INGHBS **Cyber Security**

Secure Connections to Control and DCS Systems

Engineering Office Schüppen

2018/04/16

1 Content

2 The Cyber Security Concept

2.1 General information

The concept presented here is not a specific or individual cyber-security solution, but the consistent application of the currently known and required, or at least recommended, state of the art protective mechanisms and measures, such as the globally recognized principles for the protection of IT structures and automation systems according to ISO/IEC 27002[1] and IEC 62443[2]. According to the nomenclature of the IEC-62443 layer model, this paper deals with the implementation of the outer layer: “Perimeter protection” or “Protection against unauthorized access to the control level”.

What distinguishes our concept from others is mainly the fact that to the greatest possible extent it manages to resolve the fundamental contradiction between maximum security and easy handling of the system. This is done by using a modular system and consistently applying the underlying mechanisms.

The purpose of this guide is to help decision-makers faced with the task of installing a cyber-security solution and users of an implemented solution to better understand their needs or the implemented systems and mechanisms.

First, it should be noted that there is a significant difference between the protection of “normal” IT systems and the protection of controllers and control systems. This is due to the fact that according to the German laws, standards and technical guidelines in the creation and implementation of security concepts for control systems, economic considerations should not play any role, but everything possible according to the state of the art has to be done to avert dangers to life and limb. Or legally speaking, in the words of § 8a BSIG[3]: “*Organizational and technical arrangements are appropriate if the effort required for this is not disproportionate to the consequences of a failure or impairment of the affected critical infrastructure.”*

In contrast, security concepts for “normal” IT systems are designed on the basis of economic considerations. This is permissible in the IT environment, since designing such systems with too little protection “only” has economic consequences.

[1] Issue 06/2017.

[2] Further development of the ISA 99.

[3] Act on the Federal Office for Information Security (BSI-Gesetz - BSIG) with amendments from 17.07.2015: Act to increase the security of information technology systems.

Before the details of the concept are explaned below, we will first give a few considerations on cyber security in order to familiarize even less experienced readers with the context of the concept:

1. Security is never achieved through the use of components (such as a firewall). Permanent safety according to the state of the art can only be achieved by the professional arrangement and configuration of the corresponding hardware and software components, in conjunction with the responsible use of the system.
2. Existing cyber security solutions must always be checked and, if necessary, adapted using the latest technology (§ 8a BSIG).
3. A security concept and its implementation are never stronger than its weakest part.

⚠ ***When protecting a control system, every state-of-the-art measure must be implemented to increase the level of safety, since it is a matter of protecting life and limb!***

2.2 Basics of the concept

The aim of the cyber security solution is to separate the automation system from the corporate network in such a way that unintentional manipulation of the automation system is reliably prevented[4]. In order to achieve this goal and still allow targeted communication between the levels, the necessary principles are defined within this concept. Just by using components such as a firewall, the security of a system cannot be established. Only the complete implementation of the regulations listed below, in conjunction with the necessary hardware and software components, results in a complete, state-of-the-art solution!

2.3 Basic principles

The following cyber security policies or regulations may also be used as a checklist for the analysis of existing security solutions. A security solution is safe according to the state of the art, if the implementation of each rule can be positively affirmed.

[4] It is widely acknowledged that there is no 100% protection. That makes it all the more important that we do everything possible to achieve this 100% protection.

1. The first basic principle is that in the initial state connections between the systems are not possible. In analogy to the European General Data Protection Regulation (DS-GVO), connections between the automation system and the company network are only permitted if there is a (mandatory) need for this.
2. Direct connections between computers of the corporate network and computers of the automation system are not permitted!
3. Rule 2 is realized according to the prior point in that both systems are separated by at least one neutral zone, the purpose of which is to implement connections or requests by a secure combination of hardware and software components that make it impossible for users or services of a network to reach a system of the opposite side.

   This zone is commonly referred to as DMZ (demilitarized zone).
4. Such a DMZ can be realized by a firewall with appropriate functionality. However, according to the requirement for maximum security, we recommend the use of two physically separate hardware components, or using the virtualized solution the application of two virtual firewalls. In case of a faulty temporary configuration, this prevents any possible complete penetration. In addition, it allows to delegate the responsibility for the configuration and the maintenance of the two firewalls to different areas of responsibility.
5. According to the motto “What is invisible is difficult to attack”, the concept requires that each network is invisible to any other network. This is achieved by the use of network and port address translation mechanisms (NAT, PAT), which at the same time makes it unnecessary to define routes[5] that provide potential attackers with system-architecture information. Furthermore, the DMZ, which is a separate network between the control system and the company level, has to be hidden from the corporate network in such a way that malware that has penetrated the corporate network cannot identify the DMZ and make the control system a target of its attacks.
6. VPN or SSL tunnels for accessing the automation system are only allowed if they end in a DMZ separated from the rest of the system (see Figure 13). This is not because the encrypted tunnels are not considered secure, but because the entrant at the other end of the tunnel must be considered to be part of the system. This means that a VPN connection to the control system is identical to direct access and thus inpermissible.

[5] Routes define the path through which computers and networks can be reached, thus documenting their presence.

7. Only secure protocols are allowed for the connection between the networks. Whether a protocol is to be regarded as safe in this sense must, if not known, be investigated on a case-by-case basis. In principle, protocols are not secure if they transfer user data such as user names and passwords in clear text[6], contain active content[7] or allow access to file systems and thus modification of the target systems or the introduction of malicious software. Even the outdated DCOM protocol, which is still widely in use, may not be used across networks, as the associated permission settings on the computers are often operated with default settings, they enable attacks on system components of the target system and even worse, also of other systems in the network of the target system,.

> ⚠ ***The selection of permissible ports is one of the essential building blocks of security. A release of unsafe ports calls into question the entire outcome of the cyber security solution.***

8. As far as possible, only read-accesses to the control system should be permitted, as for stations that are able to transfer information or data to the automation system, the same securing effort has to be made as for the automation system itself. Likewise, in such a case, the transmission paths have to be secured accordingly.
9. The firewalls shall not transmit any protocols and/or commands that determine the architecture or the existence of computers, networks or network components[8].
10. Passwords used for cross-network actions[9] must comply with the current complexity guidelines and be renewed at regular intervals.
11. User names of the last access should not be displayed in logon masks.
12. The same users must never be created on the computers of the control system and the company level.
13. Systems should be configurable only from the secure inner side.

To build a state of the art secure connection based on these principles, two firewalls must first be set up so that they do not allow connections between the systems. Then, for each desired communication of the enterprise level with the automation system, a port (communication channel) from the source system on

[6] Known representatives of this genus are: telnet, ftp and rsh.

[7] These are e. g. executable scripts and / or program parts.

[8] Well-known representatives are e. g. Internet Control Message Protocol (ICMP), Cisco UDLD, Simple Service Discovery Protocol (SSDP).

[9] Basically this should be kept on all computer systems.

the corporate network must be released into the DMZ and a possibly different port from the DMZ to the target system in the control system. Doing so, there is no way to get from the corporate network directly into the automation system.

To make communication between the systems happen in this constellation, it requires a corresponding instance for each type of communication within the DMZ, consisting of hardware and software that converts the requests between the ports so that it results in a secure information exchange. As part of the modular cyber security kit presented here, there is a separate architecture for each type of communication in the form of preconfigured hardware and software. The most common communication types between enterprise level and control system are:

- Database access
- Real-time data access[10] as well as access to historized information, such as measured values, messages and alarms, e.g. based on the OPC industry standard
- Access to the operating and configuration interface of the automation system
- Access to the interface of ERP and other enterprise-level systems from the automation system
- Exchange of files, such as reports or virus signatures.
- Remote access to the control system for the purpose of troubleshooting or maintenance

Based on these most common types of communication, the procedure for setting up secure connections can be explained and it can be easily derived how other communication channels can be set up safely.

