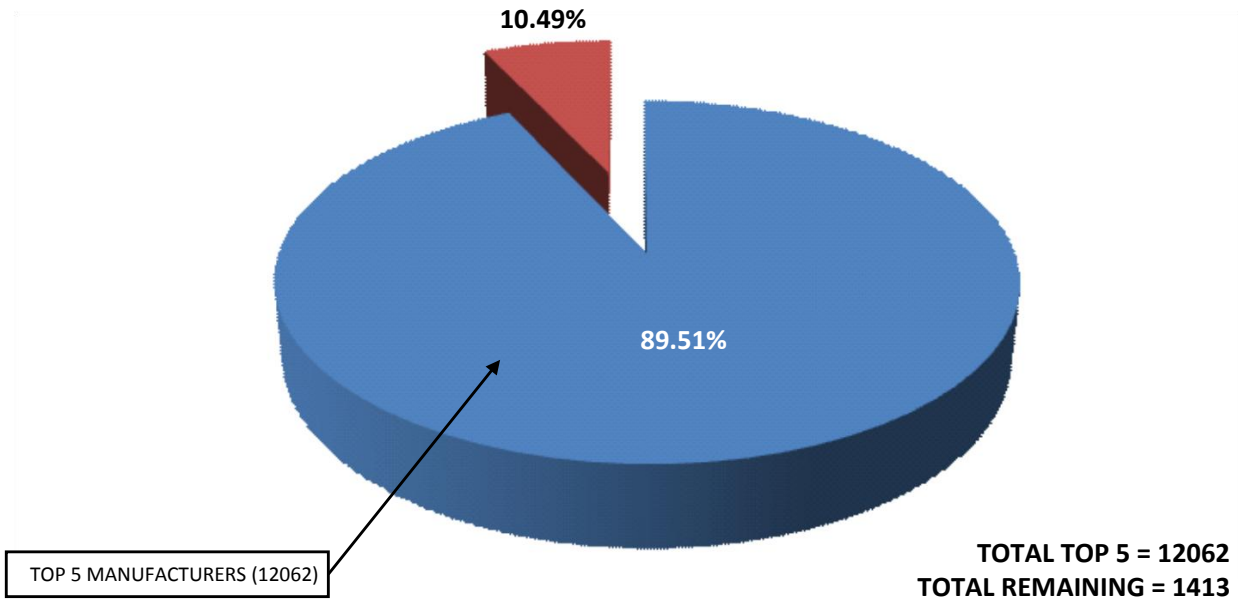

Figure 5.

[6] http://www.greentechmedia.com/articles/read/are-your-hvac-and-lighting-systems-vulnerable-to-hacking
[7] http://krebsonsecurity.com/2014/02/target-hackers-broke-in-via-hvac-company
[8] http://www.wired.com/2013/05/googles-control-system-hacked
[9] http://www.darkreading.com/risk/security-guard-busted-for-hacking-hospitals-hvac-patient-information-computers/d/d-id/1131436?

Results from Project SHINE: Serial-to-Ethernet Converters

Some statistics of SCADA/control systems are broken down based on SCADA/control systems that are accessible via serial-to-Ethernet converters. Data acquired from the SHODAN search engine confirms that SHODAN is periodically scanning and collecting metadata for these devices.

Serial-to-Ethernet Manufacturer Results

As of 31-Jan-2014, there are six (6) manufacturers currently being scanned and metadata collected for SCADA/control systems specific devices, which include:

- Allied Telesis
- Digi International
- Moxa
- ATOP Systems
- Lantronix
- Multitech Systems

The total number of serial-to-Ethernet devices discovered is estimated at approximately 204,416, roughly 9.35% of the 2,186,971 devices sampled. This does not include other manufacturers not known; therefore, the total count is estimated for this reason.

Manufacturer	Count	% Out of 100%
MOXA	78309	38.31%
LANTRONIX	56239	27.51%
ALLIED TELESYS	34573	16.91%
DIGI INTERNATIONAL	30557	14.95%
ATOP SYSTEMS	3846	1.88%
MULTITECH SYSTEMS	892	0.44%

Figure 6.

An issue that many asset owners fail to understand is that serial-to-Ethernet converters provide an easy attack vector for accessing SCADA/control systems devices, often camouflaged as something else. One method of spotting a given device utilized for SCADA/control systems are two (2) communications protocols provided by many serial-to-Ethernet converters: (1) RS-485[10]; and (2) DF1[11].

RS-485 is used as the physical layer underlying many standard and proprietary automation protocols used to implement SCADA/control systems, including the most common versions of MODBUS and PROFIBUS. This protocol may be used to control video surveillance systems or to interconnect security control panels and devices such as access control card readers, etc. DF1 (alt. "DF-1") is an asynchronous byte-oriented protocol that is used to communicate with most Allen Bradley / Rockwell Automation RS-232 interface modules. The DF1 protocol consists of link layer and application layer formats. Link layer serial frame is a composition of conventional RS-232 serial frames with parameters specified as 8 data bits, no parity, and a maximum baud rate of 19200.

[10] http://en.wikipedia.org/wiki/RS-485

[11] http://en.wikipedia.org/wiki/DF-1_Protocol

Results from Project SHINE: SCADA/Control Systems Ports

Some statistics of SCADA/control systems are broken down based on SCADA/control systems port-specific protocols. Data acquired from the SHODAN search engine confirms that SHODAN is periodically scanning and collecting metadata for these protocols. The results on a daily basis (ingested into Project SHINE) is inconsistent; therefore, any and all relevant metadata presented was taken entirely (as a sample as of 31-Jan-2014) from the SHODAN search engine.

SCADA/Control Systems Port Results

As of 31-Jan-2014, there are five (5) ports currently being scanned and metadata collected for SCADA/control systems specific devices, which include:

- Siemens SIMATIC / ICCP – Port 102
- DNP3 – Port 20000
- BACNet – Port 47808
- MODBUS/TCP – Port 502
- Ethernet/IP – Port 44818

Graphical representations are presented in 2 different graphics/pages: (1) top five (5) countries; and (2) total count (with percentage difference) for all countries captured from SHODAN, proceeding in ascending port number order, starting with port 102 (Siemens SIMATIC / ICCP).

Background Regarding the SCADA/Control Systems Protocols

SCADA/control system information and command processing is distributed across multiple stations which are connected through a local area network (LAN). Information is shared utilizing network protocols that were still not standardized. These protocols are considered manufacturer-proprietary. There are several dozen manufacturers of these proprietary protocols; however, the SHODAN search engine is scanning and ingesting metadata for five (5) popularly used protocols.

Siemens / ICCP (Port 102)

Port 102 is primarily used by Siemens Corporation to administratively control their SIMATIC control systems (which is based from the RFC 1006 design). Areas of application include: (1) remote programming; (2) ISO-on-TCP connections; and (3) S7 connections via Industrial Ethernet.

MODBUS/TCP (Port 502)

MODBUS TCP/IP is a variant of the MODBUS[12] family of simple, vendor-neutral communication protocols intended for supervision and control of automation equipment. Specifically, it covers the use of MODBUS messaging in networked environment using TCP/IP protocols. The most common use of these protocols is Ethernet-attached devices such as PLC's, I/O modules, and other field buses or I/O networks.

