Zero Trust in DNP3

The DNP Users Group addresses Zero Trust holistically: the basic tenets of a zero trust architecture are part of the design of DNP3 Secure Authentication version 6 and of the Authorization Management Protocol, and the Users Group's conformance certification program addresses important parts of supply chain security.

This white paper provides details about the DNP Users Group's approach.

Executive summary

PRINCIPLE	HOW IT IS ADDRESSED
AUTHENTICATE EVERY SESSION	DNP3-SA authenticates all sessions between Controlling Stations and Outstations
CONSTANT MONITORING	DNP3-SA continually reports statistics on each session to one or more Controlling Stations
EXPLICIT AUTHORIZATION	AMP provides authorization, not just authentication, and DNP3-SAv6 provides an authorization enrollment process
SUPPLY CHAIN SECURITY	Defines a device security upgrade process from warehouse to fully installed and authorized
REMOTE ATTESTATION	Not provided in this version but provides a framework to which it is easy to add attestation
ROLE-BASED ACCESS CONTROL	AMP Authority can assign roles at the device level. Implementing roles in devices is the responsibility of the vendor.
OBJECT-LEVEL ACCESS CONTROL	Still to be added. Can be accomplished through separate (non-DNP) configuration of access control lists.

1 Introduction

1.1 Background

In April 2021, the US Department of Energy's Office of Energy published a Request For Information (RFI) titled "Request for Information (RFI) on Ensuring the Continued Security of the United States Critical Electric Infrastructure" [1]. This RFI, which was issued to support the development of a recommendation pertaining to the replacement of Executive Order, Securing the United States Bulk-Power System [2], recognized the "immediate imperative to secure [the US'] electric infrastructure" and to "[prevent] exploitation and attacks by foreign threats to the U.S. supply chain".

The Distributed Network Protocol (DNP3) is presently used by over 90% of power utilities in North America to monitor and control the field devices in their Operational Technology (OT) networks and implement their Supervisory Control and Data Acquisition (SCADA) systems. It is also widely used in the water and waste water industries, both in North America and around world. It plays a crucial role in automating the US bulk-power system and is an integral part of the US critical electric infrastructure.

The DNP Users Group (DNP-UG) maintains DNP3 and the related specifications, technical bulletins, application notes, test procedures, and certification program. It has a role of stewardship and continuous improvement and innovation in the SCADA industry, particularly around the DNP3 protocol and its applications, and works with the IEEE to develop standards using DNP3, such as IEEE Standard 1815.1, the IEEE Standard for Exchanging Information Between Networks Implementing IEC 61850 and IEEE Std 1815(TM) [Distributed Network Protocol (DNP3)], and the upcoming Standard Profile for Communications with Distributed Energy Resources (DERs) using IEEE Std 1815 [Distributed Network Protocol (DNP3)] (P1815.2).

Within the DNP-UG, the DNP Technical Committee (DNP-TC), develops and maintains the DNP3 specifications and provides technical leadership and expertise around the protocol; the DNP Cybersecurity Task Force (DNP-CSTF) provides expertise around cybersecurity as it pertains to DNP3, and develops and maintains DNP3 Secure Authentication (see below), and the Authorization Management Protocol; the DNP Test Procedures Committee (DNP-TPC) develops and maintains test procedures for DNP3; and the DNP Test Management Committee (DNP-TMC) develops and maintains the certification program for implementations of DNP3.

DNP3 was one of the first SCADA protocols to specify application-level security with the introduction of Secure Authentication (DNP3-SA). DNP3-SA version 2 was part of the first edition of IEEE Standard 1815, in 2010, and DNP3-SA version 5 is part of the current version of that standard, released in 2012.

Since its publication in IEEE-1815 in 2012, the DNP Technical Committee has received feed-back pertaining to the implementation of DNP3-SA version 5, and the limitations of the current specification. With this in mind, in 2017, the DNP-UG stood up a Cybersecurity Task Force (DNP-CSTF) at the annual Face-to-Face meeting to study and develop the next generation of DNP3-SA. In 2018, at the next annual Face-to-Face meeting, the requirements for the next version of DNP-SA were defined.

1.2 Zero Trust

1.2.1 Overview

A "zero trust" model is a model in which there is no implied trust in any particular asset or entity based on its physical or network location, or based on asset ownership[1]. Zero trust security models assume that an attacker is present in the environment. It requires re-authentication for all sessions, and re-authorization for every action taken by the entity being trusted for that action. It focuses on protecting the assets (i.e. devices) rather than the network perimeter. The goal of a zero-trust approach to cybersecurity is to prevent unauthorized access to resources coupled with making the access control enforcement as granular as possible [3].

