

## Using Exchange Network and CDX Services: Key Steps for Exchanging Emissions Inventory Data

Roy Chaudet and Chris Clark

U.S. Environmental Protection Agency, Office of Environmental Information (OEI),  
1200 Constitution Avenue NW (2823T),  
Washington, DC 20460  
[chaudet.roy@epa.gov](mailto:chaudet.roy@epa.gov), [clark.chris@epa.gov](mailto:clark.chris@epa.gov)

### ABSTRACT

Approximately 18 agencies are currently submitting emissions inventory data to EPA's Emissions Inventory Analysis Group (EIAG) for the National Emissions Inventory (NEI) using the Exchange Network (EN). All fifty States have full EN nodes in place and are exchanging data for at least one environmental program. The remaining, State, Local and Tribal (SLT) agencies can use the services provided by the EN and EPA's Central Data Exchange (CDX) to submit and exchange emissions inventory data. Some of these agencies are receiving EN grants to include the NEI dataflow for network exchanges through their existing nodes.

Use of the EN and Extensible Markup Language (XML) is considered the main data transfer process and format for submissions to the new emissions inventory system (referred to as the Emissions Inventory System (EISys)). In addition, future benefits may include the use of a universal XML schema and the EN for exchanges between industry and States.

This paper focuses on: 1) available EN and CDX services, 2) how those services are used in exchanging and publishing emissions data between SLT agencies and EPA —as well as among SLTs, 3) steps for state air programs to use these services through their existing state nodes, 4) steps for Local and Tribal agencies not working with a state to use a Node Client for exchanges, and 5) communicating between program and IT staff. The paper will include examples of how the EN is being used to share data among States for multimedia as well as for emissions data exchanges.

### INTRODUCTION

State, Local and tribal Agencies (SLT) are using CDX and EN services to submit emissions inventory data to EPA's EIAG for the National Emissions Inventory (NEI). These agencies have submitted emission inventory data via CDX Web since 2002 and via the EN since 2005. In 2007, approximately 18 agencies submitted emissions inventory data using the Exchange Network and while the remaining agencies submitted via CDX web. Because all fifty States have fully operational EN nodes, SLTs who traditionally submitted via the web can now use EN services to submit emissions inventory data to EPA, —as well as exchanging data directly between agencies.

The new modernized NEI system, known as Emissions Inventory System (EISys), will use the EN and Extensible Markup Language (XML) as the main data transfer process and format for submissions. These submissions will make use of a consolidated emissions reporting schema that provides a single schema for one program for exchanges between industry and States<sup>1</sup>.

To leverage the use of this schema and the EN services, some SLT agencies are applying for and receiving EN grants to include the NEI dataflow for network exchanges through their existing nodes. Six states and one tribe were awarded 2007 EN grants and four states were awarded 2008 EN grants to implement the NEI dataflow<sup>2</sup>.

This paper will describe the EN services used by the NEI data flow and steps for STL air programs to use these services. Before describing these services, the Exchange Network and the Central Data Exchange are briefly described.

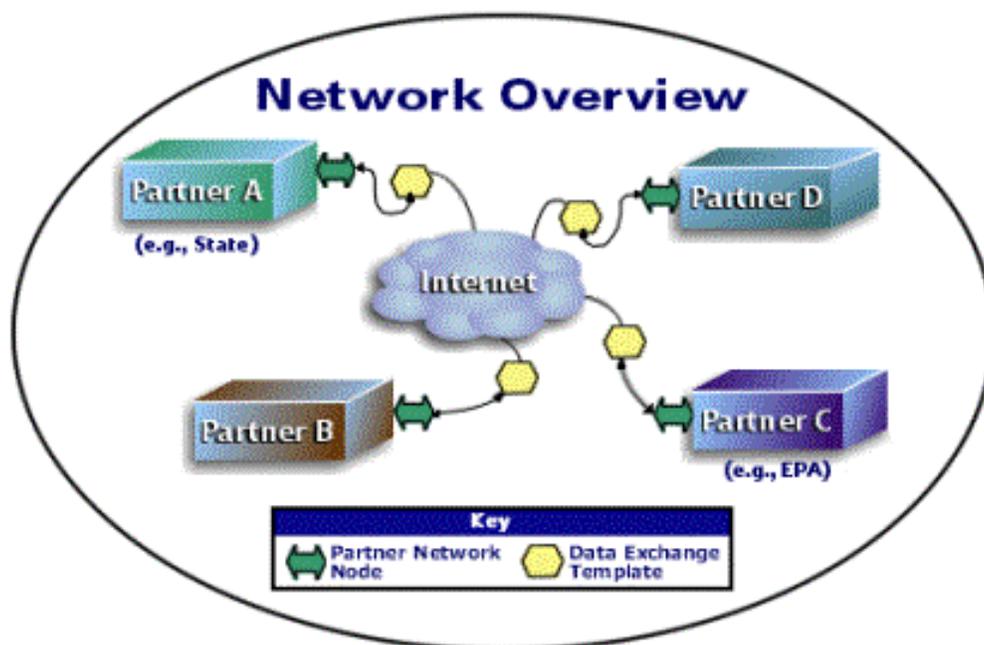
### **WHAT IS THE EXCHANGE NETWORK?**

The National Environmental Information Exchange Network (Exchange Network) is “a secure Internet- and standards-based approach for exchanging environmental data and improving environmental decisions. The U.S. Environmental Protection Agency, State environmental departments, and U.S. tribes and territories are partnering to build the Exchange Network to increase access to environmental data and make the exchange of data more efficient.”<sup>3</sup>