3 Secure Architecture – Implementation

3.1 General Information

As long as 15 years ago, in 2003, the Slammer Worm virus emerged, which, due to the frequent use of the Microsoft® SQL database in industrial environments, frightened many industrial companies because production and manufacturing systems were affected. All systems connected to the corporate level via the Mircosoft SQL Server port (1433) as shown in Figure 1, were potentially jeopardised, even if they were separated by a firewall! When the

[10] Normally with cycle times equal or greater than 1 second.

worm invaded from the infected company level PCs in the control system, it caused extreme network and CPU utilization and the affected systems could only be operated with considerable delay or could no longer be operated at all. In Germany, the consequences of the Slammer virus – so far as this is known - not as dramatic as first feared. But they could have been dramatic, if the virus had also contained a destructive component!

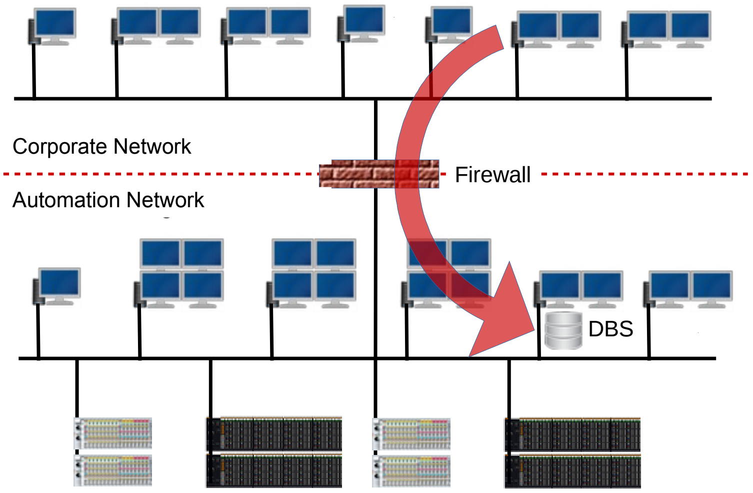


Figure 1: *Database connection with single-stage firewall.* ***Such a solution does not provide sufficient protection!***

This example, on the one hand, proves the existential importance of rule 7 in section 2.3 and that the potential danger of each port must be re-evaluated after each update of the associated software. On the other hand, it also shows that using a firewall alone cannot guarantee a secure solution. Nevertheless, this “idea” is still common in many minds and a reality in many plants.

How to configure a database connection that can be considered as safe is shown in the next section 3.2.

3.2 Secure Database Connection

In modern database systems, due to their complexity, an unforeseen but possible use of faulty algorithms or unplanned side effects can not be excluded. This is very clearly illustrated by the example "Slammer worm" in section 3.1. Therefore, in order to enable a secure data exchange of databases on enterprise level with an automation system, the architecture should be chosen according to Figure 2 and the shared resource “database” should be separated from the respective systems/networks accessing it.

The further proceedings depend on which database is actually involved. Here are the proprietary databases of the control system manufacturers such as Honeywell (PHD), Siemens (PCS 7 Process Historian) or AVEVA (Wonderware Historian) or from the process data specialist OSIsoft (PI System). These have in common that for the communication between automation and database system (DBS) in the DMZ a proprietary port (PortGreen) is available. And for the communication between company level and DMZ, either a proprietary port or, if the system is based on a standard SQL database, the appropriate SQL port of the database (Port Red) can be used.

⚠ *It is, however, always necessary to examine the risks that might result from opening a port.*

Taking into account the security aspects, the proprietary protocols and ports are generally preferable[11], as their less widespread distribution and their low level of awareness already provide their own means of protection, and thus an additional protection.

The opening of the ports must always be chosen in a way that the use of the connection is only possible from outside into the DMZ, and not from the DMZ to the outside. This makes direct access from the corporate network into the control system impossible, because even if malicious software entered into the DMZ it could not connect to the control system. The concept works even with pure standard SQL databases, which are addressed by both sides via the same port. The already mentioned Slammer worm could have infected the server in the DMZ due to an infected enterprise level client and provided for chaos and malfunction on the DMZ server, it might even have led to the loss of data, but the control system itself would not be have been affected, except for the dysfunctional database connection.

[11] Inglorious exceptions are too simple, unencrypted protocols that use dangerous mechanisms such as file transfer implementation.

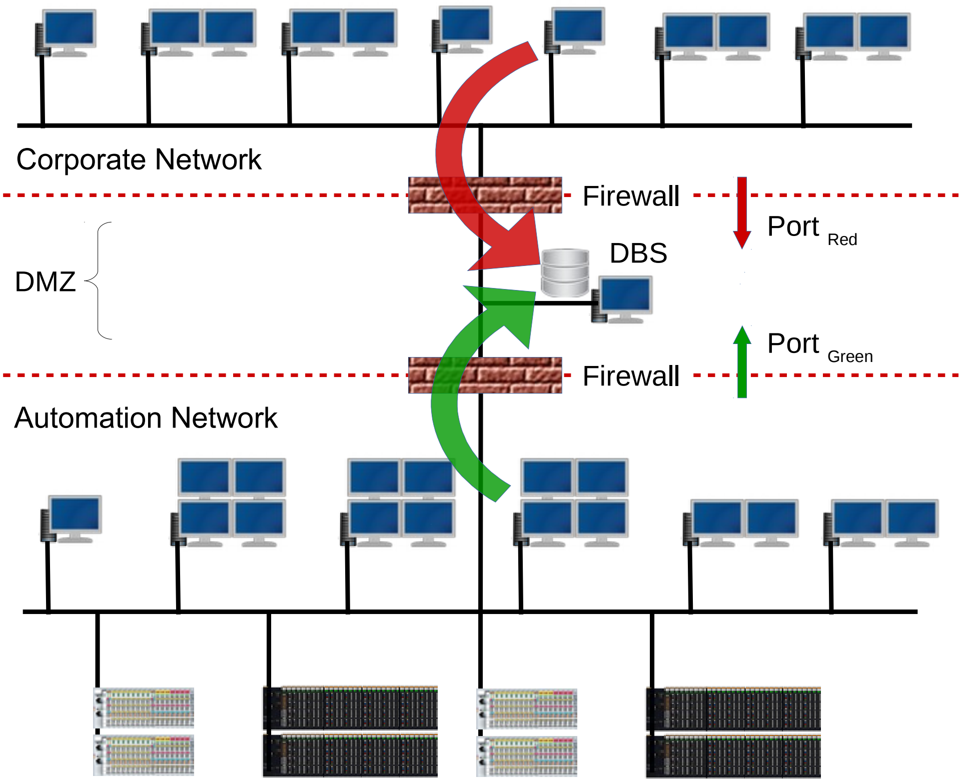


***Figure 2**: Architecture for a secure database connection. The arrows indicate the direction of access.*

Very good, in the sense of secure, connections, can be built with database systems in which the database is supplied with data via a special collector program from the automation level that does not allow communication in the opposite direction, and the evaluation of the data from the enterprise level is done by pure SQL queries[12].

The classic procedure for setting up a system according to Figure 2 would be to address the database computer from the client systems with its IP address and specify the on-the-way firewall as a router (switching station). The disadvantage, however, is that there is information on the client systems from which attackers can determine the architecture of the system. For the purpose of maximum safety this should be prevented. For this purpose, the database system is mirrored via Network Address Translation (NAT) in the corporate and the automation network. This is a performance of the firewall and causes the database system to exist virtually in both networks. The database computer now has an IP address corresponding to the address range in all three networks (company level, automation level and DMZ). In this procedure, structural information is only available in the most secure components, the

[12] If the results of such evaluations are to be made available at the automation level, a secure connection according to section 3.4 is recommended.

firewalls. For an attacker in the corporate network, the system presents itself as shown in Figure 3. The real database system is not visible, other targets can not be identified by an attacker.

