
Early version, also known as pre-print

Link to published version (if available):
10.1049/cp.2013.1720

Link to publication record in Explore Bristol Research
PDF-document

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
http://www.bristol.ac.uk/pure/about/ebr-terms
Incident Analysis & Digital Forensics in SCADA and Industrial Control Systems

T. Spyridopoulos*, T. Tryfonas*, J. May‡

*Cryptography Group, University of Bristol, Merchant Venturers Building, Woodland Road, Clifton BS8 1UB, UK
‡Systems Centre, University of Bristol, Merchant Venturers Building, Woodland Road, Clifton BS8 1UB, UK

Abstract. SCADA and industrial control systems have been traditionally isolated in physically protected environments. However, developments such as standardisation of data exchange protocols and increased use of IP, emerging wireless sensor networks and machine-to-machine communication mean that in the near future related threat vectors will require consideration too outside the scope of traditional SCADA security and incident response. In the light of the significance of SCADA for the resilience of critical infrastructures and the related targeted incidents against them (e.g. the development of stuxnet), cyber security and digital forensics emerge as priority areas. In this paper we focus on the latter, exploring the current capability of SCADA operators to analyse security incidents and develop situational awareness based on a robust digital evidence perspective. We look at the logging capabilities of a typical SCADA architecture and the analytical techniques and investigative tools that may help develop forensic readiness to the level of the current threat environment requirements. We also provide recommendations for data capture and retention.

1. Introduction

An Industrial Control System (ICS) is an information system used to control industrial processes. Industrial control systems include Supervisory Control and Data Acquisition (SCADA) systems, Distributed Control Systems (DCS), and Programmable Logic Controllers (PLC). SCADA systems historically distinguish themselves from other ICS by being the largest subgroup of ICS systems and large scale processes that can include multiple sites and large distances. Traditionally, control systems have been operated as isolated systems with no network connection to the rest of the world. Therefore, threats against these systems were limited to physical damage attacks or data tampering that originated from inside the system. However, the introduction of emerging technologies along with the connection of such systems to the Internet in order to improve in performance and effectiveness have exposed the once closed systems to the various internet threats. Additionally, the very nature of SCADA systems, that can be part of a national critical infrastructure, makes them increasingly an attractive potential target for a variety of threat agents and attack vectors, ranging from disgruntled insiders and dissident groups to foreign states.

This whole new environment in which SCADA systems operate urges the need for more secure implementations and better protection solutions. The ability to respond to security incidents along with the ability to analyse and learn from what happened is crucial. Towards this direction, digital forensics is an area of significant importance which however has been neglected in favour of operational convenience.

Collecting evidence related to the incident can reveal the actions that took place during the incident along with the incentives and perhaps the identity of the attacker. The whole process of evidence collection and analysis has to conform with certain regulations, such as the ACPO guidelines [8], in order to preserve the authenticity and integrity of the findings so that they can be used to raise legal accusations against the identified perpetrators. Identifying the actions that took place during the incident can also disclose the vulnerabilities that the system bears upon which the attack was based. Understanding these weaknesses can help in strengthening the system against future attacks. A forensic investigation gives the ability to rely on robust evidence in order to respond to the changing nature of domestic and alien threats and ensures that enough learning takes place in order to deploy resilient systems.

In this paper we investigate the digital forensic capabilities in SCADA systems. We discuss the challenges that emerge during evidence collection and analysis due to the lack of security-oriented logging mechanisms and the existence of legacy devices that may still be in use in many ICS. We also provide certain recommendations to enhance the forensic process as well as analytical techniques and tools that may help develop forensic readiness.

The paper is structured as follows: In Section 2, we provide a description of the SCADA architecture. In Section 3, we present the state of art digital forensics capabilities in SCADA systems. Section 4, introduces certain recommendations regarding actions that will facilitate a forensic investigation. Lastly, our conclusions are presented in Section 5.

2. High level SCADA system architecture

All SCADA systems base their function on measurements taken from sensors and instruments regarding actions that took place during the incident along with the incentives and perhaps the identity of the attacker. The whole process of evidence collection and analysis has to conform
logic or commands from the system operators. SCADA systems consist of two main component types: the control centre and the various field devices it controls. The two parts are connected to each other via a SCADA server.