Figure 1 provides an overview of the Exchange Network. The Network consists of trading partners that exchange data over the internet using Network Nodes to a data exchange templates in XML. Key components used for EN exchanges are:

- *Data standards* that provide a common vocabulary in the data exchanges<sup>4</sup>
- *Standard protocols* for exchanging data using web services
- *Nodes* on the network that conform to Network Specifications<sup>5,6</sup>
- *XML schema* located in the XML Schema registry<sup>7</sup>
- *Grant Program* to apply for and fund NEI implementations

The Exchange Network supports automated exchange of data that enables timely and accurate exchange. The goal is to reduce reporting burden and improve data quality through validation checking.



**Figure 1. Exchange Network Overview**

## WHAT IS CDX?

The Central Data Exchange (CDX) is the central point through which environmental data enters the Agency. It is also EPA's Node on the Environmental Information Exchange Network. EPA's Node supports both inflows and outflows or publishing data to trading partners.<sup>8</sup> Exchanges through CDX are made through the following:

**CDX Node** – EPA's Web services infrastructure on the Exchange Network that supports machine-to-machine data exchanges using standardized web service protocols and XML schemas.

## EXCHANGE NETWORK AND CENTRAL DATA EXCHANGE SERVICES

To enable data exchanges through the Exchange Network, a variety of services are available and are currently being used:

- **Registration (Network)** — A SLT Agency can register on the Exchange Network and establish an account for NEI submissions/exchanges. You must Obtain a Network Authentication and authorization (NAAS) user ID by either contacting your node administrator for States that have nodes or contacting the Node Help Desk. The account request is routed for approval by the EN and the program lead.
- **Authentication and Authorization (NAAS)** — Once an account is established, an authorized user can use NAAS to authenticate using certificate mechanisms for EN exchanges.

- **Secure Document Exchanges** —The Exchange Network Security model allows secure exchange of emissions data via secure socket layer (SSL) and use of the NASS. These secure exchanges also meet requirements beyond those required for emissions inventory submissions (e.g., for confidential business information and for meeting requirements for the Cross Media Electronic Reporting Rule (CROMERR)).
- **Data Reporting and Querying**— Standard Web Service methods are used for basic exchanges over the EN that include Submit, Solicit, Query, Notification, GetStatus, and Download.
- **Messaging (status, error)** — A variety of messaging services are available that provide users and partner nodes with information on error conditions and current processing state of submissions and exchanges (e.g., pending, complete, failed). Messaging is provided through either Notification via the node or E-mail notifications.
- **Quality Assurance (validation)** — QA Services provide interfaces for checking the validity (i.e., structural integrity and correctness) of XML document instances and improving data quality. The QA Services are provided either through the web (<http://tools.epacdxnode.net>) or when a submission is made through the Exchange Network node. CDX uses Schematron services (an XML schema language) in order to validate the content of XML instance documents by checking extended business rules that schema cannot validate. The QA service sends e-mails to the user in response to completion of processing events and system validation failures.
- **Web Service Management (Monitor Services)** —Web Service management services include (1) availability monitoring to determine the active status of other Nodes on the Exchange Network (Node Ping), (2) process and resource monitoring of resource utilization and process activity, and (3) service alerts. In the event of service interruption, CDX automatically notifies maintenance personnel and trading partners.
- **Internal Services:** Several internal CDX services support EN exchanges that include but are not limited to:
  1. *Archiving* includes the initial data submission files, translated/transformed versions, and associated signing/encryption items. CDX also archives result files, such as any data validation or QA results returned from the QA Service. The archived records can be retrieved as needed.
  2. *Logging* includes creating log files for transactions submission, QA, and distribution to EPA systems or other trading partners (e.g., for data publishing). All archived file submissions are written to a transaction log, capturing the date/time of the submission, file name, and who uploaded it.

3. *Routing/Distribution* includes routing transactions to EPA program systems or for publishing to SLTs or other trading partners. CDX has the capability to extract user or embedded data information in order to enhance submission routing, content and workflow. Submission files can be extracted from approved compression formats or compressed for reduction in data transmission size.
- **Database Extraction/Insertion Tools** are available that SLTs can use to move data between XML and SQL formats for extraction from or insertion to their databases. These tools (e.g., Velocity Mapper/Hibernate Object Relational Mapping (ORM) Tools) are configuration-driven in that the parameters for defining the mappings for extraction or insertion are in a configuration file that requires little to no actual coding. These tools can be used to generate XML files and have been integrated with different node configurations.
  - **Registry/Discovery Services** for the network include (1) Exchange Network Discovery Services (ENDS) that dynamically find and bind partner web services, and (2) Universal Description, Discovery, and Integration (UDDI) that provides a registry used to publish service listings, discover each other, and define how the services or software applications interact (e.g., detailed validation services, get network node addresses, XML parser, node queries by flow).

To utilize these services for exchanging emissions inventory data, State, local and tribal air program staff need to determine if they are using a *full network node* through their node administrator or using a *node client* (also referred to as a Network Desktop). ***A node client is recommended for this type of annual submission where the program staff can directly submit and correct errors without the assistance of the IT Node Administrator.*** To assist SLTs in using these services, different nodes and web services are described below.

