Realtime-Aware Control Unit Network Attachment Module

Rainer Falk, Steffen Fries
Corporate Technology
Siemens AG
D-81739 Munich, Germany
e-mail: {rainer.falk|steffen.fries}@siemens.com

Abstract—**Information security is gaining increasingly more importance for real-time industrial automation networks. Protection is not only needed against attacks originating from external networks connected as remote data access, but also from potential attacks originating from locally connected devices connected during regular operation. Relevant use cases comprise local service access and plug-and-play of automation components. A conventional firewall filtering is not sufficient as it filters allowed network traffic depending on the contents of data packets, but it is not aware of specific real-time restrictions that have to be obeyed within an industrial automation network. For these reasons, a solution is required allowing to connect uncontrolled resp. untrusted devices with a real-time automation network segment in a secure way. This paper describes the general concept for a network filtering device that supports real-time criteria for making filtering decisions.**

Keywords-security; device authentication; real-time; network access authentication; firewall.

I. INTRODUCTION

Standard communication technologies as Ethernet and Internet Protocol IP are increasingly used in industrial environments down to the field level. Guaranteed real-time communication plays an essential role for many industrial control applications. In contrast to audio-video communication, industrial real-time requires that the real-time constraints are obeyed under all conditions as intermediate disturbances or delays are not acceptable. Different kinds of industrial automation may be distinguished, e.g., process automation for running a chemical process or a power plant, factory automation, e.g., a conveyor belt production line, or energy automation for the energy grid. For these time-critical types of applications, different real-time variants of Ethernet are available, see [1] for an overview. A basic mechanism is to reserve time slots for real-time traffic that is not allowed to be used for non-real-time traffic.

The correct operation of a shared real-time network segment depends on the fact that all attached network devices respect the real-time medium access protocol. Different usage scenarios require, however, that uncontrolled network devices are connected to a real time Ethernet segment. A network firewall filtering the data traffic according to defined filter rules depending on data content of frames/packets does solve the problem only partly as an uncontrolled device cannot be trusted to support or obey correctly the real-time constraints valid for the real-time network segment. Exemplary usage scenarios for direct connection of an uncontrolled device to a real-time field level network are maintenance and diagnosis, where a service notebook may be connected temporarily during operation. Also, standard network equipments, like a file server (NAS / network attached storage), video camera, a data recorder etc. shall be integrated within a real-time segment in a secure way.

This paper presents a firewall-like network traffic filtering security device that performs a time-dependent filtering of data traffic such that access to the communication medium is prevented during time-periods reserved for real-time traffic. It provides a control measure to enforce network access policy for a real-time field level industrial control network segment.

The remainder of this paper is structured as follows: Section II provides an overview on real-time control networks, depicting different topology types. Section III describes the problem statement and the solution approach for a real-time network attachment module allowing phasing in no-real-time capable device into real-time networks. Section IV gives an overview about related work and Section V concludes the paper and provides an outlook.

II. REALTIME CONTROL NETWORKS

Real-time systems typically consist of hardware and software that are subject to time constraints regarding execution of commands. This comprises the initiation of a command, the execution itself and the acknowledgement of the execution. Real-time in the context of this paper refers to systems with a deterministic behavior, resulting in a predictable maximum response time. These systems will handle all events at appropriate (context-dependent) speed, without loss of events.

In industry environments, especially in automation systems (robotics, motor control systems), there exist hard real-time constraints. Here, high performance system and network technologies are required to cope with the timing requirements [1]. Characteristic features of devices used in high performance communication networks are short communication bus cycle times, low device latency and high update rates of inputs and outputs of network nodes. To support strictest real-time requirements, industrial systems

 ISBN: 978-1-61208-209-7

often apply an isochronous communication bus with very low bus cycle jitter. The process control systems synchronize measurement, control, actuation and communication, taking advantage of the isochronous bus, and providing a fix schedule for process control. Real-time requirements vary from response times < 100 ms for video and voice over IP (VoIP) to response times in the range of µseconds for motion control applications.

Characteristic for this type of networks is the engineering phase, in which all systems are configured according to their responsibility in the overall system for the support of a dedicated task. Besides the engineering phase there is also a maintenance phase, in which administrative tasks like Firmware updates or similar can be applied to the system components. This phase is typically scheduled to consider its impact in the common workflow.