The control centre’s components include operator workstations also known as Human Machine Interfaces (HMIs), engineering workstations, plant data historians, databases and various shared resources. Control centre components communicate with each other through the management network, and with the field devices and other SCADA networks utilising the SCADA server. The SCADA server functions as the sole interface between the control centre and one or more sites for which the control centre is responsible. The server is usually implemented with vendor specific software and its services are often based on the OPC (‘open connectivity via open standards’\(^1\)) standard, which is maintained by the OPC Foundation as interoperability standard for industrial automation and related domains. The physical site contains the actual process system. It has three types of components: the control devices, such as the programmable logic controllers, remote terminal units, input/output controllers (IOCs) and intelligent electronic devices (IEDs), the I/O devices such as sensors that perform measurements and actuators that perform control actions, and the SCADA gateway. A control network interconnects components within a site utilising various application specific protocols, e.g. Modbus and DPN3 [4]. Figure 1 illustrates a typical SCADA system as a block diagram, when security is also taken into account. In order to maintain security in the system, a firewall is used between the corporate network and the Internet to block unwanted traffic from the Internet. Another firewall is used between the corporate network and the control network in order to block unwanted access form inside the corporate network. Intrusion Detection Systems are deployed in the Control Network and the SCADA server in order to detect and log any unwanted access to the system.

\(^1\)http://www.opcfoundation.org

Figure 1. SCADA system block diagram (*IDSs and Firewalls are not typical parts of a SCADA system, however they are essential for security.*

3. Current digital forensic capabilities of SCADA systems

Log files constitute a critical source of evidence in a digital forensic investigation. In the traditional IT domain, keeping log files is usually enforced by security policies. However, in the field of control systems logging focuses mostly on the production monitoring and to support troubleshooting. Thus, even though control systems are engineered so that transactions and activity are closely monitored, resulting log files may lack valuable information required by an investigation [3].

Legacy equipment or conventional networked control devices do not retain network traffic that could provide valuable evidence in an investigation. Enabling network traffic monitoring and logging would require more advanced architectures and additional IT components. In [4] the authors describe an architecture that supports post mortem analysis of SCADA network traffic. In fact, it exploits current capabilities in order to provide an audit and logging mechanism. The architecture makes use of “agents” at specific locations within a SCADA system. These agents forward network packets to a central data warehouse where they are stored for future usage in case of a security incident. This way, a complete history of the network activity can be retrieved by reconstructing network events from packet fragments captured.

Depending on the nature of a SCADA system different data can be extracted from its various components during a forensic analysis. According to [3] control systems can be categorised into modern/common, modern/proprietary and legacy/proprietary systems, depending on the technology used.

3.1 Modern/Common Control Systems Technologies

**Control Centre:** In modern/common control technology systems the components of the control centre (engineering workstations, databases historians and HMIs) base their function on mostly well-known technologies. Most of them use either Windows operating systems or some sort of UNIX platform or a combination of them. Therefore, common data acquisition techniques can be applied. Data recovery on these devices can rely on contemporary digital forensic tools, such as the EnCase digital forensics suite or the FTK forensics toolkit. Information gathered from logging mechanisms and volatile data (such as memory data and data form registers) are the main sources of evidence. If the system is still running, which is the case in most instances, the investigator should also obtain valuable information from the running processes. However, special care is needed since an accidental change on the system may result not only in evidence corruption but also in the abnormal function of the whole system. Due to these restrictions, the opportunity of obtaining real-time data regarding the system can be of vital importance. There exist a variety of traditional examination tools that could be used, including network sniffers as Wireshark where a plugin for the decoding of ModBus protocols used in SCADA systems has been developed, and Snort [7], which already incorporates signatures related to SCADA. It is important to mention that HMIs in modern/common deployments can offer valuable information regarding the system’s state. It can guide the investigator.
towards the components of the system that impacted the production environment. However, in many cases even though HMIs are based on common operating systems, they are often modified in order to serve the system in the most efficient way, lacking thus non-essential services such as logging mechanisms or having modified versions of them. Furthermore, even though data can be extracted with an offline examination of an HMI, the file structure may be so different that contemporary digital forensic tools might not be able to handle. In such a case, the investigator has to come in contact with the system administrator in order to gather knowledge on the HMIs and their way of function.