## NETWORK NODES AND WEB SERVICES

Two general types of Nodes are used in the Exchange Network:

**Full Network Node** can both request data from the Network, as well as publish data to the Network in response to requests (e.g., a query or solicit) from other Network Nodes. Full nodes can potentially leverage the full capabilities of the Network for machine-to-machine interaction by sending requests for data, and publishing data for use by other Network partners. It sends *and receives* standards-based messages and returns requested information as XML. Each trading partner on the EN has only one Node. Full nodes provide machine-to-machine interactions and are available to be deployed on a number of platforms (JBOSS, Apache Tomcat, Weblogic 8.1 and 9, Oracle 10.1.3, Websphere 6.1.) To download the Full Node for your platform go to the following Website

<http://test.epacdxnode.net/ngn>

**Node Client/Network Desktop** provide essentially the same methods and functionality as a Full Node except that it *cannot listen (and respond to) requests from other nodes*. This can limit the ability of processes that require responding to query requests to serve up data for data publishing. These clients are primarily for human-to-machine interaction and are normally used by trading partners that do not publish to the network. Network Desktops are easy to use and available in .NET and JAVA versions and software developer kits are available to simplify use and functions. Several versions of Node Clients are available, —most are free and can be downloaded from EPA or vendors.

To download a node client/network desktop, go to the following web pages:

Node Clients: [http://exchangenetwork.net/node/network\\_clients.htm](http://exchangenetwork.net/node/network_clients.htm)  
 CDX Node Client (includes velocity mapper): [https://test.epacdxnode.net/cdx\\_client/](https://test.epacdxnode.net/cdx_client/)  
 Network Desktop: <http://clientcentral.epacdxnode.net/msi/setup.msi>

A customized Node Client is being developed for use with the Emissions Inventory System (EISys) for use in submitting 2008 inventory data..

**Web Services** can be defined as information sources/application components whose functionality and interfaces using *standard messaging protocols* including:

- Extensible Markup Language (XML) for the payload or file,
- Simple Object Access Protocol (SOAP) as a wrapper for the payload,
- Web Services Description Language (WSDL) for network exchange functions and services, and
- Hypertext Transfer Protocol (HTTP) for secure communication via the internet using HTTPS for secure socket layer (SSL) transmissions.

In the simplest terms, Web Services are basically a machine-readable e-mail message.

**Web Service Methods** used by the Exchange Network include:

- ***Authenticate*** for authenticating to the Network,
- ***Submit*** that for making submissions from one node to another.
- ***Query*** provides the capability in real time to query and retrieve data (synchronously) through another node to a system database.
- ***Solicit*** provides the same capability as query except that it is not in real time (asynchronous).
- ***Notification*** for messaging between nodes
- ***GetStatus*** provides information on the status of transaction (e.g., pending, complete, failed)
- ***Download*** transactions to be downloaded from the EN (e.g., XML documents submitted, processing error reports)<sup>4</sup>

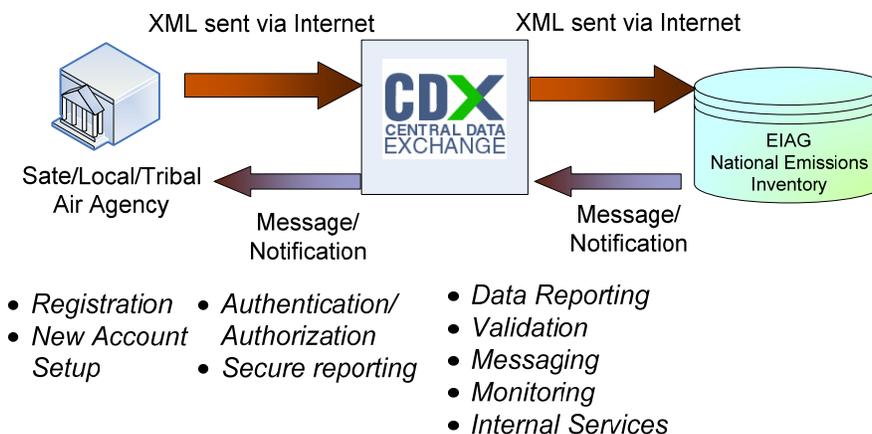
## USE OF EN/CDX SERVICES FOR EMISSIONS INVENTORY SUBMISSIONS/EXCHANGES

Figure 2 provides a simplified overview of the NEI submission process and the EN services used. The SLT uses the EN registration system for setting up a new account to obtain an ID and credential. The SLT first prepares an XML instance document or file for submission (the bulk of the work) and sets up an account if one is not already established. The authorized user authenticates, selects a file for submission/reporting and invokes a “Submit” web service call. For the current process for submitting to NEI through the EN, the SLT submitter will receive several messages:

- *Transaction ID* for authentication and submission (node) that acknowledges receipt
- *GetStatus* message of pending, failed, or complete status (node)
- *Validation* results for each XML or zipped XML file submitted (via e-mail)
- *Download* error report if the submission failed for SLT to correct and resubmit (node)

The user corrects any error until a success status is returned. The submission is distributed to the NEI for further processing.

