THE APPLICATION OF DATA DIODES FOR SECURELY CONNECTING NUCLEAR POWER PLANT SAFETY SYSTEMS TO THE CORPORATE IT NETWORK

R.T. Barker *, C.J. Cheese †

**MIET, UK, tom.barker@edf-energy.com, †MIET, UK, chris.cheese@edf-energy.com*

Keywords: data diode, nuclear, safety, security

Abstract

The complexity and frequency of cyber attacks against plant control systems is rising, as is the corporate demand for real-time access to plant control system data. When implemented correctly data diodes claim to provide the means to deliver real-time plant data to the corporate IT network and provide a barrier impervious to network based attacks. EDF Energy is exploring data diode technology and the available implementations as a potential means of providing corporate users with real-time plant data without exposing critical safety systems to an unacceptable level of risk from cyber attack.

1 Introduction

Supervisory Control and Data Acquisition/Industrial Control Systems (SCADA/ICS) perform an essential role in ensuring the safe, reliable operation of critical infrastructure. High availability requirements and often safety related functionality demand that these systems be protected against deliberate and inadvertent incidents which could compromise their operation. Historically, these systems were completely isolated from external networks with access to control functions and plant data typically requiring authorisation and physical access to the plant or facility. The ability to optimise maintenance regimes through the increased use of condition monitoring and the often remote location of plant systems has led to the increased demand for remote access to the data produced by these systems. To facilitate these requirements, modern Internet Protocol (IP) based networking technologies have been deployed accelerating the interconnectivity of these once isolated systems [2].

The connection of plant systems to external networks introduces the potential for network based threats which, due to the routable nature of IP based communications, could come from anywhere. These systems are typically not designed with security as a primary objective and standard precautions such as anti-virus are often not available or may be undesirable for performance or maintenance reasons. Furthermore, the move towards commercial off the shelf (COTS) hardware and software makes the development and deployment of these threats much simpler. This necessitates the use of new security management activities which simply did not exist in the days of legacy serial communications [2].

When considering information security for a corporate IT network, the priorities are defined by the CIA Triad; confidentiality, integrity and availability. When applying these principles to a control system, especially where the system performs a role important to safety, the priorities need to be reversed with availability taking precedence and confidentiality typically being of least significance.

2 Connecting a control system to an external network

When connecting a control system to an external network its internal processes are exposed to that network; as the control system receives and processes data from the network it is in some way influenced by it. However, this influence may not always lead to desirable behaviour of the control system and in some cases could be potentially dangerous.

There are numerous possible scenarios that could result in undesirable operation arising from network connectivity, not all of them malicious. For example, an issue could arise during a routine operation simply due to an error or shortfall in the design of the hardware or software of the system but equally, the cause could be a deliberate network attack launched internally or externally against the control system with the intention of rendering it unable to carry out its normal functions thereby disrupting or disabling operations.

Modern networks tend to be physically large and highly interconnected making the challenge of securing them electronically and physically both significant and costly. When a control system is connected to an external network the security of the network underpins its safe, reliable operation and the risks associated with network connectivity must be identified and reduced to an appropriate level.

3 Traditional solutions

A study performed by the Centre for the Protection of National Infrastructure (CPNI) identified that when connecting control systems to external networks, only a small number of architectures are typically used, these range from hosts with dual network interface cards to multi-tiered

combinations of switches and routers. The study concluded however, that the most secure, manageable and scalable segregation architectures were those based on a three zone system such as a De-Militarised Zone (DMZ) [1].

A DMZ provides a means of protecting two networks from each other whilst still allowing for the sharing of data or resources. Creating a DMZ involves placing a firewall offering three or more interfaces, or a pair of firewalls between the plant system and the external network (Figure 1). A shared system such as a data historian located within the DMZ is able to communicate with both the plant system and the external network but direct communication between the external network and the plant system (and vice versa) is not permitted. This approach enables users to access plant data (and may even permit control functions) from the external network but by preventing direct communication with the plant system, the risk of incidents that could compromise its availability or operation is significantly reduced.