Field Devices: In general, field devices do not employ any logging mechanism. However, as mentioned in [3] information regarding the network communication between the field devices and the rest of the system along with activities of the field devices can be found in the control centre part of the SCADA system. Although, the technology of the filed devices should not be neglected since, in some cases, it might bear some sort of logging mechanisms, the nature of the system that requires the continuous run of the field devices turns the investigator’s attention exclusively towards the logs that can be found in the control centre. Furthermore, data regarding the field devices’ operations can also be found in the memory or in cash files in the various parts of the control centre. However, due to the volatile nature of these data it would be beneficial if the incident response team collected volatile information available from the field devices. In case field devices are off, some valuable information may be available inside these devices. In this case, the investigator needs to have in depth knowledge of the devices’ configuration in order to be able to retrieve any information available. Thus, the investigator should be in contact with the system administrator or the engineer in charge in order to draw this information. On the other hand if the field devices are still running then there is little that can be done in order to obtain data from them. However, in case a device carries advanced computing hardware then live analysis might be applied in order to extract information, which would be lost if the device was shut off, of the running device without interrupting its normal function. According to [3] such information includes the device date and time, current active processes and current running processes. Furthermore, implementing network logging agents as described can provide valuable information, regarding network connections, open ports and running applications when the system is running.

3.2 Modern/Proprietary Control Systems Technologies

Control Centre: Even though technology has achieved huge advances, most industrial control systems base their function on modern/proprietary technologies that are 10+ years old. Many of the incident analysis methods used in modern/common control systems can also be used in modern/proprietary systems since most of them base their function on contemporary operating systems such as Windows and Unix. However, due to their proprietary nature the whole process may face difficulties. Therefore, when investigating a modern but proprietary control system interaction between the investigator and the vendor prior to the investigation is mandatory. Since these systems are still supported by the vendor, lots of valuable information can be drawn through this interaction. For instance, although the type of file system utilised by the data historian may be known, it might have undergone certain modification by the vendor in order to involve desirable characteristics. Such modification may result in evidence corruption during the collection phase, or even mislead the investigator if there is no contact with the vendor. Regarding HMIs in modern proprietary systems, although they utilise relatively modern computing capabilities, the modifications upon their services along with the unique services added by the vendor can pose huge difficulties in the process of data acquisition. These added services include fault tolerance capabilities that perform automated job killing or real-time reallocation of memory space. Such operations can lead in evidence loss. Therefore, it is imperative that the unique characteristics of an HMI be fully understood [3].

Field Devices: Modern/Proprietary field devices may incorporate computing capabilities that can help an ex post incident analysis, however, as with the components of the control centre, their unique vendor-based characteristics make the investigation process more difficult than in the modern/common control systems category. Collecting information about the devices from the vendor becomes a necessity. However, as in the former category, network and activity logs may be found on the control centre part of the system, though due to its proprietary nature such functionalities might not be enabled. An interesting point made by the authors in [3] is that such devices may incorporate embedded vendor-specific security mechanisms. In this case, the vendor has to inform the investigator about the deployed security mechanisms in the field devices since possible evidence may be found there.

3.3 Legacy/Proprietary Control Systems Technologies

Control Centre: When dealing with legacy equipment within an industrial control system incident analysis seems to be impossible. Since these systems were designed in such a way so that to ensure data integrity and availability, the notion of ex post analysis and security in general is absent. In most cases logging mechanisms are absent and databases do not follow the structure of modern databases. Therefore, traditional post-mortem analysis methods cannot be applied. Furthermore, due to the fact that such systems are not supported anymore by the vendor little knowledge can be obtained regarding their way of function and whether there are any possible sources of evidence. Only the interaction with the owner of the equipment may give the investigator some information regarding the system, which however might seem useless due to the inapplicability of analysis tools. As stated in [3] legacy HMIs run mostly on proprietary systems or operating systems that are no longer supported by the original vendor. However, in case that they are legacy but common there might be a chance to perform an ex post incident analysis. Furthermore, network activity can be tracked and captured, however such systems
also utilise serial-based communications in which case there is no way of conducting an ex post incident analysis.

