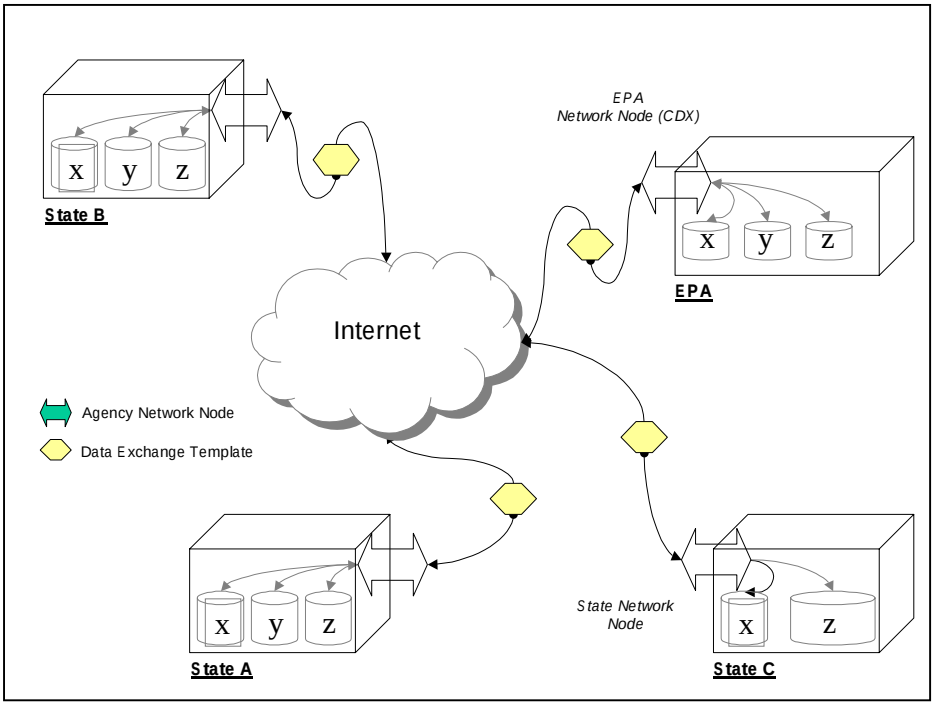
Report to the State/EPA Information Management Workgroup

Blueprint for a National Environmental Information Exchange Network



Prepared by the Network Blueprint Team
October 30, 2000

Forward

The design of the Network is based on the observations and findings of the State/EPA Information Management Workgroup (IMWG), as outlined in *Shared Expectations of the State/EPA Information Management Workgroup for a National Environmental Information Exchange Network (the Network),* June 2000 working version. (See Appendix B)

In July 2000, the IMWG formed a team charged with developing a Blueprint that would serve as the conceptual design of the Network. The intended audience for this document includes Chief Information Officers/Chief Technology Officers at state environmental agencies and their associated counterparts at EPA. The Blueprint team was asked to: 1) Describe the Network's components 2) Test and refine the Network vision itself; 3) Identify and assess the technical issues and options; 4) Identify critical policy and political issues; 5) Describe the specific, visible benefits we expect the network to produce as it is implemented; and 6) Finish as quickly as possible. This report to the State/EPA Information Management Workgroup documents the Blueprint team's analysis and recommendations with regard to these issues.

The Workgroup received, deliberated on and formally endorsed this report at its October 2000 meeting. The Workgroup authorized release of this final version of the report for wider distribution and further communication. In addition, the Workgroup charged the team to report back within six weeks of the final report with specific recommendations on:

1. Sourcing and/or establishment of the Network Administration function, including: consideration of the use of third parties, funding of this function, and its relationship to the Workgroup and the Data Standards Council.
2. Specific roles and responsibilities of the function of Network administrator and the means by which those roles and responsibilities could be fulfilled.
3. An implementation approach for these recommendations.

