EPRI ELECTRIC POWER RESEARCH INSTITUTE

Zero Touch Provisioning

An EPRI White Paper **Project Set 161G - 2018**

Introduction

Document Purpose and Scope

The purpose of this document is to present a framework and conceptual view of zero touch provisioning for utility field area network (FAN) systems. The overarching idea is that devices can be deployed in the field with no depot-level configuration. All relevant data is imported into the utility back-office systems, and devices follow a secure process to initially authenticate and provision on the utility FAN Network.

Why Zero Touch?

Intelligent devices in the FAN have traditionally been manually configured. The number of devices was small, and the provisioning process was simple. The integrated grid requires orders of magnitudes more communicating devices, ranging from distribution automation devices to new types of line sensors. Modern wireless technologies used for the FAN require additional, network-specific details to be configured in edge devices. Cybersecurity also presents a new level of provisioning challenge, to ensure that every device is securely provisioned with unique credentials.

As distribution control systems become larger and more complex, it will not be possible to manually manage configuration and provisioning of systems that scale to hundreds of thousands of devices.

Current Industry State

Challenges with Field Area Network System Vendor Options Today

Provisioning configures device settings and installs credentials that are required to enable the device to function as a part of the utility's unique network. For wireless technologies that are commonly deployed in utility FANs, it is common practice for the vendor to provide a unique, proprietary mechanism for provisioning the device. This is a consequence of the lack of standards for the provisioning process. Provisioning is necessary for the edge devices that initiate a connection to the FAN, as well as for infrastructure (base station) devices that create the FAN for devices to connect to.

The provisioning data contains two types of information—information and parameters that are common across the network and the same for all devices, and information that is unique to each device. The latter type can include security credentials and certificates, network addresses, and utility-assigned serial numbers or asset identifiers (IDs).

For many types of devices, provisioning is initiated through a browser interface that is presented at a default Internet Protocol (IP) address on the device's local network interface. This is a user-friendly method that

makes the required settings visible, but the process does not scale to provisioning large numbers of devices for deployment. Most devices also offer a method to upload configuration files or access a command line interface suitable for scripting, either over the network port with Secure Shell (SSH) or through a serial port. Although scripting can automate the provisioning process to a great extent, the development of the script (or configuration file generator) itself can become a significant software task, with its own debugging and maintenance challenges. Each device still must be handled individually, powered up, and physically connected to the system used for provisioning.

Examples of Industry Practices for Cellular and Other Systems

Utility FANs using commercial cellular networks can take advantage of a different provisioning method. The commercial cellular industry developed the Subscriber Identity Module (SIM) as a means of provisioning devices that connect to the operator's network. It provides credentials and network data needed for joining an operator network. The design limits the ability to store and access arbitrary, user-specific data, though. Although the SIM approach has benefits in separating the configuration from the modem device, it still must be initially written with the correct, unique data, and then physically installed in the device. Although the modern micro- and nano-SIM form factors are quite small, radio modules continue to become smaller. In modular modems designed to be embedded in other devices such as sensors, the SIM can be a significant part of the device size (see Figure 1).

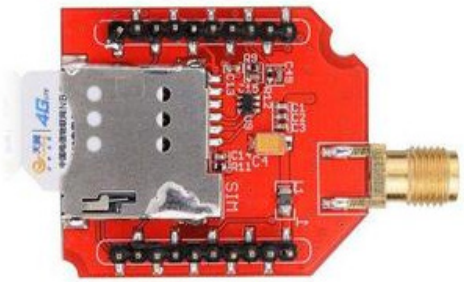

Figure 1 – Example of a narrowband Internet of Things embedded modem module with a physical SIM

The Groupe Spécial Mobile Association (GSMA), which defines SIM standards, has introduced the Embedded Universal Integrated Circuit Card (eUICC) or e-SIM. Although it is not removable, it supports other provisioning-related SIM services and enables over-the-air provisioning.

