Pacific Northwest
NATIONAL LABORATORY

PNNL-29417

Universal Utility Data Exchange (UUDEX) Workflow Design

Cybersecurity of Energy Delivery Systems Research and Development

November 2019

SA Neumann
CM Schmidt
ML Cohen
S Sridhar
SR Mix

U.S. DEPARTMENT OF ENERGY

Prepared for the U.S. Department of Energy
under Contract DE-OE0000190

DISCLAIMER

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@adonis.osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
email: orders@ntis.gov <https://www.ntis.gov/about>
Online ordering: http://www.ntis.gov

Universal Utility Data Exchange (UUDEX) Workflow Design

Cybersecurity of Energy Delivery Systems Research and Development

November 2019

SA Neumann
CM Schmidt
ML Cohen
S Sridhar
SR Mix

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99354

Revision History

Revision	Date	Deliverable (Reason for Change)	Release #
1	11/15/2019	Initial draft	PNNL-29417

Summary

This workflow design document describes the process of establishing a Universal Utility Data Exchange (UUDEX) Connection between two or more UUDEX Endpoints.

The existing processes required to establish a data link using Inter-Control Center Communications Protocol (ICCP) are very time consuming, from both the perspectives of effort and calendar time. The intent of UUDEX is to provide a more streamlined alternative.

The UUDEX Workflow is also used to establish UUDEX Connections to exchange data other than that found in traditional ICCP data exchanges such as exchanges of power system model files, security events and mitigations, disturbance reports, and market data.

Terms, Acronyms, and Abbreviations

The following terms and acronyms are relevant to this specification:

ACL	Access Control List
Alert	A notification that conveys important information, typically where an action should be taken by a user
API	Application Programming Interface
Application	A software component that provides end use functionality for a participant, which may be a source or target of information conveyed through the UUDEX Framework
ASCII	American Standard Code for Information Interchange, as defined by ISO/IEC 646
CIM	Common Information Model, as defined by EPRI, the Utility Communications Architecture Users Group and as used by IEC 61968 and 61970 series of standards
CIP	Critical Infrastructure Protection
Consumer	A subscriber for information, subscribes to a subject and consumes information that has been published
CSV	Comma Separated Values, as defined by IETF RFC 4180
DOE	U. S. Department of Energy
ICCP:	Inter-control Center Communications Protocol, also known as TASE.2 or IEC 60870-6
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IP	Internet Protocol
ISO	International Standardization Organization
JSON	JavaScript Object Notation, as defined by IETF RFC 7159
MQTT	Message Queuing Telemetry Transport
NERC	North American Electric Reliability Corporation
Notification	Any message or signal that is generated asynchronously to report a condition of potential interest to users
OE-417	DOE Electric Emergency Incident and Disturbance Report
PDF	Portable Document Format, as specified in ISO 32000
Publish	To produce information by sending it to a subject
RFC	Request for Comment – Used in reference to published IETF standards, which are called RFCs
Subscribe	To consume information by receiving the information published to a subject

TASE	Telecontrol Application Service Element
TASE.2	Synonym for ICCP
Time series	A sequence of data values captured at different points in time that are telemetered, measured or calculated that represent some aspect of the state of an object.
Topic	A virtual address in a publish/subscribe messaging system. In UUDEX these are used for UUDEX Subjects, but can potentially be used for other purposes such as administrative and monitoring functions.
URL	Universal Resource Locator
UUDEX	Universal Utility Data Exchange
W3C	World Wide Web Consortium
XML	eXtensible Markup Language, as defined by the W3C

UUDEX Roles and Definitions

(These roles and definitions will be migrated to the UUDEX Functional Requirements Design document in a future release.)

UUDEX Administrator	Administrative users that have global responsibility for a UUDEX Framework, and can authorize UUDEX Participants
UUDEX Auditor	Users that have the ability to view information for the purposes of operational and management auditing of the UUDEX Infrastructure
UUDEX Bridge	A gateway between heterogeneous UUDEX Infrastructures (e.g. UUDEX Infrastructures that use different transport technologies)
UUDEX Component	An individual hardware or software element that supports the functioning of the UUDEX Infrastructure
UUDEX Connection	A communication channel between two UUDEX Participants that conforms to all UUDEX requirements (e.g., security, performance). Formerly called a UUDEX Tunnel.
UUDEX Data Element	Any data collection that corresponds to one of the supported types of data conveyed over a given UUDEX Instance.
UUDEX Directory	A single logical entity consisting of multiple individual UUDEX Servers that would establish the UUDEX Framework.
UUDEX Cloud	One or more interconnected, homogeneous UUDEX Infrastructures
UUDEX Endpoint	A producer or consumer of information within UUDEX
UUDEX Framework	Includes the totality of UUDEX, specifically UUDEX Infrastructure, UUDEX Endpoints, UUDEX APIs, UUDEX Participants, UUDEX Protocols, UUDEX Communication Fabric, UUDEX Information Models and UUDEX Workflows
UUDEX Infrastructure	Those UUDEX Components that permit the management and flow of information between UUDEX Endpoints. These UUDEX Components provide a variety of services and are typically replicated for availability purposes.

UUDEX Instance	A collection of connected UUDEX Participants that is closed with regard to the exchange of one or more type of UUDEX Data Elements. In other words, if, for a given collection of UUDEX Participants, certain UUDEX Data Elements that are exchanged between those UUDEX Participants will not be exchanged with UUDEX Participants not in that set, then the set forms a UUDEX Instance. Note that separate UUDEX Instances may have overlapping UUDEX Participants, but UUDEX Data Elements do not cross from one UUDEX Instance to another via those UUDEX Participants. More intuitively, a UUDEX Instance is created when a group of entities deploys the UUDEX Components of a UUDEX Framework for the sharing of certain information within their own community.
UUDEX Monitor	A user with the ability to monitor the current state of the UUDEX Infrastructure, including performance metrics and informational logs.
UUDEX Participant	An organization that is a onboarded member of a UUDEX Instance.
UUDEX Participant Administrator	An administrative user that performs activities related to the publication and consumption of information for a given UUDEX Participant.
UUDEX Producer	A publisher of information, sends information to a subject (also referred to as a “publisher” by some messaging systems)
UUDEX Subject	A subject is implemented using a topic, where a topic is essentially a virtual address used in many publish/subscribe messaging systems. In UUDEX there are specific rules around the creation and use of subjects, resulting in the specific term UUDEX Subject.

Additional terms pertaining to UUDEX roles and definitions are defined in Section 4 of the UUDEX Functional Design Requirements document.

Contents

Figures

1.0 Introduction

The purpose of this document is to describe the workflow required to quickly implement a new Universal Utility Data Exchange (UUDEX) Connection or modify an existing UUDEX Connection to enable the exchange of data that may be used for applications supportive of grid operations, market operations, distribution management, incident reporting, security notifications, and a wide variety of other utility-centric purposes. The focus is to provide perspective on the differences of the UUDEX workflow over the current workflow used in industry for the Inter-Control Center Communications Protocol (ICCP)[1]. The document also covers processes for onboarding new participants, data organization, resiliency, monitoring, and user interfaces.

This document focuses on a cloud-based UUDEX deployment. Figure 1 provides an overview of the UUDEX Framework. Cloud here refers to the communications infrastructure used to connect UUDEX Endpoints whether a shared multi-point communications infrastructure or a legacy point-to-point communications infrastructure.

UUDEX Endpoints could also be hosted in a cloud environment using a variety of management approaches. Specific details of these cloud-based managmenet approaches is beyond the scope of this document.

[1] ICCP is also known as Telecontrol Application Service Element 2 (TASE.2), or IEC 60870-6, “Telecontrol equipment and systems - Part 6: Telecontrol protocols compatible with ISO standards and ITU-T recommendations”.

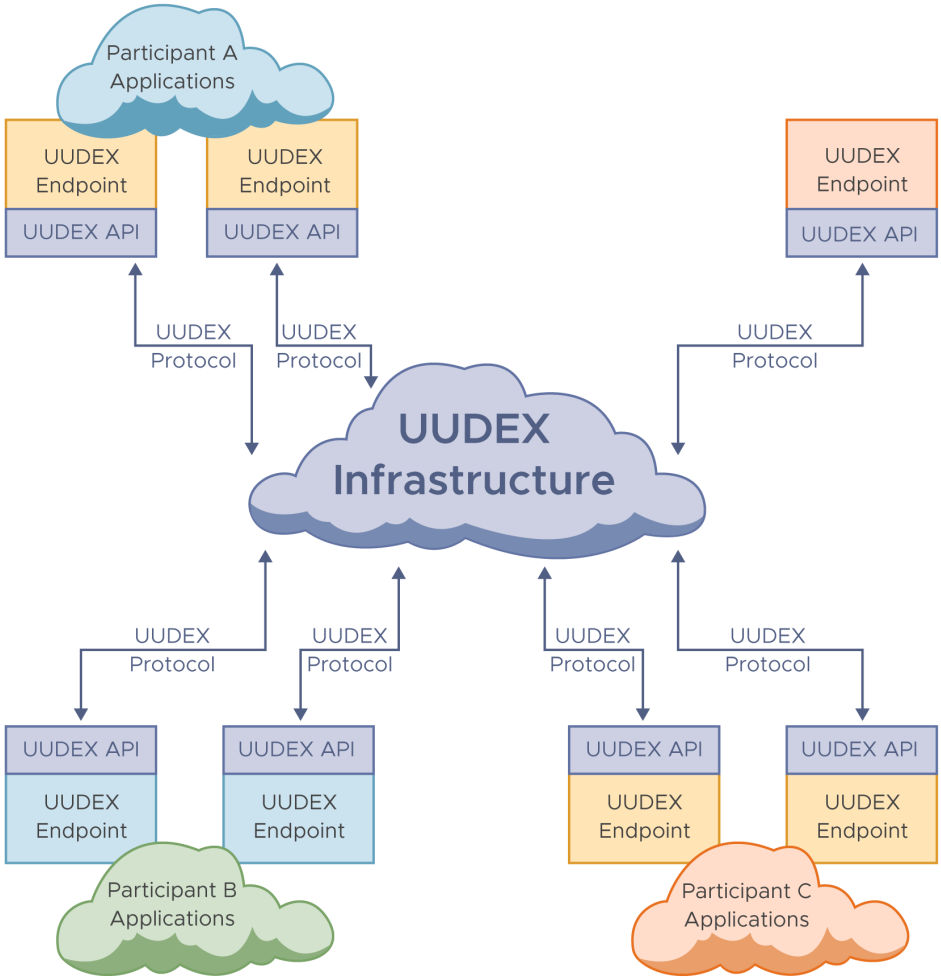

Figure 1. UUDEX Framework

The workflows needed to establish information exchanges within the UUDEX Framework can be summarized as:

- Deployment of the UUDEX Infrastructure, which involves servers, software, and a communications fabric.
- Registration of UUDEX Participants, representing companies (e.g., utilities, generator operators, independent system operators, and market participants) and government organizations (e.g., federal agencies and oversight authorities) that will allow information exchanges.
- Identification of the information a UUDEX Participant will publish for consumption by other authorized UUDEX Participants. These subjects are called 'UUDEX Subjects.'
- Registration of UUDEX Participant endpoints, where each endpoint can be assigned the rights to consume or publish specific information subjects (this includes creation and termination of subscriptions). Each of these is called a UUDEX Endpoint and may be physically deployed at a site chosen by the UUDEX Participant.
- Support real-time information exchanges, where endpoints may consume or publish specific subjects of information, and then transfer the information to or from specific applications that belong to a UUDEX Participant.

- Support the management (including revocation) of information access rights for UUDEX Participants.
- Support the management (including revocation) of UUDEX Endpoint access to the UUDEX Infrastructure.

2.0 Principles

While UUDEX is designed to be transport neutral, some aspects of this workflow may be impacted by the specific transport chosen for a specific UUDEX Implementation. It is the intent of UUDEX to leverage existing technologies where possible, avoiding a 'from-the-ground-up' implementation.

