Introduction
There has been a lot of buzz in the automation industry over the last few years regarding unidirectional gateway appliances nicknamed data diodes. You may recall from college or if you have an electrical background, a traditional diode is an electronic device that restricts the flow of electricity to one direction. Conceptually similar, a data diode is a security appliance that is designed to prevent network communications from entering a higher security zone from a lower security zone (e.g., from a plant DCS to corporate LAN/WAN).

They have been around in various forms for over fifty years and have been used extensively in military applications to protect highly sensitive data from being compromised as well as to protect the associated systems from cyber attacks. Data diodes have also been used in the nuclear power generation industry for several years and are now gaining traction in the traditional fossil fueled power industry, primarily as a means to avoid or limit an entity’s compliance burden with the North American Electric Reliability Corporation’s Critical Infrastructure Protection (NERC CIP) Standards. However, in Version 5, the data diode alone will not be sufficient to achieve compliance.

How data diodes function
Data diodes essentially allow data from a more secured zone, such as a control system network, to be communicated to the plant Local Area Network (LAN) or corporate wide area network (WAN) to meet the needs of the business. Traditionally, this data would flow back and forth through a firewall, which dictates what systems can communicate with each other and the protocol to be used. The main protocol that systems utilize to pass information back and forth today is TCP/IP, which requires two way communications through firewall appliances to the destination systems protected by firewall. Firewalls have been an extremely popular choice to protect automation systems, but there are potential risks that come with the technology.

Firewalls have vulnerabilities just like other systems and these vulnerabilities can be leveraged by hackers to penetrate the environment. They can also be fooled by knowledgeable hackers who could use encrypted tunnels, for example, to defeat the data inspection capabilities, but the largest weakness of firewalls is that they can simply be misconfigured due to lack of experience or human error. Data diodes completely eliminate the inbound threats as the architecture is based on one way communication, from the secured network to the less secure network. The communication is typically done by two servers in the appliance, one that transmits and another that receives data. The communication crosses from the higher secured network to the outside network across an optical, dedicated one way link using proprietary one way communication. A simple architecture drawing follows.
Industrial control system cyber security

Getting information out of the control systems environment for business purposes

One of the drawbacks of using a data diode is that when you deploy them, the normal communication functionality of TCP/IP is no longer available. Of course, getting the data out is the only reason not to simply “air gap” the systems (i.e., disconnect them from the corporate LAN/WAN) in the first place. To get the data out of the DCS zone (or DMZ zone) shown above, the traffic needs to be converted to the proprietary protocol of the data diode appliance, then transmitted and acknowledged, and then reconverted to TCP/IP.

This means that the data diode vendor has to develop a way to deconstruct and then reconstruct the messages, files, data, etc. This requires the data diode vendors to develop interfaces/proxies to the specific applications such as:

- HMI – DCS, turbine control, SCADA, etc.
- Historians
- Systems monitoring protocols – SNMP, syslog, log transfers, etc.
- File transfers – FTP, FTPS, SFTP, CIFS, etc.
- Other applications - databases, backups, e-mail, etc.

This places a tremendous burden on the owner to configure and maintain the systems, and makes future modifications much more laborious, and both are prone to human error.

The data diode also results in a larger technology footprint. For example, the historian will need data on both sides of the data diode and will need the data diode to replicate the data. This requires another software instance as well as additional hardware capacity. There will also be configuration required to map the data.

Why are power plants deploying or considering data diodes?

The major business driver for power plants to explore implementing data diodes is compliance, specifically NERC CIP compliance. To understand the motive to deploy data diodes for NERC CIP compliance, it is important to understand the various versions of the standard, and how that affects compliance requirements and market behavior.

As of October 1st, 2013, CIP Version 3 is still the current enforceable version of the NERC CIP regulations, as the industry waits for FERC to pass Version 5 of the standards, or for Version 4 to go into effect on October 1, 2014.