History has shown that these maintenance windows may not be sufficient and that there is need for administrative actions, e.g., due to defect field devices or the need to measure certain data, e.g., during the production cycle. Especially if there are dynamic error situations in the network that need to be tracked over a certain period of time or detected security vulnerabilities, which may be leveraged by malware, it may be necessary to introduce an additional component in the network, like a service component. It is very likely, that such a service component may not cope with the strict real-time requirements of the rest of the system. Nevertheless, the introduction of these service components shall not interfere with the real-time communication of the original industrial system.

The next subsections will provide more insight into typical network architectures used in industrial automation environments, which build the base for the problem and solution description in the next chapter.

A. *Network Architecture*

Industrial networks as used e.g., for energy automation are typically shared networks connected in a ring, star, or bus topology or a mixture of these [2]. Most often, the time critical part is performed on a dedicated network, while the rest of the communication supporting the industrial systems is performed on networks with lower performance or latency requirements. An example may be the connection of the industrial network to the office network. The real-time specific part is contained in a so called industrial automation cell as shown in Figure 1. The cell connects to the outside world through the connectivity module, which may be a switch or similar. To achieve security, this connectivity module often features a firewall to filter incoming and outgoing data communication. When, however, uncontrolled equipment is connected directly to a real-time network segment, e.g., for maintenance or diagnosis, or when plug and play of automation equipment is supported, the filtering firewall at the border of the automation cell is not sufficient when it cannot be excluded that the connected device may be compromised or malicious.

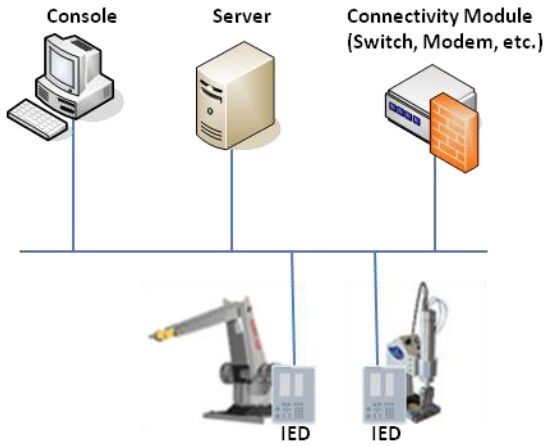

Figure 1. Example for an industrial automation cell

A bus topology connects IEDs (Intelligent Electronic Devices) via a shared communication line. There may be no further switches or hubs involved. The bus topology typically utilizes IEEE 803.3 based Carrier Sense Multiple Access with Collision Detection (CSMA/CD) mechanism [3]. Protocols deployed in this setup are for instance Profinet or Ethernet in use cases like process automation. The IEDs may belong to more complex systems like robots or melding machines. More complex setups may also provide a reason to use a substructure of the network, e.g., a robot may communicate with the outside via one dedicated connection switch, while the robot itself is build using several IEDs thus building its own automation network or automation cell. This is depicted in the following Figure 2.

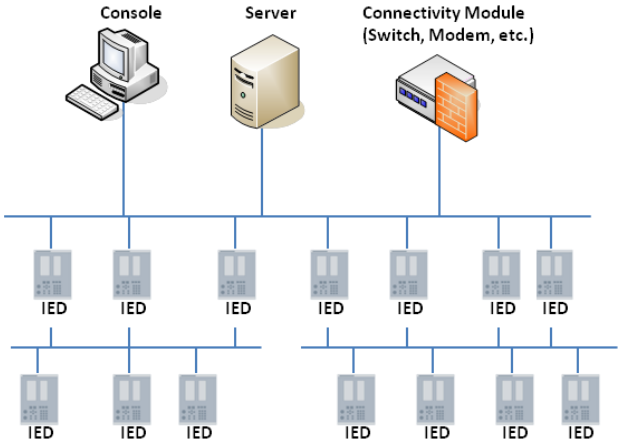

Figure 2. Bus Topology in an Industrial Network

Besides the shown bus topology, there are further network topologies used in industrial communication cells like ring or star topologies as stated above. The selection depends on the automation task to be performed. Common to all of them is the strict separation of real-time and non real-time traffic.

 ISBN: 978-1-61208-209-7

Besides these pure topologies, also mixture resulting in combinations of star and bus or star and ring are deployed in the field.