Utilities that deploy private Long-Term Evolution (LTE) networks obtain the benefits and challenges of SIM management. A comprehensive SIM management solution is required to provision and manage a variety of devices, ranging from conventional data modems to mobile workforce phones and tablets. Commercial operators often have internally developed SIM management systems that are integrated with their customer management and billing systems. Utility LTE networks may have different requirements for SIM and e-SIM management. Solutions may be offered by the infrastructure vendor or third parties such as Gemalto's On-Demand Connectivity.

What Is the Ultimate Goal?

What does *zero touch provisioning* mean? The ultimate goal is to be able to receive and inventory FAN devices from the vendor without any manual intervention before installation. A zero touch provisioning system would allow the installer to select the proper device from inventory, take it to the installation site, unbox it, and power it up for the first time. For devices that have an integrated global positioning system (GPS), all relevant data is in the back-office and a device can simply power up, find the network, and provision itself based on location. For devices without integral GPS location capability, the installer would scan a serial number from the device with a handheld tool that is network connected, matching the device to be provisioned with the location. The device would establish a connection (perhaps locally or through a restricted sub-network on the FAN), be authenticated, and then the appropriate operational provisioning data would be pushed into the device, enabling it for permanent installation. The following section explores more of how this could work.

This section describes a conceptual process for the way that a zero touch provisioning system could work (see Figure 2). Such a system would require support by device manufacturers and possibly commercial network operators (where used for the FAN). Ultimately, these systems would become part of telecom network management and planning systems. This conceptual process is intended to outline the process and become the subject of further evaluation and possible prototyping. Potentially, a standards development effort could be initiated.