CIP Version 3

From a power generation perspective, Version 3 allowed the plant owners to decide whether their generation plants were critical assets based on their own risk based methodology. If they determined they had a critical asset, they would have to determine their critical cyber assets that communicated outside their Electronic Security Perimeter (ESP) using a routable protocol and apply the controls that the CIP standards specify. Very few of the generation owners determined that their plants were critical, which led to a change in the regulations that produced Version 4.

CIP Version 4

Version 4 introduced a few new concepts. The most major being that the risk based methodology for determining if a generation plant was a critical asset was replaced by specific “bright line” criteria. From a generation perspective, this included:

- Each group of generating units at a single plant location with an aggregate highest rated net Real Power capability of the preceding 12 months equal to or exceeding 1500 MW in a single Interconnection.
- Each generation facility that the Planning Coordinator or Transmission Planner designates and informs the generator owner or generator operator as necessary to avoid Bulk Electric System (BES) Adverse Reliability Impacts in the long-term planning horizon.
- Each Blackstart Resource identified in the Transmission Operator’s restoration plan.

But the only systems that were considered to be Critical Cyber Assets (systems that needed protection) were defined by NERC as “those shared Cyber Assets that could, within 15 minutes, adversely impact the reliable operation of any combination of units that in aggregate equal or exceed 1500MW.”

Generation Critical Cyber Assets are further qualified to be those that use a routable protocol to communicate outside the Electronic Security Perimeter, similar to Version 3. The three new criteria resulted in many more plants that were going to be Critical Assets, which meant that companies were now being mandated to develop a costly NERC CIP compliance program, create a cyber security leadership and awareness program, implement ongoing cyber and physical security measures, conduct vulnerability assessments, collect a lot of compliance data, prepare for and be subject to audits and potential fines.
Generation owners and operators were presented with a choice when it came to Version 4 compliance. They could start a new compliance program, expand their existing one or cleverly take advantage of how a critical cyber asset is defined. A Critical Cyber Asset by definition has to communicate outside of an electronic security perimeter using a routable protocol. This gave generation owners a couple of options if they wanted to escape the compliance requirements:

- Disconnect any network connections from the plant control systems to the plant LAN, Internet, Corporate WAN, etc., essentially creating an air gap.
- Communicate outside of the ESP with a non-routable protocol.

The first choice is difficult as there is a lot of data that various internal and external entities are looking for that resides in the control system. Some plants made that choice, but now must use manual processes to move data to other systems.

The second option that has been explored by many generation owners, and at the time seemed to be a very attractive choice based on the NERC CIP version 4 standards, was to deploy a data diode, which essentially exempts the plant from having to comply with any of the CIP version 4 standards.

There were a lot of generation owners heading down that path: determining if they could live without the traditional TCP/IP communications by utilizing a data diode. They had to evaluate the costs associated with purchasing, implementing and maintaining the system, as well as understanding if they would lose any current interfaces or critical information by deploying the data diode. However, the situation will change again with the advent of CIP Version 5.

CIP Version 5
The Federal Energy Regulatory Commission (FERC) was never close to being satisfied with Version 3 and directed NERC to make numerous changes to the standards in what is referred to as FERC Order 706. While they approved the minor changes that became Version 4, they were looking for over 100 additional changes, many of which appeared in the next CIP standard, Version 5, which was approved by NERC in November of 2012 and is awaiting FERC’s approval.

Version 5 results in a major overhaul of the standards, as terminology was changed, additional standards were carved out for Change Management and Information Protection, and the blanket exemption a data diode could provide and the benefit of not having to comply with the security standards in NERC CIP was removed.

This means that even if a plant subject to Version 5 installs a data diode, they still have to have a NERC CIP compliance program if the plant is classified as having Medium Impact cyber systems. To be Medium Impact, a system would have to be able to impact 1500 MW in 15 minutes similar to version 4, as well as a plant that was deemed to be critical for the reliability of the grid by the transmission planner. Once Version 5 passes, the generation owners that have Medium Impact cyber systems will have to establish a leadership program, develop security policies and procedures, implement security controls, backup their systems, have scheduled audits and potentially face fines.