[12] http://en.wikipedia.org/wiki/Modbus

DNP3 (Port 20000)

The Distributed Network Protocol (DNP3)[13] is a set of communications protocols used between components in process automation systems. Its main use is in utilities such as electric and water companies; usage in other industries is not common. As it was developed for communications between various types of data acquisition and control equipment, this protocol plays a crucial role in SCADA/control systems, where it is used by SCADA master stations, Remote Terminal Units (RTUs), and Intelligent Electronic Devices (IEDs). It is primarily used for communications between a master station and RTUs or IEDs. ICCP, the Inter-Control Center Communications Protocol (a part of the IEC 60870-6 protocol), is used for inter-master station communications.

Ethernet/IP (Port 44818)

The EtherNet/IP™[14] protocol is an industrial Ethernet network that combines standard Ethernet technologies with the media-independent Common Industrial Protocol (CIP). EtherNet/IP™ is one of several industrial Ethernet network protocols utilized throughout the world, and is widely used in a range of many industries including manufacturing, hybrid and process control. The EtherNet/IP™ and CIP technologies are managed by ODVA, Inc., a global trade and standards development organization founded in 1995 whose 300+ corporate members are comprised of the world leading automation device suppliers.

BACNet (Port 47808)

The BACNet is a communications protocol for building automation and control network, and is an ASHRAE, ANSI, and ISO standards[15] protocol.

BACNet was designed to allow communication of building automation and control systems for applications such as heating, ventilating, and air-conditioning control, lighting control, access control, and fire detection systems and their associated equipment. The BACNet protocol provides mechanisms for computerized building automation devices to exchange information, regardless of the particular building service they perform. Proper communication between building automation devices is critical for maximizing building energy efficiency, indoor air quality, and other aspects of "green"[16] buildings.[17][18][19]

[13] http://en.wikipedia.org/wiki/DNP3

[14] http://en.wikipedia.org/wiki/EtherNet/IP

[15] Standard 135-2012-- BACnet--A Data Communication Protocol for Building Automation and Control Networks (ANSI Approved). 2012.

[16] http://en.wikipedia.org/wiki/Green_building; A green building (also known as “green construction” or “sustainable building”) refers to a structure utilizing a process that is environmentally responsible and resource-efficient throughout a building's life-cycle: from sitting to design, construction, operation, maintenance, renovation, and demolition. This type of building requires cooperation between the design team, architects, engineers, and the client at all project stages.

[17] http://hpac.com/building-controls/bacnet-technology-key-world-s-greenest-office-building?NL=HPAC-04&Issue=HPAC-04_20130905_HPAC-04_741&YM_RID=bdorsey@kmccontrols.com&YM_MID=1420553&sfvc4enews=42v

[18] http://www.facilitiesnet.com/buildingautomation/article/BACnet-Helps-Make-Buildings-Intelligent-Meet-Control-and-Data-Analysis-Goals--14237?source=part

[19] http://www.businessenergy.net/DE/articles/25502.aspx

Siemens SIMATIC / ICCP (Port 102) – Top 5 Countries

[Sampling data collected as of 31-Jan-2014]

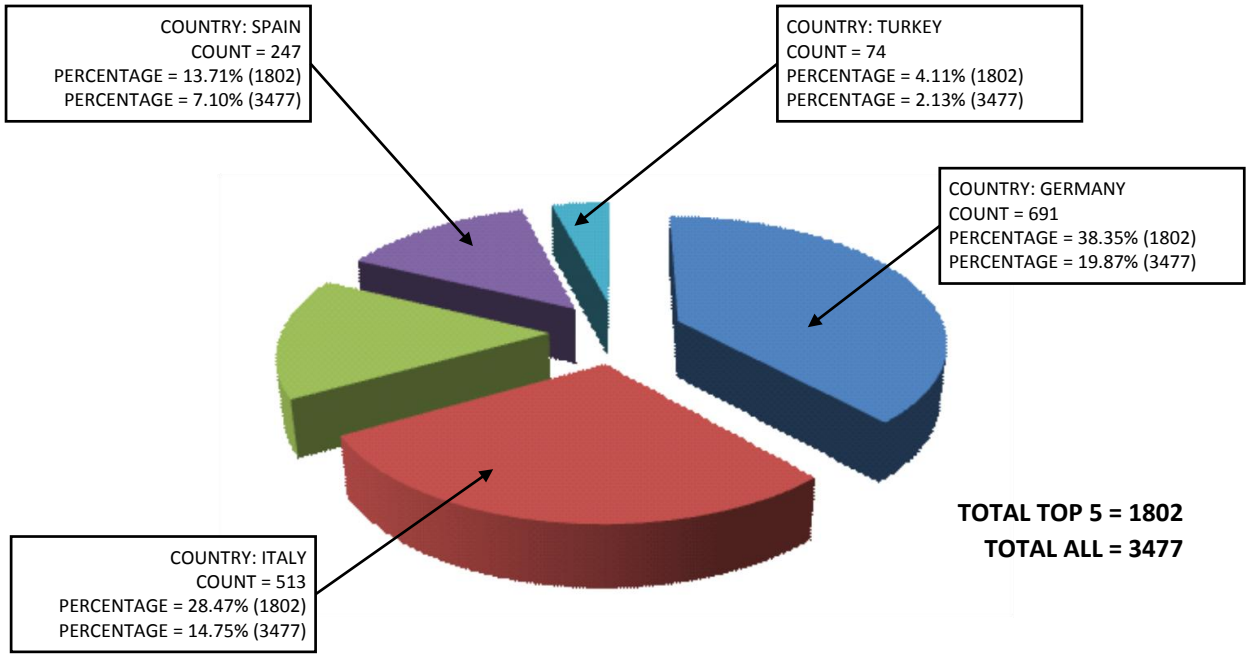

Figure 7.

Position	Country	Count	% Out of 1802	% Out of 3477
1	Germany	691	38.35%	19.87%
2	Italy	513	28.47%	14.75%
3	United States	277	15.37%	7.97%
4	Spain	247	13.71%	7.10%
5	Turkey	74	4.11%	2.13%
		1802 (total)		

Figure 8.

Siemens SIMATIC / ICCP (Port 102) – All Countries

[Sampling data collected as of 31-Jan-2014]

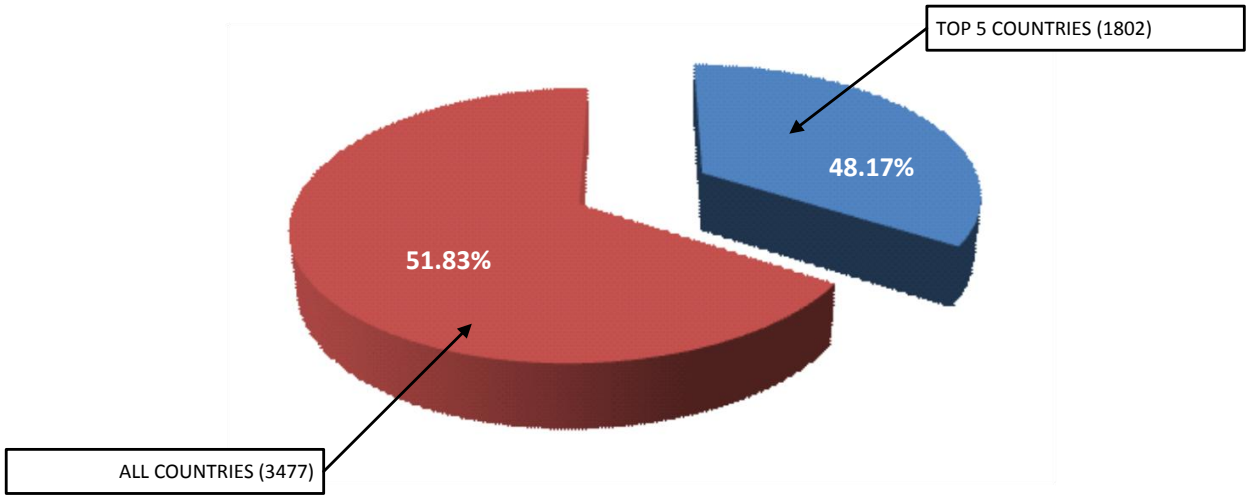

Figure 9.