The following are key requirements that are effectively design principles for the workflow:

- Support key information exchanges currently supported by ICCP
- Support additional information exchanges not currently supported by ICCP
- Allow a variety of different types of documents to be exchanged
- Provide secure, reliable, performant transmission of information
- Provide rapid deployment of new UUDEX Participants and UUDEX Endpoints.

3.0 Scope

The scope of this document are workflows related to establishment of UUDEX Connections and information exchanges using the UUDEX Framework.

Areas that are outside the scope of this specification include:

- The means by which organizations contractually agree to exchange information and associated controlled usage.
- Workflows related to software development.
- The details of the administrative user interfaces, beyond the description of their general usage within the UUDEX workflows.
- The details of the processes by which potential UUDEX Participants and UUDEX Administrators are vetted for inclusion in a specific UUDEX Instance may vary between UUDEX Instances and normatively dictating the nature of such vetting policies. However, the sensitivity of the data exchanged over UUDEX may necessitate strong vetting practices in certain instances, often in accordance to industry standards and regulations. Appendix A provides guidance for establishing vetting processes for organizations and individuals that participate in UUDEX Instances.

4.0 Processes

4.1 Configuration

The process of configuration involves establishing a UUDEX Framework. Details of this process are dependent on the transport technology used to implement a specific UUDEX Instance.

Note that different UUDEX Instances could be implemented using a variety of underlying transport technologies. These transport technologies may change and evolve over time. While a specific instance of UUDEX would use a specific transport technology, different UUDEX instances can interoperate with each other if they use a common transport technology or if they are linked using a UUDEX Bridge.

One important aspect of UUDEX configuration is the selection of a network infrastructure. Organizations implementing a UUDEX Infrastructure should select a network infrastructure based on their specific needs and risk profile. For many uses of UUDEX, the internet may be generically leveraged. However, for other uses it may be desirable to leverage a software defined network in order to achieve higher levels of security combined with improved quality of service. The basic principle is to avoid requiring the complexities and costs related to private network infrastructures, such as Multiprotocol Label Switching networks. However, if private network infrastructures already exist, UUDEX can use them to transport data.

The next aspect of configuration would be the deployment of the UUDEX Directory. The UUDEX Directory is a single logical entity consisting of multiple individual UUDEX Servers. These servers could be deployed in a variety of ways:

- Within a cloud infrastructure hosted and managed by a third party, with one or (preferably) more physical locations
- Within an enterprise network, with one or (preferably) more physical locations
- Using a federated, heterogeneous infrastructure at multiple locations, where each server could be managed by a different entity.

Given the anticipated uses of UUDEX, the servers that comprise the UUDEX Directory should optimally be deployed at a minimum of three physical locations, ideally each in separate geographic regions in a way that ensures that the loss of one or two locations does not result in data loss and minimizes the degree of service degradation. This allows for survival of the UUDEX Directory service in the event of network segmentation events, node failures, site failures, and some types of denial-of-service attacks.

A single organization might be a UUDEX Participant in more than one UUDEX Instance. In such a case, it could provide a ‘data transfer’ function between these UUDEX Instances, consuming data from one UUDEX Instance and then publishing it in another UUDEX Instance. These data transfers could occur in either direction. As data are published in the target UUDEX Instance, the organization performing the data transfer would be responsible for establishing the relevant UUDEX Subject and assigning access controls to that UUDEX Subject. Such behavior would need to be contractually constrained, stipulating which UUDEX Participants would be allowed to transfer data, what data could be transferred, and what access controls would need to be applied in the destination UUDEX Instance.

4.2 Initialization

The process of initialization involves the set of activities required to enable the onboarding of UUDEX Participants. This primarily involves seeding the UUDEX Directory with the information required to enable administrative activities. It also defines the standard set of data element types that may be conveyed using UUDEX.

The key data structures and their relationships in the UUDEX Directory are illustrated by the diagram of Figure 2.

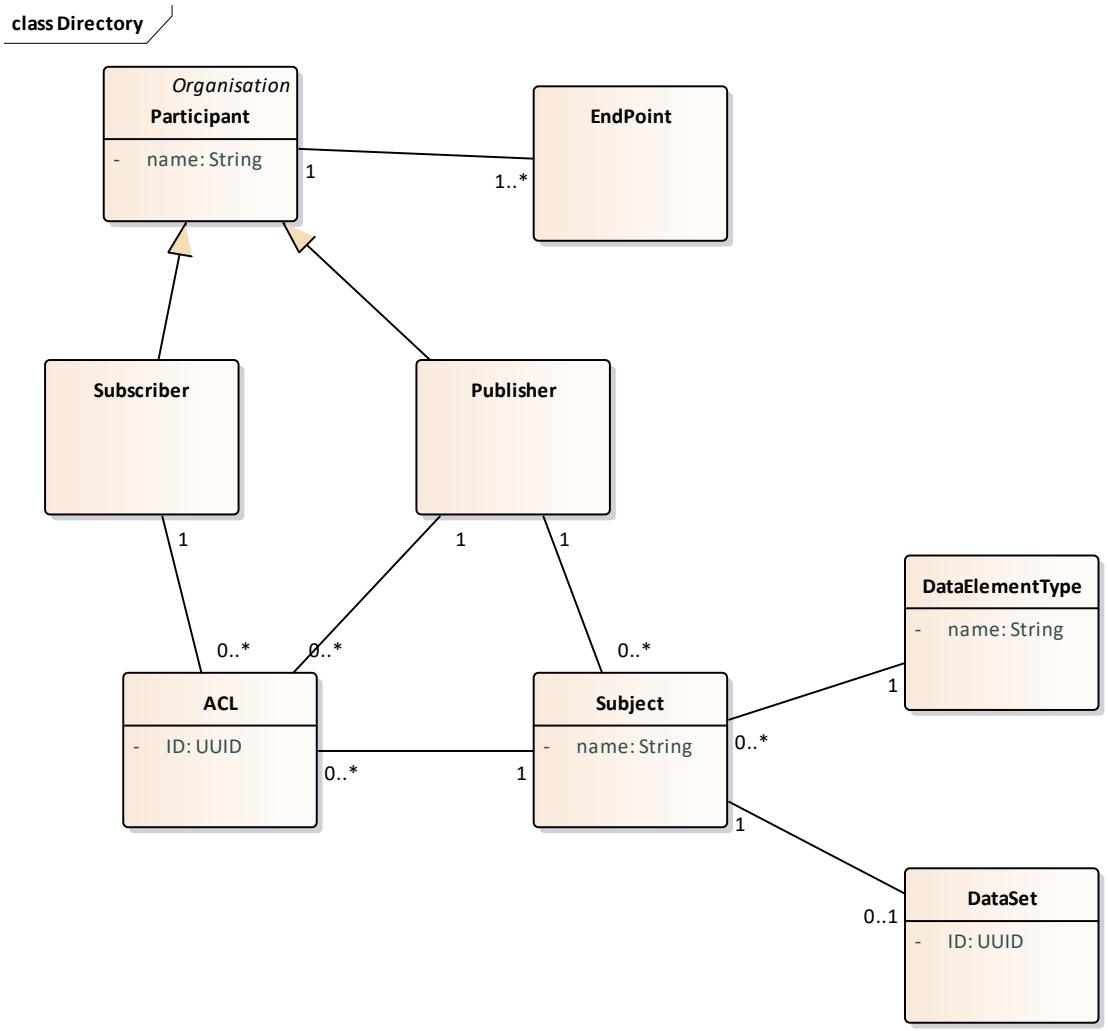

Figure 2. UUDEX Directory

UUDEX will define a 'standard' set of DataElementTypes, where there may be an expanded set of DataElementTypes that may be reflective of additional needs. Each DataElementType has a defined data format and optionally a defined schema. A key DataElementType is a DataSet, which uses a JavaScript Object Notation (JSON) format to convey measured and calculated values. Other DataElementTypes defined may include ‘documents, such as for OE-417 forms. The set of DataElementTypes can be extended as a consequence of standardization efforts or regulatory requirements. Defining new DataElementTypes is an administrative activity.

4.3 Security configuration

The process of security configuration involves establishing access control lists (ACL) that would be applied to UUDEX Subjects to constrain their access by UUDEX Participants and UUDEX Endpoints. The administrative users associated with given UUDEX Subjects are able to change these controls at any time, including rescinding access to individuals who previously had access. Such an action would not impact any data those individuals had gathered already but would apply to all attempts to discover or collect data after the ACL change was applied.

As shown in Figure 3, there are two exchanges of interest regarding ACLs. The first involves the direct setting of an ACL attached to a UUDEX Subject. In this case, the authorized UUDEX Participant Administrator sends a message identifying the UUDEX Subject, the action to take (add or remove the indicated permissions), and a set of permissions expressed as combinations of identities and access rights. The smallest unit of identity that can be specified would be an individual UUDEX Participant, although groups of UUDEX Participants might also be specified. The latter will depend on technical choices of access control paradigms. If the UUDEX Directory authenticates the UUDEX Participant Administrator and recognizes them as authorized to alter the ACLs of the UUDEX Subject, the action is executed. The updated ACL governs all actions UUDEX Participants have with the UUDEX Subject from then on. This includes fulfillment of subscriptions.

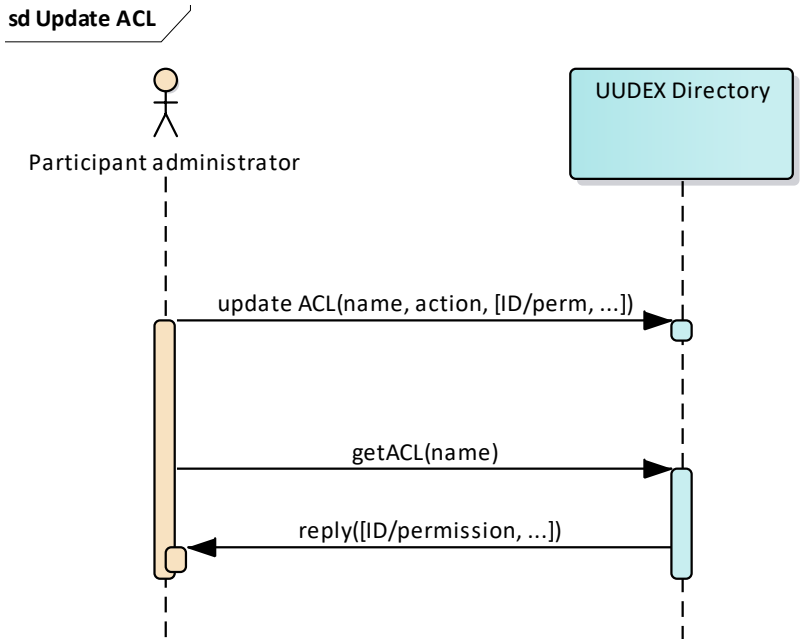

Figure 3. Update of ACLs

The second exchange involves the UUDEX Participant Administrator collecting the complete ACL associated with the named UUDEX Subject. In this case, if the UUDEX Participant Administrator is authenticated and authorized, the UUDEX Directory returns the complete ACL associated with the target of the query. This exchange allows the UUDEX Participant Administrator to verify that their understanding of a UUDEX Subject's ACL meets their expectations. If it does not, they can use the first exchange to make appropriate changes.

Some aspects of this are dependent upon the transport technology used for the UUDEX implementation.

4.4 UUDEX Participant Onboarding

Within UUDEX, the onboarding of a UUDEX Participant involves the processes listed below and shown in Figure 4.

- Out-of-band vetting of would-be participants. Appendix A provides guidance for vetting organizations and individuals.
- Registering an organization as a UUDEX Participant.
- Identifying and vetting designated UUDEX Administrators that can configure UUDEX Endpoints, UUDEX Subjects, or data element types on behalf of a UUDEX Participant. The guidance in Appendix A may be applicable to this vetting process.