**Figure 2. NEI Submission Processes Using EN Services**



### Validating Emissions Inventory Data Before Submitting

To validate your XML instance document prior to submission, you can go the QA Test site at <http://tools.epacdxnode.net> and shown in Figure 3. This QA Test Tool provides rapid processing using the same QA service you use during submission. The Basic Format Content Checker (BFCC) that some agencies use, may vary slightly since business rules may be changed in the CDX/EN QA Service between updates to the BFCC.

Figure 3. Web QA Tool

**SOAP Web Client - Microsoft Internet Explorer provided by EPA - version 6**

Address: <http://tools.epac.danode.net/soapclient?CacheWSDL=0&template=%2Fclientform.html&mf=soapform&SoapTemplate=%2FSoapResult.html&SoapWSDL=https%3A%2F%2Ftools.epac.danode.net%2Fxml%2Fvalid...>

**Network Node Client Forms**

Service Documentation: A set of xml validation services for the National Environmental Information Exchange Network (NEIEN)

**SOAP Method : SchemaValidate**

Documentation: Validate an XML document using schema files  
 Server-Address: <https://tools.epac.danode.net/xml/validator.wsdl>

userId:  xsd:string  
 password:  typens:PasswordType  
 documentType:  typens:DocumentType  
 xmlDocument:   xsd:base64Binary  
 docFormat:  typens:DocumentFormat  
 sendResultTo:  xsd:string  
 Show:   Format:

**SOAP Method : SchematronValidate**

Documentation: Validate an XML document using schematron file  
 Server-Address: <https://tools.epac.danode.net/xml/validator.wsdl>

userId:  xsd:string  
 password:  typens:PasswordType  
 documentType:  typens:SchematronType  
 xmlDocument:   xsd:base64Binary  
 docFormat:  typens:DocumentFormat  
 sendResultTo:  xsd:string  
 Show:   Format:

**SOAP Method : ValidateDocument**

Documentation: Validate an XML document using xml schema and schematron  
 Server-Address: <https://tools.epac.danode.net/xml/validator.wsdl>

userId:  xsd:string

The submission is made through the CDX Test environment until it passes validation, at which time your Production account will be activated for final submission. Depending on whether the type of node being used, the messaging received will be by program staff or IT staff. The process steps and for each type of node are discussed below.

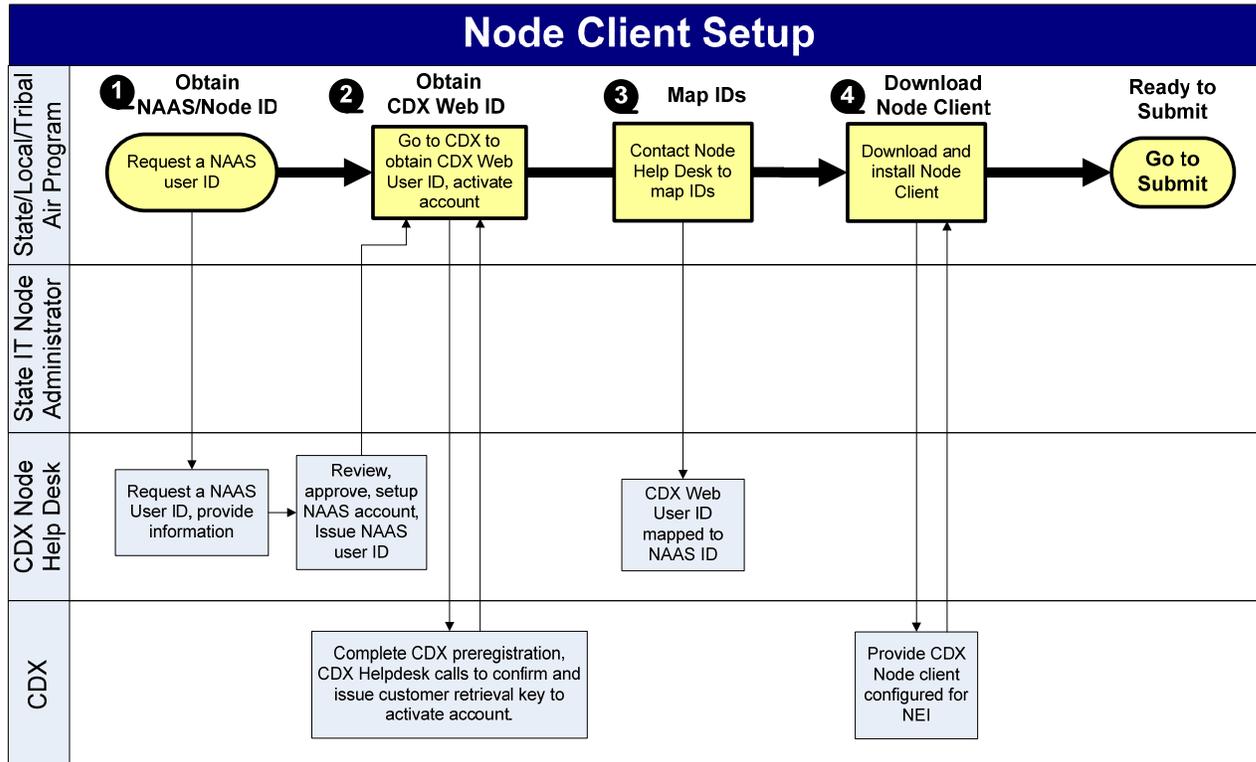
## SETTING UP AND USING EXCHANGE NETWORK SERVICES

Depending on whether an air agency uses a Full Network Node or Network Desktop, the following steps need to be taken to setup and use the EN services.