Automation network cells often realize the concept of a security cell. Here, a plant network is divided into so called security cells, which span a dedicated part of the network belonging to a distinguished administrative domain and most importantly, belong to a physically separated part of the network in which real-time requirements have to be met. A firewall at the border of a security cell does not address the case when uncontrolled and potentially corrupted or malicious devices are connected to a real-time network segment directly.

B. Real-time Communication

A real-time system can be described as showing deterministic behavior, meaning that the maximum response time of a system can be predicted and guaranteed. To describe the latency resulting from communication, the transfer time is typically used, which includes the communication processing on the sending and the receiving node as well as the network transport time. Another property often used is the cycle time, which describes the time need to complete a defined set of actions from one start to the next.

For real-time communication there are several protocols on different layers of the OSI communication stack defined, which provide the functionality stated above. In industrial communication this is often realized using field busses. For field bus communication there exist a complete family of protocols, which has been standardized within ISO-IEC 61158 and IEC 61784. Among these protocols are protocols like PROFIBUS [4], Interbus [5], PROFINET IO [6], EtherCAT [7], and SERCOS [8]. An overview about these protocols can be found in [1].

To support real-time requirements, industrial Ethernet based systems use an isochronous communication bus with very low bus cycle jitter. Isochronous communication buses are typically applied when process control systems have with very high performance requirements to synchronize measurement, control, actuation and communication. Some of protocols stated above provide support for isochronous operation, such as PROFINET-IO. To enable synchronous operation of different nodes, even below 200 milliseconds, the Precision Time Protocol (IEEE 1588) [9] is used to synchronize the involved nodes.

The basic idea of many real time Ethernet variants is to have time periods that are reserved for real time traffic, see Figure 3. These periods are called also channels or phases.

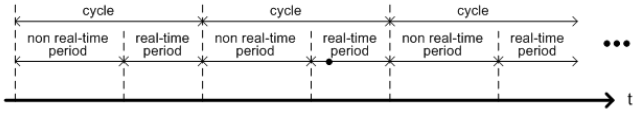

Figure 3. RT and Non-RT Periods of a Real-Time network Segment

Within these real-time periods, a strict scheduling is being applied. Non real-time capable devices should not interfere with these periods to avoid functionality loss through communication disturbances.

III. Real-Time Network Attachment Module

In the following, the problem statement and a solution is described targeting the connection on uncontrolled and potentially malicious non real-time capable devices to real-time networks. The reason for a device being malicious may be an intentional attack, but as well the fact that the device may be infected unintentionally by malware.

A. Problem Statement

The correct operation of a real-time network segment depends on the fact that all attached network devices conform to the real-time protocol. This is an additional requirement to the known network traffic filtering as performed by packet filters or application layer firewalls. It is not sufficient to limit the maximum data transfer rate, but to allow or discard data traffic depending on whether it is sent ensuring a period reserved for real-time data.

Different usage scenarios require that network devices shall be connected to a real time Ethernet segment that is not supporting the real-time enhancements. Examples comprise a service notebook connected temporarily during operation for maintenance or diagnosis. Further examples are provided through attachment of standard network equipment like a file server (NAS / network attached storage), video camera, a data recorder, etc., to be used within a real-time network segment. For the future, connecting new devices in a plug-and-play manner during the operation of a real-time automation system is envisaged. In particular when a device is connected to different networks over time or used as general PC, it is not possible to guarantee that the device has not been infected by malicious software.

In all these scenarios, it shall be ensured – in addition to ensuring that only allowed network traffic is occurring – that the real-time communication in the network segment is not affected. A real-time aware filtering of data communication will in practice be combined with well-known firewall-like filtering depending on the contents of data packets and on the connected device.

B. Real-Time Aware Network Access Control

This section first describes the new concept of real-time aware network access filtering. This targets different properties as there are the adherences to the real-time requirements, the detection of unauthorized components on the network segment as well as unexpected traffic induced by one of the network components.

To solve the described problem, a real-time network attachment module (RTAM) is proposed that prevents medium access to the real time network segment during time periods that are reserved for real-time traffic. While being a mechanism that could be used as such, preferably RTAM will realize filtering depending on contents of data packets and depending on an authentication of the connected device described in the following subsections.