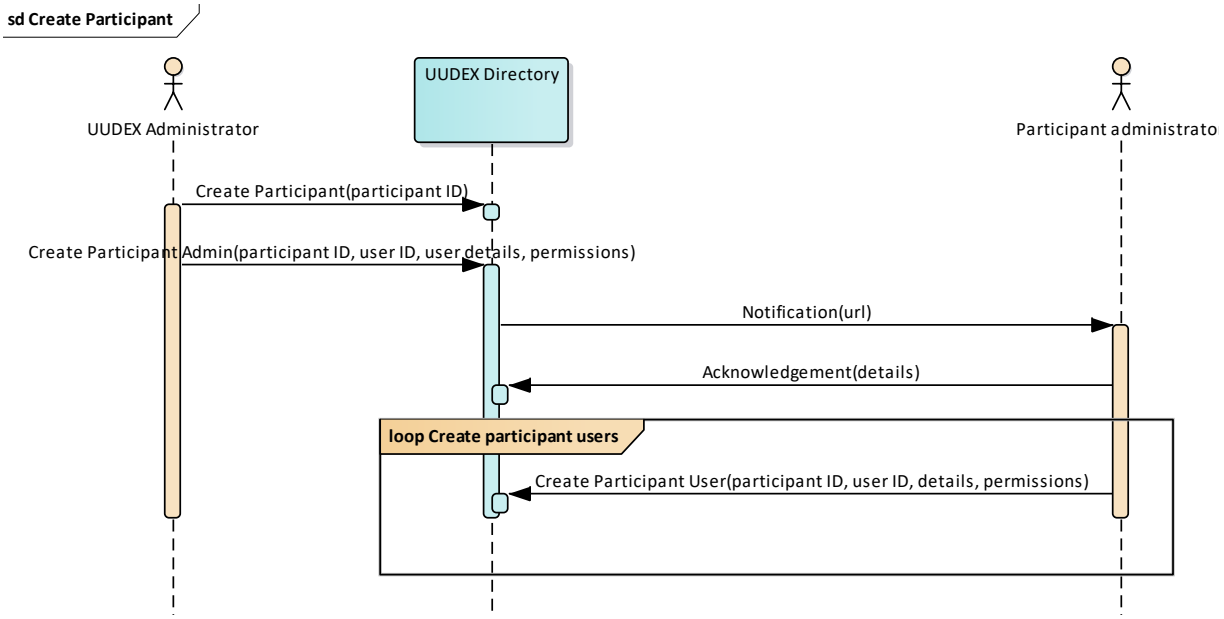

Figure 4. UUDEX Participant Onboarding

The initial steps of onboarding need to take place out-of-band from a UUDEX information flow, potentially over email or other mechanisms. Guidance provided in Appendix A might apply to this vetting process. This is because, until a UUDEX Participant is onboarded, they will not have any credentials that allow them to engage in authenticated exchanges over UUDEX.

Organizations implementing a UUDEX Instance will need to develop procedures for handling requests to become a UUDEX Participant in that UUDEX Instance. In the process of onboarding a UUDEX Participant, the UUDEX Participant (organization) is identified. UUDEX Participants will need to determine what users are permitted to administer their interactions with a given UUDEX Instance. These administrative users are responsible for all configuration of UUDEX Endpoints and related information flows within the UUDEX Framework. The notification for creating a UUDEX Participant could be in the form of an email. After receiving a notification, the designated primary technical contact (i.e., the UUDEX Participant Administrator) for the UUDEX Participant would then provide additional details and register other users for access to the UUDEX Directory.

When a UUDEX Participant is registered, it is provided with credentials that will authenticate it to the UUDEX Directory. The UUDEX Participant will need to register specific UUDEX Endpoints

with the UUDEX Identity Authority employed by the UUDEX Directory before it will be able to engage in regular UUDEX activities, such as publishing or subscribing to UUDEX Subjects.

The UUDEX Identity Authority is implemented using a secure, replicated database that is required for the operation of the UUDEX Framework. Only a UUDEX Administrator can make changes to the UUDEX Identity Authority.

4.5 UUDEX Subjects and DataSets

Prior to enabling UUDEX Endpoints for a UUDEX Participant, the UUDEX Subjects that a UUDEX Producer publishes must be defined. Each UUDEX Subject is associated with a specific DataElementType. In the case of UUDEX Subjects used for DataSets, an initial definition of the data points within each DataSet must be defined.

This is shown in Figure 5.

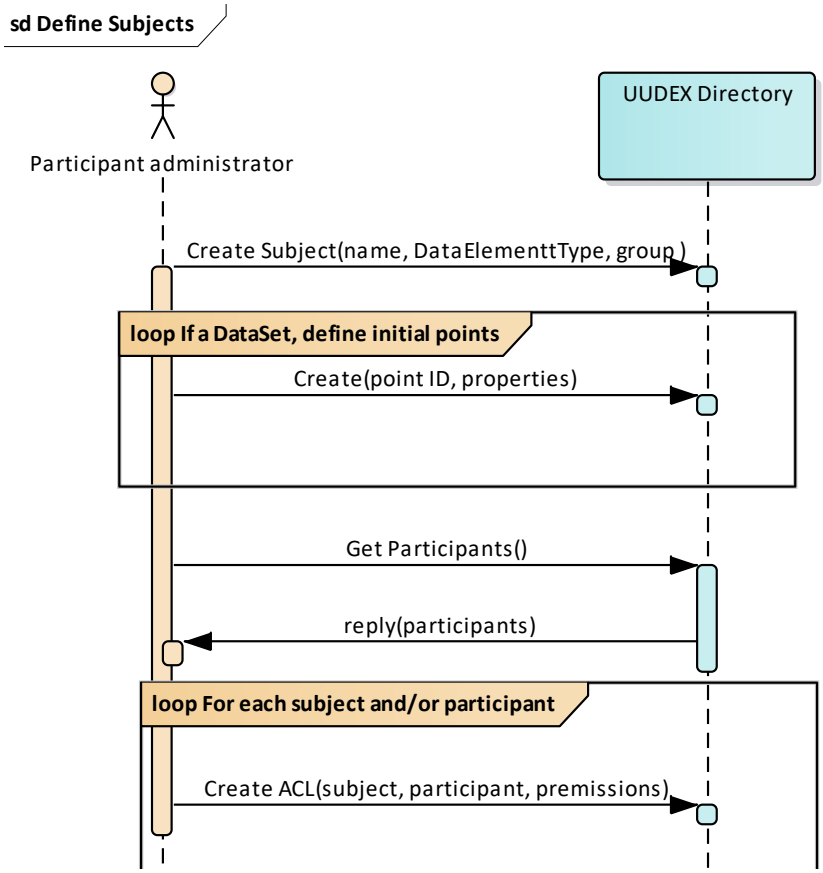

Figure 5. Defining UUDEX Subjects

The user would be provided a user interface that interacts with the UUDEX Directory. For defining data points within a DataSet, an import format for loading the data points could be provided. This would be especially useful when migrating legacy ICCP implementations.

When defining a UUDEX Subject, the following can be specified:

- Persistence, with a specified maximum lifetime

- Access controls, where subscriptions are allowed to specific UUDEX Subjects either as public or by specific UUDEX Participants.

4.6 Endpoint Definition

UUDEX Endpoints use the UUDEX application programming interface (API) to publish information (as UUDEX Producers) and receive information (as UUDEX Consumers, based on subscriptions). UUDEX Endpoints connect to the UUDEX Infrastructure, not directly to other endpoints (with the exceptions of peers as discussed in Section 4.6.1). Endpoints may be the result of:

- Third-party products that directly integrate with other components in a vendor's product suite
- A custom integration that uses any one of a wide variety of integration technologies.

Defining UUDEX Endpoints involves:

- Designating the UUDEX Endpoint, where the UUDEX Endpoint can be a UUDEX Producer or a UUDEX Consumer (or both) of information conveyed using UUDEX Subjects
- Defining UUDEX Subjects that will be used by producers
- Identifying UUDEX Subjects of interest to consumers (i.e., discovery)
- Physically deploying and configuring the endpoint, which includes UUDEX credentials and UUDEX Connection information
- Connecting the UUDEX Endpoint to the UUDEX Infrastructure.

The process for defining UUDEX Endpoints is shown in Figure 6.

Passing of credentials to a UUDEX Participant for use in configuring a UUDEX Endpoint would involve the user securely interacting as needed with the UUDEX Directory. This could involve notifications as described in Section 5.3.

UUDEX Endpoints are validated for connections to the UUDEX Infrastructure by a valid UUDEX Participant. The result of the validation is a set of credentials that is used by the UUDEX Endpoint when connecting to the UUDEX Infrastructure. Each UUDEX Endpoint may have peer instances, which are accessible at specific network addresses. In the case of peer instances, it is the responsibility of those peers to coordinate with each other for aspects of operation such as:

- Detection of startup or failure of peers
- Leader election
- Load balancing.

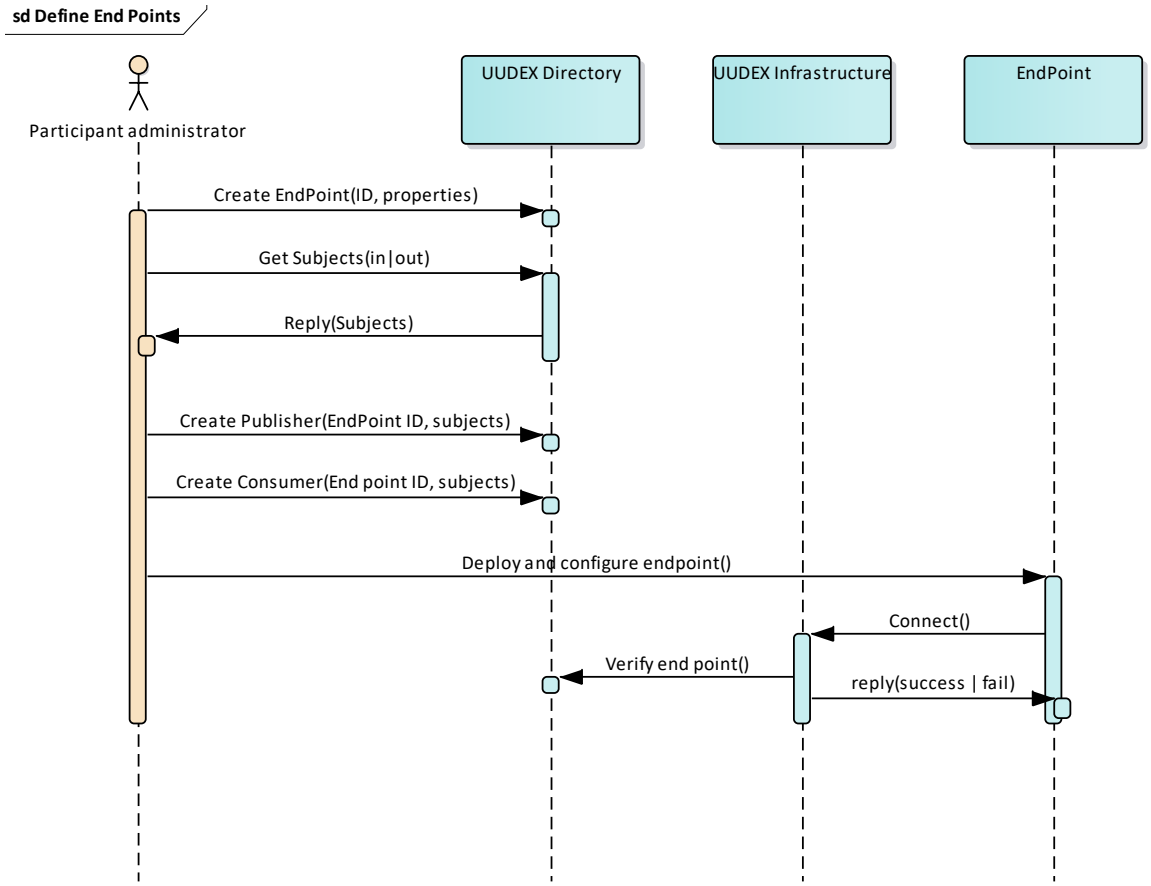

Figure 6. Defining Endpoints

The user would be provided a user interface that supports interactions with the UUDEX Directory. This would allow the user to:

- Locate UUDEX Subjects 'owned' by the UUDEX Participant, and assign them to a UUDEX Endpoint
- Identify UUDEX Subjects owned by other UUDEX Participants, where ACLs permit them to be consumed by UUDEX Endpoints belonging to the UUDEX Participant.