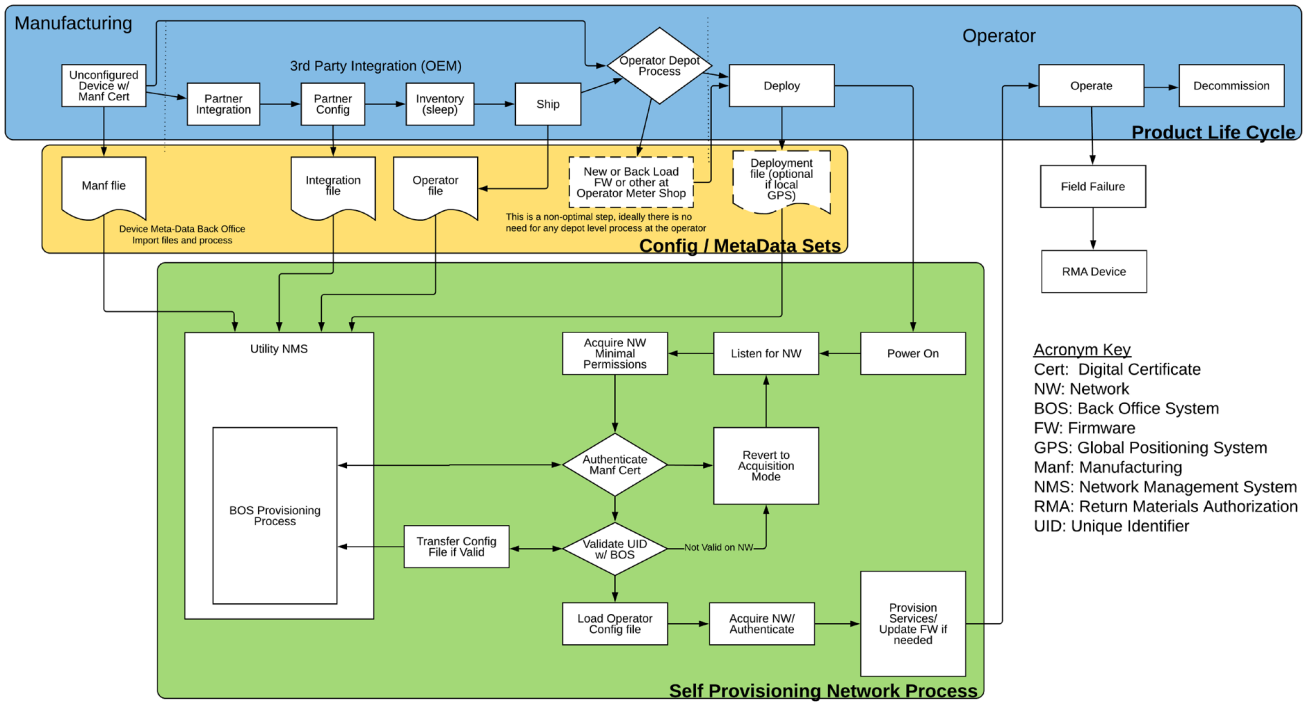

Figure 2 – Zero touch provisioning conceptual process flow

The three colored process groups are described in the following sections.

Product Life Cycle

The product life cycle section of the diagram depicts the flow of a device from manufacture to end of life. At manufacture, a device is provided a manufacturing certificate and a unique identifier that will later be used as a first level of authentication to allow the provisioning process. The device-specific metadata and private keys are all provided to the end user through an out-of-band process. After a device is manufactured, it may go directly to the field for installation (for example, a distribution automation (DA) communications unit) or be integrated into a third-party device for later deployment (for example, a smart meter). At the time of installation, the device is powered on and begins the provisioning process. When completed, it moves into an operational mode. The device may continue to run until it reaches its end of life and is decommissioned. If it fails in the field, it might go through an return merchandise authorization (RMA) process, resulting in either being discarded or repaired and then re-provisioned into the network and operate until end of life.

Configuration and Metadata Sets

As a device progresses through the initial phases of its life cycle, it is configured with a base set of data and an initial manufacturing certificate. If a device is a module that is later integrated with a meter or a particular DA device, additional metadata is generated that includes the relevant details of the device integration and operational parameters for the parent system (such as AMI meter system integration and metadata for meter ID, as well as operational data for the AMI head-end system). It is also possible that the time from manufacture to actual deployment could be relatively long, and either the firmware or other base parameters of the system could have changed. It should not be required but, if desired or more efficient, a depot level upgrade may be performed. Finally, at the time of deployment, the locational information and any other information such as DA device IDs or other intelligent electronic device (IED) data must be sent to the back-office system (BOS). Ideally, the device incorporates a GPS,[1] which greatly simplifies the transfer of locational data and provides a highly accurate local time reference both for logging and system feature performance.[2] DA device data should also be established by the BOS systems through a work order process, alleviating the need for any out-of-band data transfer at the time of deployment. Figure 3 illustrates this process.