 ISBN: 978-1-61208-209-7

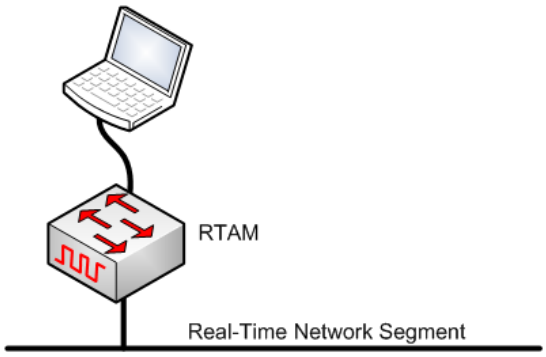

Figure 4. Real-Time Attachment Module

The real-time attachment module (RTAM) allows connecting an arbitrary network device with a real-time network; see Figure 4. It is ensured that the real-time traffic is not disturbed by the connected device. Towards the network device, a standard network interface is provided so that the network device does not require any adaptations or be aware of real time traffic.

There exist different realization options to prevent disturbance of real-time traffic:

- The simplest option it to interrupt physically the connectivity to the network segment during real-time periods. This has however the disadvantage that frames that happen to be sent during a real-time period are lost.
- A pseudo-carrier may be provided by RTAM towards the connected device during real time period independently of whether there is traffic on the real-time network segment or not. This prevents that the connected device transmits data during these time periods.
- A more complex option would be to buffering frames transmitted by the connected device and resend them during an allowed non-real-time period. So basically, a network switch is realized.

It depends on the target use case, which of the realization options is the most appropriate. While the first option is relatively simple, it has the problem of packet loss, which is addressed by the latter two options. The third option is the most flexible one, which basically works in a store and forward operation, ensuring that the real-time periods are obeyed and avoiding packet loss by buffering communication from the attached device to be sent during the non real-time periods. As typically the real-time network access will be combined with packet filtering that requires packet processing anyhow, the third option of buffering frames is the preferred option.

A preferred realization is a stateless or stateful packet filter that performs filtering decisions depending on both the packet contents the real-time period. The information whether a packet has been received during a real-time period or a non-real-time period is used as a filter criterion to allow or discard a packet. Furthermore, the actions to allow or discard a packet may be extended with a “delay” action so that traffic not permitted for a real-time period can either be discarded or delayed.

The configuration of filter rules can be organized in different ways:

- Configure separate filter rule sets for contention period and non-contention period.
- Single filter rule set with filter criteria “contention period” / “non-contention period” and actions extended from simple “allow/deny” to “delay-non-contention”. Allow traffic, but delay it to a non-contention period.

Configuring the filter rules appears to be simpler with a single filter rule set supporting additional filter criteria (real-time filter period) and actions.

In an extension, the connected device may be identified and authenticated, e.g., based on the IEEE 802.1x standard [10]. This is required in particular to support plug-and-play of network devices of different kind. After the device has been authenticated, a device-specific filtering policy is determined locally or using an authentication, authorization, and accounting (AAA) server [10]. While a real-time capable device may be granted direct, unlimited access to the real-time network segment, a generic device as an office notebook gets only limited access to non-real-time periods. Unknown or unauthenticated devices can be rejected, i.e., they do not get access to the field level network.

In a further extension, the RTAM module may provide access to the real-time network based on priorities of the non-real-time devises. The priority of a non-real-time device may be a matter of policy. As an example, non-real-time devices acting on the detection of security breaches may have a higher priority for changing the configuration of the real-time network to ensure reliable operation. Monitoring, e.g., queries from the non-real-time network (like MES systems) regarding process monitoring data recorded from the devices on the real-time network may get a lower priority.

C. Real-Time Aware Intrusion Detection / Intrusion Prevention

Intrusion prevention systems (IPS) or intrusion detection systems (IDS) monitor network traffic to detect anomalies [11]. RTAM functionality may be integrated into the IDS / IPS monitoring a real-time automation network.

The IDS/IPS analyzes communication on the network segment. In addition to the contents of data communication, it analyzes whether a data packet has been observed during a real-time period or a non-real-time period. A security event is detected when data traffic allowed only for a non-real-time periods is observed during a real-time period.

Since industrial automation networks are typically engineered, the communication behavior of the connected nodes under normal circumstances is known. This information can be used to support the analysis, which can either increase the efficiency or decrease the requirements on the hardware of an IDS/IPS system.