For answers to some Frequently Asked Questions about the Network, please refer to Appendix A: Network FAQs. A list of acronyms used in this document and a glossary of terms are included as Appendix C and D.

Acknowledgements

This document has been developed through the support and analytic contributions of a number of individuals and programs within EPA, ECOS, and several state agencies. These individuals offered valuable insight, lessons learned from their experience in various agencies and organizations, and hard work in creating the blueprint.

State Participants

David Blocher, Maine Department of Environmental Protection
Dennis Burling, Nebraska Department of Environmental Quality
Troy W. Delung, Virginia Department of Environmental Quality
Ken Elliott, Utah Department of Environmental Quality
Irene Kropp, New Jersey Department of Environmental Protection
Renee Martinez, New Mexico Environment Department
Mark McDermid, Wisconsin Department of Natural Resources
Melanie Morris, Mississippi Department of Environmental Quality
Kimberly Nelson, Pennsylvania Department of Environmental Protection
Lynn Singleton, Washington State Department of Ecology
Ron Tuminski, New Jersey Department of Environmental Protection
Cathy Wagenfer, Maryland Department of the Environment
Mitch West, Oregon Department of Environmental Quality
Bob Zimmerman, Delaware Department of Natural Resources

EPA Participants

Mark Badalamente, Office of Environmental Information
Andy Battin, Office of Water
Chris Clark, Office of Environmental Information
Connie Dwyer, Office of Environmental Information
Steve Goranson, Region 5/Office of Information Services
Lisa Jenkins, Office of Environmental Information
Chuck Spooner, Office of Environmental Information
John Sullivan, Office of Environmental Information

Environmental Council of States

Mary Blakeslee

Support Contractors

Matt Nielson, Concurrent Technologies Corporation
Andrea Reisser, Concurrent Technologies Corporation
Angela Jones, Ross & Associates Environmental Consulting, Ltd.
Lewis McCulloch, Ross & Associates Environmental Consulting, Ltd.
Louis Sweeny, Ross & Associates Environmental Consulting, Ltd.

Table of Contents

1. Executive Summary

A. Introduction

Information is fundamental to the work of environmental protection. State environmental agencies and U.S. EPA depend upon the rational flow of quality information for every aspect of their work, as individual agencies and collectively. Yet, many of the current systems and approaches to information exchange are ineffective and burdensome. This Network Blueprint describes a practical vision for an alternative to the current approach. It outlines a National Environmental Information Exchange Network (Network) that applies the technologies and approaches that have transformed the Internet to the exchange of data between environmental agencies. The specific technologies, and their application, are detailed in this blueprint document. The core of the Network, however, is not technology: it is a commitment to change the way data is exchanged.

The Network will depend on the ability of environmental agencies to negotiate and then define the exact format in which data will be exchanged (data exchange template), to document the agreement in a trading partner agreement (TPA) and to hold parties responsible for fulfilling these agreements. Responsibility for data quality, timeliness, format and availability will be explicitly defined, documented and agreed to by a designated individual for each party. Data originators will fulfill these agreements by maintaining information sources (nodes) on the Network that can provide this information upon authorized request. Once established, these data exchanges will replace (and be superior to) the traditional approach to information exchange that relied upon states "feeding" information directly to EPA's national data systems. Those agencies that choose to utilize the Network would do so in place of their traditional “feed the system” uses of national systems at EPA.

B. Background

The analysis and discussion reflected in this blueprint involved a team of more than 40 State and EPA staff, as well as associated contractors and technical experts Given the complexity and diversity of existing flows, this transition will be gradual, but accelerating. New and old approaches will necessarily exist side by side for many years. Guidance for managing these transitions will emerge only through actual experience. The recommendations at the end of this document constitute a proposal from the Network Blueprint Team to the IMWG to begin this joint effort now.