Field Devices: In most cases legacy field devices do not have any inherent mechanism that could aid an incident analysis. The investigator has to understand each specific device based on the knowledge conveyed by the vendor (if the vendor still supports these devices), otherwise it is practically impossible to obtain any evidence. Furthermore, legacy field devices communicate through serial connections, which make it impossible to capture network traffic. These devices pose considerable challenges to ex post analysts as their rapid rate of sampling and data override in combination with the limited amounts of memory and the very specific, fault-driven nature of the retained data make it difficult to extract very meaningful information. It is generally advised when faced with the challenge of working with legacy and proprietary field devices that the vendor should be contacted and an experienced engineer should be made available to support the investigation [3].

3.4 Digital forensic process

There are five basic steps when it comes to performing an ex post incident analysis of any device [5,6], the examination phase, the identification phase, the collection phase, the analysis phase and the documentation phase.

Examination: In the examination phase the investigator has to understand all the potential sources of evidence in a SCADA system. In addition, any other system related to the SCADA system under investigation also needs to be taken into account. This includes access terminals, logging servers and routers. Since the configuration of a SCADA system can vary significantly even across similar devices, information regarding the system and its components has to be gathered prior to the investigation. At a minimum the following information should be obtained:

(1) Network diagrams
(2) Configuration details
(3) Change logs if available
(4) Authentication credentials

Identification: The starting point of the identification stage is the identification of the type of system under investigation. Once the type of system has become known, the next step is to identify the operating system that is used, the types and manufacturer of the PLCs, and the network design and implementation. The manufacturer’s documentation, the design specifications, network diagrams can assist the identification process. In a more general sense the identification step encompasses the correlation of the information gathered from the examination process in order to find the proper tools that can be used, based on the hardware and software specifications, to the next step of data collection. In general, in digital forensics it is common that the examination and the identification step be considered as one phase in the whole process.

Collection: The collection phase involves the collection of data from all the memory systems that have been identified in the previous step. Network traffic between the identified system’s components, such as network traffic between the control network and the management network, and between the SCADA system and the Internet should also be captured. It is important that the investigator collects all types of information including both volatile and dynamic data. Information with higher volatility should be given higher priority. Some of the most crucial areas to check for volatile data include registers, caches, physical and virtual memories, network connections, running processes, and disks. Captured data must be saved in external devices, in a secure and safe place, so that they may be safely removed and kept offline. In [1] the authors list the order of volatility in a computer system as:

- Registers, cache
- Routing table, address resolution protocol (ARP) cache, process table, kernel statistics
- Memory
- Temporary file systems
- Disk
- Remote logging and relevant monitoring data
- Physical configuration, network topology
- Archival media

Analysis: In the analysis phase evidence is identified in the data collected. Eventually, a timeline of activities based on the data that was gathered in the collection phase is created. The major categories of ex post incident analysis can be defined using the notion of abstraction layers [2].

Physical Media Analysis: The analysis of the physical media translates the contents of a storage layout to a standard interface (e.g. IDE or SCSIs). Examples include a hard disk, compact flash, and memory chips. The analysis of this layer includes processing the custom layout and even recovering deleted data after it has been overwritten.

Media Management Analysis: In the analysis of media management, evidence sources are organized based on certain data structures criteria. Examples of this layer include dividing a hard disk into partitions, organising multiple disks into a volume, and integrating multiple memory chips into memory space. This process may not be applicable for all type of media. For instance, a database may access an entire hard disk without creating partitions.

File System Analysis: The analysis of the file system layer of abstraction, which translates the bytes and sectors of the partition to directories and files, involves viewing directories and file contents leading to the recovery of deleted files.

Application Analysis: Analysis in this layer includes viewing log files, configuration files, images, documents and reverse engineering executables. The input data will typically come from the file system, but applications such as databases may read directly from the disk.

Network Analysis: Analysis in this layer includes managing network packets and IDS alerts. Analysis of logs generated by network services, a firewall or web server for instance, falls under the Network Analysis.