Tag	Count	% Out of 100%
Top 5 Countries	1802	48.17%
All Countries	3477	51.83%

Figure 10.

MODBUS/TCP (Port 502) – Top 5 Countries

[Sampling data collected as of 31-Jan-2014]

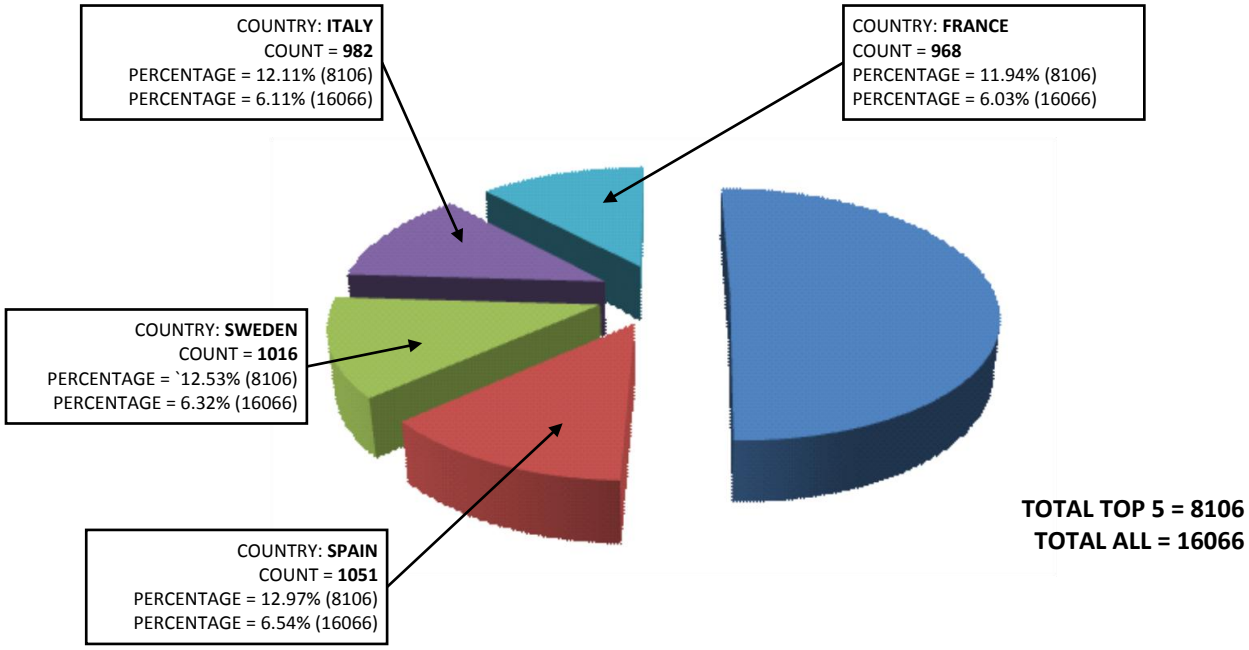

Figure 11.

Position	Country	Count	% Out of 8106	% Out of 16066
1	United States	4089	50.44%	25.45%
2	Spain	1051	12.97%	6.54%
3	Sweden	1016	12.53%	6.32%
4	Italy	982	12.11%	6.11%
5	France	968	11.94%	6.03%
		8106 (total)		

Figure 12.

MODBUS/TCP (Port 502) – All Countries

[Sampling data collected as of 31-Jan-2014]

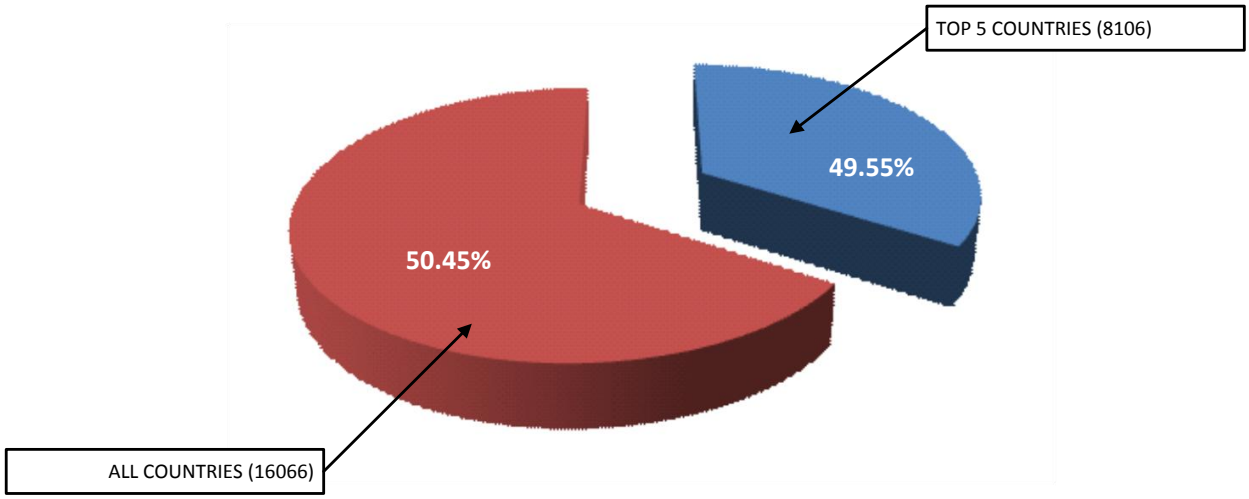

Figure 13.

Tag	Count	% Out of 100%
Top 5 Countries	8106	49.55%
All Countries	16066	50.45%

Figure 14.

DNP3 (Port 20000) – Top 5 Countries

[Sampling data collected as of 31-Jan-2014]