C. Challenges and Opportunities

A joint commitment to implement this Network clearly carries challenges and risks: these are described in the document. Inaction also carries risks. Regardless of this Network, states, EPA and other potential partners are making, and will continue to make, investments in new systems designed to fit their business needs. In most cases, this will mean that EPA national systems will

no longer be primary operational systems for states (and others). Without a compelling and credible organizing framework for how to share information in this new world, the quality and reliability of those collective efforts will be at risk and a unique opportunity for joint progress will have been missed.

Within this Network Blueprint document, the key remaining issues to be resolved cluster around the administration of the Network itself and the logistics of converting historical system-specific flows to Network flows.

D. Network Design

The Network is based upon four basic principles. These principles were developed in the Shared Expectations document and have remained intact. These principles are:

- Stewardship of specific data will be established by mutual agreement between two or more trading partners.
- Stewards, through their node, are directly responsible for the quality and availability of this data.
- Network members, whose use of stewarded data necessitates the maintenance of local copies, are responsible and accountable for ensuring the integrity and currency of those copies.
- Network members agree to use the Network technology standards, as described (and refined) in this blueprint and as documented in their individual trading partner agreements. These principles are implemented through five components: 1) Data Standards, 2) Data Exchange Templates, 3) Trading Partner Agreements, 4) Technical Infrastructure and Network Administration and 5) Member Organizational Infrastructure.

E. Recommendations

The Blueprint Team ultimately envisions a broad and diverse membership, including local, state, Federal and tribal agencies. The Blueprint Team also envisions the Network beginning with states and EPA and expanding as fast as experience and the interest of others allow. This Network is expected to dramatically improve the quality and availability of environmental data to environmental agencies and the public. The Blueprint Team recommends that the IMWG formally and fully endorse this Blueprint. Further, the IMWG should charge the Blueprint Team with developing and forwarding a specific proposal for how the network administrative function, including financing options, should be established. Finally, the IMWG should identify its next steps in advancing the Network, including a plan for outreach, and recognize that these steps should begin immediately.

Note: this blueprint report was endorsed by the IMWG at its October 2000 meeting. Per this recommendation, the Workgroup has charged the Team with developing a specific recommendation for how the network administration function should be established. The team will deliver this final product in December 2000.

2. Introduction

Information is fundamental to the work of environmental protection. State environmental agencies and U.S. EPA depend upon the rational flow of quality information for every aspect of their work. Yet, many of their current systems and approaches to information exchange work ineffectively and are overly burdensome, with obsolete and expensive computer systems that satisfy neither staff nor external users (e.g., the public, regulated industries). At the same time, two significant trends exacerbate the need for a new approach to environmental information systems. First, environmental protection agencies collect, access and utilize increasingly more environmental data, as the scale and complexity of the problems addressed has grown. Second, a widening system of environmental information exchanges has already evolved with the devolution of management from the federal to the state and local levels.

In response to these trends, and to the growing expectation that this information and government services themselves be available online, EPA, states and others are making major new investments in information systems. The pace and intensity of these changes have brought the problems with the traditional system-to-system approach into clear view. As states and EPA make these new investment decisions, they have asked for a framework that can coordinate their efforts and build on a common vision. This Blueprint is intended to provide this framework. Specifically, state environmental agencies and the U.S. EPA have struggled with modernizing systems at different paces, making it difficult to maintain the traditional direct system-to-system exchanges.

The rapid growth of the Internet and electronic-commerce (e-commerce) now provides a solution–an Internet-based voluntary National Environmental Information Exchange Network (Network) for state, federal and tribal environmental agencies. A Network based on standardized Internet language allows individual agencies to invest in internal data storage systems of their choice at a pace they can afford, while also supporting easy exchange of environmental data. Although the drivers and capability to create such a Network are already in place, its development will require *deliberate and collaborative* design and work. These areas are the focus of this document.