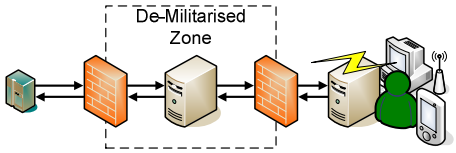

Figure 1: A control system protected by a DMZ.

The use of a DMZ is not without its drawbacks. Vulnerabilities can exist within the firewalls and the shared systems that could allow an attacker to gain control of the DMZ and launch further attacks against the control system. These vulnerabilities can arise due to incorrectly configured firewall rule sets, unauthorised administrative access or from errors in the design or implementation of the various components of the DMZ.

In order to realise the maximum security benefit, a DMZ requires continual maintenance to ensure that firewall rule sets and any shared systems remain correctly configured and up to date. This places a significant burden on finances and resources and requires suitably qualified and experienced personnel.

Errors in the configuration and maintenance of a DMZ can open holes in the defensive barrier that could be exploited by an attacker to gain control of the DMZ providing them with a direct communication path to the attached control system. Even a properly configured and maintained DMZ is not immune from the possibility of 0-day (previously unknown) exploits that could provide a means of attack.

4 Data diode technology

As its name suggests, a data diode is a network device which allows data to pass through it in one direction only.

The most common form of a data diode is simply an optical link. The physical difference between transmission and reception hardware, one being a laser emitter and the other a light detector prevents operation in the opposite role. Thus one way transmission is guaranteed at the physical layer (layer 1 of the OSI model) and there is only a very small possibility of error in the design and implementation.

Data diode technology was developed to facilitate the sharing of data between networks of different security classification; data flow from a lower security network to a higher security network is permitted but data flow in the opposite direction is physically impossible. This application of the technology has typically been employed within government and military network infrastructures to prevent the transmission of protected information from secure networks while still allowing new data to be received (Figure 2).

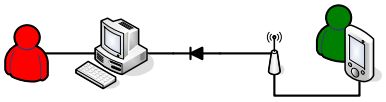

Figure 2: A secure network receiving information through a data diode.

Recent developments have seen this technology applied in a different manner for connecting critical SCADA/ICS to potentially hostile external networks. By connecting a control system to an external network using a data diode, control system data can be made available to the external network in real-time while guaranteeing that all access to the control system from the external network, whether inadvertent or malicious is impossible (Figure 3).

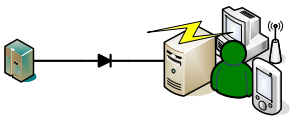

Figure 3: A control system connected to an external network through a data diode.

Despite the simplicity of the concept there is however a significant difference in the capability and complexity of the various implementations and more importantly, in the level of assurance of unidirectional operation.

Data diodes do have a number of disadvantages when compared with firewalls. Modern networked computer systems including plant control systems are typically designed to operate on a network with bidirectional communication. This is because most modern software routing and communication protocols require bidirectional communication at multiple layers. For example, Transmission Control Protocol/Internet Protocol (TCP/IP) requires bidirectional communication and many applications that use TCP/IP as a basis for network connectivity would have an application layer protocol on top of TCP/IP that also requires bidirectional communication. Simply placing a data diode between the source and destination will immediately break any protocols that require bidirectional communication.

For some applications, such as remote control, this means that a data diode simply cannot be tolerated and another security solution must be implemented. However in applications that simply provide data, that is, no control function is required, unidirectional communication is acceptable. In the likely case that any control system protocols are designed for bidirectional communication, gateway servers upstream and downstream of the data diode (Figure 4) can perform the necessary handshaking with both the control system and the destination application, effectively masking the unidirectional communication across the data diode.

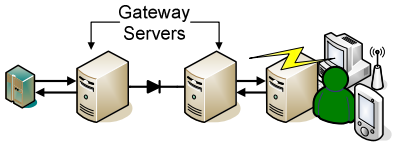

Figure 4: Gateway servers handle two-way communication.

