What Protection Engineers Need to Know About Networking

ANCA CIORACA, ILIA VOLOH, MARK ADAMIAK
Markham, ON, CA King of Prussia, PA
GE Digital Energy

1 INTRODUCTION

Retrieving data available in substations, converting it into information for improved decision making and delivering it to control applications and key staff is essential [1]. Under these circumstances, the security and quality of communications between substations and control centers are very important. Towards providing these qualities, many features already in use in the Information Technology paradigm may be adapted and others specifically defined for the electric grid. In this paper we are going to present some of these features.

2 Keywords

P&C = Protection and Control
RBAC = Role Based Access Control
SEM = Security Event Management
AAA = Authentication, Authorization, Accounting
RADIUS = Remote Authentication Dial In User Service
LDAP = Lightweight Directory Access Protocol

3 ROUTING CAPABILITIES IN PROTECTIVE RELAYS

3.1 General

Routing is the process of selecting best paths in a network. In packet switching networks, routing directs data packets from their source toward their ultimate destination through intermediate nodes, such as routers or firewalls [3].

Routers are devices that forward data packets between networks. A router is connected to two or more networks using Ethernet network adapters and they operate at the IP layer of the TCP/IP stack. When a data packet comes from one network, the router reads the destination IP address to determine its ultimate destination and, using information from its routing table, it directs the packet to the next network towards its destination. This way a data packet is forwarded from one router to another until it reaches the destination.

To keep their routing tables up to date in large networks, routers use dynamic routing protocols, as the networks have complex topologies which can change rapidly. Such routing protocols are Routing Information Protocol (RIP) and the Open Shortest Path First protocol (OSPF). In small local networks, such as home networks, a router may be configured with static routes, which are manually added.

Note that the Ethernet network adapter of a device, be it router, a regular computer or any other type, is the one responsible for transmitting data onto the network and receiving data from the network. For this reason the network adapter is usually referred to as a network interface, or a network port. In this paper we will use the term network interface and not network port, to ensure no confusion with TCP/IP ports is possible.

3.2 Routing in Protective Relays

Protective relays are not intended to ever become full routers, in the sense of forwarding packets received through one network interface towards a network connected to another interface.

However protective relays may be equipped with more than one Ethernet network interface these days, either for the purpose of providing the means for redundancy or for separating various types of traffic

and they may also be connected to more than one router for this purpose. When this happens, relays need to know how to send data to destinations reachable through each router.

When such a topology is desired, some basic routing capabilities in relays are essential. A few static routes configured in the relay will ensure that the various types of traffic are directed towards the networks they are intended for.

Figure 1 is an example of such topology, in which the relay has 3 Ethernet interfaces and each interface is used for a different purpose and communicates through a distinct router: Interface 1 passes GOOSE traffic between relays located in the same substation through switch Sw1, but it also passes IEC 61850 Client/Server towards the substation gateways through router R1. Interface 2 communicates with a local control centre where a laptop L1 for remote management of the relay is located and it passes traffic through router R2. Interface 3 passes PMU (Phasor Measurement Unit) data via router R3 to a PDC (Phasor Data Concentrator) device in a nearby substation.

Since in this example the various destinations may not be reached through the same router, three static routes and a default route are used. Two static routes for gateways on networks 10.2.1.0/24 and 10.2.2.0/24 are set through R1, a third static route to PDC is set through router R3 and the default route is set through R2.

Table I – Static Routes Required for Figure 1

Route Name	Route Destination		Route Gateway (Next hop)
	IP Address	IP Mask	
Route 1	10.2.1.0	255.255.255.0 (/24)	10.1.1.1
Route 2	10.2.2.0	255.255.255.0	10.1.1.1
Route 3	10.2.3.0	255.255.255.0	10.1.3.1
Default route			10.1.2.1

In the routing process of a device, the default route is used as the last resort, when no matching route towards a given destination is found. The default route does not specify a destination, but only the address of the next hop, which must be a router on a connected network. It is assumed in this case that the connected router has enough information in its routing table to make a decision on how to send the packets to their destination.

Until recently, relays did not offer static routes functionality, but only a default route, which limited the types of topologies they were able to participate in, for they could communicate outside their network through maximum one router. Consequently, a topology such as the above was not functional. With the introduction of static routes, relays become more flexible and powerful, easy to deploy in various complex topologies.