In overview, the Network facilitates information exchanges between “nodes” maintained individually by participating partners (initially envisioned as state environmental agencies and EPA). These nodes use the Internet to exchange information via standardized data exchange templates (DETs), using common (Internet-based) protocols. Exchange of data is governed by trading partner agreements (TPAs) between the partners. TPAs document the agreed upon data, exchange format, frequency of exchange and related issues. For example, a state and its EPA Region negotiate a Performance Partnership Agreement (PPA) that includes a TPA for the exchange of permitting, enforcement and compliance data for the National Pollution Discharge Elimination System (NPDES) program. This TPA explicitly defines the quality, timeliness and format of the data, binding the state and EPA Region in a "data-centered" agreement.

Held together by such agreements, the Network will bring clear and measurable benefits:

- A common approach to environmental information exchange that is manageable by an agency as an agency, and not a collection of stovepiped systems, loyalties and approaches.
- A transition from traditional information exchange approaches, which are rife with management and data quality problems, to a data-centric approach focused on data and data quality.
- Enhanced potential for data integration.
- Lower cost to exchange data.
- More agency control over its own data, especially in light of public and legislative trends driving all public data onto the Internet.

The approach and benefits envisioned for the Network have already been validated in the private sector, such as RosettaNet (see reference document on RosettaNet).

The Network approach also explicitly recognizes the ownership and responsibility of agencies for their data; and the responsibilities of participants who aggregate that data. By moving pro-actively to create this Network, participants can establish their nodes as the sources of record rather than have piecemeal or prescriptive approaches legislated or otherwise mandated. Although not a panacea for all existing problems, the Network allows more focus on interpretation of the data and, in turn, enables better environmental decision-making.

Initially, the scope of the Network will be limited to information that partners are already exchanging on a formal basis (e.g., states with EPA); vastly more agency data may be available on public access websites, state clearinghouses, and other informal arrangements than on this Network. As indicated above, flows of environmental information involve an ever-increasing number of governmental agencies (local and international). While this Blueprint focuses on state, EPA and tribal information flows as a starting point, it will expand to these participants as their interest and the capacity of the Network allow. The ultimate vision is a broad and diverse web of quality information, but the design begins small.

Figures 1 and 2 compare the more complex and disjointed process of data flow typical today with a more streamlined and efficient process that would occur on the Network. The most important aspect to note about these figures is the shift from the use of many transfer mechanisms between the states and EPA today to a much more standardized mechanism envisioned on the Network. Beyond improved data quality, consistency and coverage, this change will allow all Network participants to achieve economies of scale as they consolidate the function of information exchange and standardize the format of data to exchange.

Figure 1: Overview of the current information reporting relationship between states and EPA.