4.6.1 UUDEX Producers

A UUDEX Producer is a UUDEX Endpoint that can publish data elements to defined UUDEX Subjects. *A UUDEX Producer publishes data elements to UUDEX Subjects.* A UUDEX Producer is defined with the following properties:

- A set of one or more network addresses, allowing for 'replicas' of each producer as needed for availability purposes.
- Designation of the endpoint as a valid UUDEX Producer for a defined UUDEX Subject.
- If the subject is used to convey DataSets, the set of data points must be defined.
- Control of which UUDEX Participants can access which types of UUDEX Subjects through use of ACLs.

The basic processing of a simple UUDEX Producer is shown in Figure 7. This example has no redundancy. A more typical example of processing in which a UUDEX Producer has peers is shown in Figure 8, where one instance will be active and responsible for publishing messages to one or more subjects. In the event of a failure, a peer instance will take over (usually through a consensus mechanism) and be responsible for publishing messages to reflect updates to data elements from a data source.

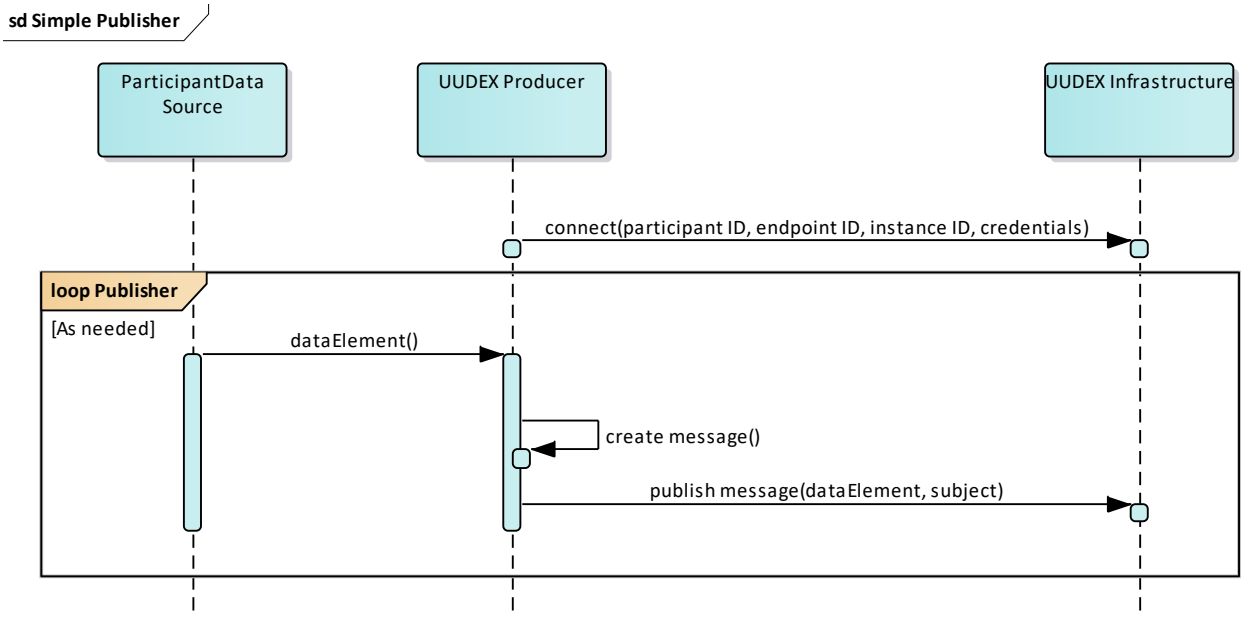

Figure 7. Simple UUDEX Producer

Where a UUDEX Producer endpoint is implemented using a set of peers for availability purposes, it is the responsibility of the endpoint implementation to properly address coordination issues so information that must be published is published only once and in the proper order.

This process involves no user intervention.

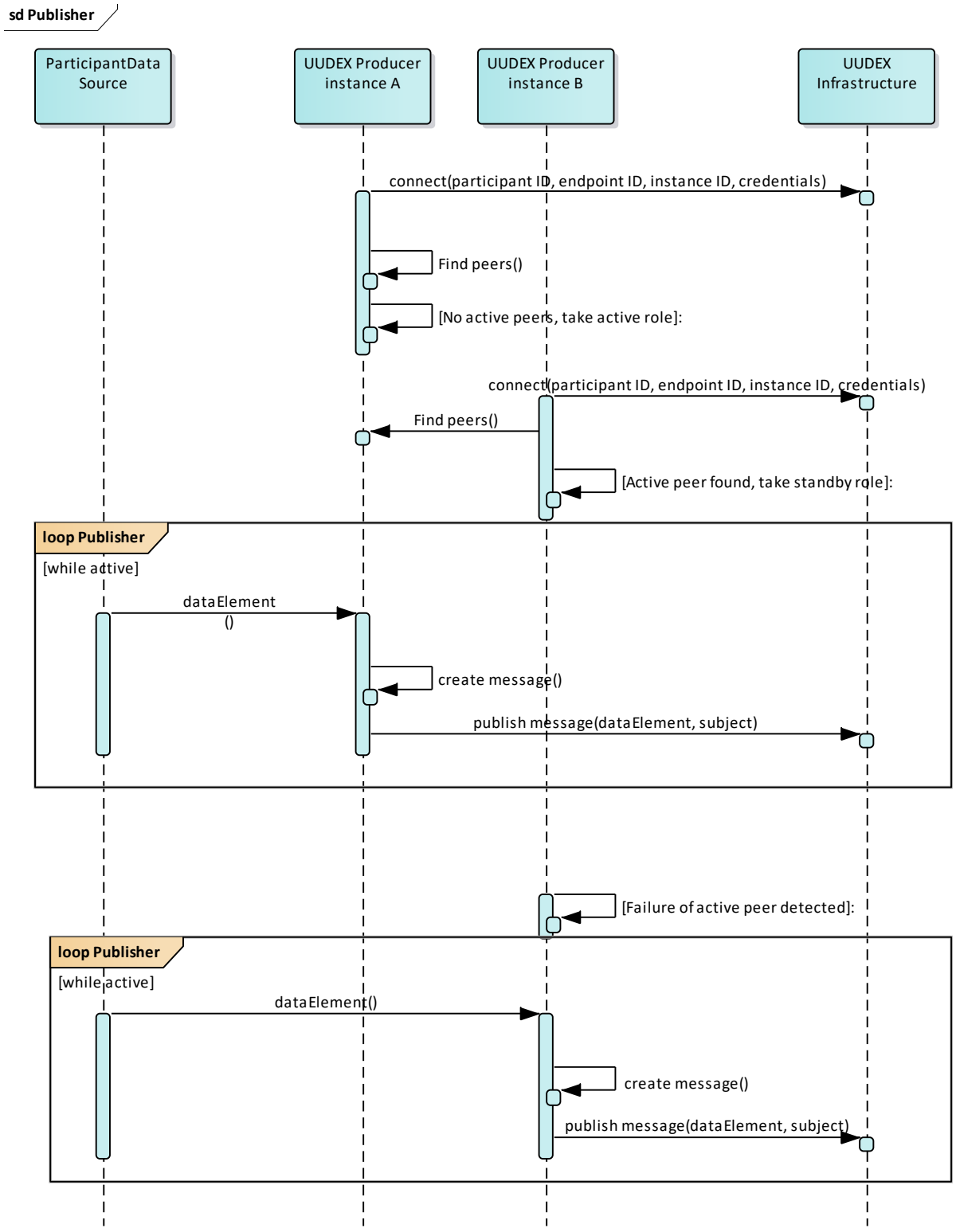

Figure 8. Publication by Peer Producers

4.6.2 UUDEX Consumers

A UUDEX Consumer is a UUDEX Endpoint that can subscribe to and consume information that was published to a UUDEX Subject. *A UUDEX Consumer subscribes to UUDEX Subjects.* A UUDEX Consumer is a UUDEX Endpoint with the following defined properties:

- A set of one or more network addresses, allowing for 'replicas' of each UUDEX Consumer as needed for availability purposes.
- Designation of the UUDEX Endpoint as a valid consumer for a defined UUDEX Subject, where the UUDEX Subjects are selected as a result of the discovery process.
- If the UUDEX Subject is used to convey DataSets, the currently defined set of data points within the DataSet can be obtained as a part of discovery.

Figure 9 shows the basic sequence for a UUDEX Consumer to subscribe to a UUDEX Subject and consume messages that were published to that subject.

Implementation of a UUDEX Consumer allows several options:

- A standalone UUDEX Consumer, without peer instances
- A UUDEX Consumer with peers that all subscribe simultaneously, but potentially coordinates their actions
- A UUDEX Consumer with peers that coordinate their subscriptions, where only one peer consumes at a time.

In all cases, a UUDEX Consumer is responsible for forwarding the information to upstream applications within its local enterprise as needed.

Where a UUDEX Consumer endpoint is implemented using a set of peers for availability purposes, it is the responsibility of the endpoint implementation to properly address coordination issues, so that information that must be retrieved is made available to all UUDEX Consumer endpoints.

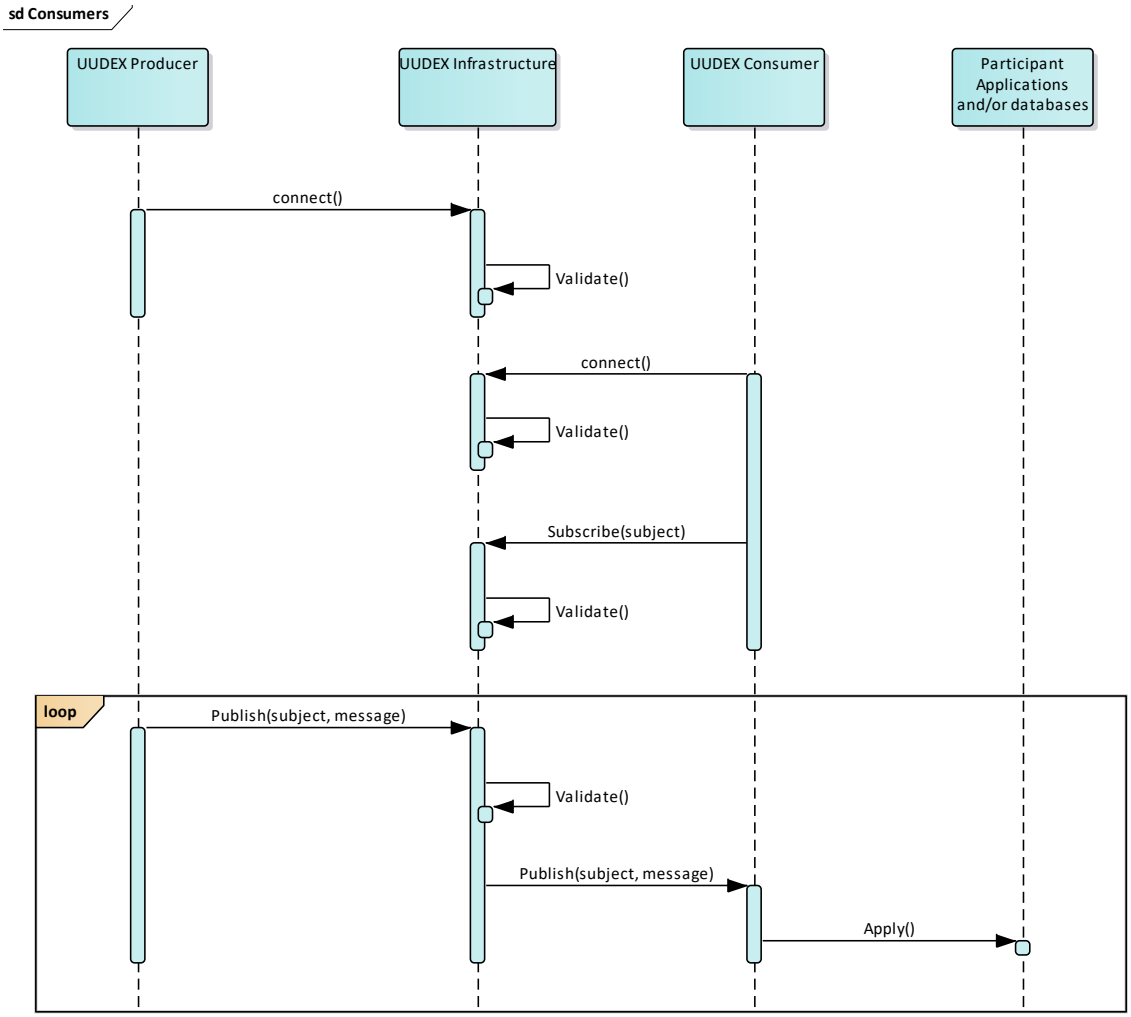

Figure 9. UUDEX Consumers

4.6.3 UUDEX Infrastructure

A prerequisite for a UUDEX Participant to deploy UUDEX Endpoints is to establish connectivity through the chosen network infrastructure to be used by the specific UUDEX Framework deployment.