1.2.2 Structure

An enterprise zero-trust architecture consists of four principal components: a *subject*, which is the entity trying to access a resource; a *resource*, which may be a computing service, data, a device, etc.; a *policy decision point (PDP)*, which decides whether the subject is allowed to access the resource; and a *policy enforcement point (PEP)*, which enforces the PDP's decision [3]. An enterprise architecture based on zero-trust principles requires both the resource and the subject to be able to access the PEP (and may embed the PEP in the resource to be protected), and requires the PDP to be accessible to the PEP (and only to the PEP) at all times.

Typically, SCADA systems and OT networks cannot meet these requirements: in SCADA systems, access to resources on the network is generally sparse and the network is generally organized in a hierarchical structure that does not allow end-devices to communicate directly with enterprise-level services. The approach to cybersecurity has generally been to create security perimeters and protect those perimeters, while access within the perimeter has remained open and access control has been practically non-existent.

As described below, the restrictions that lead to this approach for SCADA networks are addressed by the Authorization Management Protocol (AMP), allowing individual resources to enforce access policies that are centrally managed, and to validate trust before access is allowed on a per-session basis, while retaining the sparse access typical of these networks.

1.2.3 Monitoring

One of the basic design principles of zero-trust is the ability to continuously monitor the system, both for the observable state of the subjects within the system, and for on-going threats [3]. Enterprises are expected to be able to observe all traffic within the network. This same requirement applies generally to SCADA networks and has been the driving reason to eschew encryption in point-to-point communications in SCADA systems. While this remains the case, DNP3-SA version 6, currently being

[1] Or, according to the replacement Executive Order, a Zero-Trust model is "a security model, a set of system design principles, and a coordinated cybersecurity and system management strategy based on an acknowledgement that threats exist both inside and outside traditional network boundaries" [6]

developed by the DNP-CSTF does define an encrypted mode for use when packet inspection is not a requirement.

Both AMP and DNP3-Sav6 provide for monitoring through logging and statistics aligned with the requirements defined in IEC 62351-14 and related standards.

2 How we address zero-trust

2.1 Tenets of zero-trust-based design

One of the central tenets of zero-trust architecture is that access to resources be granted on a per-session basis: trust in the requester is evaluated before access is granted. It further requires all communications to be secured, regardless of network location [3].

The general idea is to assume the attacker is already on the network, and therefore to limit lateral movement within the network as much as possible, requiring authentication and authorization to access any resource on a per-session basis. All the current work of the DNP-CSTF is based on these tenets.

2.2 On-going work in the DNP-CSTF

2.2.1 SAv6: session-based authentication of all application messages

Contrary to previous versions of Secure Authentication, SAv6 is built into the protocol as a separate protocol layer, between the Application Layer and the Transport Layer[2]. Because of its new position in the protocol stack, SAv6 no longer has the concept of "critical" messages that need authentication as opposed to messages that do not: all messages exchanged by the application layers of devices that implement SAv6 are authenticated.

This authentication is done using Session Keys. These keys are symmetric keys that are established during session negotiation. From a cybersecurity perspective, session establishment in SAv6 can be seen as analogous to TLS session re-negotiation with the distinction that SAv6 uses an extra set of keys (the Session Keys) while TLS only uses a single set of symmetric keys (its session keys, which are analogous to the SAv6 Update Keys). Aside from this mechanism, which is intended to be used by devices on a much more regular basis than TLS session resumption would be expected to, SAv6 provides similar guarantees for confidentiality, integrity and availability as TLS does.

The SAv6 Update Keys are established when the Controlling Station-Outstation Association is first established and, while not required to time. out, can be replaced at any time. Association establishment uses a similar key negotiation based on the same cryptographic primitives as used by TLS 1.3: a shared secret is established using Elliptic Curve Diffie-Hellman (ECDH) and HMAC-based Extract-and-Expand Key Derivation Function (HKDF). This requires valid X.509 identity certificates for both the Controlling Station and the Outstation, tying the authentication of every message exchanged between the two applications back to those certificates.

[2] The DNP3 Transport Layer is called the "Transport Function" in IEEE 1815 for historical reasons.

Hence, SAv6 establishes trust prior to allowing access to system resources on a per-session basis, and secures all application-to-application communications between all DNP devices.

2.2.2 Enabling secure monitoring

In a zero-trust architecture, it is important to be able to monitor the state of individual assets, as well as their observable behavior, to determine the security of the system in real time [3].

While SAv6 introduces an encrypted mode to ensure the confidentiality of exchanged application messages in critical infrastructure if needed, the encrypted mode remains optional: the ability to perform deep inspection of control commands and real-time data often trumps the confidentiality requirement in grid-edge systems.