 ISBN: 978-1-61208-209-7

D. Real-Time Aware Network Access Control

Network Access Control (NAC) technology provides the functionality of verifying that a network accessing host complies with a dedicated policy [12]. Typical application examples are given through enterprise networks, were mobile clients like laptops, tablet PCs, or personal digital assistants, are frequently connected for local or remote access to the network.

Using NAC functionality it may be checked, if the client (device) has a certain operating system and what the patch status of this operating system is. Moreover, additional features like the availability of a virus scanner and the pattern status may be checked as well. If the client complies with the given policy, access to the network is granted. If not, the client may be put into a quarantine zone, in which he can be updated with the appropriate software packages. This can be depicted as multi-step approach consisting of:

- Detection of devices connecting to the network
- Identification and authentication of clients/devices
- Validation of device compliance to a given policy
- Authorization of compliant devices or
- Remediation of clients to comply to policy

The NAC functionality is standardized by the Trusted Network Connect Working Group (TNC WG) of the Trusted Computing Group (TCG) [11]. The IETF also addresses this functionality in the Network Endpoint Assessment (NEA) Working Group [13].

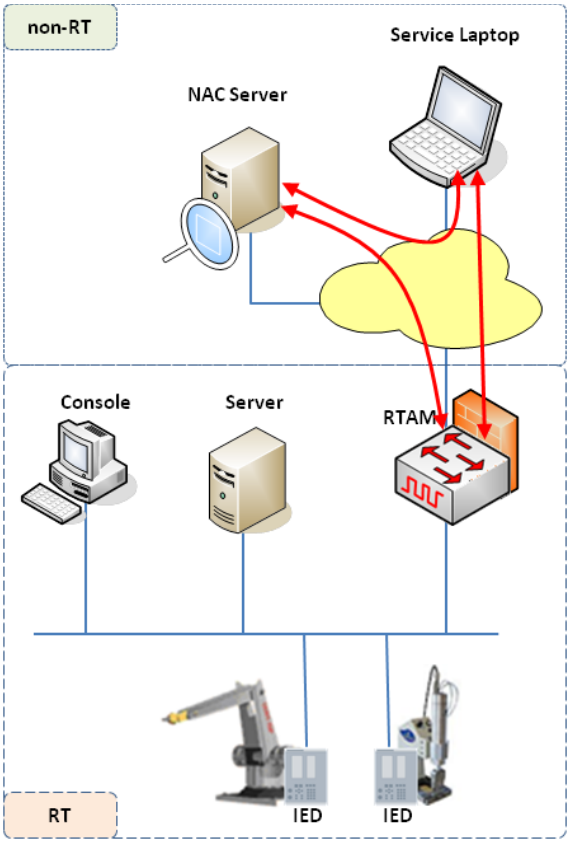

Figure 5. RTAM-NAC-FW Integration

While the network setup in industrial automation used to be static, future deployments, as currently discussed in ongoing funded projects, may be more dynamic. Providing the opportunity in industrial automation for quick replacements of IEDs, attachments of service laptops, remote administration and also agile production requires the compliance of the device compliance to a local policy to ensure reliability of the industrial automation. Moreover, this approach may also be used to realize a secure inventory management, which collects information about all nodes in a network, targeting the observation and maintenance of system integrity. Hence, NAC functionality is increasingly required in industrial automation, too.

RTAM may benefit from a NAC to ensure that only devices complying with a dedicated policy connect to the module. One realization option here is to combine the RTAM module with the NAC functionality of a switch as depicted in Figure 5. Before granting access to the industrial automation cell, the service laptop has to be checked regarding compliance to the operator's policies. Once this process has finished, the RTAM gets the confirmation from the NAC server. Depending on the confirmation, the RTAM may provide the service laptop with an unrestricted or restricted access option to the real-time network. The restricted access variant combines the firewall approach additionally with RTAM to, e.g., check the commands intended for the real-time network before actually submitting them.

IV. Related Work

Firewalls for filtering network traffic are a widely deployed security technology. While firewalls are available supporting real-time applications like VoIP (i.e., SIP and RTP), they do not support industrial real-time Ethernet layer 2 communication [14]. Also protection mechanisms against denial-of-service attacks are known that limit the allowed data rate, but these do not address specifics of real-time Ethernet as well [15].

Intrusion Detection and Prevention functionality is already known from and successfully used in telecommunication networks to detect and defeat unexpected traffic, e.g., through network externally launched denial of service attacks or network-internally deployed malware. An overview about IDS and IPS systems is given in [16].