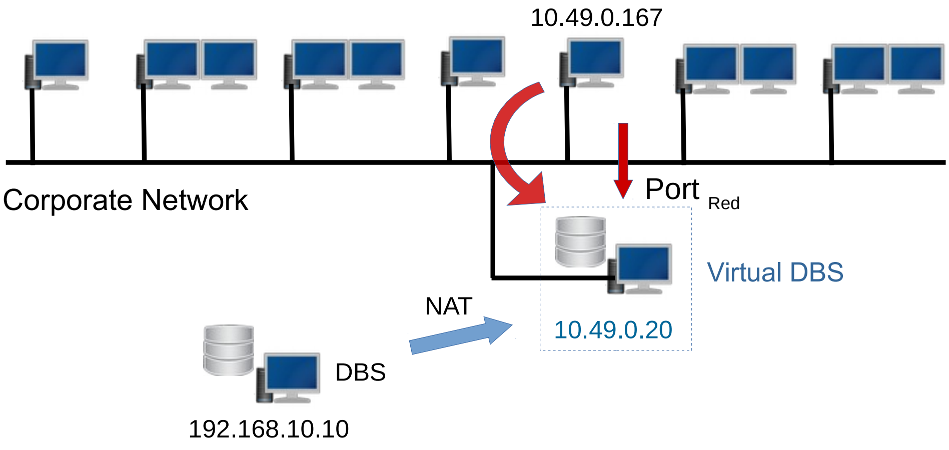


***Figure 3**: The DMZ network is rendered invisible by mirroring the DBS into the corporate network.*

By the automation level firewall the database system is mirrored in the same way.

Once you have got used to this somewhat unusual procedure, there are great advantages for maintenance because the networks - including the virtual systems - are self-contained and can be managed independently. Cross-network considerations only have to be done when configuring the firewalls.

3.3 Secure OPC Connection

OPC[13] is today the defacto standard for data exchange in the automation industry. The most widely used substandard today is OPC-DA for the exchange of real-time values, followed by OPC-HDA for accessing historical values and OPC-A&E for the transmission of alarms and messages. The latest substandard OPC-UA is designed for the exchange of all automation data types including the corresponding metadata. Most manufacturers of automation systems have already implemented implementations, but in practice they are

[13] OPC stands for "OLE for Process Control", where OLE again stands for the Microsoft® standard "Object Linking and Embedding", the precursor of COM/DCOM. For long this has been the most important part of the Microsoft® Windows system and software development and is still widely used in todays software. Originally OPC was based on DCOM only, the newer standards such as e.g. OPC-UA don't need it anymore.

less common because of the complexity of making 1:1 connections. These 1:1 connections are the main application for automation system interconnections.

From a safety point of view, the more common older OPC standards have significant disadvantages:

- They are based on the DCOM standard, which holds a great danger potential.
- Usually they require dynamic port shares, which is why they can not be used with a firewall.
- The DCOM standard does not allow network address translation.

Secure OPC communication across network boundaries is not possible for these reasons. However, since a direct connection between automation and company level is not allowed anyway, this does not represent an additional restriction.

The solution is to limit the DCOM or OPC communication to the DMZ itself. An architecture according to Figure 4 should be chosen. The OPCGate is a computer with a specially hardened operating system on which both the OPC client and the OPC server are installed. A typical application would be the connection between a Siemens PC7 system and the PI process database from OSIsoft. In this case, the OPC server would be from Siemens and the OPC client from OSIsoft. The decisive factor for a secure connection is that the participating systems have OPC server and client applications that are executed as stand-alone programs and communicate with the main system via a proprietary port, since otherwise they cannot be installed on a computer in the DMZ. The company OSIsoft for example provides such OPC clients for their system, while with Siemens and other control system manufacturers, depending on the used system, there are also OPC servers running only on the automation system stations, sometimes even as controller plug-in cards. In these cases a so-called OPC tunnel software such as Softing's dataFeed tunnel is used. This is a client-server software that communicates with the automation system server as an OPC client and transmits the requests via a dedicated port to the server side of the tunnel, which then in the DMZ talks to the installed OPC client as a server. Good OPC tunnels work completely transparent, so no duplicate configuration is necessary.

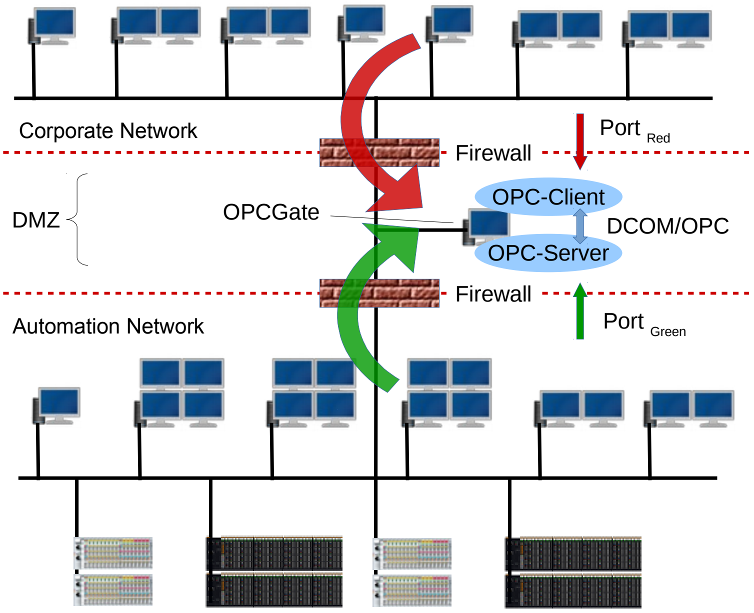


Figure 4: *Architecture for a secure OPC connection.*

Although the use of the OPC interface, especially the older standards, requires a great effort to make secure connections, it is the first choice from connectivities point of view, since all manufacturers of automation systems or components provide proven OPC (DA) servers and/or clients and system connections can be made quickly and without “experimental phase”. Examples of such OPC connections are the horizontal integration (control system - control system), the connection of the automation system to databases that are located in other security zones (DMZ) of the corporate network, as is often the case with the OSIsoft PI system, Advanced Control applications (e.g. Aspen DMCplus®) or information systems such as from AspenTech or Matrikon.

⚠ It is quite common today to transmit more than 10,000 values per second over such links. A sufficient bandwidth of all involved components is therefore to be considered!

The use of the OPC standard is a good example of the fact that the perimeter protection described in this document is not sufficient for the safety of an automation system alone, as shown by the following example from practice:

In a plant of the process industry, an OPC connection has been established to a process database in the enterprise network according to the rules described here. Three days later, the entire system fails. The initial cause is the fact that the system was no longer operable because the connections between the operating stations and the controllers failed. Subsequent investigation revealed that obviously many connections between the controllers of the system did not work anymore. What happened?

The process database administrator had started configuring the database. He wanted to gain an overview of the system to be configured via the browser functionality implemented throughout the system. It is not known if he used the function “*. *” (= Everything) knowingly or accidentally. The decisive factor is that subsequently the OPC server of the automation system tried to access all the real-time values and parameters of the system and thus overloaded the controllers of the system. The system became inoperable and had to be shutdown.

From this example, the following clues can be derived.

> ⚠ An automation system can be overloaded with read operations. Therefore, anyone who accesses an automation system through interfaces needs to know the system and it’s behavior in detail.

> ⚠ Every access to the automation system consumes resources in this system, which are finite in any case. Anyone who wants to read information from the automation system not only has to know these limits, but also the current degree of resource utilization.