Memory Analysis: Analysis in this area includes identifying the code that a process was running and extracting sensi-
tive data that was stored in this code. Traditional digital forensics applications such as the EnCase and the FTK can be used in each abstraction layer in order to extract the available evidence. However, unsupported data structures that may be used in legacy control systems can raise significant challenges in the analysis of the collected evidence.

Documentation: In every investigation process, it is imperative to maintain comprehensive documentation. Detailed notes have to be kept with records of time, date, and the person responsible plus other essential information. This way it is assured that no evidence has been tampered with by someone from inside during the forensic analysis.

4. Recommendations

Based on our review of the state of art capabilities we have identified the following key areas where action should be taken in order to develop investigative readiness that matches the level of perceived risk. We describe each recommendation in further detail in this section.

Identify gaps in digital investigation skills: It is important to understand the available level of (or the lack of) skill and knowledge of investigative expertise among existing staff. There may be adequate understanding of incident responding procedures but not a well-developed sense of the impact of mitigating actions to potential evidence sources. Such a review will facilitate the integration of investigative and analytical skills with the existing structure for incident handling and ensure that where possible pointers to resources on ex post analysis will be taken into account.

Identify physical and cyber response interfaces: A review of organisational roles and responsibilities involved in incident response, including operational, physical security and cyber incidents, may facilitate the integration of the responding capability from both physical and cyber perspective. Although each role will have dedicated responsibilities, several outcomes may be related, or the impact of an incident may be cross-domain (as in the example of stuxnet). In such cases a well-coordinated response is required, including personnel from all affected functions. Unless there is effective interfacing of all these, it will be difficult to assess the true dimensions of a security breach and to ensure that there is enough evidence retained. Such a review will at the very least enable the personnel involved to be identified and brought together, which in itself may improve the coordination of the response function. Further to this other synergies may be developed such as harmonising of the related procedures, more accurate impact assessment and improved cross-domain risk understanding.

Understand where evidence may be found: As part of the traditional risk assessment process, it may be beneficial to consider along with the scenarios of security breaches where evidence is crucial and to identify where this evidence could be found. In following best practice an analysis of potential attacks would need to be performed anyway, so the consideration of the associated data that could be generated in the course of such scenario unfolding is something that is feasible. This exercise will increase confidence in both the operating personnel that evidence acquisition may be possible and unobtrusive and to the IT specialists who will have a clearer picture of the impact of responding actions.

Understand the impact of data retention: It is recommended that some form of impact assessment of data retention policies is performed on a test infrastructure that resembles the operating environment. It is essential to develop an understanding of whether any overhead is introduced (and how much if so), when enabling more advanced logging features over and above the traditional fault recording and performance-tracking paradigm of operation. If evidence retention is deemed essential, then the cultural barrier of the end user community created by the focus on operational need will need to be overcome based on sound testing evidence. On the other hand, where the absolute priority is the safety of the system there may be limited opportunity to enable logging and data retention features.

Manage obsolescence and the IT/Ops interface: Albeit not directly related, a structured plan for obsolescence management, where applicable, will ensure that adequate knowledge of legacy systems exists and that access to the appropriate facilities for their management is possible. Such plan should identify gaps in skills and knowledge for operating the legacy system and provide the means for continuity of operating experience, e.g. through mentoring schemes of younger engineering personnel by experienced senior technical managers and principal engineers. More broadly, the management of the interface between IT and ops planning is also essential in order to ensure visibility of the concerns of both communities (operational and technical personnel vs IT specialists). The development of shared understanding of key issues is essential in order to overcome cultural barriers and the silo mentality that both could develop. Filling the gaps that are created because of this and ensuring smooth interaction is essential in order to ensure an integrated approach in responding to security incidents and be able to analyse the available evidence.

Deploy adequate security controls – firewalls and intrusion detection systems: The cornerstone of effective security management, which will lend itself in credible capability for ex post incident analysis, is the implementation of appropriate and well-measured controls able to balance the risk and provide mechanisms to counter and follow up incidents. These controls include firewalls and intrusion detection systems and, although they are not formally a part of the typical architecture of a SCADA system, we feel that they should be viewed as essential to implement by end users.