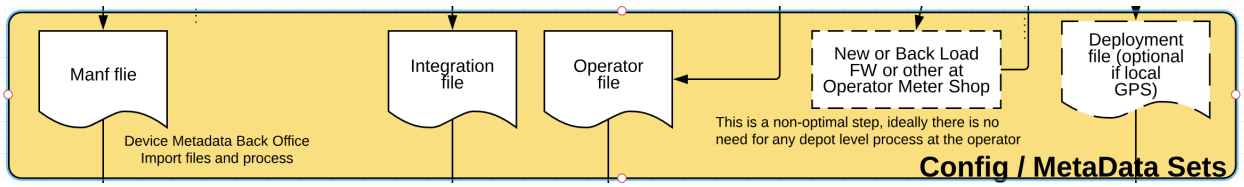

Figure 3 – Configuration and metadata sets

Self-Provisioning Network Process

After a device has been deployed and powered on, it will begin the provisioning process. For this process to complete, all relevant metadata and a security credential must have been imported into the BOS. The general process is depicted in Figure 4. The detailed transactions are much more complex, and different vendors may use a variety of mechanisms and architectures for both acquisition and security exchanges. When a device initially powers up, it has a credential that will allow it to authenticate with the network management system (NMS) for initial provisioning. It will first listen for a network. This can be complicated, as there are scenarios such as two utilities with adjacent territories using the same system, which must be taken into account in the system design. After a network is acquired, the device attempts to authenticate and, if validated by the NMS, it is provisioned with operational configurations, given a new certificate, and possibly migrated to the utility certificate authority. If the device is not validated, it reverts to acquisition mode, continues to look for networks, and periodically attempts to initiate the provisioning process. This can happen if the BOS metadata has not been transferred at the time of deployment or if the network has been only partially deployed.

[1] GPS should be an RFP requirement, and many systems do not incorporate this functionality

[2] A local GPS provides stratum 1 time reference. This greatly improves the accuracy of local time stamping of events and allows the device to more readily target and acquire neighbor devices. It also allows for more advanced quality-of-service (QoS) mechanisms operating in the media access control (MAC) and data link layer, such as identifying and synchronizing to specific time slots for transmissions.

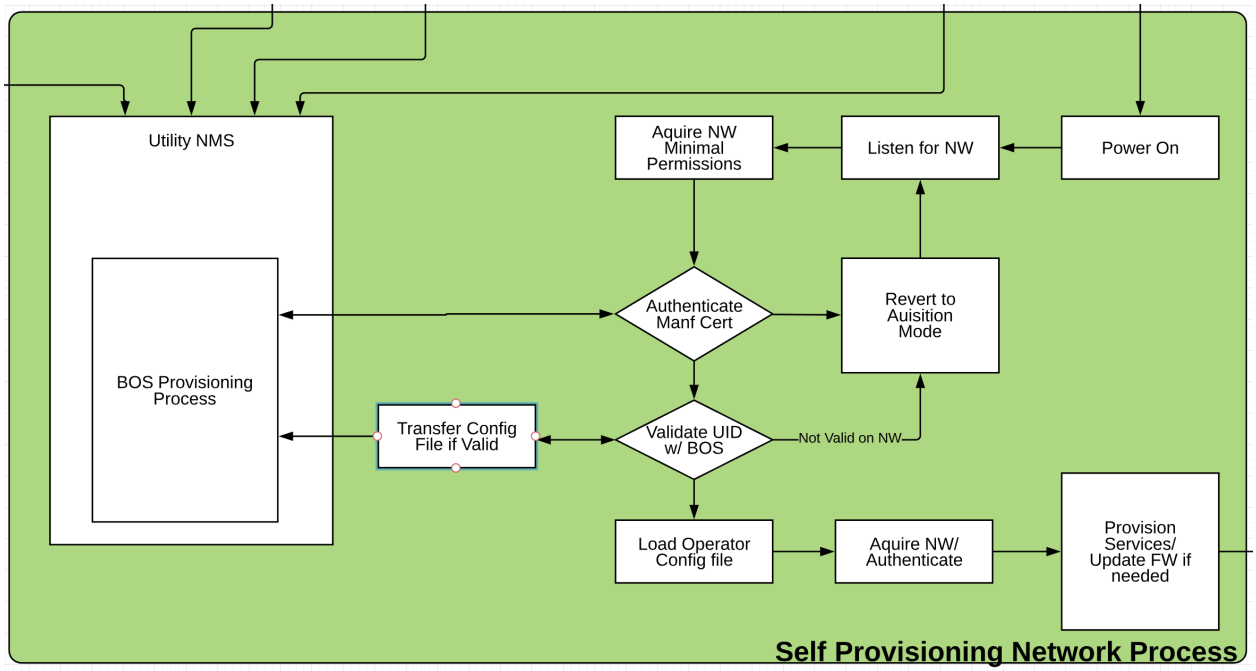

Figure 4 – Self-provisioning network process

Security Design and Issues

The provisioning process is inherently coupled to the security process through the life of the device. The mechanisms used to initially allow devices to access the network must not limit the functions needed to manage the security life cycle. The process must consider all elements of the system, from the NMS to the edge devices, and include aspects such as authentication, enrollment, revocation, and eviction of devices. There is no one best architecture as vendors will support different technologies, and individual utilities will mandate different security standards. There is a set of common practices and scenarios that should be accommodated in the design listed below.