Another issue is with regards to data loss or corruption, which in the case of bidirectional networks can be handled by the protocols themselves (this is standard behaviour in TCP/IP). However where data is lost or corrupted when transmitted along a path which at some point contains a unidirectional link, it is not possible to request that the data is re-sent. Thus the error must be logged, operators alerted and a manual operation undertaken to retransmit the data if it is still available.

5 Data diode solutions

When considering the connection of particularly high risk systems to control system networks, the security of all parts of the network which could be used to access the control system must be considered. Even where strong barriers to attacks from outside networks exist, these barriers may be bypassed by mis-configuration of, or physical access to, the control system network. For this reason, where high risk systems are connected to large control networks, the data diode should be located within the physical security perimeter of the control system or the entire network infrastructure upstream and including the data diode must all be maintained to the security level of the control system. In the case of large distributed control system networks, maintaining the physical security of the network to a high standard may be a significant challenge.

There are a number of decisions to be made when considering the deployment of data diodes for securing control systems. The choices available will depend upon the architecture of the network that the control system is connected to. Architectures could vary from a single control system connected directly to an external network, to an entire sub network of control systems with a gateway to an external network. The first step is to assess where the true boundary between safe and potentially hostile networks really is.

The location of the data diode can be thought to be at some distance both physically and in terms of network hops from the control system. The closest possible point is internally to the control system, for example if it had a built in data diode; then immediately upstream of the control system, perhaps within its physical security perimeter; then further out into the network architecture as far as the boundary of the control system network. Figures 5, 6 and 7 show three possible configurations. The dashed links represent areas of network infrastructure that could be used to launch a network based attack on a control system; the solid links represent areas of network infrastructure from which it is impossible to attack the control systems.

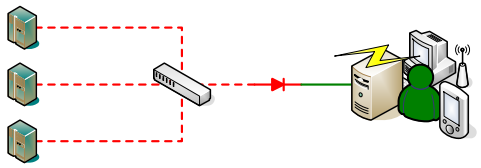

Figure 5: A network of control systems protected by a unidirectional gateway.

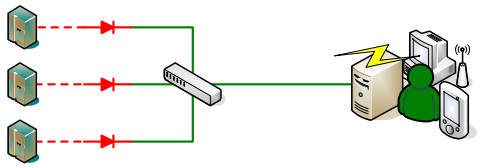

Figure 6: A network of control systems individually protected by data diodes.

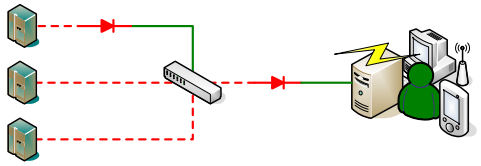

Figure 7: A network of control systems protected by a unidirectional gateway and one critical system with an individual data diode.

In terms of security, the optimum would be a data diode internal to the control system itself specifically for secure data output. This removes any possibility of mounting network based attacks on the control system from any location (assuming 100% assurance of unidirectional operation). Implementing an assured internal unidirectional link and associated protocol would require vendor support and technical competence; at this time equipment with inbuilt unidirectional network data output ports is not commonly available. It would also be impossible to use the network interface for any control type functions even from within the local secure control system network.

The benefits of sitting further from the control system are apparent when for example using a unidirectional gateway which can forward data from multiple control systems.

However this approach does not protect the control systems from threats originating within the control system network (Figure 8). The control system network must be adequately managed and physically secure such that the risk of it being accessed by attackers or incorrectly configured in a way that bypasses the data diode is adequately low. If the existing system uses a DMZ type approach then there should be no additional risk from installing a data diode as it would be expected that the control network is already adequately managed and physically secured. When connecting a new control system to a network, especially a safety related system, this must be considered.