> ⚠ Before using OPC servers, it is recommended to monitor the performance of the affected automation system and calculate the additional load of the system in advance, based on the - hopefully existing - system manufacturers informations. The decisive parameters for this are typically the number of measured values, the recording rate and the dead band.

> ⚠ The load introduced into the automation system by the OPC servers today is often by far the highest resource load of the system.

For completeness, it should be noted that the OPCGate, as shown in Section 3.2, is mirrored by the NAT mechanism (Figure 3) in the other networks in order to limit the structural information to the firewalls.

3.4 Secure Access to User Interfaces

To not call a program's user interface locally on the computer on which it is executed, but on another computer connected via network, is a feature that has become indispensable ever since the introduction of virtual systems. UNIX or

Linux-based systems provided this mechanism “from the beginning”, named X11 protocol; Microsoft Windows implemented a similar mechanism using the Remote Desktop Protocol (RDP). The more efficient RDP protocol is now disclosed and thus can be implemented freely. In addition to these two protocols, there are more, like VNC, NoMachine NX and others, but the RDP protocol has the great advantage that it is today freely available on any computer architecture (Microsoft Windows, Linux, Apple iOS and Android). Often there are even several implementations, it is eihter already preinstalled, or can be installed via the appropriate software marketplace.

There are a number of reasons for remote operation of programs:

- It is a virtual system that does not have a physical screen.
- It is a computer to which no monitor is connected in normal operation.
- Several computers in distributed systems should be served from on place.
- The computer systems to be operated are spatially far away.
- For security reasons physical access to the computer should not be possible for the operator.

In everyday business it is commonplace for plant managers and MSR staff to access the control and configuration interfaces of the control systems from their office workstations to avoid permanently changing rooms between the office and the control room and/or control rooms and to minimize pathways.

⚠ Without adequate security, however, such a connection between office PCs and the automation system is considered negligent, if not even grossly negligent!

Access from the office PC to the terminal server of a control system violates rule 2 of section 2.3 that a system must never be accessed directly.

In contradiction to this is the fact that remote access to computers from a security point of view offers a number of advantages:

- The user has no physical access to the system he is working on, so he can not inject viruses via floppy disk, CD or USB.
- For the same reason, he can not gain control of the system through direct access. An example of this would be to boot and manipulate systems from CD or USB stick. A possible manipulation of the system is difficult to prevent in a direct access, as it gives attackers a number of additional possibilities.
- If configured correctly, the system can not be rebooted or shut down by the user.

However, these safety aspects only apply if solely graphic data is transmitted and the basic requirement of Rule 2 in Section 2.3 is met. The fulfillment of these conditions is by no means self-evident. Here, in particular, the actually advantageous RDP protocol from Microsoft presents as problematic, since it also allows the integration of drives and the linking of other services of the accessing system in addition to the function of graphical information transfer. In plain language, this means that a USB stick, which is integrated in the file system of the accessing system, can also be integrated into the file system of the control system terminal server. This can be set on the client side when calling the RDP client and works wherever it has not been deactivated on the server side. There are a number of reasons for such an omission in practice. The most important are:

- **Lack of knowledge.** The corresponding settings are not made because the problem is not known at all. For example, it’s erroneously assumed that it is sufficient to secure the connection through a firewall.
- **A change to the system policies is not permitted.** Many control system manufacturers deliver preconfigured system stations with defined policies. Deviating settings may not be made without the manufacturer's warranty becoming void. This is understandable while by applying wrong settings may break the functionality of a system.

From the above one could easily conclude: “If everything is configured correctly - and that should be the case with cyber security solutions anyway - a direct access would be safe to arrange”. But this is a big mistake. The RDP protocol is so powerful that the real implementation repeatedly contains errors that can be exploited as security vulnerabilities to attack the target system. An Internet search with the keywords “RDP”, “Protocol” and “Vulnerability” gives an insight here. Most recently, in March 2018, such a serious vulnerability was discovered that allowed program code injected on the target system to be executed. All Windows versions from Windows 7 to Server 2016 (CVE-2018-0886) were affected.

Therefore, for safe access to user interfaces via the RDP protocol an architecture according to Figure 5 must be selected. With this arrangement, the computers on the corporate network can only access the HMIGate in the DMZ via RDP. This HMIGate is a hardened machine with Linux operating system running special software (ScreeKnox) whose task is to control the incoming RDP calls depending on the user and related settings, and connect the session via another RDP session (or other protocols such as ssh or X11) with the permitted target systems. The configuration of the ScreeKnox[14] software

[14] ScreeKnox is an artificial word from Screen and Fort Knox to symbolize the purpose of the software.

determines to which target systems a user is allowed to connect and which software he can use.

An essential component of the system security in the context of this concept is the use of the Linux operating system on the HMIGate. As a result, two fundamentally different implementations of the RDP protocol are used on the way to the target system. Vulnerabilities, no matter in which implementation end at the other. This creates a safer connection than using two times the same implementation.

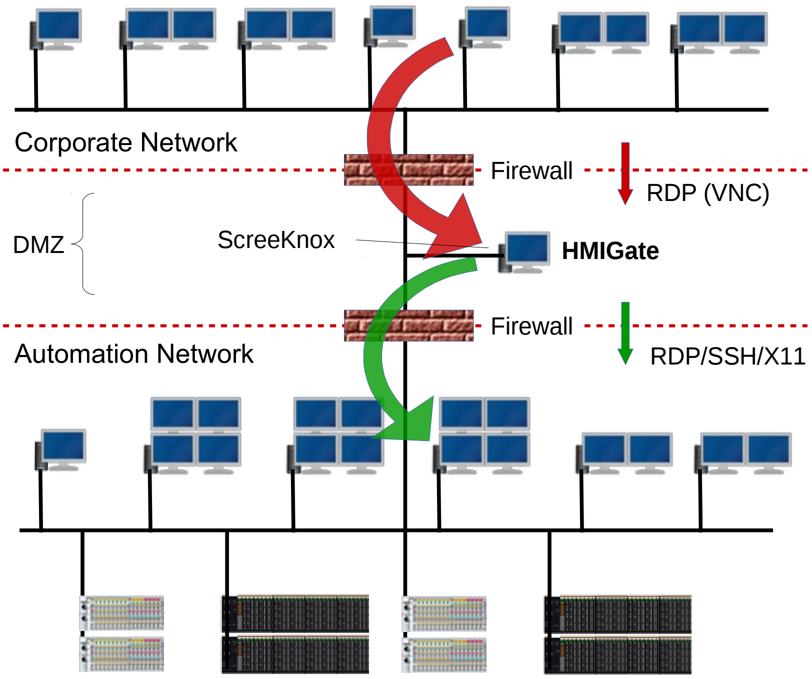


Figure 5: *Architecture for secure access to automation system operating or configuration interfaces.*

For the enterprise-level PCs, RDP was chosen as the default[15], as no additional software needs to be installed on the enterprise PCs. For users, there is little difference to a normal remote desktop session.

The HMIGate must also be mirrored into the corporate network through the NAT mechanism, as shown in Section 3.2, Figure 3, in order to limit the structure information to the firewalls. While in opposite to the RDP example the terminal server session in the automation system is called from the DMZ, the HMIGate is not mirrored into the automation network, but the stations to be accessed are mirrored into the DMZ.

When accessing automation system user interfaces, it should be noted that perimeter protection alone is not sufficient, as is the case with the OPC

[15] The concept also works with VNC, NoMachine NX or similar protocols.

connection. Access always means allocation of finite resources in the automation system. General rules can hardly be set up here because there are big differences between the systems. But the following instructions should always be followed, as even an “unrestrained” access to user interfaces can lead – and has led - to overload and failure of components. It should be remembered that there are companies that allow 100 concurrent accesses to the automation system's HMI while only 20 accesses are made in the control room. The main burden is therefore no longer in the system itself, but is imprinted from “outside”.