### Account Setup and Submissions Using Network Desktops/Node Clients

Figure 4 provides an overview of the steps for Local and Tribal agencies not working with a state to set up a Node Client. Figure 5 provides an overview of the steps for making emission inventory submissions using a node client.

Figure 4. Setting Up a Network Desktop/Node Client



- 1. Obtain NAAS ID** by contacting the Node Helpdesk at [nodehelpdesk@csc.com](mailto:nodehelpdesk@csc.com) or call 1-(888)-890-1995 Option 2. The Node Help Desk will provide a list of information required to be submitted tot the EN and the program lead for approval of your account. Only pre-registered users have access to the NEI dataflow.
- 2. Obtain CDX Web ID** Complete CDX pre-registration by contacting the CDX Help Desk to confirm and issue customer retrieval key to activate account.
- 3. Map IDs** Once you have your CDX Web ID, contact the Node help Desk to map it to your NAAS ID
- 4. Download Node Client/Network Desktop** Node clients can be downloaded from the following web sites:

Node Clients: [http://exchangenetwork.net/node/network\\_clients.htm](http://exchangenetwork.net/node/network_clients.htm)

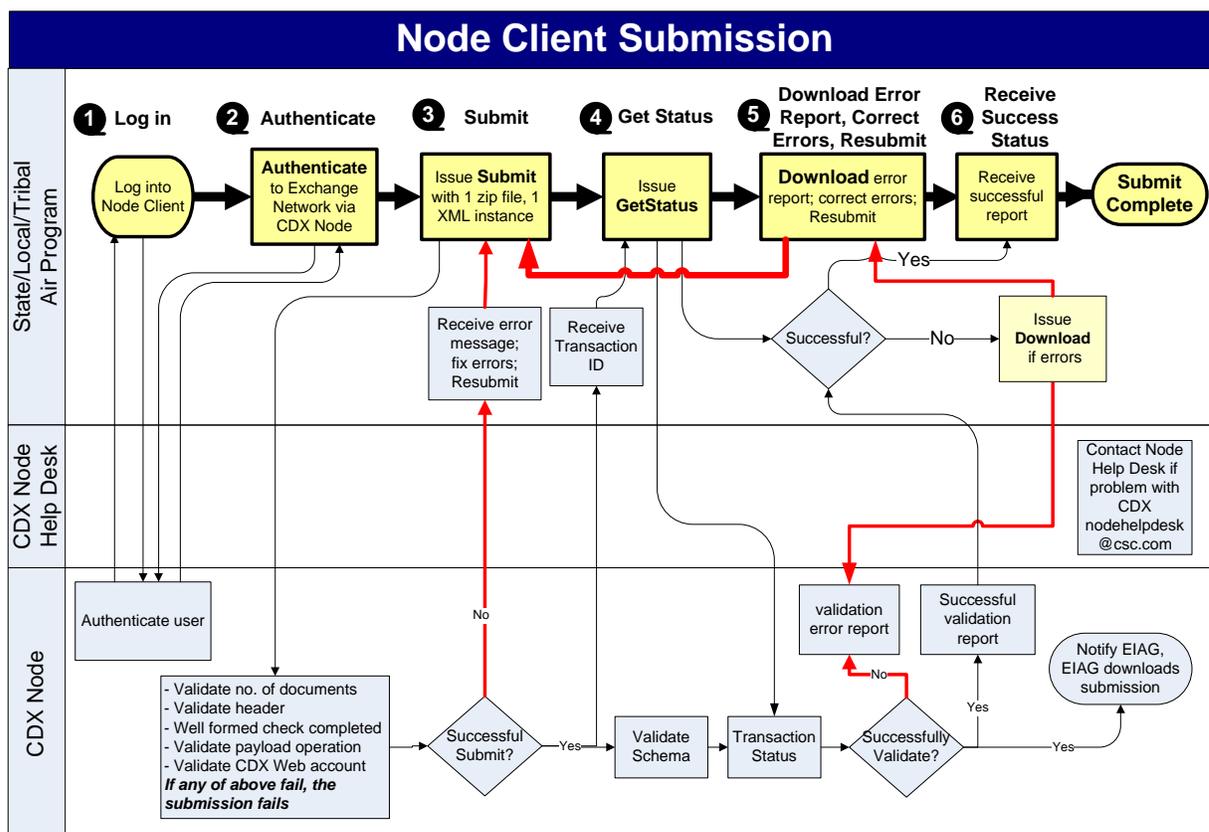
CDX Node Client (includes velocity mapper): [https://test.epacdxnode.net/cdx\\_client/](https://test.epacdxnode.net/cdx_client/)

Network Desktop: <http://clientcentral.epacdxnode.net/msi/setup.msi>

For the EPA distributed Node Client/Network Desktop, you can contact the Node help Desk for assistance if you encounter problems in configuring or setting it up.

Once you have completed these steps your account is setup and your Node Client/Network Desktop is downloaded, installed and configured, *and you have generated your XML instance documents for submission*, you are ready to submit your emissions inventory data.