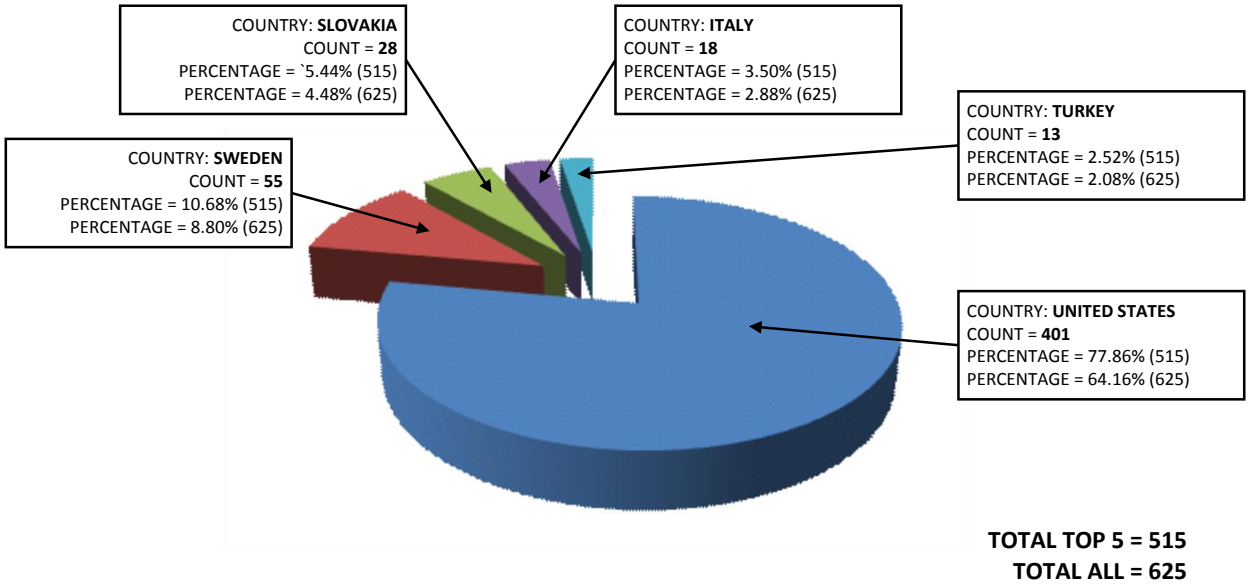

Figure 15.

Position	Country	Count	% Out of 515	% Out of 625
1	United States	401	77.86%	64.16%
2	Sweden	55	10.68%	8.80%
3	Slovakia	28	5.44%	4.48%
4	Italy	18	3.50%	2.88%
5	Turkey	13	2.52%	2.08%
		515 (total)		

Figure 16.

DNP3 (Port 20000) – All Countries

[Sampling data collected as of 31-Jan-2014]

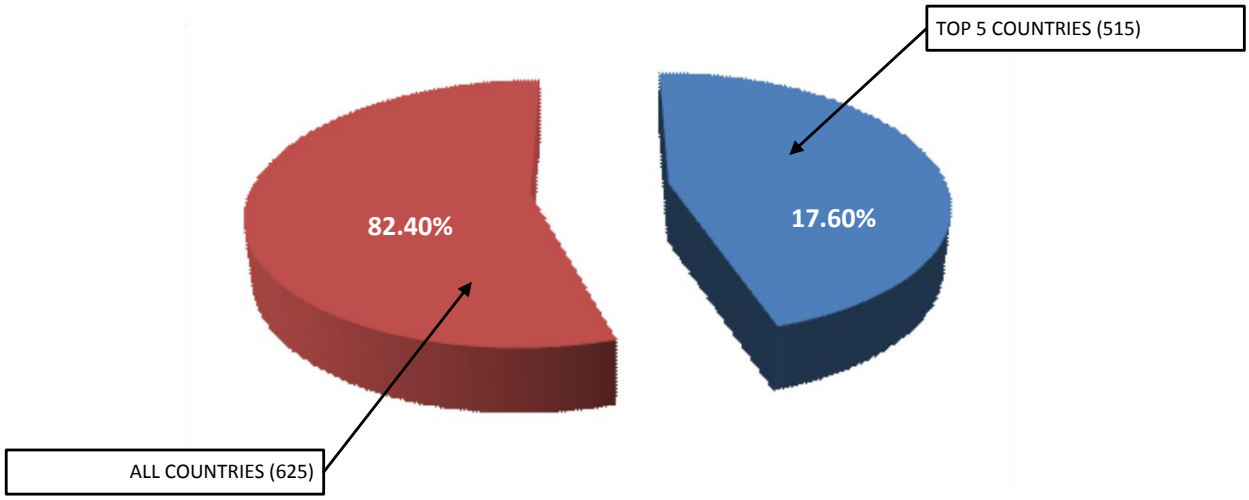

Figure 17.

Tag	Count	% Out of 100%
Top 5 Countries	515	17.60%
All Countries	625	82.40%

Figure 18.

Ethernet/IP (Port 44818) – Top 5 Countries

[Sampling data collected as of 31-Jan-2014]

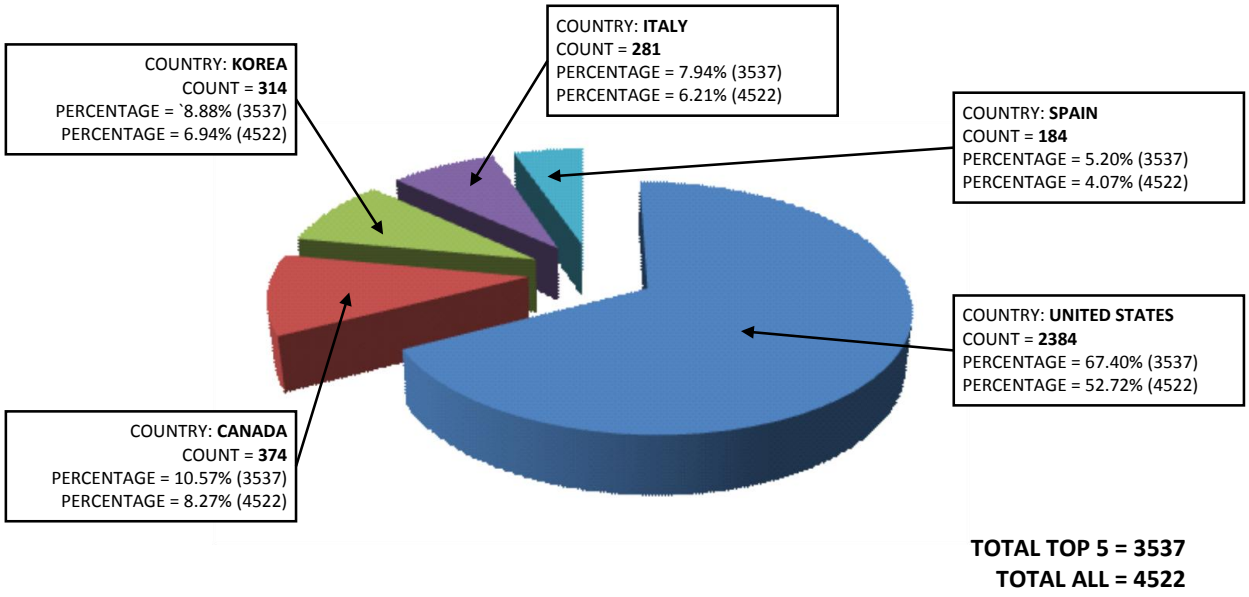

Figure 19.

Position	Country	Count	% Out of 3537	% Out of 4522
1	United States	401	67.40%	52.72%
2	Canada	55	10.57%	8.27%
3	Korea	28	8.88%	6.94%
4	Italy	18	7.94%	6.21%
5	Spain	13	5.20%	4.07%
		3537 (total)		

Figure 20.