⚠ The maximum number of sessions that can be opened in the automation system must be limited depending on the available resources of the control system!

⚠ In some control systems, the additional accesses can also lead to overloading of individual controllers!

The above-mentioned case of access to the user interface of automation systems is hereby representative of any other secure access to user interfaces of programs or systems. Examples are:

- Display of trends and other process data evaluations from information systems in the corporate network in the automation system. Here the access chain is reversed.
- Display of PDF documents such as instructions, standards, system drawings and others which are located on a server in the corporate network. Since modern PDF documents can contain active content, it is a good idea not to transfer them to the control system at all. This is especially true for permanently automatically generated documents such as reports.
- Access to SAP or other ERP systems to track orders, stock data, consumption data and other information.
- Display of Internet browsers that run on computers in the corporate network and thus the ability to securely access content from browser-based systems. This even allows secure Internet access to be set up in the control room.

The HMIGate can provide secure connections in both directions. For the participating computers corresponding reflections must be made in the firewall.

⚠ From the examples given above for secure connections we can see, that the DMZ often has numerous virtual (mirrored) systems and thus a great need for additional IP addresses. In order to reduce this number, as part of this concept in addition to the NAT mechanism and Port Address Translation (PAT) is used. Here, the "real ports" of several computers are mirrored on a single virtual machine (an IP address).

3.5 Secure File Transfer

The file exchange between enterprise and automation level has the highest requirements in terms of security. Although it would be most convenient to completely ban the file exchange, this is not possible in practice. Examples are:

- Virus signatures that must be regularly made available to the control system computers.
- Files for system updates or patches.
- Protocols generated by automation systems and sent to the corporate network.
- Instructions, standards or similar documents to be sent from the corporate network to the control system.

⚠ Before setting up a file transfer from the corporate network to the automation level, the first thing to investigate is whether there is no other way to achieve the underlying goal.

If transferring files is essential after discussing all alternatives, there are a number of requirements to be met in order to get a secure transfer:

- No files with active content may be transferred. Exceptions to this rule are patches and updates that require special care in their storage and use. The application of mechanisms that detect changes to such files, such as using checksums, should be a matter of course to exclude tampering.
- Even when transferring files, there must be no direct contact between enterprise-level and automation-level computers.
- Only files of known types should be transmitted, in which the danger potential is known and can thus be detected and banned during transmission. An example are Microsoft Excel files, which may contain macros and thus malicious code. With the newer XML-based file

formats, this is stored separately and can be detected during a transfer and completely removed. The older formats, where this is not possible, may not be transferred according to point 1.

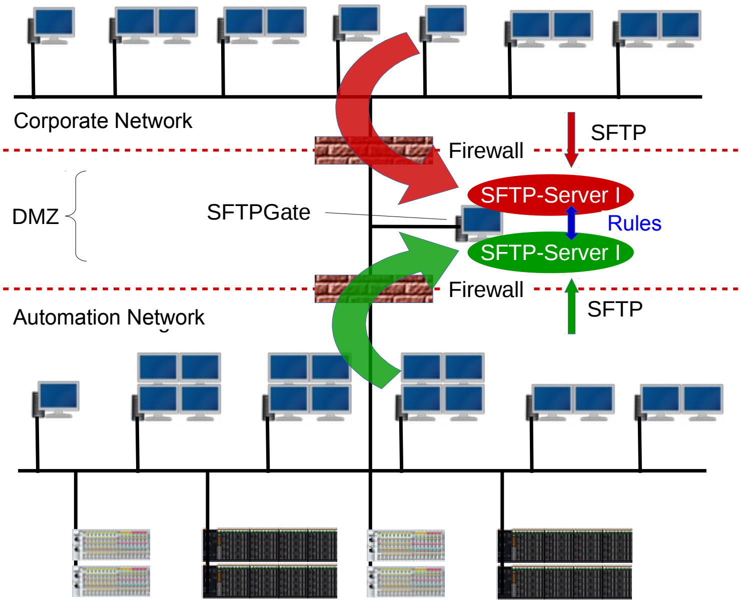


Figure 6: *Architecture for secure file transfer. The arrows indicate the direction of access, not file transfer.*

From the above rules, the basic design for secure file transfers can be derived (see Figure 6). As with the other couplings, a computer with a special function, the SFTPGate, is used in the DMZ.

The SFTPGate is a computer with a hardened Linux operating system. On this two SFTP servers are installed, which can be reached only from one level (company or automation) and in which the authorized users are isolated in a “container” with additional mechanisms, such as “chroot”. Users of one level can only ever read or write files within their SFTP server. The exchange between the two servers is carried out by a set of rules, which monitors the respective directories of the server and carries out the change driven file transfer. The actions of the rules are essentially:

- Determine the file type of new file.
- Delete files with illegal file types.
- Remove active content from Excel files (only possible with XML-based formats).
- Transfer (move) files to the other SFTP server according to the configured target directory (if allowed).

- Logging of the actions taken (delete, transfer).

⚠ Special handling of virus signatures and updates as well as software updates is always necessary if the consistency checks can not be carried out by the rules and no harmlessness has been determined.

The above note could be interpreted to mean that these types of files should not be transferred into the automation system at all. But the question is not, whether or whether not, but only by what means, because the application of patches to correct software errors for example is essential. The only alternative to file transfer via network would be a transfer via CD or USB stick. However, these mobile media are classified as very dangerous due to the associated operating system-side automatisms. For example, the major cyber attacks against industrial companies have primarily used USB sticks and drives and not the network to introduce and distribute the malware into the systems. The first well-known representative of this kind was StuxNet. From this one can conclude::

⚠ **Mobile data carriers should not be used in the area of the automation system!**

⚠ **By using mobile data carriers, the entire perimeter protection of a system can be compromised!**

To access SFTP servers from Windows-based enterprise-class or enterprise-level PCs, open-source programs such as Filezilla or WinSCP can be used, which provide graphical user interfaces similar to the Windows File Explorer.

If file transfers are to be automated, command line tools such as ncftpput/ ncftpget or scriptable programs like the aforementioned WinSCP or PSFTP must be used.

⚠ **Passwords should never be specified in the command line or in scripts, as often shown in examples!**

Instead of using passwords, an automated transfer requires the setup of certificates, eliminating the need to enter a password.

⚠ Computers that are used for automatic data exchange should be protected against the normal corporate network as well as the stations of the automation system (e.g. by a separate DMZ).

The convenience of secure file transfer can be further enhanced by the establishment of corporate-level and enterprise-level interfacing stations that use CIFS or SMB file systems and can therefore be directly integrated as network drives from Windows PCs. These coupling stations monitor changes to their CIFS or SMB file systems and automatically exchange with the SFTPGate. For the user the whole transfer then looks like a simple network drive.

3.6 Implementation of further secure Connections

After the principle of secure connections between enterprise and automation level has been presented, using the most common types of connections in sections 3.2 - 3.5, the framework for other connections can be easily derived. One of the “routine activities” is the establishment of the firewall rules, with port approvals and the mirroring of the computers involved, whereby it is always necessary to investigate which dangers are associated with a port release. The less trivial part is the realization of the gate computer, which transforms the transactions between the levels so that a state-of-the-art secure connection is established.

4 The Cyber Security Kit

The efforts to standardize the above-described connection types according to the concept presented and to implement corresponding components have led to a modular system with which cyber security solutions can be effectively implemented. Components of this kit are as already mentioned:

- The firewalls
- The VPNGate
- The HMI*Gate*
- The SFTPGate
- The OPCGate
- The DBStation

When implementing our concept, we prefer to rely on open source components, i.e., we only use proprietary components if there are no equivalent open source solutions or if certain proprietary systems are required by the customer.