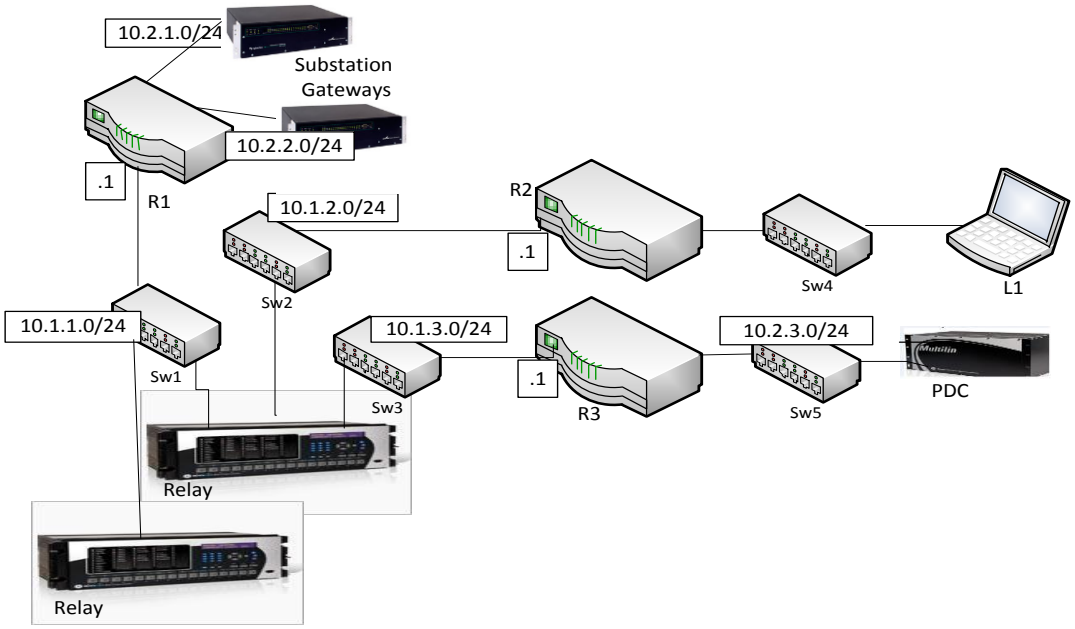

Figure 1 – Example of Topology that Benefits from Routing

4 REDUNDANCY IN PROTECTIVE RELAYS

4.1 General

Towards the goal of achieving self-healing highly available systems within the electric grid, redundancy in the network and redundancy in the devices has been given a lot of attention in the recent years. The IEC 62439 standard suite [4] offers several methods for both types of redundancy. In the first category we have switches and bridges implementing redundancy, by using protocols such as Rapid Spanning Tree Protocol (RSTP). In the second category we have devices, such as protective relays, with two network interfaces attached to redundant networks and sending the same traffic simultaneously through both interfaces. Thus this second type of redundancy is also known as parallel redundancy and it offers zero time recovery, essentially not interrupting the traffic at all, even if one of the paths fails, providing that connectivity still exists on the second path. The two parallel redundancy protocols are Parallel Redundancy Protocol (PRP) and High-availability Seamless Redundancy (HSR), described in IEC 62439-3.

4.2 Characteristics of PRP and HSR

Both PRP and HSR provide redundancy at the Ethernet layer of the TCP/IP stack and they are beneficial for protocols that operate directly on top of the Ethernet layer, such as GOOSE, which must deliver data in real time, hence zero time recovery is essential for it.

Both PRP and HSR require special changes on the participating devices, as these devices must have two Ethernet interfaces that operate in parallel. The two Ethernet network adapters attach to the upper layers of the TCP/IP stack through a Link Redundancy Entity (LRE) module responsible for duplicating outgoing packets and sending them over the two redundant networks, as well as for dropping one of the duplicates at the receiving device. The devices that implement the LRE are called DANPs (Doubly Attached Node that supports PRP) or DANHs (Doubly Attached Node that supports HSR) depending on the protocol they implement, whether PRP or HSR. Frames sent over the dual networks are augmented with an additional 6 bytes, which is either a trailer added after the payload (PRP) or inserted before the payload (HSR). These six bytes are inserted by the LRE and they contain information for the destination doubly attached node, which needs to detect duplicate packets of the same stream and drop one of them, while sending only one to the upper layers of the TCP/IP stack. If one path fails, the destination will still receive one of the two duplicated packets.

PRP may be used with any topology, such as tree or ring, but it requires two independent networks and switches for connecting the devices.

HSR does not need switches and two independent networks when used in a ring topology and in its default mode of operation, Mode H. With HSR the network size is reduced to half compared to PRP. A disadvantage of HSR however is that it cannot be mixed with regular devices on the same ring, as regular devices cannot interpret the inserted HSR tag. Regular devices may still be used in HSR rings if they are attached through a device called “red box”, which is essentially a converter from a singly attached network to a doubly attached network.