UUDEX Participants also are responsible for establishing a secure hardware and network infrastructure for the deployment of endpoint software.

4.7 Publication

Publication involves a UUDEX Producer and one or more UUDEX Consumer endpoints. The UUDEX Producer will send a data element of a given data element type to a defined UUDEX Subject. Where a UUDEX Producer may have a set of peer instances, only one instance of the UUDEX Producer should publish a data element instance; that is, information published to a given UUDEX Subject is unique and is not sent redundantly.

The process flow for publication is shown in both Figure 6 and Figure 8.

4.7.1 Models

Models are descriptions of an electricity network that are consumed by applications and simulate, analyze, and or assist the operation of the real-world electricity network. Industry standard data formats are used to convey these models. These are typically defined in non-JSON formats. They will be compressed and encoded for conveyance using the UUDEX JSON message envelope. The payload metadata will describe the specific format.

4.7.2 Data Sets

A data set is a type of UUDEX data element that is used to convey time series data. The definition of a data set involves defining a set of data points and identifying the nature of the point and typically its relationship to objects defined in a model. Data points may represent measured, calculated, or entered values. Publication of a data set involves sending the most recent values for each data point, which includes:

- The measured, calculated, or entered value
- A timestamp to indicate when the value was last obtained
- A quality code.

A DataSet is essentially a container for a snapshot of values for a given period of time. It is expected that DataSets will make up the majority of all messages. DataSets are defined using JSON, with two variations:

1. The detailed definition of data points, describing their data types, nature of the data. and relationships to other objects (e.g. model objects that might be defined by flows in Section 4.7.1. These are normally conveyed to establish mappings of data points between UUDEX Participant models.
2. The current value for each data point with timestamp and quality code. These typically will be conveyed in real time to reflect the current state or may be sent periodically or on an exception basis.

4.7.3 Documents

The definition of documents involves defining a data element type with the following properties:

- A data element type name
- A data format
- A schema specification (optional, but common to structured documents).

Documents can be classified as structured (where specific data elements can be readily extracted (e.g., JSON, XML or CSV files) or unstructured (e.g., image, Microsoft® Word document, text file, or binary file).

Documents can often be conveyed through UUDEX with source formatting being non-JSON, in which case they will be compressed and encoded for conveyance using the UUDEX JSON message envelope. The payload metadata will describe the specific format so the object can be readily reconverted to its original form by the receiver. This flow is shown in Figure 10. Upon receipt, the documents will be decoded and uncompressed to return them to their original format.

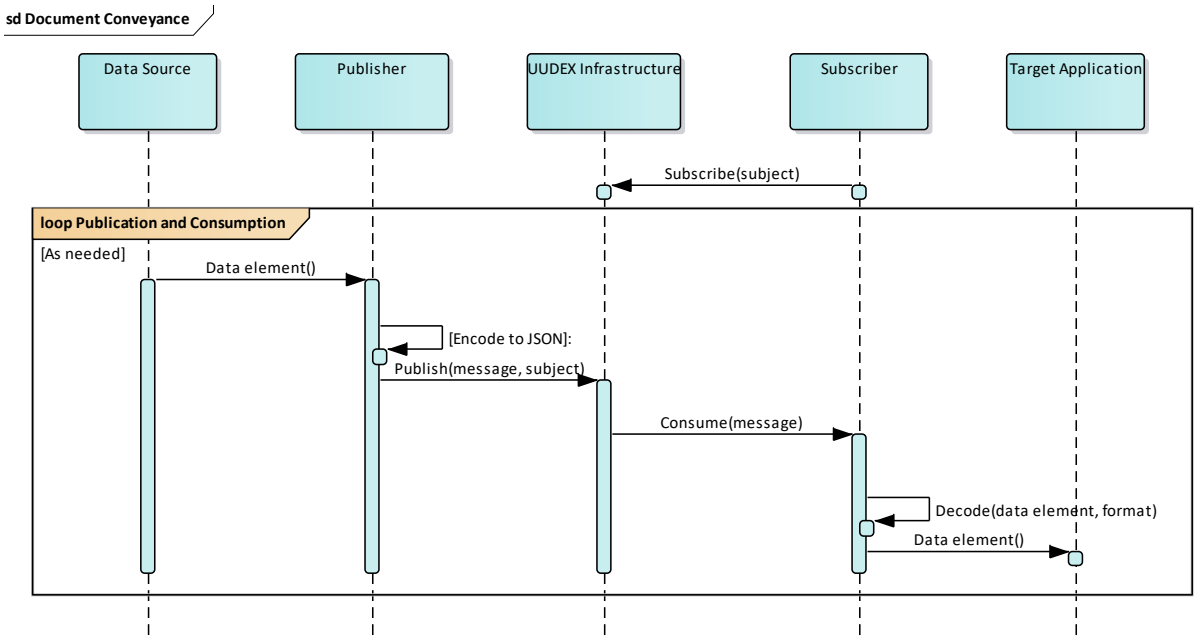

Figure 10. Document Conveyance

4.8 Discovery

Discovery occurs when a UUDEX Participant wants to access information that may be provided by other UUDEX Participants. ACLs defined by the 'owning' UUDEX Participant will identify whether or not a UUDEX Participant is able to:

- Discover the existence of a subject
- Subscribe to a specific subject
- View details of the DataSet related to a subject (if one exists)

4.9 Subscriptions

Subscriptions are the mechanism by which one UUDEX Participant collections information from UUDEX Subjects. The following are basic statements about subscriptions:

- UUDEX Endpoints subscribe to UUDEX Subjects.
- Many UUDEX Endpoints may subscribe to the same UUDEX Subject and a single UUDEX Endpoint may subscribe to multiple UUDEX Subjects.
- Each subscribing UUDEX Endpoint will receive a copy of any message published to a specific UUDEX Subject.
- The UUDEX Subject's publisher's ACLs control which UUDEX Participants may configure UUDEX Endpoints to subscribe to a given UUDEX Subject.

Figure 11 shows the subscription to a UUDEX Subject by UUDEX Endpoints, after which DataElements published by a UUDEX Producer to the subject can be consumed by those UUDEX Endpoints.

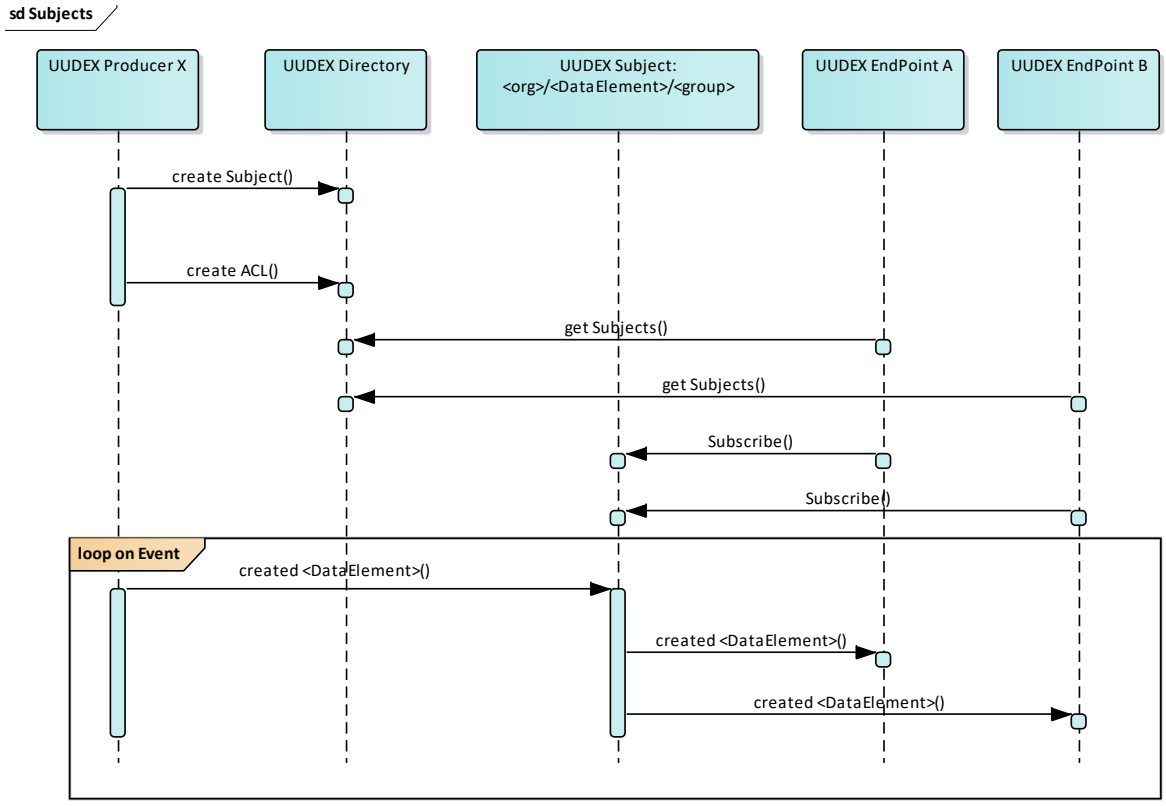

Figure 11. UUDEX Subscriptions

As noted in Section 4.6, UUDEX Endpoints may have peers to support redundancy and resiliency. As such, when a particular UUDEX Endpoint submits a request to create a subscription, there may be other UUDEX Endpoints belonging to the associated UUDEX Participant that are treated as peers of the requesting UUDEX Endpoint for the purpose of this subscription. Should a given UUDEX Endpoint become unavailable, its peers need to be able to receive subscription fulfillment messages. It is the responsibility of the UUDEX Directory to ensure this happens, and as a result, when a subscription is established, the requesting UUDEX Endpoint will identify the UUDEX Endpoints that should be treated as peers from the perspective of the subscription. Similarly, all peer UUDEX Endpoints would be considered 'owners' of the subscription, with the power to manage that subscription as described below. Regarding Figure 11 and the following descriptions, we will refer to 'the UUDEX Endpoint' that created and manages the subscription. However, note that it may be the case that this 'logical UUDEX Endpoint' is actually a peer group of UUDEX Endpoints, all with equivalent rights and roles with regard to the subscription.

Subscription fulfillment refers to the process by which new data is added to a UUDEX Subject, and the logical UUDEX Endpoint is alerted to this action. As shown in Figure 11, the UUDEX Framework supports two methods of fulfillment—Data Push and Data Notification.