4.1 Firewalls

Following the open source philosophy, the concept prefers open source firewalls based on Linux. The basic system is a hardened Debian and/or IPCop. The actual firewall functionality is implemented in each case on the basis of iptables/netfilter with a cross-system configuration program. The firewalls can either be redundant or clustered to increase availability. We use industrial embedded DIN rail PCs EACIL20 or DR2100 from TL-Electronic or UNO-2271G from InoNet/Advantech (see Figure 7), if no special hardware is desired. GB bandwidth and sufficient computing power is a matter of course with these firewalls. In the choice of hardware there is extensive freedom.

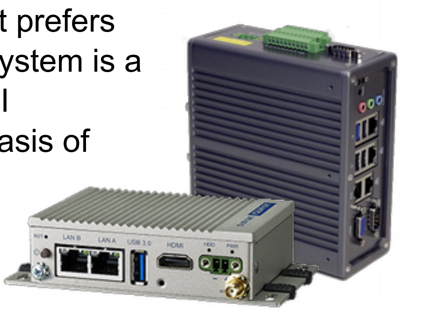


Fig. 7: DR2100 from TL-Electronic (background) and UNO-2271G from Advantech.

Essential for the correct understanding of our solution are three points:

1. The range of firewall functions is not subject to high demands, because in addition to the packet filtering only the basic mechanisms such as NAT/PAT[16] and VLAN[17] are used in this concept. More complex mechanisms such as intrusion detection, application control, web content filtering, malware protection, antivirus, anti-bot, URL filtering, which make up the majority of today's firewalls and which are indispensable for the protection of enterprise networks, are not used in this framework. There is a need for these algorithms whenever the transmission of active content, such. B. JavaScript in HTML / XML streams or PDF documents can not be avoided. In this case you have to constantly examine these contents with the help of current signatures for dangerous components and to eliminate them. The problem here is, that a hazard can only be banned after it has been recognized. Thus, in simple terms, even with the greatest protection effort the systems can fail when attacked with new patterns. When transitioning from the corporate network to the automation level, connections with potentially active content are generally not allowed - as per rule 7 from section 2.3. The firewalls can be compact and robust.
2. VPN connections are not established via the firewall, but always via a separate VPN gate, which is typically installed in its own DMZ, formed by the outer firewall. As a result, the accesses from the VPN gate to the

[16] NAT (Network Address Translation), PAT (Port Address Translation): wird verwendet, um Systeme zu spiegeln und damit Strukturinformationen auf die Firewalls zu beschränken.

[17] VLAN: Virtual Local Area Network. Mechanismus, um physikalische Netzwerke in mehrere virtuelle Netzwerke zu untergliedern.

"actual" DMZ must be enabled through the VPN gate and independently through the firewall. This eliminates a common firewall vulnerability (VPN) and further increases security. As long as the configuration is carried out by a cross-system configuration program, the additional effort is kept within manageable limits.

3. A "simple" firewall as described in point 1 is a pure packet-switching station, there are no users on this computer. Viruses like Meltdown and Specter, which have recently alarmed IT worldwide, lack the level of attack. The use of an "old" firewall such as IPCop as a basic system can even be of decisive advantage.

At this point, a fundamental discussion could follow, whether it is safer to use a modern day-to-day, but complex, or an old compact proven system. Reports such as Heise Verlag's January 2018 "Highly Rated Vulnerability Threatens Cisco Appliances and Firewalls" clearly demonstrate that industry leaders' systems can fail as well. It follows:

⚠ The highest safety standard is achieved with a mixture of the described mechanisms and components.

The motto is: It is primarily the unexpected that stops an attacker.

Our cyber security solutions are designed according to these criteria. On request, we also use firewalls from the ASA series from CISCO or the SRX series from Juniper. Others are theoretically conceivable, but in our understanding it is not sufficient to be able to deposit rules in the firewall, there should also be experience with manufacturers and devices ("How good is the support?", "How vulnerable are the devices?" , "Where are the special problems?", "What's up?", "What's wrong?" ...).

In the case of the Cyber Security Appliance (see Section 4.7), the above-mentioned open source firewalls are run as virtual machines or, if desired, Juniper vSRX virtual firewalls can be used.

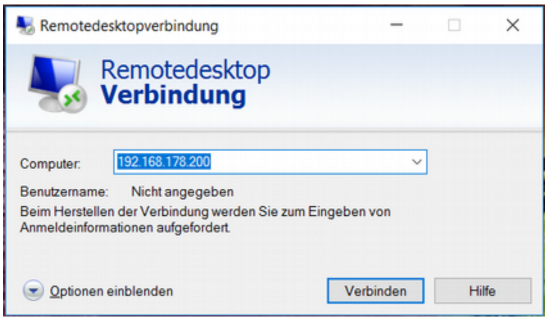
4.2 HMIGate

The HMIGate is a computer with a specially hardened Linux version. Hardened simply means that all unnecessary protocols and services are disabled.

The only allowed accesses to the HMIGate are the Remote Desktop protocol to fulfill the function and ssh access for administration purposes. The ssh access is allowed from the secure side only and limited to specific computers and users. HMIGate's RDP server disables file system, clipboard, and peripheral access. In order to make the system even more resistant against attacks, the users of the terminal server functionality[18] on the HMIGate are not assigned a shell so that they have no access to the computer. By configuring the terminal server session, the application ScreeKnox, which was already mentioned in Section 3.4, is started instead of a user interface. Essentially, the function of this software is to provide the user with a selection of target stations and to connect to the target system based on RDP or ssh/X11 after selection. Functions such as the suppression of hotkeys (Task Manager, Explorer ...) are required to prevent the accessing user from taking control of the system and restricting him to the default programs.

For the user of the solution, access looks like the following:

On the enterprise-level PC - under “Start → Windows → Accessories → Remote Desktop Connection”[19] - the Remote Desktop Client is started and a dialog box appears as shown in Figure 8. In the dialog shown, options for access to the server’s periphery can be set and saved. These settings only have an effect if they are allowed on the HMIGate.



Remotedesktopverbindung

Remotedesktop Verbindung

Computer: 192.168.178.200

Benutzername: Nicht angegeben

Beim Herstellen der Verbindung werden Sie zum Eingeben von Anmeldeinformationen aufgefordert.

Figure 8: *Establishing a Remote Desktop session (options hidden).*

The essential settings are the specification of the target system and the definition of the resolution of the session to be opened. The color depth of the session is never more than the maximum color depth specified on the HMIGate, regardless of the settings.

⚠ If the IP address is specified here instead of the name, it must be remembered that the mirrored HMI gate in the operating network must be addressed, i.e. the virtual address in the company network must be specified.

When saving the options (only possible with visible options), a shortcut is created which then can be copied to the desktop or to the quick launch bar.

[18] The terminal server functionality is provided by the software packages xrdp and Xvnc.

[19] These specifications refer to the operating system Windows 10 Professional.

Even if this is provided in the RDP start dialog, the password should never be saved with the link (check box "Save password" must not be activated!).

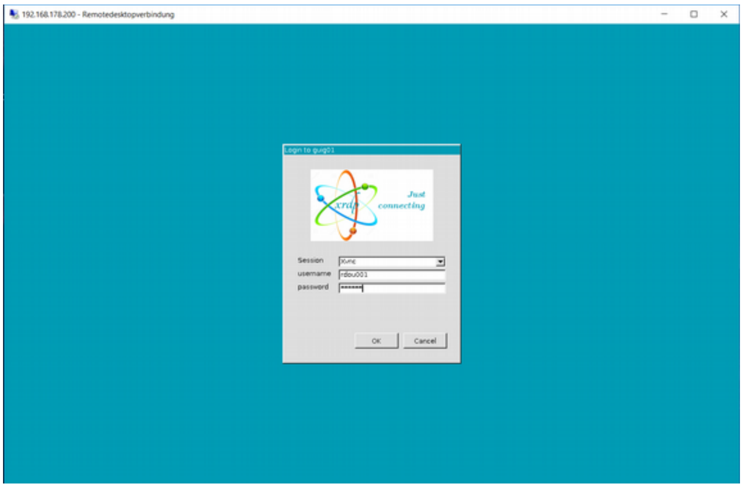
In general:

In shortcuts, scripts and program logins, user names and passwords must not be saved, otherwise a simple but effective protection element will remain unused.