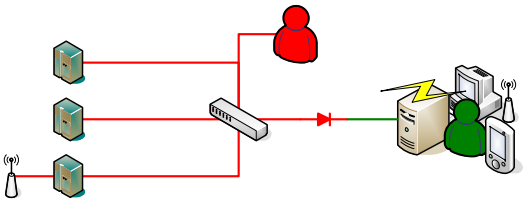

Figure 8: A large distributed network with an unauthorised 3G modem installed and physical access by a potential attacker.

6 Handling bidirectional protocols

One of the primary issues with the installation of data diodes within what is designed to be a bidirectional network is the handling of bidirectional protocols. Control systems and their associated client applications are typically designed to work with bidirectional protocols and these protocols must be handled correctly to enable data to be transmitted over the diode.

Figure 9 shows a simple application where a control system is interfaced to a client application using a bidirectional network and associated protocol. This could be a custom control system protocol sitting on top of an Ethernet network using TCP/IP. The client application will use the control system protocol to request information from the control system which in turn sends back the relevant information. Should the bidirectional link be replaced with a unidirectional link, there would be no way to establish a TCP/IP communication channel and even if there were, there would be no way for the client application to make a data request to the control system.

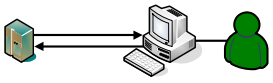

Figure 9: A control system interfaced via a bidirectional network providing control and viewing facilities to an operator.

To enable this type of system to function with a unidirectional link between the control system and its client, some form of data concentrator with unidirectional forwarding is required. This system needs to be capable of interfacing with the control system to extract the required information and forwarding that data over the unidirectional link. A second system downstream of the unidirectional link receives the data and forwards it to any clients (Figure 10).

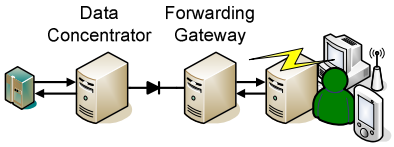

Figure 10: A data concentrator extracting information from a control system and sends it over a unidirectional link to a forwarding gateway which forwards it a user.

In order for this arrangement to work, the specific protocol implemented by the control system must be supported by the data concentrator. Due to the number of available control systems protocols there is no single solution. However there is vendor support for many of the available protocols in use within control systems networks. A control systems network with a standardised protocol will make implementation of unidirectional links far simpler than in networks with many different protocols.

It is also possible to have the data concentrator separate from the forwarding gateway. For example a data historian can receive data from multiple control systems. The entire data historian is then replicated across the unidirectional link to create a duplicate downstream. This replicated historian can then be accessed by clients on the external network (Figure 11). This approach has the benefit of securing the both the control system and the upstream data without any detriment to the user.

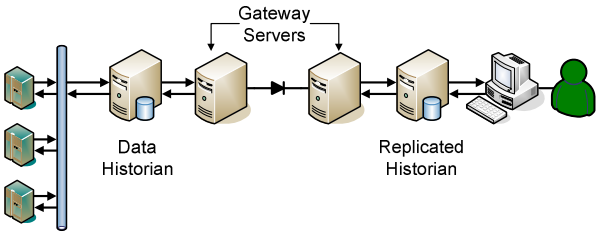

Figure 11: A data historian receives data from multiple control systems and is then replicated across the unidirectional link.

7 Selecting and evaluating a range of data diodes

In today's changing cyber threat environment there is a need to provide confidence in the company's ability to provide safe and secure generation to meet the expectations of the government and public as part of our role within the Critical National Infrastructure. The increase in the number and sophistication of the threats has also resulted in the Office for Nuclear Regulation (ONR) taking a greater interest in the security and management of Computer Based Systems Important to Safety (CBSIS). The changing cyber

environment along with the regulatory concerns has led EDF Energy to take the decision to assess the feasibility of deploying data diodes for securely connecting CBSIS to the corporate IT network.