- Data Push – This mode of fulfillment causes new data to be sent to the logical UUDEX Endpoint directly and is useful in sharing information such as measurement sets to which the UUDEX Endpoint is subscribed.
- Data Notification – In this mode of fulfillment, the UUDEX Directory sends a notification message alerting the logical UUDEX Endpoint to the presence of new data and providing a

pointer to the data so the logical UUDEX Endpoint can directly query and retrieve that data. This method is useful in conveying changes or updates to data, such as power system models, that may then be retrieved by the UUDEX Endpoint.

Each UUDEX Subject will consist of one or more ‘topics.’ A topic is a combination of subject and fulfillment procedure. A given UUDEX Subject might have separate topics to support both Data Push and Data Notification fulfillment, or it might just support one. (For example, Data Notification would not make sense for UUDEX Subjects that provide data requiring low latency delivery, so the UUDEX Subject might only provide a Data Push topic.) A logical UUDEX Endpoint subscribes to a specific topic, thus specifying both the UUDEX Subject of interest and the manner of subscription fulfillment desired.

Logical UUDEX Endpoints control their subscriptions. They can Create, Pause, Resume, and Delete each subscription for which they are the subscriber. These controls are described below:

- Create – Requests that the UUDEX Directory establish a subscription to the named UUDEX Subject. If the logical UUDEX Endpoint's UUDEX Participant is allowed to consume the data from the UUDEX Subject according to the UUDEX Subject's ACL, the subscription will be permitted.
- Pause – Temporarily pause notifications/delivery related to subscription fulfillment. The subscription remembers the last data that the logical UUDEX Endpoint received.
- Resume – Pick up subscription fulfillment, starting with the first available data added to the UUDEX Subject following the Pause request. (Note, however, that if a subscription is paused for an extended period of time and if the UUDEX Subject's data is of ephemeral value and periodically deleted, some data between the Pause and Resume action might be lost.)
- Delete – remove the subscription from the UUDEX Directory. If, at some later point, the logical UUDEX Endpoint creates a new subscription to the UUDEX Subject, fulfillment will start with the first material to arrive at the UUDEX Subject following the new subscription.

After connecting to the UUDEX Infrastructure, UUDEX Consumer endpoints will subscribe to one of more UUDEX Subjects, where each subscription is validated based on ACLs defined for that UUDEX Subject by the ‘owning’ participant.

There might be situations that cause a subscription to be cancelled other than through a direct request from a logical UUDEX Endpoint. In such a case, the UUDEX Directory must immediately inform the logical UUDEX Endpoint of this cancellation. This avoids a situation in which the logical UUDEX Endpoint believes it will receive timely messages, but this is not the case. It is possible that a single UUDEX Participant might have multiple subscriptions to the same UUDEX Subject and topic. These subscriptions will be identified differently and managed separately (i.e., pausing one subscription would not pause the other). The UUDEX Directory must not try to unify multiple subscriptions from the same UUDEX Participant into a single subscription for either fulfillment or subscription management.

4.10 Resiliency

4.10.1 Data Source Selection

UUDEX Endpoints are responsible for the selection of data sources. This is usually done through the process of development or configuration of endpoint software. Typically, redundant

data sources would exist that could be leveraged for an endpoint that is responsible for publishing information on behalf of the data source.

4.10.2 Redundancy

Redundancy is accomplished in several ways:

- Redundancy within the UUDEX Infrastructure, where UUDEX services are provided by a set of servers and a primary network that has at least one backup network to which the primary network fails over
- Definition of peer instances for UUDEX Endpoints (producers or consumers)
- Existence of redundant data sources and sinks from which UUDEX Endpoints read or write data.

Redundancy within the UUDEX Infrastructure is accomplished by replicating UUDEX Components, where those UUDEX Components form a coordinated cluster that ensures continued operation through a variety of failure and recovery scenarios. It is the responsibility of a UUDEX Endpoint and other uses of the UUDEX API to use a specified set of candidate connection addresses when connecting to the UUDEX Infrastructure, at which point the active set of underlying sites, servers and services are transparent to the UUDEX Endpoints. Ideally, to achieve the highest levels of availability, the underlying servers that support the UUDEX Infrastructure would be deployed across a minimum of three geographically separated physical locations.

UUDEX Participants can decide on the level of redundancy that is needed for given information flows and can deploy an appropriate number of peer endpoints. Where a high level of availability is needed, peer endpoints would typically be deployed at different sites and would use more than one physical network path to the UUDEX Infrastructure.

The proper access and update of redundant data sources and sinks is the responsibility of the UUDEX Participant Infrastructure and UUDEX Endpoint implementation.

4.10.3 Consensus

Redundancy brings on the need to avoid confusion. A key concern is to avoid a ‘split-brain’ scenario, where UUDEX Components within the UUDEX Framework continue to operate while isolated from other UUDEX Components, resulting in a diverging view of state. This scenario is avoided by having UUDEX Components operate through ‘consensus.’ There are two key situations for which a consensus mechanism is needed:

- Where a UUDEX Participant has peer UUDEX Endpoints that are configured for redundancy
- By replicas of servers and their associated services within the UUDEX Infrastructure.

In the case of a UUDEX Endpoint that is configured with peer UUDEX Servers for redundancy, it is the responsibility of those peers to employ an appropriate consensus mechanism such as a quorum calculation. Within the UUDEX Infrastructure, there should be *N* UUDEX Servers where *N* is an odd number, minimally three of each UUDEX Component, to achieve some level of fault tolerance while avoiding the ability to segment the peers into two partitions each assuming it is is functional and the other has failed. Ideally, *N* should be an odd number greater than four. The servers should be configured to avoid a segmentation or other failure event that would leave an

equal number of UUDEX Servers in each resulting partition. This allows the calculation of a quorum, where quorum is achieved when the number of UUDEX Servers aware of each other in a UUDEX Instances is greater than half of the number of UUDEX Servers that are configured in the UUDEX Instance.

In the event of a loss of quorum (e.g., a transition from three peers to one, or five peers to two), the remaining peers should transition to a down state in order to avoid a split-brain scenario. Transitions that involve an increase in the number of peers (e.g., from three instances to four) or a loss of a minority of peers (e.g. transition from three peers to two or five peers to three) would not impact UUDEX functionality. When a quorum is lost from the perspective of a given peer, there is no ability to participate in the determination of a consensus, which then requires the peer to transition to a down state.

Peers in a down state should continually monitor other peers to determine if quorum can be regained, and the peers can exit the down state. UUDEX Producer peers entering the up state should resynchronize with other UUDEX Producers before fulfilling a subscription by allowing existing UUDEX Producers to fulfil subscription requests.

4.11 Monitoring and Management

Each UUDEX Endpoint should periodically publish a health and statistics message that can be captured and recorded by the UUDEX Infrastructure for the purposes of monitoring. This message would provide information such as:

- The time duration the endpoint has been active (a measure of the reliability of the endpoint)
- A timestamp[1] with the current time (useful to identify message latencies)
- The time of last report
- The number of messages/data elements published and consumed since last report
- The number of bytes published and consumed since the last report.

Figure 12 shows the process for a UUDEX Endpoint to periodically report its health and other operational statistics. The *Monitoring Agent* function shown in the figure is a service deployed within the UUDEX Infrastructure that is responsible for capture of UUDEX Endpoint health and related statistics. The *Monitoring Tool* function shown in the figure would provide the ability for a user to view UUDEX Endpoint current state and performance history. This is further described in Chapter 7.0.

[1] Note – this assumes that time is synchronized between the two endpoints

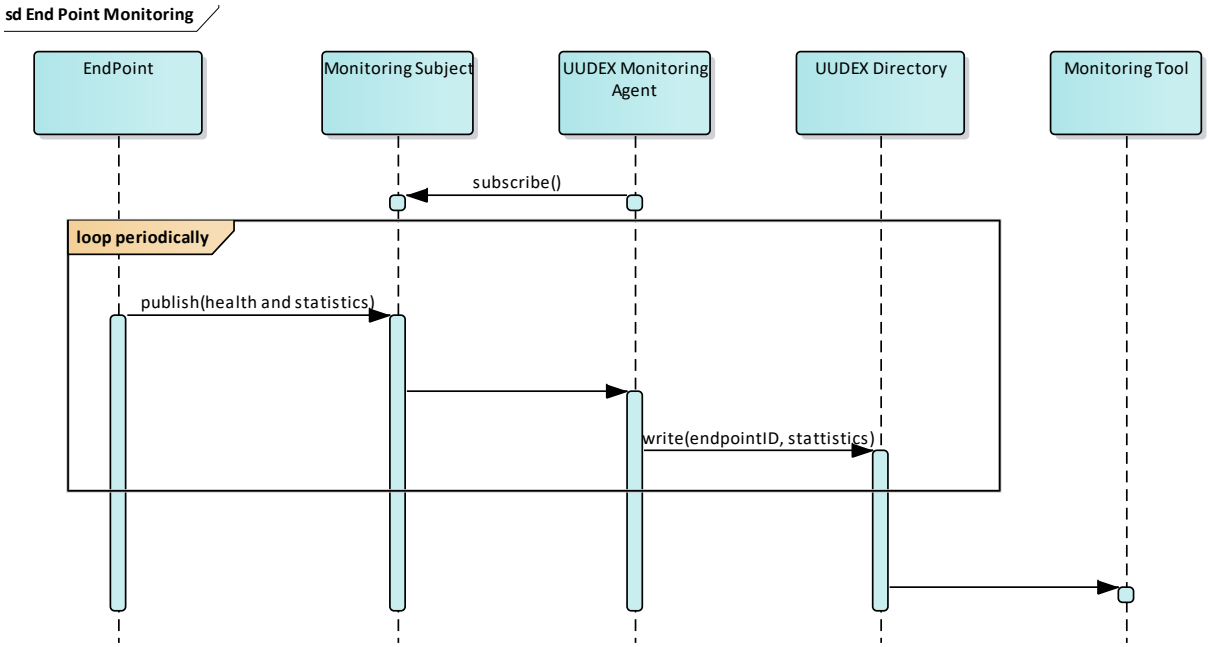

Figure 12. UUDEX Endpoint Monitoring

4.12 Testing and Verification

When a new UUDEX Endpoint is connected to the UUDEX Infrastructure, some level of testing and verification is needed. Key aspects of this are essentially automated, as UUDEX Endpoints are required to issue periodic messages for the purpose of monitoring. When a UUDEX Endpoint is deployed, the following should be visible:

- The UUDEX Participant can access the UUDEX Endpoint and determine if it has successfully connected to the UUDEX Infrastructure using interfaces provided by the specific UUDEX Endpoint product.
- UUDEX Administrators and the UUDEX Participant Administrators can see that the UUDEX Endpoint is connected to the UUDEX Infrastructure and is providing periodic health and state information.

Provisions would also be put in place to ‘test’ UUDEX Subjects, where information could be published and subscribed for testing and verification purposes.

5.0 User Interfaces

The nature of UUDEX user interfaces is described in this chapter. The user interfaces are not intended to provide any end-user functionality, rather they provide administration and monitoring capabilities.

These user interfaces are not to be confused with the API used to publish and subscribe to UUDEX Subjects using the UUDEX Framework.

5.1 Roles

The two primary roles for UUDEX administrative users are UUDEX Administrators and UUDEX Participant Administrators. Their roles are briefly described below:

- UUDEX Administrators have global responsibility for the UUDEX Framework and can authorize UUDEX Participants.
- UUDEX Participant Administrators can perform activities related to the publication and consumption of information for a given UUDEX Participant and also can access monitoring tools.

There may be needs to define other roles over time. The following are examples of potential additional roles:

- A UUDEX Auditor who could view the information needed to adequately audit the operation and management of the UUDEX Infrastructure
- A UUDEX Monitor who could view the current state of the UUDEX Infrastructure, logs, and related performance metrics.

An individual can be assigned to one or more roles.

Individuals designated to fulfill these sensitive UUDEX Infrastructure roles should be vetted. Guidance provided in Appendix A might apply to this vetting process.

5.2 Core functionality (minimum requirements)

5.2.1 UUDEX Administrator

The user interface should provide the following capabilities for a UUDEX Administrator:

- Create, authorize, and de-authorize UUDEX Participants
- Update UUDEX Participant details
- Define valid DataElementTypes, beyond the standard set defined by UUDEX.