Password complexity and expiration policies should be used in any case as proliferation of passwords can never be avoided completely (in emergency or stress situations, when reading a keyboard ...). The probability of distribution increases with the lifetime of a password.

For low bandwidth connections, e.g. via an ISDN line as they are often used for remote maintenance, the tab “Advanced” in the RDP start dialog is important. Here you can set several parameters that lead to a lower bandwidth consumption. This can be further supported by setting the lowest possible resolution and color depth under the heading “Display”.

When the “Connect” button is pressed the HMIGate’s login screen appears, as shown in Figure 9.



192.168.178.200 - Remotedesktopverbindung

Login to guig01

Session: Xvnc

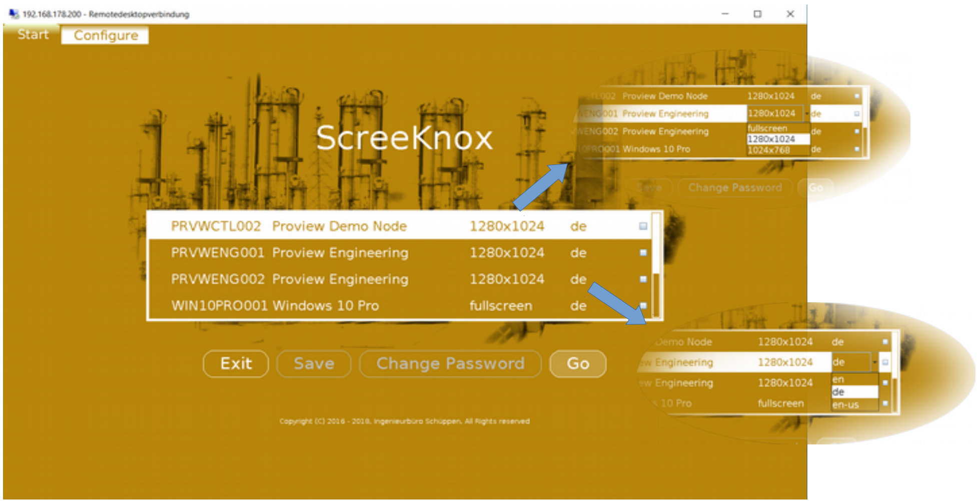
username: rdou001

password: *******

Figure 9: *Login mask of the HMIGate.*

After entering the user name and password, the application ScreeKnox is started. The startup screen of the software appears (Figure 10) and the user can select one of the configured target computers.

The screen resolution and the keyboard layout on the target system can be changed directly in the selection list. The resolution of the target system must bes elect from the predefined values specified by the configuration. This is needed because some program systems - and especially control system surfaces - only work properly with certain resolutions. The same applies to the definition of the keyboard layout. Only layouts the target system needs to work, should be configured. This is usually important in older systems that for example only allow an English keyboard layout.



192.168.178.200 - Remotedesktopverbindung

PRVWCTL002	Proview Demo Node	1280x1024	de
PRVWENG001	Proview Engineering	1280x1024	de
PRVWENG002	Proview Engineering	1280x1024	de
WIN10PRO001	Windows 10 Pro	fullscreen	de

Copyright (C) 2016 - 2018, Ingenieurbüro Schüppen, All Rights reserved

Figure 10: *Startbildschirm des HMIGate.*

Normally the parameters for resolution and keyboard layout are preset depending on the user, so that a change of the settings is usually not necessary.

If the user presses the “Go” button, the session is started on the target system. Depending on the configuration settings, user and/or password for the target system must be entered.

⚠ In this case, the variant with user/password query is not necessarily the safer variant.

After the session is connected, the user can hardly make a difference to a directly connected RDP session (see Figure 11).

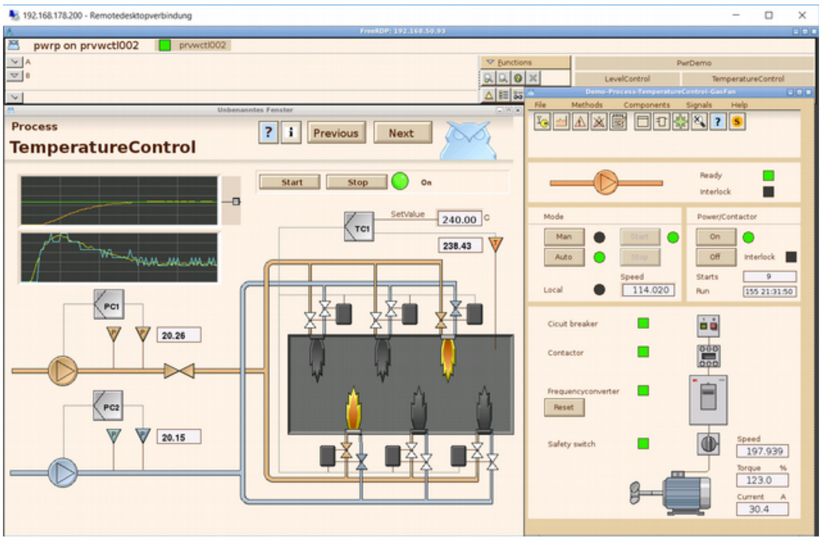


Figure 11: *Screen output of the selected target system.*

User interfaces that are remotely accessed should only allow read access at startup. Operator interaction should only be possible - if at all - after a login/user change within the target application.

Not all programs running locally on a machine will run in a remote desktop session without modification. This sometimes requires adjustments to program and/or program access permissions.

A practical side effect of the concept over direct corporate-level access to Linux, Solaris, or other Unix-based systems that do not provide RDP terminal server functionality results from the fact, that on the PCs no special X11 server software (e.g. Exceed on Demand from OpenText, formerly Hummingbird) has to be installed. Access is always via RDP.

Fig. 12: *CONCEPION®-hX from InoNet.*

As hardware for the HMIGate, the rail PCs from Figure 7 can be used, but we prefer to use the CONCEPION®-hX series from InoNet (see Figure 12) with i5 or i7 processors and due to the integrated IntelAMT they can be maintained remotely, enabling a screen-less and keyboard-less system design.

Because it is so easy to set up and use automation system stations from the company level, it remains to be noted:

⚠ Company-level computers that can access the user interface of an automation system must be specially protected against unauthorized access (secure passwords, locked offices or cabinets ...).

4.3 VPNGate

As described in section 4.1, point 2 on page 22 this concept never uses the firewalls' VPN functionality. The VPN functionality is transferred to a stand-alone system, which does not take on any additional function apart from the VPN functionality and is completely isolated in its own DMZ. A typical architecture is shown in Figure 13.

The VPNGate is a computer with a hardened Linux operating system. The open source package OpenVPN is used for the VPN functionality. The top hat rail PCs shown in Figure 7 on page 22 can be used as hardware. The VPN server can be configured from the secure side via ssh, the access rules are set via the cross-system configuration tool, also from the secure side. In the demo system according to figure 13 this is the station CONF01.