Ethernet/IP (Port 44818) – All Countries

[Sampling data collected as of 31-Jan-2014]

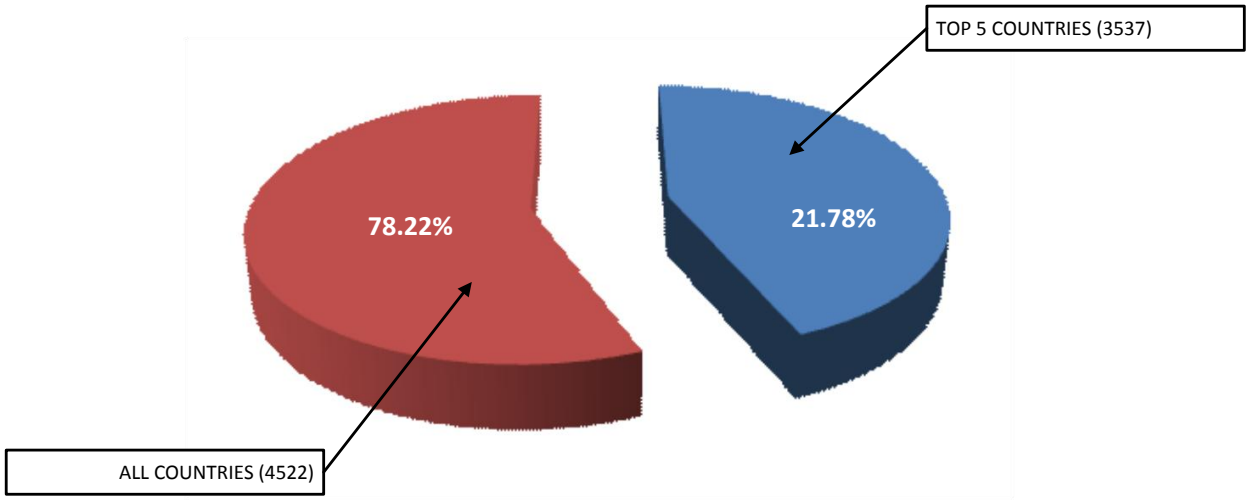

Figure 21.

Tag	Count	% Out of 100%
Top 5 Countries	3537	21.78%
All Countries	4522	78.22%

Figure 22.

BACNet (Port 47808) – Top 5 Countries

[Sampling data collected as of 31-Jan-2014]

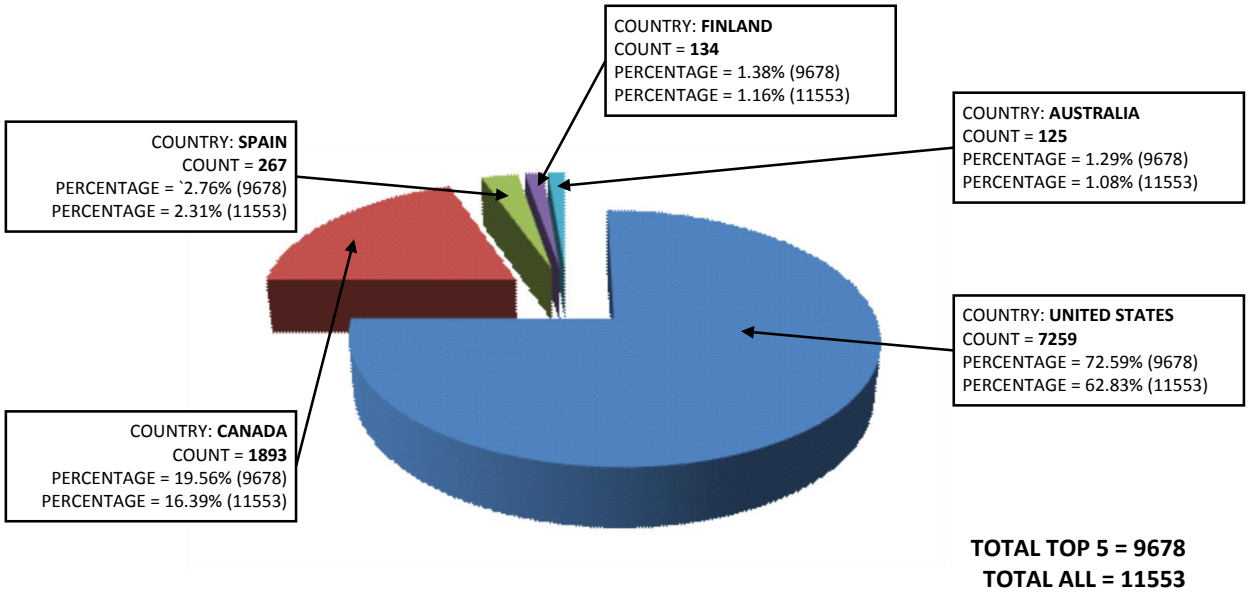

Figure 23.

Position	Country	Count	% Out of 9678	% Out of 11553
1	United States	7259	75.01%	62.83%
2	Canada	893	19.56%	16.39%
3	Spain	267	2.76%	2.31%
4	Finland	134	1.38%	1.16%
5	Australia	125	1.29%	1.08%
		9678 (total)		

Figure 24.

BACNet (Port 47808) – All Countries

[Sampling data collected as of 31-Jan-2014]

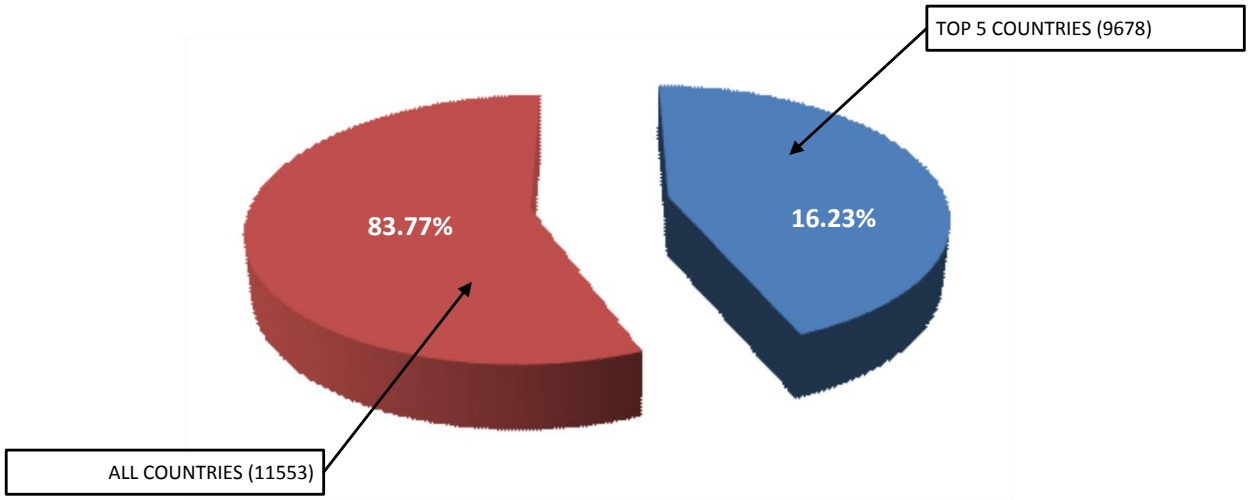

Figure 25.

Tag	Count	% Out of 100%
Top 5 Countries	9678	16.23%
All Countries	11553	83.77%

Figure 26.

Comparison of All Protocols

A comparison horizontal bar graph (shown below) provides the total count per protocol.