These capabilities are shown in Figure 13.

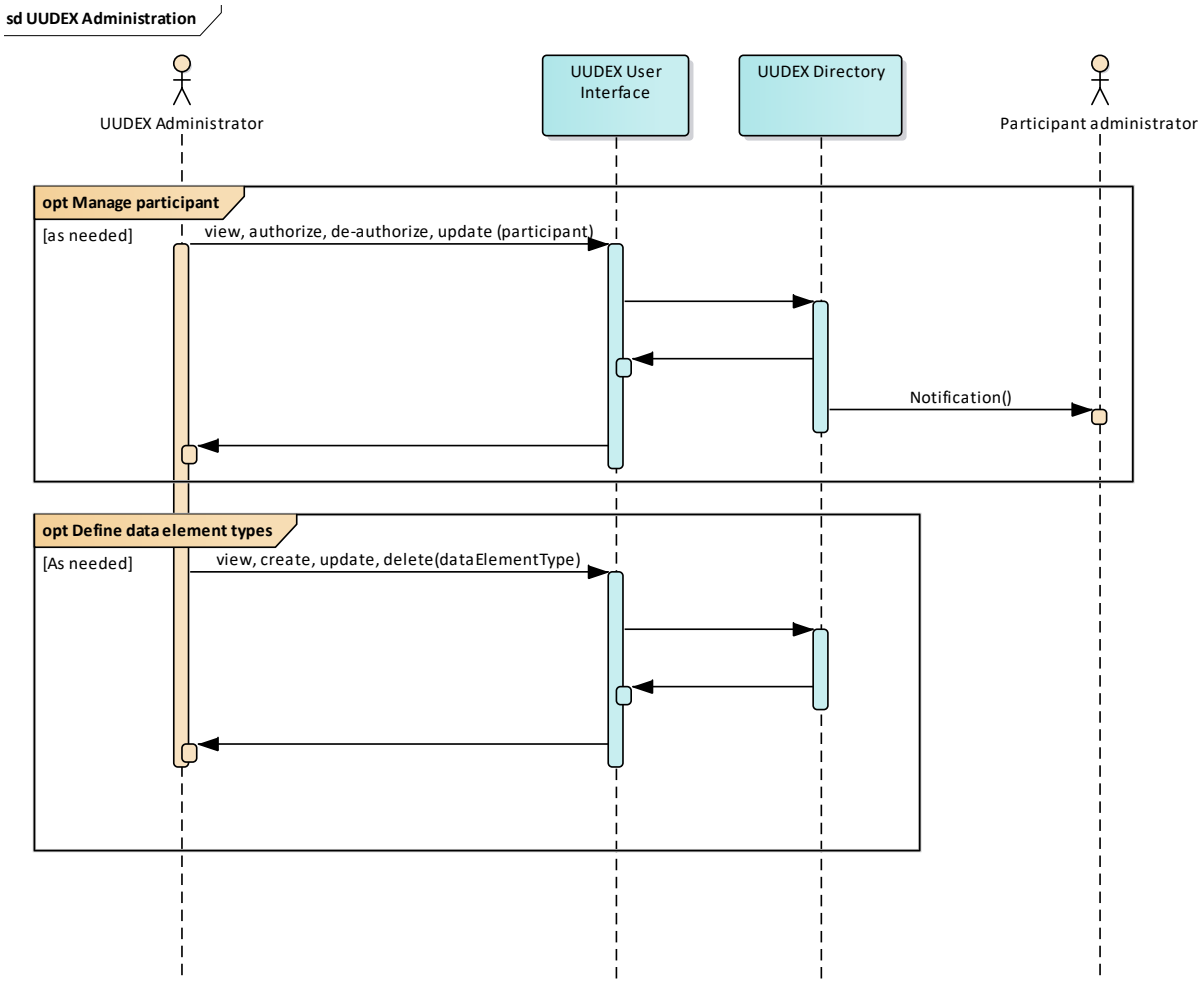

Figure 13. UUDEX Administrator Actions

5.2.2 UUDEX Participant Administrator

The user interface would provide the following functionality for a UUDEX Participant Administrator whose role is different from that of a UUDEX Administrator:

- Create and delete UUDEX Subjects (for publications)
- Discover (view) available UUDEX Subjects (for subscriptions)
- CreateDataSets and populate them with point definitions
- Edit the point definitions in DataSets
- Delete DataSets
- Create, authorize, and de-authorize UUDEX Endpoints
- Associate UUDEX Subjects to UUDEX Producer UUDEX Endpoints
- Set subscriptions for UUDEX Consumer UUDEX Endpoints.

These capabilities are shown in Figure 14.

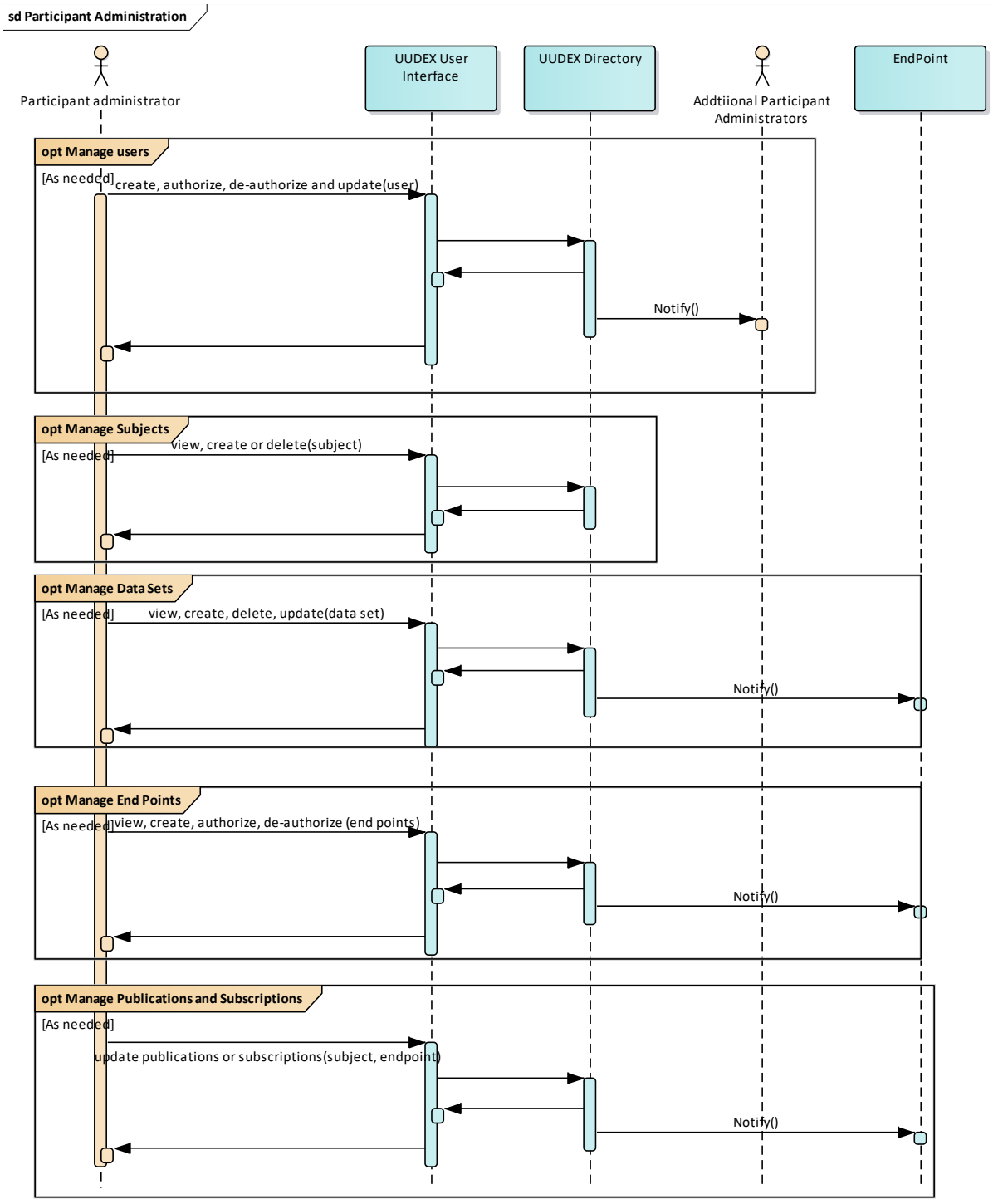

Figure 14. UUDEX Participant Administrative Actions

5.3 Notifications

A key aspect of the user interface is to provide the means to issue notifications and alerts to users, such as UUDEX Administrators or UUDEX Participant Administrators. These notifications would normally be issued 'out-of-band,' using mechanisms such as email or text messages.

When a condition of interest is detected by the UUDEX Infrastructure, it is reported to the UUDEX Directory for recording and generation of notifications to potentially interested users. Conditions of interest can be defined in two ways:

- Hard coded into UUDEX Components, typically as a part of exception handlers
- Rules that would identify thresholds (e.g., message rates) upon which notifications are issued.

Given security concerns, only a summary of the notification is provided with a URL that would enable the user to securely authenticate in order to get details or take consequential actions. This is shown in Figure 15.

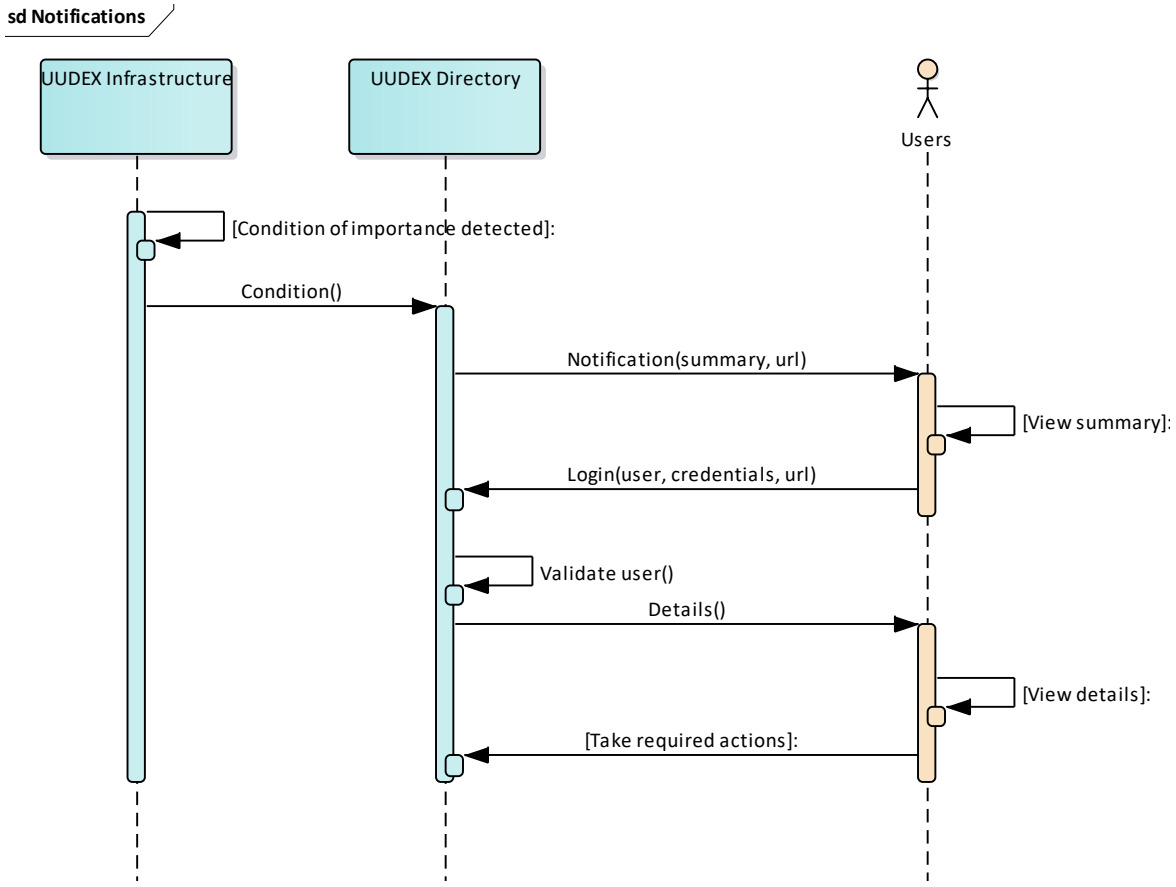

Figure 15. Notification Processing

As examples, notifications could be generated as a consequence of:

- User authorization
- UUDEX Endpoint authorization
- UUDEX monitoring (e.g., alerts of UUDEX Endpoint or UUDEX Infrastructure 'health' issues, see Chapter 7.0)

It is important to note that alerts are notifications that typically require some action by a user. For example, a notification would be generated if a UUDEX Endpoint was created or a

configuration updated, but an alert would be generated to indicate that a UUDEX Endpoint has failed.