In addition to allowing authenticated cleartext messages to be exchanged, DNP3-SA further defines a number of statistics to monitor the state of the different associations a device is part of, and can report the statistics of downstream devices. Additionally, SNMP MIBs will be defined for AMP, and certain events are recommended in the specification to be reported over syslog.

Hence, AMP and SAv6 provide the tools to monitor the security posture of devices, and do not impede monitoring the operational behavior of devices while retaining authentication.

2.2.3 Using AMP for trust validation

In order to establish trust, the credentials used by both parties in mutual authentication need to be validated when the association is created, and regularly afterwards. Public Key Infrastructures (PKI) are an essential part of cyber security and zero-trust architectures: they provide a root of trust for cyber security, based in the organization's root Certificate Authority (CA). In order to implement this correctly, a device that uses X.509 certificates must be able to authenticate those certificates linking it back to the trusted root CA.

In a system that implements AMP, the Authority acts as an intermediate CA that signs the certificates of all devices in the system, thus establishing itself as the authority for all authentication within the system. The Authority's CA is intended to be an intermediate CA to allow the PKI itself to be rooted with a trusted certificate that does not need to be used on a regular basis to sign device certificates: it is considered good cybersecurity practice to have at least one level of intermediate CA in a PKI for security and operational purposes. The Authority's CA certificate is also distinct from the Authority's Identity Certificate, which is signed by its CA certificate and used to sign its AMP messages. Again, this follows a good cybersecurity practice known as the "one key, one purpose" principle [4].

In order for any device in the system to be able to validate the authenticity and validity of the X.509 certificates being used to authenticate and authorize the Controlling Station-Outstation Associations, AMP implements specific messages to allow devices to request the current status of the certificate, and implements messages for the Authority to revoke certificates. It further uses OCSP stapling to reduce the The Authority will not distribute Certificate Revocation Lists (CRLs) because of the communications overhead for certificate validation, and implements this entails, but with AMP, it does implement an

equivalent of the Online Certificate Status Protocol to avoid having to distribute complete Certificate Revocation Lists.

AMP policies can further be used to manage devices' configuration w.r.t. certificate validity, e.g. by configuring devices to regularly request the status of the certificates used to authenticate a MOA, without interrupting the application-to-application communications as long as the certificates remain valid.

Thus, AMP allows devices to maintain trust in the peer device throughout the association's life-cycle which, in SCADA systems, can be very long, while also retaining efficient use of the system's communications capabilities.

2.2.4 AMP: authorization of Controlling Station-Outstation Associations and least privilege

Access to resources should be granted with the least privilege needed. To implement this, AMP and DNP3 provide the means to implement Role-Based Access Control.

Controlling Station-Outstation Associations are authorized using Authority-signed X.509 Attribute Certificates. These certificates identify both the Controlling Station and the Outstation by name, and may authorize more than one Controlling Station-Outstation Association (i.e. by naming more than one target Outstation in the certificate).

The certificate further encodes a role, according to IEC 62351-8, to assign a role to the Controlling Station on the Outstation, thus allowing utilities to implement Role-Based Access Control (RBAC) on a per-association basis.

Essentially, from a Zero Trust perspective, it establishes the link between the Policy Decision Point (the Authority) and the Policy Enforcement Point (the AMP Controller within the device), and can do so over an un-trusted network.

2.2.5 Enabling better testing of critical infrastructure resources

The 2021 DOE RFI requested information to aid in the development of a long-term strategy to "ensure [utilities'] procurement practices and requirements evolve to match changes in the threat landscape and best protect critical infrastructure". One way to do this that the DOE specifically called out was by "[enabling] better testing of critical grid equipment" [1].

The DNP Users Group has two committees that specifically address testing: the DNP Test Procedures Committee develops test procedures for DNP3 Controlling Stations, DNP3 Outstations, and DNP3 Secure Authentication. The Users Group is also looking into developing test procedures for the DER implementation defined in AN2018-001. Additionally, the AMP and SAv6 specifications are both designed with validation and verification in mind, to prepare for the development of test procedures as soon as the specifications are ready.

The DNP Test Management Committee oversees the conformance certification program and the conformance review process. In order to obtain a certificate of conformance for a DNP3 implementation

from the DNP Users Group, vendors must have their devices tested, must provide a DNP XML Device Profile that accurately describes the device, and must have both the DNP XML Device Profile and the test results reviewed by an accredited independent expert using the conformance review process.

Utilities should require conformance certification from vendors in order to ascertain the quality of the devices they install as part of their critical infrastructure, and the DNP Users Group has one of the most robust conformance certification programs in the industry.