There are an increasing number of commercially available data diode products originating from numerous places around the globe and carrying a broad range of accreditation and certification from various bodies. The general approach taken by all manufacturers is similar, with a hardware data diode device connected between upstream and downstream gateways which provide the interface connections to the trusted (plant) and un-trusted (corporate IT) networks. Some products include the two gateways with the data diode in a single package while others provide gateway software that must be installed on separate servers. All of the products being assessed support a version of Microsoft Windows, Linux or both.

The immediately apparent difference between all of the various data diode products is the number of protocols and applications that are natively supported. Typically TCP, User Datagram Protocol (UDP) and file transfer are supported by all products but in some cases it stops there. At the other end of the spectrum, support is provided for a wide range of protocols as well as applications such as data historians and databases; perhaps unsurprisingly, this additionally functionality comes at a cost premium over the more basic products.

Another significant difference between the products was the support offered by the suppliers or manufacturers. In some cases, the product was simply being sold as an appliance, in others, a full product solution was offered including the ability to develop bespoke software to interface with legacy systems or custom applications.

Before commencing with the assessment, a set of objectives were established identifying the qualities considered important for deployment on a nuclear power plant. The strictly controlled nature of the nuclear industry means that the installation of new plant and changes to existing plant are not undertaken lightly. As such, the required service lifetime of any installed data diode products is likely to be a significant number of years. For this reason, non-quantifiable qualities such as ease of installation, configuration and management as well as the cost and availability of long term support were deemed more significant than performance and unit cost. Obviously performance is still important and it was decided that measurements of bandwidth, latency and integrity would be taken, not for the significance of the actual values obtained but in order to compare the results of the different data diodes.

It should be noted that it was never the intention to confirm with any level of integrity that the data diodes under test were truly unidirectional. The guaranteed unidirectional operation of an individual data diode can only be assured if the entire lifecycle of design, manufacture, testing, transport, storage and operation by all parties involved is adequately controlled.

The process of taking performance measurements proved challenging. The unidirectional nature of the data diode meant that regular network test software wouldn't work. As a result, a suite of applications were developed in-house to perform the required measurements and enable a fair comparison to be made. A pair of applications is required for each measurement; one running on the trusted side of the data diode, the other on the un-trusted side. Data packets of configurable size and quantity are sent across the diode using either TCP or UDP and the appropriate measurements taken. This approach also enabled details such as the maximum supported packet sizes to be determined for each data diode.

Another important aspect was the provision of any fault detection mechanisms. The unidirectional nature of the data diodes again preventing standard retransmission protocols from functioning. Typically, a 'heart beat' is sent across the diode enabling the downstream identification of any communication faults in addition to the detection of missing or corrupt data. Identification of any of these faults can be configured to trigger a variety of actions such as emailing the administrator. Recovering from a fault is likely to require human intervention though, since remote access to the trusted system is impossible.

It is important to remember that there are significant differences in the architectures of a plant system to that of an IT network. In the IT environment the main servers are generally located in a series of rooms and deployment of a DMZ can fairly easily be achieved. In the plant environment a large number of systems may be separated by significant distances making the task of securing them more complex. It is unlikely that a single architecture will be suitable for all applications and in many cases, installing a data diode may not be possible.

When considering the implementation strategy for securely connecting plant system to the corporate IT network, a balance must be struck between the cost and complexity of installing and managing the deployment of data diodes and the required level of security. Supplier trust is also a key element. The assurance of both unidirectional operation and the security and reliability of any gateway software connected to a control system is essential. Without this trust the risk introduced by connecting a previously standalone control system to a network using this technology may be significant. Development of an implementation strategy will be the next step in the deployment of data diodes on an EDF Energy nuclear facility.

References

[1] CPNI Good Practice Guide: Firewall Deployment for SCADA and Process Control Networks http://www.cpni.gov.uk/documents/publications/2005/2005022-gpg_scada_firewall.pdf

[2] CPNI Viewpoint: Securing the Move to IP-Based SCADA/PLC networks www.cpni.gov.uk/documents/publications/2011/2011034-scada-securing_the_move_to_ipbased_scada_plc_networks.pdf