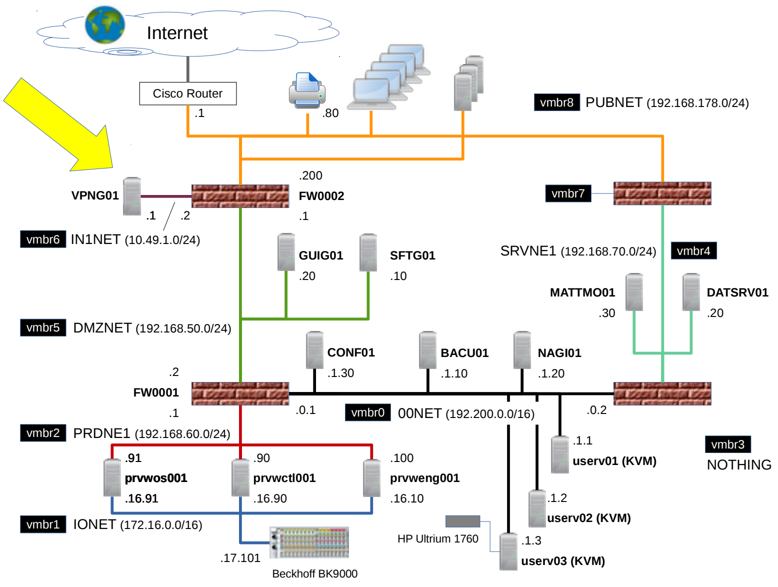


Figure 13: *The VPNGate (arrow) is installed in a seperate DMZ.*

Typical applications for VPN connections are:

- Remote Maintenance
- Secure connections between different areas in the enterprise
- Connections between different locations of a company

4.4 SFTPGate

Function and structure of the SFTPGate have already been explained in detail in Section 3.5 so that only a few additional notes should be given here. No high demands are placed on the hardware of the system, so that one of the DIN rail PCs from Figure 7 on page 22 is normally used here. It uses a standard SFTP server, only the configuration of the system is special.

⚠ When specifying the system, keep in mind that the size of the "hard disk" should correspond to the desired amount of data.

4.5 OPCGate

The OPCGate is - almost always - Windows-based, since the manufacturers of automation systems/components at the moment exclusively provide Microsoft Windows based OPC servers and clients. At present, the operating system versions Windows 10 Professional and Server 2016 Standard are used. If the OPCGate is executed as hardware[20], the server operating system is generally used as it better supports remote maintenance.

The system is hardened by disabling all unnecessary services and protocols and making adjustments to the security policies, such as DCOM connections can only be made locally on the computer.

It is always a stand-alone system that may not be included in any corporate domain.

On the OPCGate the software Filezilla is installed to load the installation software for OPC client and server software as well as subsequent updates - if possible - via an SFTPGate (see section 3.5). The RDP access to the OPCGate should take place in the sense of the concept via the HMIGate. This means that all maintenance work can be carried out without physical access to

[20] For the implementation as a virtual machine, see section 4.7 on page 30..

the computer. Due to the screen and keyboard-less system design, we prefer to use the CONCEPION®-hX from InoNet (see Figure *12* page *27*).

4.6 DBStation

When using databases in the DMZ, three cases can be distinguished:

1. Use of small Windows-based databases
2. Database computers provided, configured and maintained by the company IT
3. Linux-based process databases (e.g. ProviewR-Historian)

For smaller Windows-based databases, we use the same hardware and software systems that are used for the OPCGate (Section 4.5), normally equipped with larger hard drives and more RAM. For maintenance access and the adoption of software on the system, the rules of the concept should be adhered to (use of HMI and SFTPGate).

Often, database computers are provided by the company IT. In this case, the adoption of software updates and RDP access should also be set up to be done through SFTP and HMIGate. If the operating system of the database computer, e.g. due to specifications of the system manufacturer, cannot be hardened, is always to think about whether the computer has to be isolated within the DMZ by an additional firewall.

When using Linux-based databases, we use the same hardware that is used for HMI, SFTP and OPCGate (Figure 12 on page 27). They are also hardened according to all the rules of the art.

For more powerful systems, we prefer to use Dell PowerEdge servers, which can also be preinstalled with the Linux operating system.

4.7 Cyber Security Appliance

The implementation of the presented concept for a secure connection of automation systems to the company level entails a not insignificant amount of hardware. To implement a secure OPC connection it takes:

2 x Firewall,

1 x HMIGate,

1 x SFTPGate,

1 x OPCGate.

Considering that the hardware has to be replaced approximately every 5 years and calculates the associated total effort, the search for suitable alternatives quickly leads to the topic of virtualization.

The implementation of this idea as part of our concept is the *Cyber Security Appliance*. It is a Linux system that is set up as a KVM[21] host with libvirt / qemu support[22]. The hardware systems listed above are now installed as virtual machines on the host machine (or on a NAS[23]). A complete DMZ with all the components described here[24] can be operated on a CONCEPION®-hX with i7 processor and 32 GB of RAM.

Figure 14 shows the structure of our demo system. Only the computers in the PUBNET (upper area, orange network) and the three computers userv01, userv02 and userv03 (in the lower area) exist as hardware.

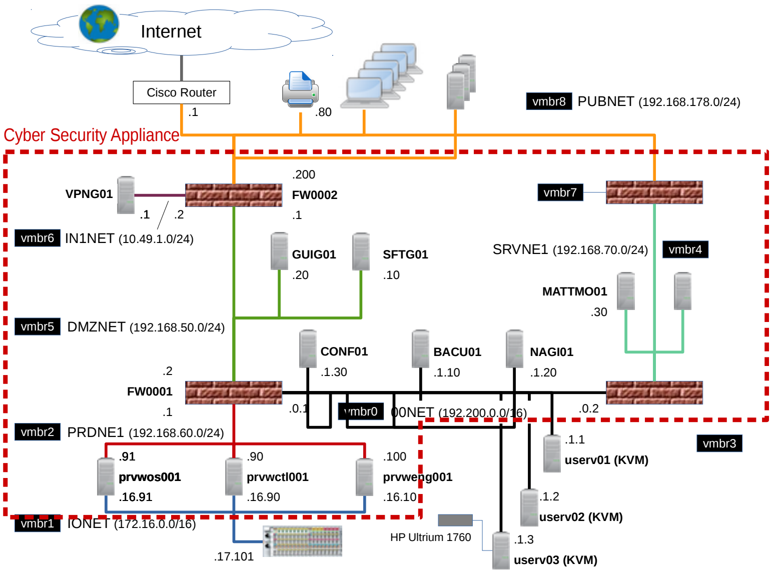


Figure 14: *Cyber Security Appliance Demo System.*

All (!) other computers, including the firewalls, run as virtual machines on the userv01, a CONCEPION® jX with i7 processor and 16 GB of RAM (red-rimmed area). The computer userv02 is identical and is used for redundancy. Because KVM handles live migration, the running virtual machines can be moved in between the two stations while running, allowing the host machines to be maintained and booted without interruption of the running systems and

[21] KVM: “Kernel-based Virtual Machine” is a virtualization mechanism implemented in the Linux kernel, based on the hardware virtualization techniques of the processor manufacturers.

[22] Put simply, tools for configuring and operating KVM (and other virtualizers).

[23] NAS: Network Attached Storage oder Netzwerkspeicher.

[24] Exception would be the DBStation in case of complex databases.

processes. The third processor, userv03, is a CONCEPION® bX with i5 processor, 4 GB of RAM and tape drive connection. It runs the backup system Bareos (as a virtual machine), which makes daily backups of the entire system.

Using the demo system as an example, it can be shown that virtualization not only reduces hardware costs, but also opens up completely new options for dealing with redundancy, data protection and virus protection.

When designing a *Cyber Security Appliance*, these options should be considered and, if necessary, implemented.

5 Conclusion

With the implementation of the concept described here, a perimeter protection according to the prior art for automation systems is achieved.

To consider remains:

What is still considered safe today can be unmasked tomorrow as unsafe.

It is therefore important to always stay up to date and react quickly to changes!