They are typically not used in industrial automation networks. As stated in section III, the engineering of industrial automation networks can be considered in the definition of rules for IDS/IPS systems. This can even be automated if the engineering data is provided to these systems.

V. Summary and Outlook

This paper described a new real-time security mechanism that filters network communication depending on whether the data affects a time period reserved for real-time communication. The mechanism allows connecting different kinds of network devices to a real-time control network segment securely by ensuring that the real-time communication within the network segment is not affected.

 ISBN: 978-1-61208-209-7

It goes far beyond well-known protections against denial of service attacks that do not respect the specific operation characteristics of real-time Ethernet networks, and filtering of data communication by a firewall depending on the contents of data packets.

This is a basic security feature enhancing known network security mechanisms to support various use cases in which network devices that cannot be trusted to respect real-time restrictions can anyhow be connected securely to a real-time network segment. Examples include temporarily connected monitoring, maintenance and service devices as well as dynamic plug-and-play of different automation devices.

REFERENCES

[1] M. Felser, "Real-time Ethernet – industry prospective," Proc. IEEE, vol. 93, no.6, June 2005, pp. 1118-1128, http://www.felser.ch/download/FE-TR-0507.pdf [retrieved: June 2012].

[2] T.S. Sidhu, M.G. Kanabar, and P. Palak, "Implementation issues with IEC 61850 based substation automation systems," Proc. Fifteenth National Power Systems Conference (NPSC), Dec. 2008, http://romvchvlcomm.pbworks.com/f/p274.pdf [retrieved: June 2012].

[3] IEEE, "IEEE standard for information technology-specific requirements - part 3: carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications," IEEE standard 802.3-2008, 2008, http://standards.ieee.org/about/get/802/802.3.html [retrieved: June 2012].

[4] Profibus Nutzerorganisation, "Profibus system description – technology and application," Nov. 2010, http://www.profibus.com/nc/downloads/downloads/profibus-technology-and-application-system-description/download/12821/ [retrieved: June 2012].

[5] Interbus Club, "Interbus basics," http://www.interbus.de/get.php?object=497 [retrieved: June 2012].

[6] Profibus Nutzerorganisation, "Profinet system description - technology and application," http://www.profibus.com/nc/downloads/downloads/profinet-technology-and-application-system-description/download/11864/ [retrieved: June 2012].

[7] EtherCAT Technology Group, "EtherCAT the Ethernet fieldbus", May 2012, http://www.ethercat.org/pdf/english/ETG_Brochure_EN.pdf [retrieved: June 2012].

[8] Sercos international, "Sercos the automation bus," April 2012, http://www.sercos.de/sites/default/files/sercos_en_april_2012_v1.pdf [retrieved: June 2012].

[9] IEEE, "IEEE Standard for a precision clock synchronization protocol for networked measurement and control systems," IEEE standard 1588-2008", 2008, http://standards.ieee.org/findstds/standard/1588-2008.html [retrieved: June 2012].

[10] IEEE, "Port based network access control", IEEE standard 802.1x-2004, Dec. 2004, http://standards.ieee.org/getieee802/download/802.1X-2004.pdf [retrieved: June 2012].

[11] A. N. Mahood, C. Leckie, J. Hu. Z. Tari, and M. Atiq, "Network traffic monitoring: application to SCADA security," Springer Handbook of Information and Communication Security, Springer, 2010, pp 383-405.

[12] TNC – Trusted Network Connect, http://www.trustedcomputinggroup.org/developers/trusted_network_connect/ [retrieved: June 2012].

[13] IETF NEA WG, http://datatracker.ietf.org/wg/nea/ [retrieved: June 2012].

[14] R. Falk and S. Fries: Voice Security: Sichere Sprachkommunikation in Unternehmen unter Benutzung aktueller Technologien, vde Verlag, 2008.

[15] John Ioannidis and Steven M. Bellovin, "Implementing pushback: router-based defense against DDoS attacks," Proc. Internet Society Symposium on Network and Distributed System Security, 2002.

[16] Karen Scarfone and Peter Mell, "Guide to intrusion detection and prevention systems," NIST SP 800-94, Feb. 2007, http://csrc.nist.gov/publications/nistpubs/800-94/SP800-94.pdf [retrieved: June 2012].

 ISBN: 978-1-61208-209-7