Generally speaking, PRP networks make more sense if the size of the network is not a concern. If it is important to keep the size of the network down and if there is no need for regular nodes and the ring topology is preferred, then HSR is a very good alternative, more compact and in no need of switches.

Figure 2 is an example of the PRP network and Figure 3 an example of the HSR network.

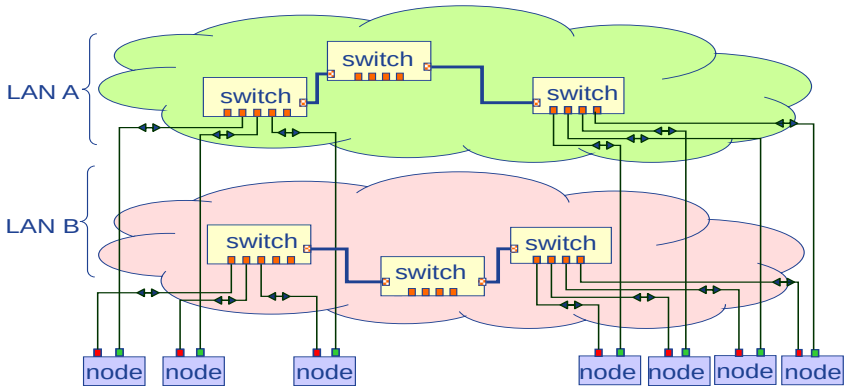

Figure 2 – Example of PRP network

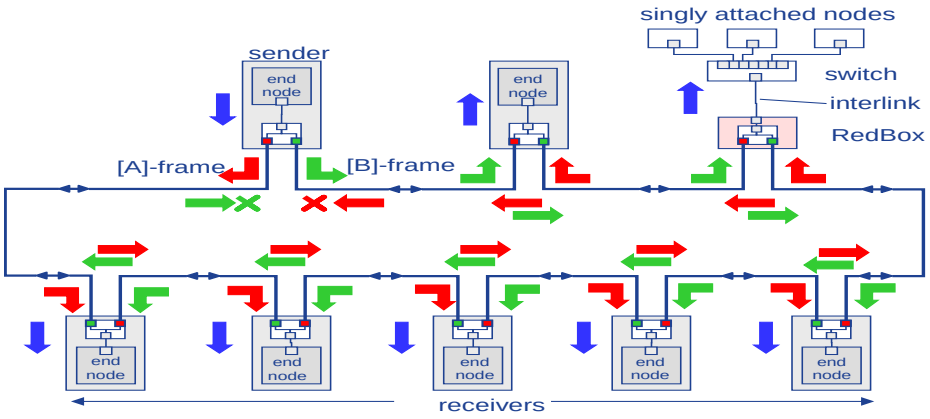

Figure 3 – Example of HSR network

5 SECURITY CONSIDERATIONS

5.1 Standards and Guidelines

Once protection and control devices, such as protective relays, are connected to Ethernet local and wide area networks, the security risk is raised dramatically, as these devices become suddenly exposed to a variety of cyber-attacks for purposes such as industrial espionage, financial or political gain (through blackmail), denial of service and destruction of expensive equipment. The last two types of attacks are especially important in an electric grid, as cyber war is used more and more frequently as a valuable addition to a physical war.

The North American Electric Reliability Corporation recognized this emerging threat and it established and enforced security standards for Critical Infrastructure Protection (CIP) in the electric

grid, which are known as NERC-CIP standards. Also the National Institute of Standards and Technology NIST came up with guidelines for smart grid cyber security, as part of NISTIR7628.

5.2 Security Principles Applied to the Architecture of P&C Devices

Most security principles defined for embedded devices connected over Ethernet are applicable to smart grid devices as well [2]. Among them we mention:

1. Establishing secure defaults, such as in enforcing password complexity by default and disabling all TCP/IP ports not required for the functionality of the device. The OWASP (Open Web Application Security Project) group provides good guidance on password complexity. The password complexity may be optionally disabled and a configuration parameter provided for this purpose on the device, but made available only to a high responsibility role, who would decide to disable it on a temporary base, such as when the device is in service mode. In regard to disabling unused TCP/IP ports, this requirement is specifically mentioned in NERC-CIP 7-5 [5], as requirement 1 (R1): "enable only logical network accessible ports that have been determined to be needed".
2. Restricted system access, such as through the establishment of a Role Based Access Control [6] scheme (RBAC), in which roles are defined and a set of rights associated with each one. When a user is added, a role is attached to it. So users are not given rights directly, but through their roles. A hierarchical scheme of roles may be chosen, in which rights are progressively added as you navigate upwards through roles. For example an Observer may only view records, an Operator may additionally execute commands on the device, an Engineer may execute commands and change settings, except for security related ones, while an Administrator has also access to change security related settings. Figure 4 is a graphic representation of this example.
3. Separation of duty, such as through defining an Administrator role and a Supervisor role, which complement each other. An Administrator may change settings, but only the Supervisor commits them, while the Supervisor is not allowed to make changes. This ensures that someone who has no right to change settings verifies and approves what changes have been made and it offers one more level of security.

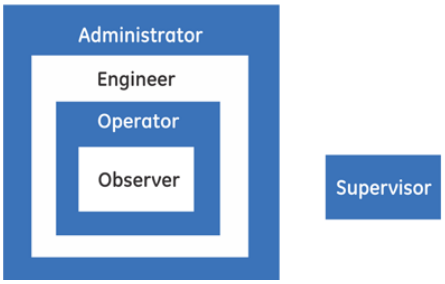

Figure 4 – Example of a Hierarchical Role Scheme with Separation of Duty

4. The principle of least privilege, which recommends that the default access is based on the least privilege possible, such as the one allowed by an Observer role.
5. The "fail secure principle" ensures that failure to perform a required function of the device is dealt with in a secure manner. As an example, a failure to log in because of wrong user id or password or because the maximum number of sessions has been reached must always result into minimum permissions, such as only for viewing. If the failure modes of a software module embedded in the device are not properly dealt with, chances are that the device may be left widely open when the failures occur, and anyone would be able to freely execute commands and change settings.

5.3 A Holistic Approach to Security in P&C Devices

The approach towards strong security in P&C devices should be multidimensional, in the sense that it should address several aspects and be applied both to the devices themselves as well as to the system they are deployed in. This section will elaborate on these aspects.

1. Secure Design and Implementation

The secure design and implementation is based on the principles described above. Specific software development processes must be defined and enforced for this to happen. Also automatic tools, such as Coverity static analysis, may be employed for checking the sanity of the code.

2. Security through Prevention

The security through prevention aspect focuses on implementing controlled user access to the device and data authentication/encryption. A role based access control may be implemented on the device for this purpose with users being allocated roles and very strong password rules. Managing authentication and authorization from a central Authentication, Authorization and Accounting (AAA) system, using standard services such as RADIUS or LDAP is advisable when many devices need to be kept in sync and updated often. In its latest version, version 5, NERC CIP imposed a new requirement for immediate revocation of cyber asset physical and electronic access of an individual upon transfer, retirement or termination. This may be an extremely time consuming activity when no centralized management is available. For situations where concern is expressed in regard to the availability of the central system, a redundant AAA server may be deployed and installed in the substation where the serviced devices reside. Figure 5 is an example of deployment with redundant AAA servers.

Options may be provided on P&C devices to temporarily disable password complexity and even total authentication in situations such as initial commissioning of devices, as long as this options are available only to a higher authority role, such as an Administrator.

3. Security through Detection

The security through detection aspect focuses on detecting and reporting problems and potential cyber-attacks. Detailed device logging of internal activities and failures, combined with the capability of sending this information to a central Security Event Management (SEM) system capable of analysing and reporting on this data are features in this category. Security events, such as logins and outs, user lockouts due to multiple login failures, setting changes, firmware upgrades, server access failures are examples of activities that should be monitored and made available to SEM for analysis and report.

As in the case of centralized authentication, it is advisable to use standard protocols, such as Syslog [7.8] for logging security events. A syslog client would reside on each device, while the Syslog server will be embedded into the SEM system. Also, as with centralized authentication, SEM systems and Syslog servers may also be setup in a redundant scheme, with one server located in the central control centre and the backup in the substation, so that events are still monitored even if the network connection towards the central control centre is lost. The same Figure 5 exemplifies this, by adding redundant SEM servers on the same physical machines where AAA servers reside. Note that the SEM servers and AAA servers are not required to necessary reside on the same computers.

In a redundant scheme such as this, the main servers and the backups are kept in sync through a protocol that runs between the two, to ensure that the backup servers are kept up to date.