2.3 Future work

One missing ingredient for a complete zero-trust SCADA architecture is application object-level access control and RBAC (which AMP prepares for, SAv6 allows for, but DNP3 does not implement): while AMP provides for the means to assign a role to a Controlling Station, DNP3 can only restrict access to objects in the device's DNP3 interface using point mappings. The DNP Cybersecurity Task Force has discussed implementing access control lists (ACLs) on a per-object basis, but this is currently in the task force's backlog.

Remote Attestation, another missing ingredient, is discussed above and also on the Task Force's backlog.

Zero-trust also calls for a heuristic trust algorithm which requires system monitoring and continuous diagnostic and mitigation. While AMP and SAv6 do not impede such monitoring, the monitoring itself is beyond the scope of a protocol and is therefore beyond the scope of AMP and SAv6.

2.3.1 Using AMP for remote attestation

No asset in the system is inherently trusted: in order to maintain trust, a device should be able to show that its firmware and its configuration have not been compromised. One way to implement this is by using remote attestation.

When using remote attestation, a trusted module within the device – i.e. a TPM – is used to register the boot loader, firmware, applications and configurations used by the device in cooperation with the device's CPU and firmware, using the hashes of the binaries being loaded. An authenticated and duly authorized remote auditor can then request the TPM to prepare an attestation of the running firmware, which includes a hash over the loaded binaries and a nonce provided by the auditor. This attestation is signed using a private key only the TPM has access to. This list is then returned to the remote auditor who can verify its authenticity using the TPM's public key.

The fact that the TPM's public key can be used to verify the authenticity of the attestation shows that the TPM generated it. The fact that the attestation includes the remote auditor's nonce shows that it was created after (and in response to) the request. In order to validate the state of the device and the firmware it is running, the expected hash for the device must also be known by the auditor, and compared with the one received.

AMP is, by design, an extensible protocol. During the DNP Cybersecurity Task Force discussions in 2020 and 2021, remote attestation was discussed as one of the applications for AMP. While it is not currently within the scope of the protocol, we have made sure that it is possible to implement remote attestation

using AMP, provided the device contains a TPM and implements the necessary mechanisms in its firmware. Including remote attestation in the AMP specification is currently in the task force's backlog.

3 References

[1] Office of Energy, Department of Energy, *Request for Information (RFI) on Ensuring the Continued Security of the United States Critical Electric Infrastructure,* 2021.

[2] Office of the President of the United States, *Executive Order 13920: Securing the United States Bulk-Power System,* vol. 85, 2020.

[3] S. Rose, O. Borchert, S. Mitchell and S. Connelly, *NIST Special Publication 800-207 - Zero Trust Architecture,* NIST, 2020.

[4] E. Barker, *NIST Special Publication 800-57 Part 1, Revision 5 - Recommendation for Key Management: Part 1 – General,* 2020.

[5] Office of the President of the United States, *Executive Order 13990: Protecting Public Health and the Environment and Restoring Science To Tackle the Climate Crisis,* vol. 86, 2021.

[6] Office of the President of the United States, *Executive Order 14028: Improving the Nation's Cybersecurity,* vol. 86, 2021.

[7] G. M. Køien, "Zero-Trust Principles for Legacy Components," *Wireless Personal Communications,* 2021.

Revision history

Revision	Author	Comments
0	Ronald Landheer-Cieslak	Initial version
1	Ronald Landheer-Cieslak	Integrated comments from Andrew's review
2	Ronald Landheer-Cieslak	Add some background material as suggested from feed-back
3	Grant Gilchrist and Ronald Landheer-Cieslak	Addition of the Executive Summary table. Replace the term "Master" with "Controlling Station".

DNP Users Group Policy on Intellectual Property Ownership, Document Use and Standards Interpretation in Reference to the IEEE Sharing Agreement (2024 Amendment)

- The DNP Users Group (DNP-UG) has unique and deep expertise with DNP3 and related standards, specifications, notes, guides, XML profile and other documents and material.
- The use of DNP-UG member-only documents is reserved for DNP-UG members who are granted a license-to-use for the active member period. Member-only documents include but are not limited to standards, specifications, profiles, application notes, technical bulletins, test procedures, guides, notices, tutorial materials, workshop presentations and training materials.
-
- The DNP-UG retains all intellectual property (IP) ownership and licenses our work to the IEEE by way of a Sharing Agreement for specific standards and purposes. The DNP-UG's work of interpretation, clarification and guidance is solely the work of the DNP-UG and not the IEEE.
- DNP-UG documents (other than those published by IEEE) solely represent the views of DNP-UG and do not necessarily represent a position of either IEEE, IEEE Standards Association (IEEE SA), the IEEE Power System Communications and Cybersecurity Standards Committee, or the applicable IEEE Working Groups.