**Figure 5. Submitting using a Network Desktop/Node Client**



1. **Log in** by providing your NAAS ID and credential (password)
2. **Authenticate** to the EN by invoking authenticate
3. **Submit** by identifying the flow (NEI), selecting file format and a file, and invoking submit. Remember that the current system only allows 1 zip file that contains only one XML instance document. Typically these are submitted by emission source type (point, area, on-road, non-road, biogenic). If there is more than one xml instance in the zip the submission will fail. Your submission will be validated for the following:

- Number and type of documents in zip
- Header elements
- Well formedness check completed when the header elements are validated

- NEI schema and a limited set of business rules
- CDX Web account

If any of the above fail, the submission will fail. The majority of the errors will occur when validating against the schema and business rules.

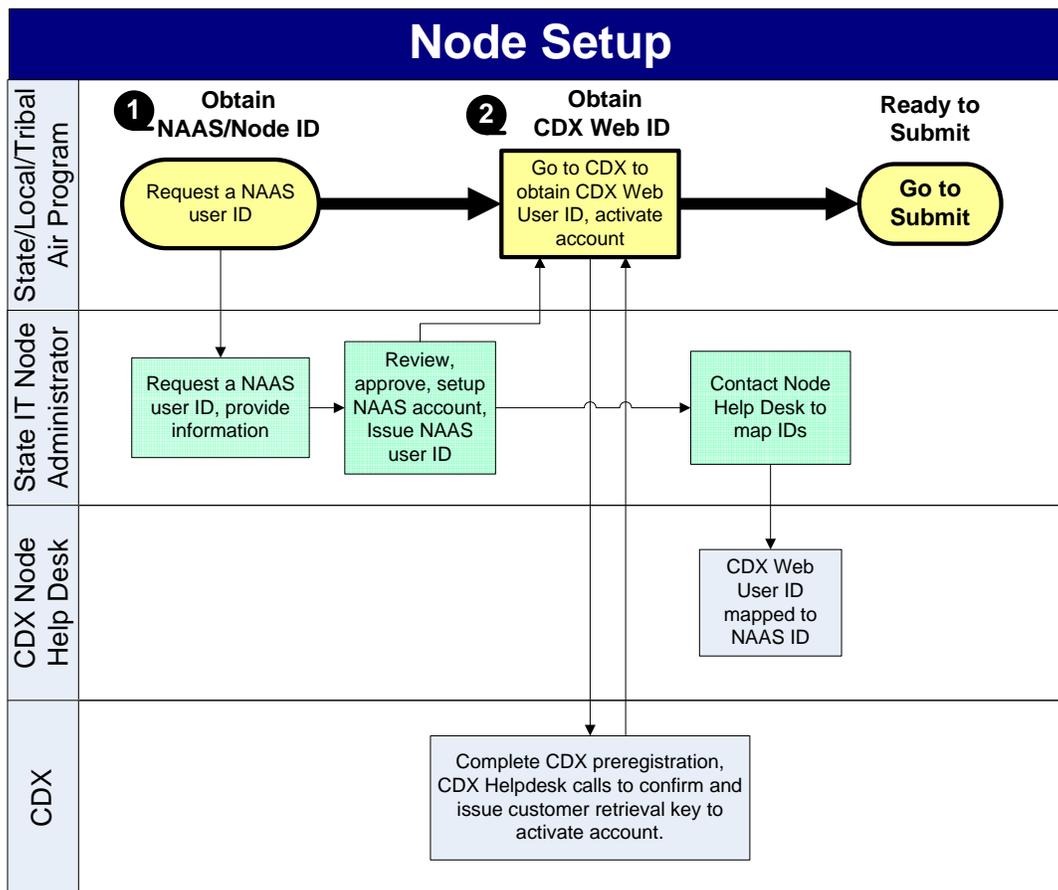
**QA Test Tool.** You can first validate your XML instance document against them prior to submission, you can go the QA Test site at <http://tools.epacdxnode.net> as mentioned earlier.

4. **Get Status** can be invoked to see determine the status of your submission (pending, completed, failed).
5. **Download Error Report Correct Errors, Resubmit** if you have determined that your submission failed to pass validation. The error report is provided in XML instance and reports the first 100 lines of errors. If the numbers of errors exceed 100 lines, correcting the first set of errors will correct repeating fields that occur throughout your XML submission. *If you encounter a large number of errors, it is recommended that you access the QA Test Tool to resolve your errors before submitting again.* Once your errors have been corrected, resubmit for final validation of the file, header and well-formedness of the XML instance document.
6. **Success status** should be received up successful completion of validation and processing through CDX. At this point, your submission is complete.

### **Account Setup and Submissions Using Full Node**

Figure 6 provides an overview of the steps for air program staff who are working with their IT staff to make emission inventory submissions using a full network node to set up accounts.