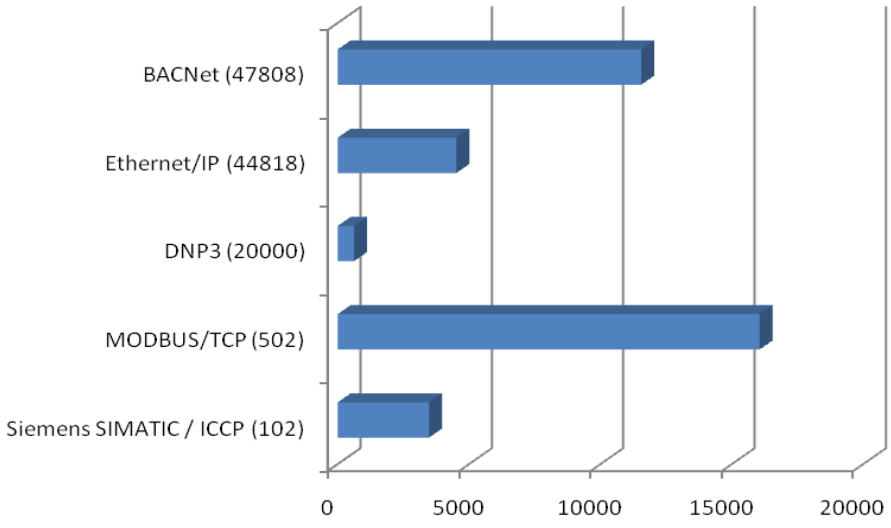

Figure 27.

Results from Project SHINE: Master Counts

Some statistics of SCADA/control systems were broken down based on SCADA/control systems master counts since Project SHINE's inception on 14-Apr-2012. Data acquired from the SHODAN search engine confirms that there are gaps in counts throughout the period sampled; the sampled count data taken was between 14-Apr-2012 through (and including) 31-Jan-2014. Some of these gaps may include:

- The SHODAN search engine was unavailable for maintenance reasons (scheduled outage);
- The SHODAN search engine was unavailable for unexplained reasons (un-scheduled outage);
- Networking errors/issues with Project SHINE's network (ISP outage); or,
- Erroneous conditions within the Project SHINE application during its initial development.

Over the past several years, SHODAN has undergone several modifications, and has been unavailable (usually) during late-night timeframes. As the search engine application has been modified over the years, the number of un-scheduled outages has reduced significantly in number.

SCADA/Control Systems Master Count Results

There are 654 entries from 14-Apr-2012 through (and including) 31-Jan-2014. This includes two fields:

- Search terms found; and,
- Total counts.

Graphical representations are presented in 2 different graphics/pages: (1) search terms found (per day); and (2) total count (per day) captured from SHODAN, proceeding in ascending port number order, starting with 14-Apr-2012 onward; the last day ingested is 31-Jan-2014.

Search Terms Found (Per Day)

[Sampling data collected between 14-Apr-2012 through (and including) 31-Jan-2014]

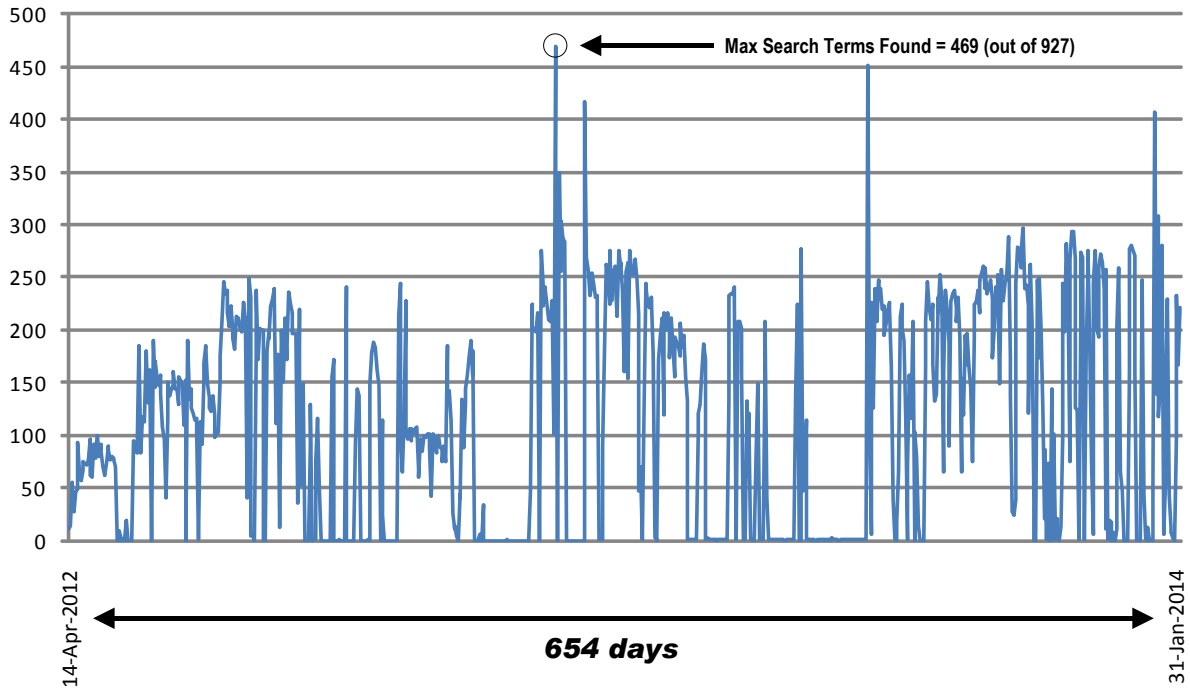

Figure 28.

This graphic represents the total number of search terms discovered through the SHODAN search engine; meaning, that for each day, if search term <*x*> is found, this represents a summarized count of a uniquely identifiable device (via its IP address) for that particular search term. The total count (performed daily) represents the total number of search terms discovered per day.

Total Counts (Per Day)

[Sampling data collected between 14-Apr-2012 through (and including) 31-Jan-2014]

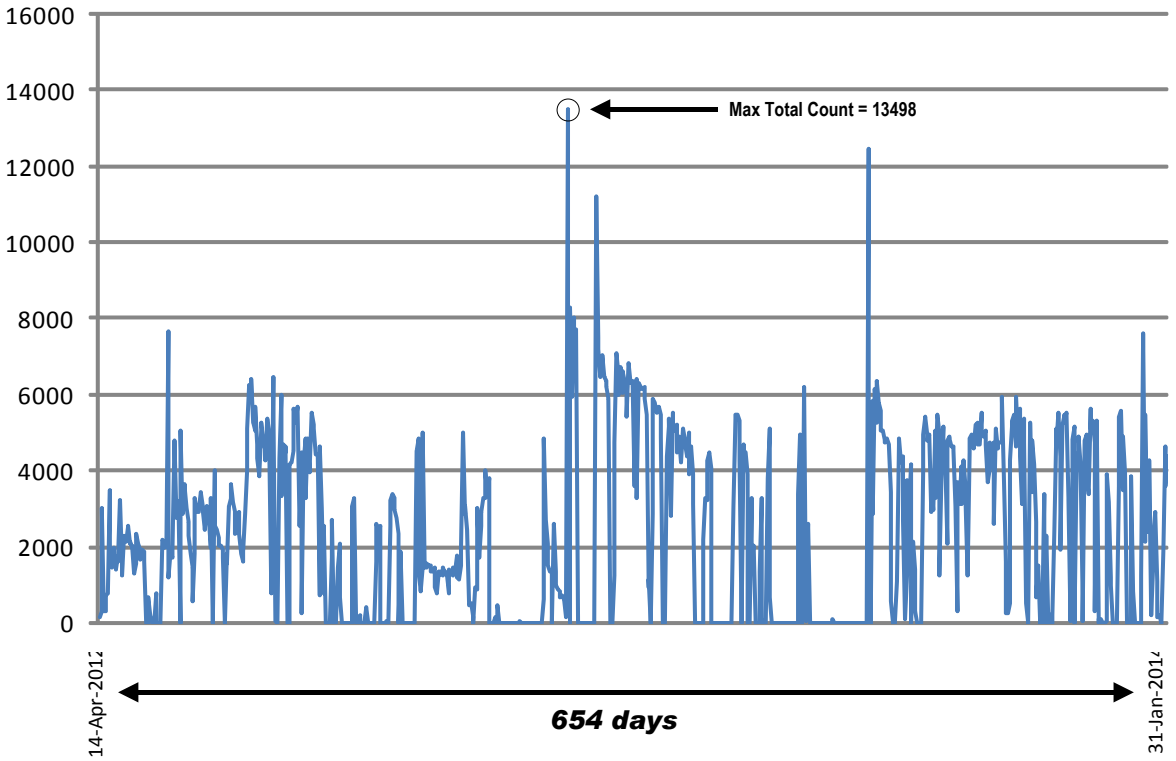

Figure 29.