Design systems with evidence protection in mind: Adequate protection of data historians is essential for forensic-grade evidence retention. Contemporary systems may be able to log a variety of events but if the access to the logs is compromised an attacker could easily erase their tracks. Therefore one needs to ensure that access controls and strong authentication are in place so that accesses of vital evidence on the data historians are monitored and recorded. Consideration for secure storage of log files needs to be given, for instance to write once read many (WORMs), external discs or tape, or validation of their integrity via cryptographic means of hash functions (such as MD5) if the former is not possible.

Enable logging of common events across the system as...
a minimum: Most contemporary control systems and equipment at device level are capable of producing and retaining a wealth of information related to their operational status and also to contextual events. However what events can be logged and the exact form of the data may vary tremendously from one equipment vendor to another. Even worse, some vendors may only record incidents in proprietary forms that may be difficult to integrate with the evidence sourced from more common devices (that may be in the form of raw text, comma separated values, spreadsheets etc.). In order to make the most of evidence collection a twofold approach should be considered:

1. Enabling the logging of common features, including as a minimum across components:
   - Event time(s) and date(s).
   - Process IDs.
   - Error codes, where applicable.
   - Host and connecting machine IDs, where applicable.
   - User IDs, where possible (e.g. process owner)
   - Any other relevant data (e.g. exploit reference etc.)

2. Developing a unified framework for event recording and incident sharing that facilitates common understanding between stakeholders and data analytics and correlation of the retained data. This is a matter of significance for both intra- and inter-organisation collaboration and can be built on established approaches such as the VERIS community framework. This provides a structured way for populating logged data into a predefined schema that describes comprehensively an incident.

Enable forensic-grade evidence acquisition processes: Response procedures could be modified to enable forensic-grade evidence acquisition where needed. This may require greater involvement and collaboration of various people from different teams. The latter also implies the need to train people to understand their role in preserving evidential integrity when handling potential evidence and to be able to make decisions on whether to proceed with mitigation procedures, or to stop in order to preserve and collect incident related data.

Inter-organisational and interstate cooperation: Enabling inter-state collaboration is critical, as attacks may be targeted across a number of sites, from a number of foreign jurisdictions. Therefore in order to develop an accurate picture of an incident there may be the need for multiple organisation involvement, and the legal side of the analysis of relevant evidence may be complex. As with every other aspect of complex problems, experience sharing and multi party collaboration may enhance the chances of delivering a solution that is broader and comprehensive. However the sharing of security experiences is understandably hindered by many factors. There exist many reasons for this, including fear of negative publicity, unclear underpinning legislation, lack of communication between stakeholders and peer roles across organisations, lack of technical understanding of the feasibility of anonymised data exchange etc. In the absence of structured frameworks for exchanging related information, various differences in culture, operational practice, risk perception, status of technology adoption and maturity of technological implementations may render ineffective any attempt to collaborate on the issue. Approaches like VERIS may thus facilitate both inter-organisation and inter-state collaboration, as they provide the grounds to develop shared understanding via safe information exchange. Another dimension that could promote community development would be a coordinated approach at inter-state level, underpinned by some enabling mechanism of guideline provision. The US example of NIST providing technical advice and developing guidelines in a federal capacity for all institutions and organisations that are part of critical national infrastructure is an example of this.

5. Conclusion

IP connectivity and machine-to-machine communication may change fundamentally the way control processes are instrumented and deployed, so it is important to be able to understand their potential impact. The management of legacy systems is a significant issue, as well as understanding and bridging the gap generated by the fundamentally different life cycles of typical control equipment and standard information technologies. Such issues are not necessarily directly related to forensic readiness of an organisation, but may have profound implications to the generation, retention and analysis of digital evidence. Finally, skilled personnel and deep expertise are top priority assets that an organisation should aim to develop and maintain in order to ensure capability of high standard. Investments in developing both, through appropriate recruitment and training programmes, technology acquisition implementation of relevant frameworks ought to be considered seriously by Executive Boards. Without their support, no effort to develop this capability will have the chances to meet its full potential.

Acknowledgements

Our research on incident responding and analysis has been supported by a study commissioned by the European Network and Information Security Agency. The first author would like to acknowledge the financial support of the Systems Centre.

References