The good news for generation plants that have deployed a data diode is that the CIP requirements are less stringent for cyber systems that do not communicate outside of the electronic security perimeter using a routable protocol (i.e., those with a data diode). Table #1 shows the difference between the requirements for a plant that uses a data diode to form their ESP versus a traditional firewall that uses routable protocol.
As shown in Figure #1, there are many requirements in CIP version 5 that Medium Impact plant owners still need to address. While there is certainly a benefit to those that had installed data diodes to exempt themselves from having to comply with CIP Version 3 or in anticipation of Version 4, it will not be sufficient to avoid Version 5 compliance, and the benefits gained by getting out of some of the CIP requirements may not be worth the cost of procuring, implementing and maintaining a data diode.
Do data diodes make a generation plant more secure?
The answer to this is both yes and no. Data diodes do one thing exceptionally well: they block all traffic from coming into the controls environment through that access point. They do it better than a firewall, because a firewall can be misconfigured, fooled and potentially exploited. A data diode cannot be fooled and is unidirectional. They are a proven technology and do what they are supposed to do very well, which is to block any malware or other form of malicious attack.

The downside from a security perspective is that they force plant owners that need to get information or personnel into the control system (e.g., AGC, historical information, performance data, etc.) to use the “sneaker-net” with their laptops, USB drives, CDs and other media, rather than leveraging the network. This increases the risks of spreading of malware (e.g., inserting an infected USB drive) more so than if they could simply access the systems through a firewall or better still, a Unified Threat Management device which has firewall, antivirus and intrusion prevention capabilities.

Security researchers have catalogued security incidents and it has been stated that 70% of security incidents happen from within the network. Stuxnet is the most famous example of malware that a data diode would not have prevented.

The other troubling aspect of deploying a data diode is that it gives the generation owners a false sense of security. Since more cyber incidents occur from within the network than through a firewall, responsible entities need to employ a defense-in-depth strategy that includes a strong perimeter protection (which a data diode does provide), and also intrusion detection, patch management, antimalware measures, security event monitoring and configuration change monitoring (none of which are provided by a data diode).

Conclusion
Data diodes are a viable solution to assist power generation owners in protecting their plants from being attacked or having malware introduced from a corporate LAN or WAN, but not in and of themselves sufficient to address cyber security concerns comprehensively and are not sufficient for CIP Version 5 compliance.

CIP Version 5, which is awaiting FERC approval, removes the blanket exemption granted to data diode users in Version 3 and 4. Now a plant that is Medium Impact must have a compliance program and adhere to a significant number of the Version 5 requirements. This lessens the attractiveness of the solution.

Data diodes do a great job of blocking inbound communications, but due to the elimination of routable protocol, make it tougher to get data out. Interfaces and proxies, additional hardware and services are required to get information, such as historian data and security events, to corporate users. They also make it harder to make changes to the protected control systems. Patches, antivirus updates, application updates, and other changes need to be introduced into the environment via removable media or laptops that could potentially introduce malware to the systems.

Data diodes are a powerful security tool that block network based threats, and when used as part of a defense-in-depth approach (including patch management, malware prevention, security event monitoring, intrusion detection/prevention, configuration monitoring and effective backup and restoration programs), allow power plants to be more safe, available and reliable.
Contact us

Ed Dupre  
Sr. Marketing Consultant  
ABB Inc.  
Power Generation  
Wickliffe, Ohio, USA  
Phone: +1 440 585 7877  
Email: ed.dupre@us.abb.com

www.abb.com/powergeneration

© Copyright 2013 ABB  
All rights reserved. Specifications subject to change without notice. Pictures, schematics, and other graphics contained herein are published for illustration purposes only and do not represent product configurations or functionality. User documentation accompanying the product is the exclusive source for functionality descriptions. Symphony is a registered or pending trademark of ABB S.p.A.