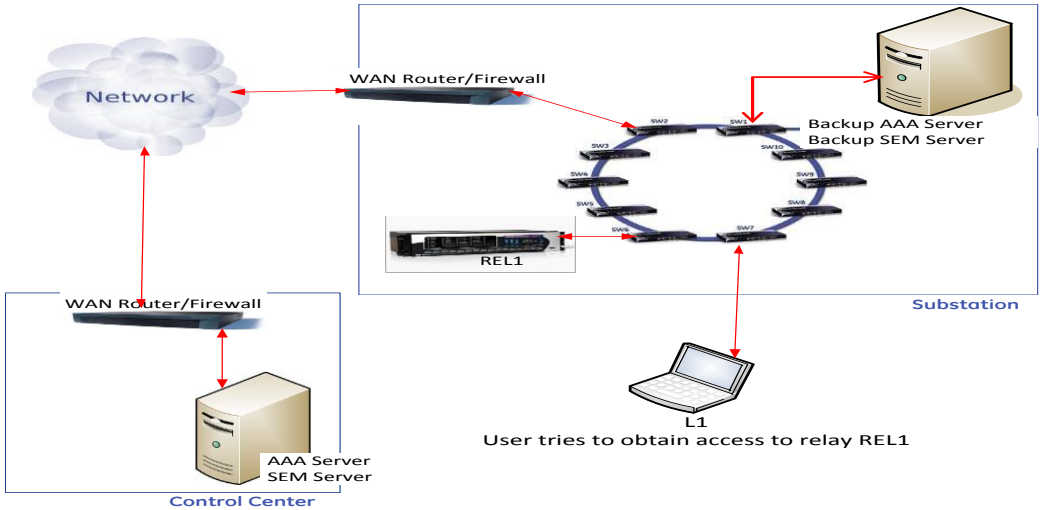

Figure 5 – Example of Deployment with Redundant AAA and SEM Servers

The operation is as follows: An operator of the protective relay REL1 tries to login from the laptop L1. The laptop L1 is used for remote configuration of the relay, using an application that communicates with the relay over the TCP/IP stack. L1 does not need to be physically present within the physical perimeter of the substation, if communications means between the two exist and adequate protection of the communications channel is ensured, through secure tunneling.

The login credentials (user and password) are received on REL1, which forwards them to the AAA server located in the Control Center, using a secure communication channel, such as EAP-TTLS in the case of RADIUS [9]. The server authenticates the user based on the information it retrieves from its database and authorizes it at the level permitted by the role associated with this user. If the main server is not reachable, REL1 will try to connect to the backup server located in the substation. The relay must know the IP addresses of both servers in order for this to happen and configuration parameters on the relay would be provided for this purpose.

For event management, REL1 may implement Syslog client functionality and send defined events to a Syslog server embedded in the SEM server. SEM will need to parse the received events and report, as needed. For better security, the Syslog messages may be tunneled, as defined in [8].

6 CONCLUSSION

Since Ethernet has become a major communication media for P&C devices and the protective relays are not isolated any more, but potentially connecting to centralized management systems over wide area networks, the P&C engineers need to familiarize themselves with a whole range of new concepts and technologies that are part of this new paradigm. While this paradigm provides our industry with a low-cost and very flexible infrastructure, P&C engineers need knowledge in this area if they want to take full advantage of its capabilities. They also need to be aware of the potential dangers that are inherent when connecting over Ethernet networks and be prepared to take actions towards preventing these dangers.

This paper aimed to give solutions to some of the typical problems faced when integrating a protective relay in an Ethernet based network, also highlighting some of the best practices when using the Ethernet.

BIBLIOGRAPHY

[1] MacDonald, J. D. et al., “Electric Power Substations Engineering, 3rd Edition“ (Chapter 7, CRC Press, May 12, 2012).
[2] D. Thanos, “P&C Engineering Concepts Applied to Cyber Security of the Power Grid” (Texas A&M Protective Relaying Conference, 2012)
[3] Cisco Networking Academy, “Routing and Switching Essentials Companion Guide” (2014)
[4] IEC 62439-3, “High availability automation networks, Part 3 - Parallel Redundancy Protocol (PRP) and High-availability Seamless Redundancy (HSR)“ (ed. 2.0, 2012)
[5] NERC-CIP version 5, http://www.nerc.com/pa/Stand/Pages/CIPStandards.aspx (Filing of revised standard in January 2015 and effective in April 2016)
[6] IEC 62351-8, “Data and communications security – Part 8: Role-based access control”
[7] IETF Syslog: http://tools.ietf.org/search/rfc5424
[8] IETF TLS Syslog : http://tools.ietf.org/html/rfc5425
[9] IETF Radius Extensions : http://www.faqs.org/rfcs/rfc2869.html