This graphic represents the total number of all devices discovered through the SHODAN search engine; meaning, the total count (performed daily) represents the total number of devices discovered per day.

Results from Project SHINE: Counts by Country

Some statistics of SCADA/control systems are broken down based on country. Data acquired from the SHODAN search engine is then compared against a geographical locator. Graphical representations are presented showing the top twenty-one (21) countries; these countries are identified by their total count that is above 20,000.

Country Summarized Results

Of the twenty-one (21) countries identified, the top five (5) countries represent 53.48% against the remaining countries identified (**bolded** text), with the United States' count at 616,994, slightly double with the next highest country, Germany, at a count of 280,248; combined, both the United States and Germany represent 49.08% of the top twenty-one (21) countries identified, with respect to 41.03% of all 211 countries identified.

Tag	Count	% Out of 100%
United States	**616994**	**33.75%**
Germany	**280248**	**15.33%**
China	**112114**	**6.13%**
Korea	**99856**	**5.46%**
United Kingdom	**66234**	**3.62%**
Canada	62712	3.43%
Brazil	62376	3.41%
Italy	62266	3.41%
France	56827	3.11%
Taiwan	46836	2.56%
India	41309	2.26%
Spain	40911	2.24%
Mexico	39904	2.18%
Thailand	39027	2.13%
Russian Federation	38395	2.10%
Japan	34013	1.86%
Netherlands	29349	1.61%
Sweden	28641	1.57%
Norway	28544	1.56%
Poland	21721	1.19%
Australia	20083	1.10%
	1828360 (total)	

Figure 30.

The count for the top twenty-one (21) countries identified is 1,828,360, from a total count of 2,186,971, with a count difference of 358,611 (or 16.40%); the top twenty-one (21) countries represents 83.60% against the remaining countries identified (out of 211 countries total).

Country Count – Top 21 Countries

[Sampling data collected as of 31-Jan-2014]

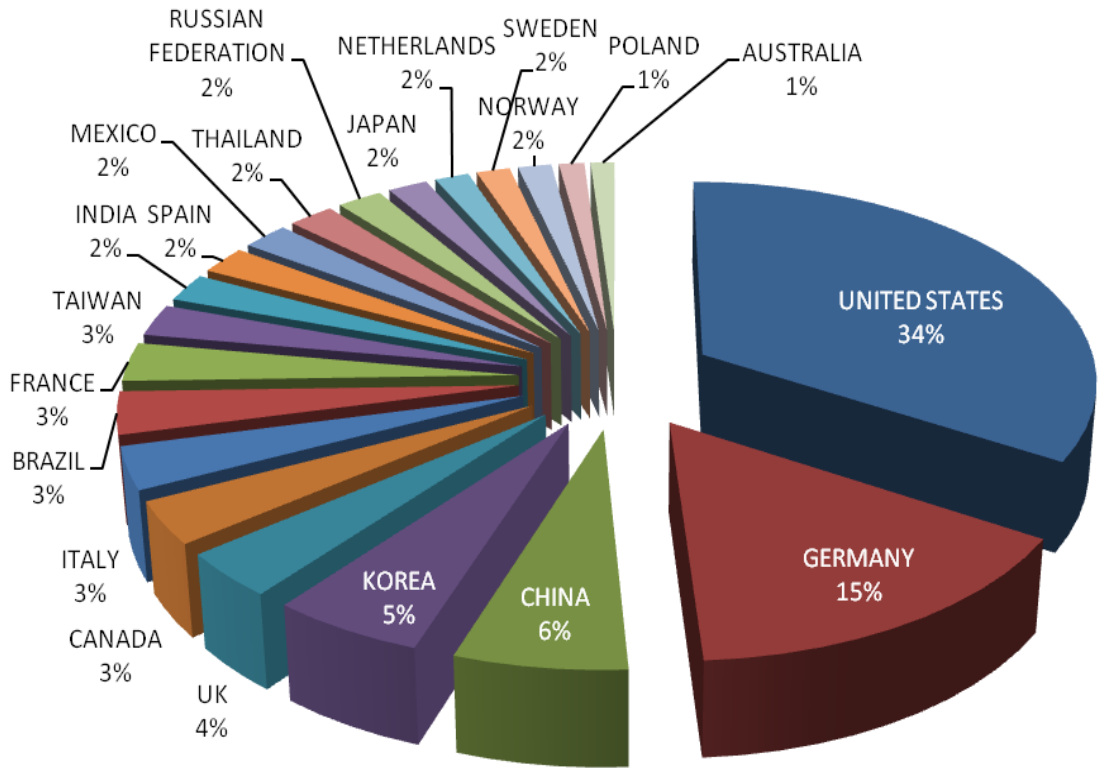

Figure 31.

Countries Identified with Total Count Results

There were 211 countries identified based on the IP addresses harvested from the search terms discovered. Outlined below is each country with total count per country. The total count is 2,186,971.

Afghanistan	165	Cameroon	136	Ghana	300	Lichtenstein	0
Aland Islands	79	Cape Verde	22	Gibraltar	312	Lithuania	1951
Albania	192	Cayman Islands	59	Greece	9287	Luxembourg	1172
Algeria	1068	Central African	4	Greenland	175	Macau	472
American Samoa	74	Chile	5993	Grenada	37	Macedonia	616
Andorra	49	China	112114	Guadeloupe	129	Madagascar	83
Angola	241	Colombia	15801	Guam	141	Malaysia	8832
Anguilla	40	Comoros	7	Guatemala	508	Malawi	68
Anonymous Proxy	88	Congo	89	Guernsey	77	Mali	92
Antigua and Barbuda	63	Cook Islands	10	Guinea	126	Maldives	105
Argentina	12634	Costa Rica	6477	Guyana	17	Martinique	135
Armenia	2487	Cote d'Ivoire	81	Haiti	144	Malta	363
Aruba	375	Croatia	11439	Honduras	413	Mauritania	30
Asia/Pacific Region	74	Cuba	143	Hong Kong	11702	Mauritius	179
Australia	20083	Curacao	209	Hungary	6871	Mayotte	7
Austria	14027	Cyprus	1948	Iceland	1159	Mexico	39904
Azerbaijan	922	Czech Republic	13955	India	41309	Moldova	416
Bahamas	196	Denmark	9992	Indonesia	7182	Monaco	88
Bahrain	275	Djibouti	32	Iran	3260	Mongolia	209
Bangladesh	955	Dominica	1410	Iraq	362	Montenegro	121
Barbados	204	Dominican Republic	1402	Ireland	4891	Morocco	1459
Belarus	807	Ecuador	2663	Isle of Man	79	Mozambique	271
Belgium	8863	Egypt	5342	Israel	11894	Myanmar	8
Belize	418	El Salvador	338	Italy	62266	Nambia	0
Benin	33	Equatorial Guinea	12	Jamaica	1186	Nepal	112
Bermuda	262	Estonia	1571	Japan	34013	Netherlands	29349
Bhutan	51	Ethiopia	28	Jersey	91	New Caledonia	82
Bolivia	3344	Europe	338	Jordan	2106	New Zealand	2783
Bonaire	29	Faroe Islands	56	Kazakhstan	2111	Nicaragua	204
Bosnia & Herzegovina	772	Fiji	77	Kenya	1282	Nigeria	1850
Botswana	109	Finland	14444	Korea	99856	Norway	28544
Brazil	62376	France	56827	Kuwait	1203	Oman	230
Brunei Darussalam	121	French Guiana	35	Kyrgyzstan	338	Pakistan	2980
Bulgaria	3369	French Polynesia	218	Lao People	77	Palestinian Territory	765
Burkina Faso	100	Gabon	60	Latvia	2093	Panama	2427
Burundi	10	Gambia	54	Lebanon	779	Papua New Guinea	90
Canada	62712	Georgia	483	Liberia	22	Paraguay	228
Cambodia	321	Germany	280248	Libya	55	Peru	2708