Access needs must be accommodated, including the following:

- NMS access through role-based access control.
- Field engineering access for debugging and testing using a field test tool tied to some form of authenticated access mechanism for a period of time, such as a work order or timed two-factor authentication.
- Apparatus engineering access to IEDs through a tool that can access the IED control through the communication device. Multiple current use cases define a need to be able to locally read, control, and test devices using DNP3 or other control protocols.

Specific scenarios must be addressed when devices lose trust, including the following:

- The NMS is not trusted
- A FAN access device (also known as a *collector*, *edge router*, and the like) is not trusted.
    - Revocation will disconnect only northbound (toward the NMS) interfaces, and device will still be connected to the southbound network (lower network tier, such as an AMI network).

- The endpoint is not trusted
    - The device should be evicted from the network

Other operational use cases must be supported, such as the following:

- Firmware updates
- Access to (and pushing of) distributed applications to embedded-compute environments on edge devices
- RMA processes, and ensuring that all access for RMA devices is removed and such devices must go through a new provisioning process
- Hardware security modules incorporated into cryptographic material management for user access
- Transfer of certificate authority to utility from vendor

A variety of standards and protocols are used in such architectures, including but not limited to the following:

- Certificate revocation list or Online Certificate Status Protocol (OCSP)[3], certificate revocation checks
    - Can severely impact network performance at large device counts, so may not be used for edge devices
- Enrollment over Secure Transport[4] (EST), secure enrollment
- Simple Certificate Enrollment Protocol (SCEP) enrollment

Next Steps

Further research in this area in 2019 will identify practical, existing systems and implementations and provide guidelines for their use with the varying types of FAN technologies. The process described here will be reviewed by utility stakeholders and refined in the ongoing development of this topic. There may be investigation into adapting eUICC and e-SIM approaches into equipment designed for non-cellular (non-3GPP) networks to better support hybrid network architectures. This type of alignment could provide a similar set of capabilities across a broader set of technologies. If warranted, a prototype or reference design for the device and provisioning back-end could be undertaken.

Appendix A

Sample RFP Language for Zero Touch Provisioning

The FAN system **shall** support secure zero touch provisioning of devices. Every device **shall** have a unique identifier and a mechanism to obtain initial permission to be authenticated on the network for automated back-office provisioning. This mechanism **shall** support both newly installed devices and replacement of existing devices. An example of such an exchange would include a device with a unique, secure identity installed at the time of manufacture. The device **shall** be shipped from the manufacturer directly to the utility's inventory. Relevant device metadata for proper configuration and operation **shall** be imported to the utility's back-office systems through a process to be determined at a later date. The desired goal is that devices **should** be installed in the field location and securely run through a process such that it is fully operational within 2 minutes of installation 99.99% of the time or better. To help achieve this goal, all devices **shall** incorporate a visible identifier such as a light-emitting diode (LED) or display that unambiguously indicates to a field installer the successful completion or failure of the installation process. The identifier **should** be either externally visible or easily accessible to the installer.

[3] https://en.wikipedia.org/wiki/Online_Certificate_Status_Protocol
[4] https://tools.ietf.org/html/rfc7030

Together...Shaping the Future of Electricity®

EPRI | 3420 HILLVIEW AVENUE | PALO ALTO, CA 94304 | WWW.EPRI.COM

3002013395

Electric Power Research Institute

3420 Hillview Avenue, Palo Alto, California 94304-1338 • PO Box 10412, Palo Alto, California 94303-0813 USA
800.313.3774 • 650.855.2121 • askepri@epri.com • www.epri.com