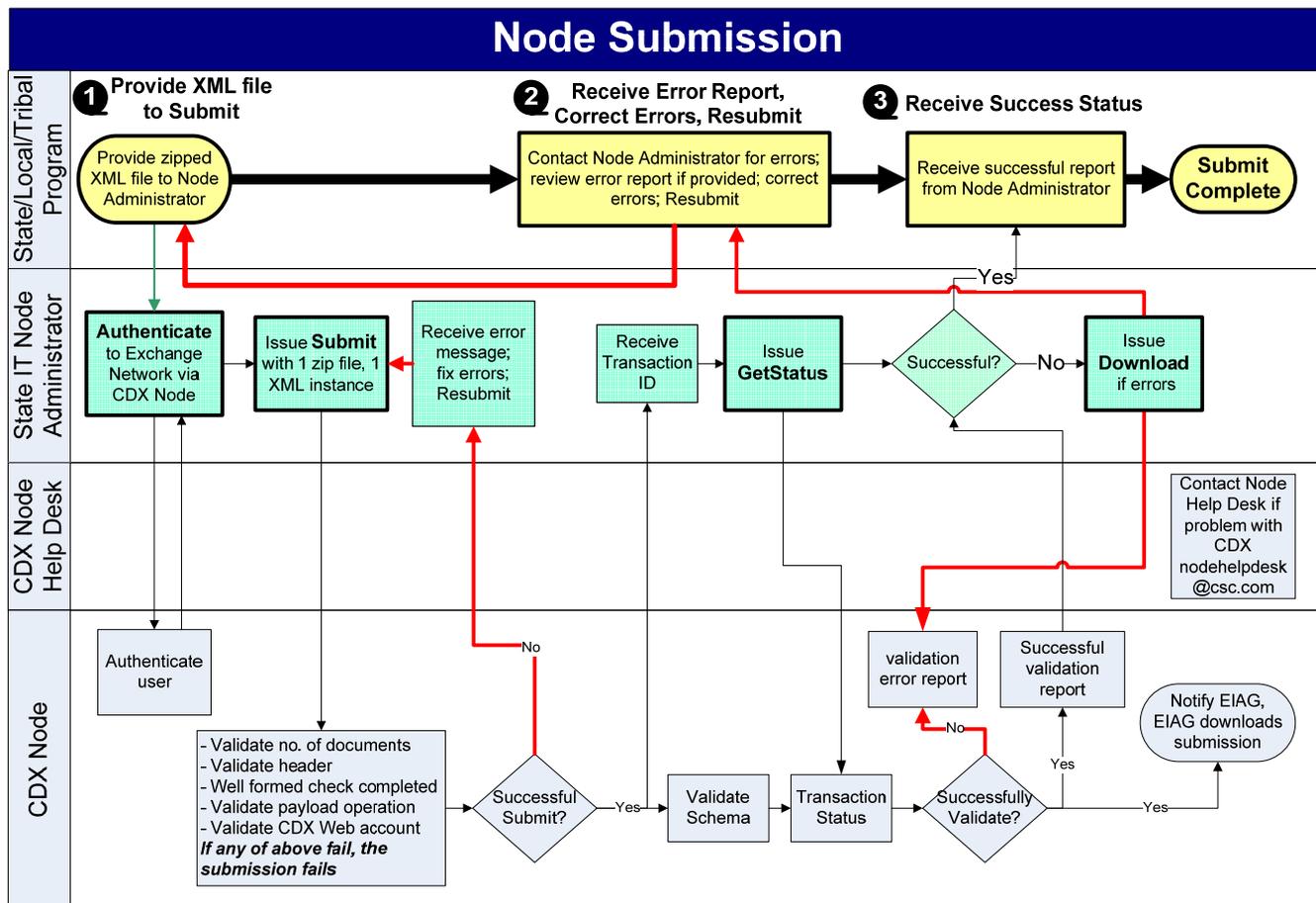
Figure 6. Setting Up an Account for a Full Node



The setup for the Full Node is the same as for the Node Client/Network Desktop *except* the SLT IT Node administrator (which may be the same as the air program staff for some SLTs) will provide your information for setting up an account and mapping your CDX Web ID. Once you have completed these steps your account is setup, *and you have generated your XML instance documents for submission*, you are ready to submit your emissions inventory data via your agency’s network node.

Figure 7 provides an overview of the steps for air program staff who are working with their IT staff to make emission inventory submissions using a full network node.

Figure 7. Submitting Using a Full Node



Although the process looks more simplified than that for a Node Client/Network Desktop for the air program staff, it still requires similar basic steps:

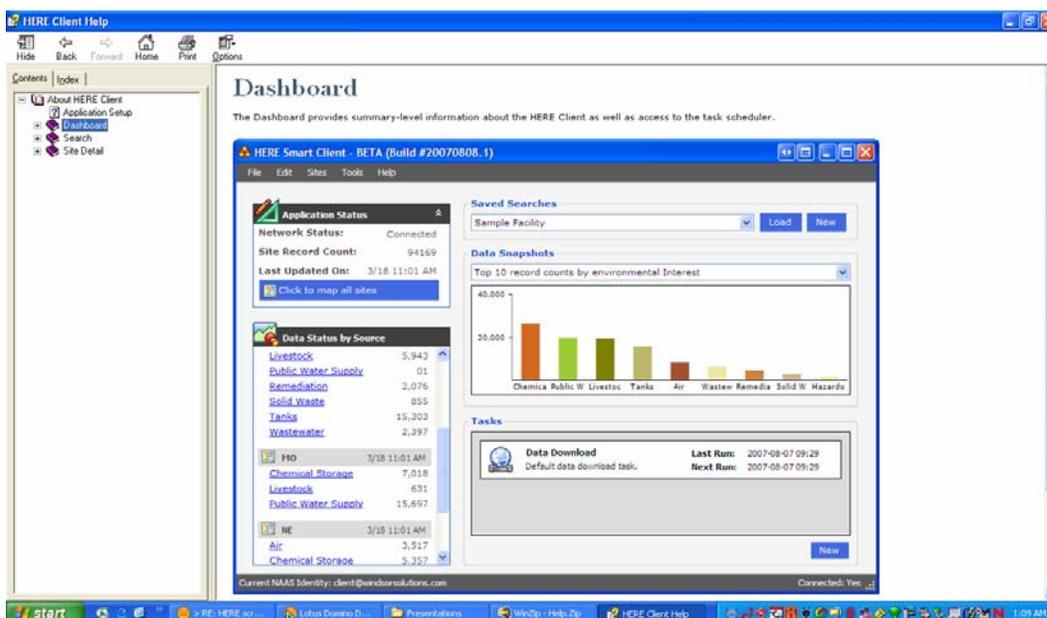
1. **Provide the XML File** to your Node Administrator for submission. The Node administrator and full node will perform the web service calls.
2. **Receive Error Report, Correct Errors, and resubmit** if you have determined that your submission failed to pass validation. **CONTACT YOUR NODE ADMINISTRATOR TO MAKE SURE YOU HAVE NO ERRORS.** Do not assume that once you provide the XML submission file to the Node Administrator that you are done with your submission. *It is recommended that air program staff establish a process for communicating when errors are received for how, who, and when those errors will be corrected.*
3. **Success Status** should be received up successful completion of validation and processing through CDX. The Node administrator should also contact you to confirm successful completion. At this point, your submission is complete.

## USE OF CDX/EN SERVICES TO PUBLISH AND USE DATA

An example of an exchange that leverages the full capabilities of the Network to publish and use data is shown in the Homeland Emergency Response Exchange (HERE)<sup>9, 10</sup>. Four states in EPA Region VII have created the HERE to share available environmental, health, and natural resource information with state agencies involved in emergency response planning and implementation. The exchange allows partners to retrieve data stored by each state (e.g., air, public water supply, livestock, remediation, solid waste, tanks, wastewater, chemical storage). Users retrieve data through a HERE client that aggregates data from multiple Exchange Network Nodes into a single merged view used on the “Dashboard” for generating queries, spatial views and reports. Local copies can be stored when an internet connection is not available. Queries can be performed and displayed on a map using Google Earth and NASA’s WorldWind. The data can also be exported. Figure 8 shows a dashboard screen that has retrieved data from multiple sources and displays the top ten records for environmental interest for a particular facility.

Using this powerful tool, emergency planners can quickly identify potential dangers posed by chemical inventories or hazardous waste storage. In minutes, response teams can assess threats to drinking water infrastructure or other environmental interests. In short, HERE provides access to better information that can offer opportunities for more comprehensive planning and more informed responses. Region V is also developing a HERE.

**Figure 8 HERE Dashboard**



## SUMMARY/CONCLUSIONS

Exchange Network and CDX services are being used by 50 States and 12 tribes for environmental data exchanges. These services are being used for emissions inventory submissions to EPA, data publishing to EN trading partners and consumption of this data for critical exchanges (e.g., EPA to SLT; state-to-state, locality or tribe to state, locality-to-locality, tribal organization-to-tribal organization). Exchange Network and CDX Services will be used in the future in the modernized EISys for emissions reporting based on a consolidated XML schema.

You can submit emissions inventory data through the CDX node on the Exchange Network using:

- *Full Network Node* being used by your state
- *Network Desktop/ Node Client* if you are not working with a state agency

If you are using a full node, the program and IT staff (which in some cases may be the same person) need to establish lines of communication and roles when it comes to the submitting and correcting submission errors. ***A node client is recommended for this type of annual submission where the program staff can directly submit and correct errors without the assistance of the IT Node Administrator.***

Although EN grants may not be available to Local Agencies, they can still use free node clients for submitting data. Free node software (node client and full node) is available for download and installation. Help is available from the Node help Desk if you have question or need assistance downloading and configuring Full Nodes and Network Desktop/Node Clients.

## REFERENCES

<sup>1</sup>Husk, M. “Building a Consolidated Emissions Reporting Schema”, Presented at the, Exchange Network National Meeting, Nashville, TN, April 30 2008.

<sup>2</sup>EPA Grants and Funding: Exchange Network Grants Website  
<http://www.epa.gov/exchangenetwork/grants/index.html>

<sup>3</sup>Exchange Network Website. <http://www.exchangenetwork.net/basics/index.htm>

<sup>4</sup>Exchange Network Website for data standards.  
<http://www.exchangenetwork.net/standards/index.htm>

<sup>5</sup>*Network Node Functional Specifications, Version 1.1*, September 17, 2003

<sup>6</sup>*Network Node Functional Specifications, Version 2.0*, June 2008

<sup>7</sup>Exchange Network Website for XML registry.  
[http://iaspub.epa.gov/emg/xmlsearch\\$.startup](http://iaspub.epa.gov/emg/xmlsearch$.startup)

<sup>8</sup>EPA CDX Website. <http://epa.gov/cdx/>

<sup>9</sup>Exchange Network Website for HERE  
<http://www.exchangenetwork.net/exchanges/cross/here.htm>

<sup>10</sup> HERE Website <http://herenetwork.org/>

## KEY WORDS

Central Data Exchange

Extensible Markup Language

Exchange Network

Network Desktop

Node

Web Services