5.4 Third-Party User Interfaces

Given that the UUDEX Framework is intended to be transport neutral, there may be user interfaces that are needed to manage and monitor aspects of the transport layer as well as the underlying communication fabric. These would be supplemental to UUDEX-specific user interfaces as described earlier in this chapter. These would likely be leveraged by UUDEX Administrators as opposed to UUDEX Participant Administrators.

6.0 Application Programming Interfaces

APIs are used to construct UUDEX Endpoints and administrative interfaces. The API may be versioned, where the UUDEX Infrastructure will respect a given set of versions. Implementation of an API would have the following characteristics:

- A transport technology
- A version identifier
- Bindings for one or more programming languages.

The APIs would be leveraged by:

- Administrative user interfaces, as used by the UUDEX Administrator or UUDEX Participant Administrator
- UUDEX Endpoints at UUDEX Producers
- UUDEX Endpoints at UUDEX Consumers.

Before using an API, the user or application needs to be authenticated. Once authenticated, only authorized API calls or data may be accessed.

Each of the workflows described in the preceding sections leverages APIs. Workflows related to the development of UUDEX software components are outside the scope of this specification.

A key aspect of APIs is to hide as much of the underlying implementation details as possible. This allows increased simplicity when implementing clients and enables the underlying implementation to evolve more easily.

7.0 Monitoring/Diagnostics/Testing

The UUDEX Framework will provide mechanisms for monitoring, testing, and general problem diagnosis. Some of these mechanisms may be specific to the transport layer or communication fabric. Other mechanisms will be more specific to the tracking and analysis of UUDEX Connections and information flows. This is shown in Figure 16.

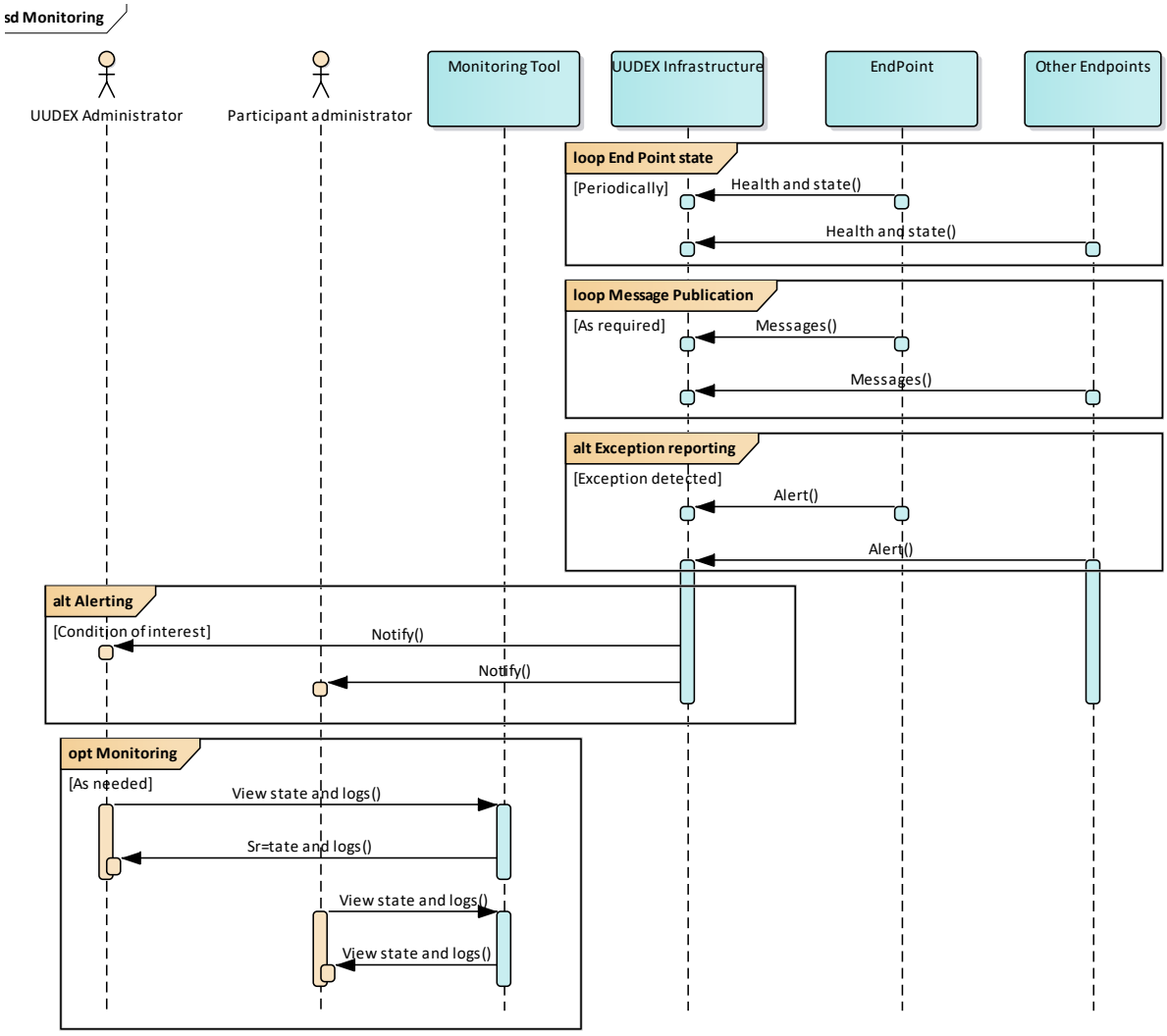

Figure 16. Monitoring

Notifications (including Alerts) are discussed in Section 5.3.

Additionally, notifications could be sent as simple network management protocol (SNMP) traps or logged via Syslog messages for use by third-party management tools and enterprise event management software.

8.0 Extensions

UUDEX is designed to allow for both evolution and extension. In the case of evolution, there could be new versions of data element types that need to be accommodated. For extensions, new data element types might be defined or information might be added for existing data element types. Extensions also could be provided in the form of APIs that would be used by endpoints.

9.0 References

The following informative references are applicable to this document:

- UUDEX Functional Requirements
- UUDEX Protocol Design
- IEEE Std. 1003.1 (POSIX.1-2008) (includes Uuencode and Uudecode)
- IEC 61970-301 (CIM)
- IEC 60870-6 (TASE.2, a.k.a. ICCP)
- IETF RFC 4180 (csv) (https://tools.ietf.org/html/rfc4180)
- IETF RFC 6713 (gzip) (https://tools.ietf.org/html/rfc6713)
- IETF RFC 7159 (JSON) (https://tools.ietf.org/html/rfc7159
- ISO 32000 (portable document format - PDF)
- XML Standard (http://www.w3.org/standards/xml/)
- Java Message Service (JMS) 2.0 (https://jcp.org/en/jsr/detail?id=343)
- Data Distribution Service (DDS) (https://www.omg.org/omg-dds-portal/)
- MQTT Standard (https://docs.oasis-open.org/mqtt/mqtt/v5.0/mqtt-v5.0.html)
- L. Pesonen, D. Eyers, J. Bacon, “Access Control in Decentralized Publish/Subscribe Systems”, Journal of Networks, April 2007
- L. Pesonen, “A capability-based access control architecture for multi-domain publish/subscribe systems”, University of Cambridge Computer Laboratory, June 2008
- G. Hohpe, B. Woolf, “Enterprise Integration Patterns”, Addison Wesley, 2004
- K. Birman, “Guide to Reliable Distributed Systems”, Springer, 2012
- North American Electric Reliability Corporation (NERC) CIP-004-6: Cyber Security — Personnel & Training (https://www.nerc.com/pa/stand/Pages/ReliabilityStandardsUnitedStates.aspx?jurisdiction=United%20States)

Appendix A – Existing Industry Organizational and Personnel Vetting Procedures

Before two Universal Utility Data Exchange (UUDEX) Endpoints can establish a connection to share information, a reason for sharing information needs to be established. In some cases, the reason is based on regulatory obligations (e.g., in the electricity sector sharing of information between a transmission owner and a reliability coordinator as mandated by North American Electric Reliability Corporation [NERC] Standards), contractual obligations (e.g., sharing of market-related data among market participants), or for public benefit (e.g., sharing of outage information to a public website or with first responders). In each case, the information publisher needs to be sure that the consumer: 1) needs to have access to the information, and 2) will use it only for the purposes for which it is shared. This requires that the two organizations either have established a level of trust or are mandated to 'trust' each other under penalty of regulatory sanctions. The specific procedures for establishing trust are beyond the scope of this document.

Individuals designated to fulfill roles of high responsibility within a UUDEX Instance need to be thoroughly vetted to ensure that they are trustworthy. UUDEX Administrators will be able to view all data in the UUDEX Instance and control who receives information exchanged over UUDEX. UUDEX Instances might exchange information that is the subject of national regulations, commercial contract agreements, and/or other limitations on use. The sensitivity of this data might require the application of certain vetting requirements for UUDEX Administrators (e.g., access to information controlled by national regulation might require, by law, background checks performed by a national law enforcement body).

The identification of parties to serve in sensitive UUDEX roles needs to be informed by an understanding of relevant regulations and other agreements that might stipulate limitations on access and ensure that the vetting process accounts for these limitations. Failure to adequately vet individuals for these roles could be very costly and damaging to all the organizations using the UUDEX Instance in question. Even if an inadequately vetted administrator is not actively seeking to leverage their authority to compromise the security of the UUDEX Instance, they are unlikely to be aware of data-use constraints that were not identified as part of the vetting process, which could lead to accidental misuse of data. For these reasons, the vetting process for individuals with authority over aspects of UUDEX Instances, especially the UUDEX Administrators, needs to be thorough and comprehensive, and including an ongoing and periodic training component so that vetted individuals understand their role in protecting the UUDEX Infrastructure and the data communicated through it.

In the United States, there are several existing frameworks for performing this increased scrutiny. These frameworks are briefly described below:

- For North American organizations that support reliable operations of the bulk electric system and are required to comply with NERC Standard CIP-004, the Personnel Risk Assessment process provides a framework for establishing and conducting a risk assessment for staff who have access to the control systems that operate the bulk electric system. This framework could be adapted and adopted by other organizations.

- The Transportation Worker Identification Credential[1], also known as TWIC®, is a U.S. government identification credential and is required for workers who need access to secure areas of maritime facilities and vessels. Many of these maritime facilities are located at power plants, so many electric utilities are familiar with the credential and the approval process. The process involves filling out an application, and after the application has been processed, an in-person interview to complete the process.

Organizations may also have additional vetting procedures that can be adapted or used—for example, as part of the hiring process or for internal transfers. Governmental organizations such as municipal utilities may have more stringent requirements that include criminal background checks.

Individual organizations that are establishing a UUDEX Infrastructure for internal purposes (e.g., synchronizing data between a primary and backup control center) may have specific requirements or procedures that are different than those for administrators of external UUDEX Connections.

[1] See https://www.tsa.gov/for-industry/twic

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99352
1-888-375-PNNL (7665)

www.pnnl.gov