Figure 32.

(continued)

There were 211 countries identified based on the IP addresses harvested from the search terms discovered. Outlined below is each country with total count per country. Total count is 2,186,971.

Philippines	5255	Togo	10				
Poland	21721	Trinidad & Tobago	384				
Portugal	7223	Tunisia	741				
Puerto Rico	1485	Turkey	16348				
Qatar	369	Turkmenistan	8				
Reunion	93	Turks and Caicos	244				
Romania	10577	Tuvalu	3				
Russian Federation	38395	Uganda	1995				
Rwanda	68	Ukraine	5493				
Saint Lucia	152	United Arab Emirates	1608				
Saint Kitts and Nevis	50	United Kingdom	66234				
Saint Pierre and	7	United States	616994				
Samoa	79	Uruguay	759				
Saint Vincent & Gren	48	Uzbekistan	71				
San Marino	44	Vanuatu	10				
Satellite Provider	845	Venezuela	2500				
Saudi Arabia	3026	Vietnam	10586				
Senegal	114	Virgin Islands	194				
Serbia	1921	Yemen	68				
Seychelles	62	Zambia	241				
Sierra Leone	20	Zimbabwe	160				
Singapore	6127						
Slovakia	4816						
Slovenia	2023						
Somalia	9						
South Africa	11587						
Spain	40911						
Sri Lanka	832						
Sudan	141						
Swaziland	20						
Sweden	28641						
Switzerland	10423						
Syrian Arab Republic	153						
Taiwan	46836						
Tajikistan	57						
Tanzania	268						
Thailand	39027						
Timor-Leste	4						

Figure 33.

Lessons Learned from Project SHINE

One of the purposes of this document is to outline some of the practices that lead to trouble. This discussion of security is not exhaustive. In fact, Mr. Radvanovsky and Mr. Brodsky have co-edited an entire handbook on the subject (it will not be reproduced here).

First, from the spot-check cases the team has studied, most exposures appear to be accidental. In other words, it is the result of poorly configured network infrastructure. The asset owners may not realize the problem until perhaps someday someone takes the time to locate the asset owners to advise them that their infrastructure is exposed. Of course, it is also entirely possible that someone with ill intentions will find such assets first.

This is one very strong reason to avoid the practice of lightly configuring equipment.

Contrary to the practice of lightly configuring devices, anything not absolutely necessary should be shut down. For example, serial port servers are one of the secret work-horses of many SCADA/control systems. The use of such devices is almost unavoidable. It is not wise to leave all configurations to the defaults it came with. Thus, turn off the web interface; turn off SNMP; turn off the Telnet interface; turn off the configuration interface (e.g.; is there really a need to dynamically reconfigure the serial interface remotely?) Then, use firewalls and virtual private network (VPN) appliances to block traffic from any but those places that absolutely must be able to "see" these devices.

When scanning industrial protocols, it is important to consider whether something is configured to respond automatically to a query. Most industrial protocols will yield some results, but not all.

For example, some DNP3 implementations are not easily interrogated if one does not know the addresses to use. Unfortunately, many system integrators start with low addresses and work their way up. Scanning for the first 10 addresses may yield many DNP3 devices. However, the address space goes all the way to 65519. Choosing "random" addresses in the middle of the address space, such as 10548, would make finding a DNP3 RTU difficult. Furthermore, many RTUs can be configured to respond to only a handful of master addresses. Make those addresses obscure and these devices will be moderately difficult to find even if the IP address is known. Some will argue that this is "security through obscurity" -- but it is at least effective enough to keep such devices out of common search engines. The DNP Users Group has application notes that further discuss this issue in greater detail.

From a perspective of project integration, there probably should be policies, if not laws, outlining each and every default password and service an embedded device has. At some point in which the project's implementation is turned over to the customer, there needs to be a documented ceremony identifying each and every such device and account. After it is turned over, the customer should be responsible for changing every single account except perhaps those previously agreed to for warranty remote access assistance by the implementer.

Too many devices appear to have such default accounts in them. Once these practices become known, it then becomes possible to scan the Internet for them. This is one very strong reason to stop building back-doors (such as factory-based passwords) into products. The authors feel that such undocumented "back-door" accounts should not be legal for use within critical infrastructure applications.

Some sought to use Project SHINE as a compliance tool. There is no way this can work. Project SHINE data does not indicate who the site belongs to, and even if Project SHINE does find something belonging to an asset owner, there is no guarantee that anyone can use that data to trace back to the owner. Often the exposure appears to happen accidentally. For example, someone configures a firewall and a VPN through that firewall to another site. Then one day, there is a problem with the connection. What's the first thing a technician would do? Possibly remove the VPN from the equation. Do they bother to restore it? Not always.

During validity checks, the Project SHINE team found what appeared to be translated addresses from behind a firewall. There were entire local area networks there waiting to be hacked. It is important to keep in mind that hiding behind a firewall with a single data concentrator is no guarantee against a determined effort to expose a site. Once the data concentrator is compromised, everything else behind it is at risk. Do not assume that a data concentrator can protect anything.

Conclusion

The Project SHINE team presents this data in the hope that individuals (and their organizations) might begin to understand the sheer scope of this problem. New regulations and legislation are needed to curb this behavior. Industry practices need to be modified. Diagnostic practices and configuration management schemes need to improve dramatically.

It is worth noting that many of these sites may technically be in compliance with regulatory -- only because the asset owners have no idea that they really are exposed. The industry must get past this terrible practice of compliance-based security and focus instead on an attitude of safety, vigilance, and performance awareness.

Within the community of aviation, there is a catch-all regulatory clause warning pilots against "careless and reckless" operations (14 CFR 91.13). This is the regulation that the United States Federal Aviation Administration (FAA) uses when pilots do something that everyone should have been aware of -- even though there may not have been explicit regulations against it. There is no analogous regulation for any other industry, particularly electric utilities. As an industry, the community needs to break this terrible "compliance" based approach towards security, and instead use a catch-all phrase somewhat similar to what the FAA has done for pilots. Doing so will put a stop to a many ignorant and dangerous practices that are presently getting the World at large into trouble.