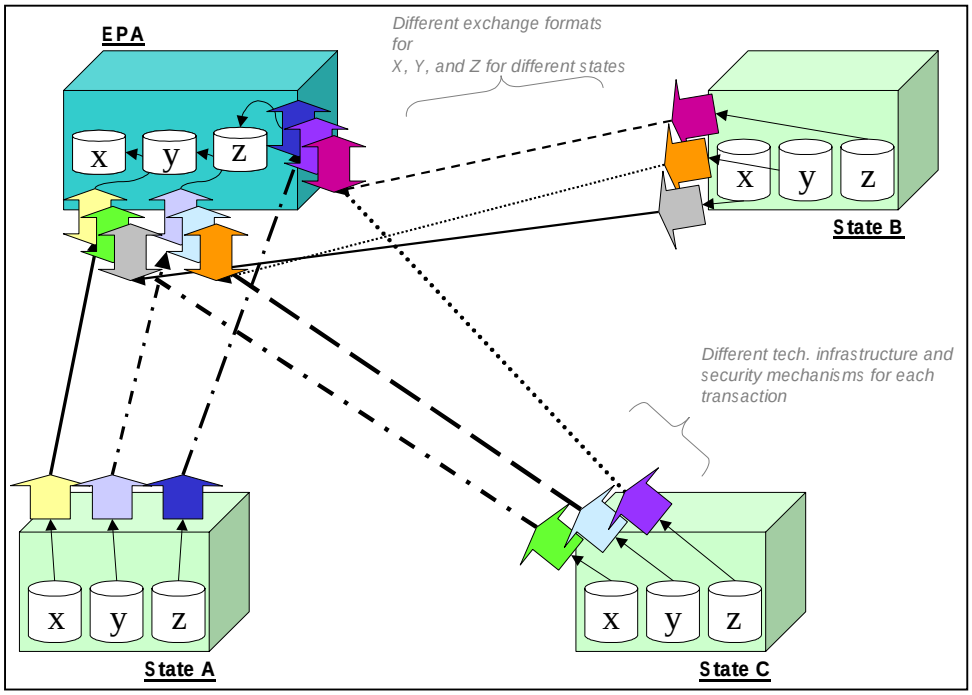
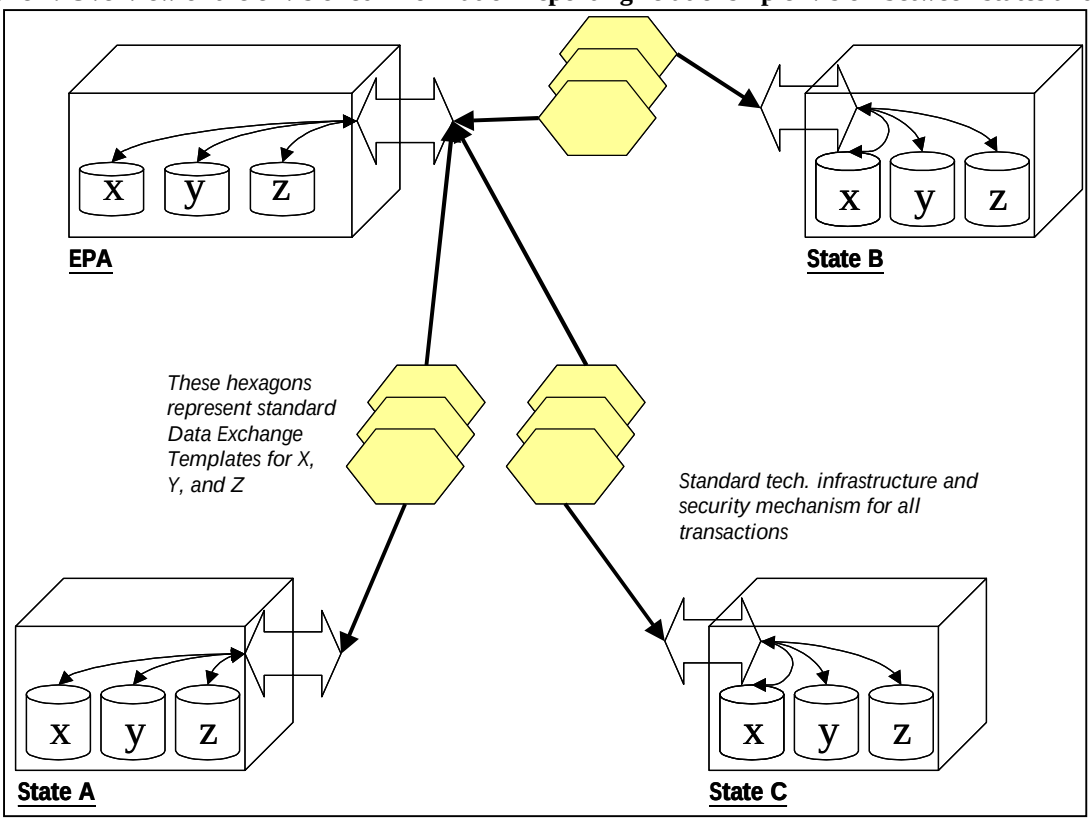


Figure 2: Overview of the envisioned information reporting relationship envision between states and EPA



The design of the Network is based on the following observations and findings of the State/EPA Information Management Workgroup (IMWG), as outlined in *Shared Expectations of the State/EPA Information Management Workgroup for a National Environmental Information Exchange Network*, June 2000 working version:

- Information, especially integrated information, is an increasingly important environmental management tool.
- Currently, this information is widely dispersed across state and EPA departments and locations and yet is increasingly demanded by a wide and diverse audience in an integrated fashion.
- Many states are investing in their own information systems and migrating away from use of EPA national systems.
- EPA faces the challenge of an increasing diversity of state and other data partner systems, ranging from those who have built integrated modern systems to those who continue to rely on EPA-sponsored systems.
- The current discussion concerning data among states and EPA is nowhere near as productive as it could be. The current collective approach leaves much to be desired in establishing clear accountability and responsibility for data quality, stewardship and management on all sides. Often these debates fail to even escape from disagreements over the definition of basic terms, or the currency or authority of given data sets or reports.
- There has been a revolutionary convergence of technologies around the Internet, World Wide Web (WWW) and e-commerce, especially the establishment of secured networks of standards-based information flows, which use the Internet as its infrastructure.
- Governments can apply these technologies to data they exchange with their partners, but governmental and inter-governmental coordination presents unique challenges to their use.
- A network blueprint is needed to allow shared and clearly defined terminology in addition to accountability and responsibility for elements such as data quality, timeliness and authority, exchange formats and methods, and access. This will allow each partner to operate independently on internal matters and in a coordinated fashion on external issues.

A. Why a Network Blueprint?

The Shared Expectations document raised both significant interest and questions among IMWG members and their staff. "How would this work?", "Who would do this?" "If we are taking this seriously, what must we start doing now?" In response, the IMWG commissioned an ad-hoc team of state and EPA staff to develop a conceptual Network design, the need for which became especially acute as the IMWG itself, EPA and individual states began incorporating these concepts into their own investment and management decisions.

Ultimately, the Network will be whatever those who build and use it create. The pace of its evolution will be uneven among users.

The blueprint is designed to support two essential next steps, without which the Network will not evolve (at least not from this effort):

1) A vigorous dialogue on the merits of and approaches to growing such a Network among states and EPA and tribes (to start).
2) Immediate support for those who will start building the Network. These efforts will start small, beginning with single data flows between two parties.

Within the context of the IMWG, this blueprint is designed to support dialog and implementation at several levels:

- EPA, as it continues to refine its information strategy and near-term investments. If EPA accepts these Network concepts, their investments will form a core strategic principle of its information strategy.
- Individual states, as they accelerate investments in information interchange, portals and e-commerce.
- The State EPA Information Management Workgroup, which seeks to coordinate state and EPA efforts.

The level of detail in this document varies widely from section to section, providing only enough detail to establish the plausibility and desirability of the Network parts. Substantial revision is expected before the design can be considered complete. Furthermore, the programmatically challenging aspects of the Network (e.g., the details of trading partner agreements) will require on-the-ground experience before refinement is possible.

B. Overview of the Organization of the National Blueprint Document

At its simplest, any network is made up of nodes and relationships (data flows and agreements) between those nodes. All the elements of the network–its infrastructure, policies and technologies–can be related back to these two fundamental parts. The Network blueprint is organized as follows:

Section 3 provides a very high level overview of the Network concept and its parts.

Section 4 defines a Network node and describes its operation.

Section 5 describes stewardship of the data and the Network.

Sections 6-10 describe the components of a network flow.

- 6: Data Standards
- 7: Data Exchange Templates
- 8: Trading Partner Agreements
- 9: Technical Infrastructure
- 10: Organizational Infrastructure.

Each section follows a common organization:

- A. Background: Basic context for the component.
- B. Definition: A brief definition.
- C. Business Case and Critical Features: The rationale for that component(i.e., why it is needed and what it does).
- D. Government Issues: Specific governmental issues raised by the component, especially where a private sector concept is being adapted to a government context.

Section 11 describes how the various network components relate to each other.

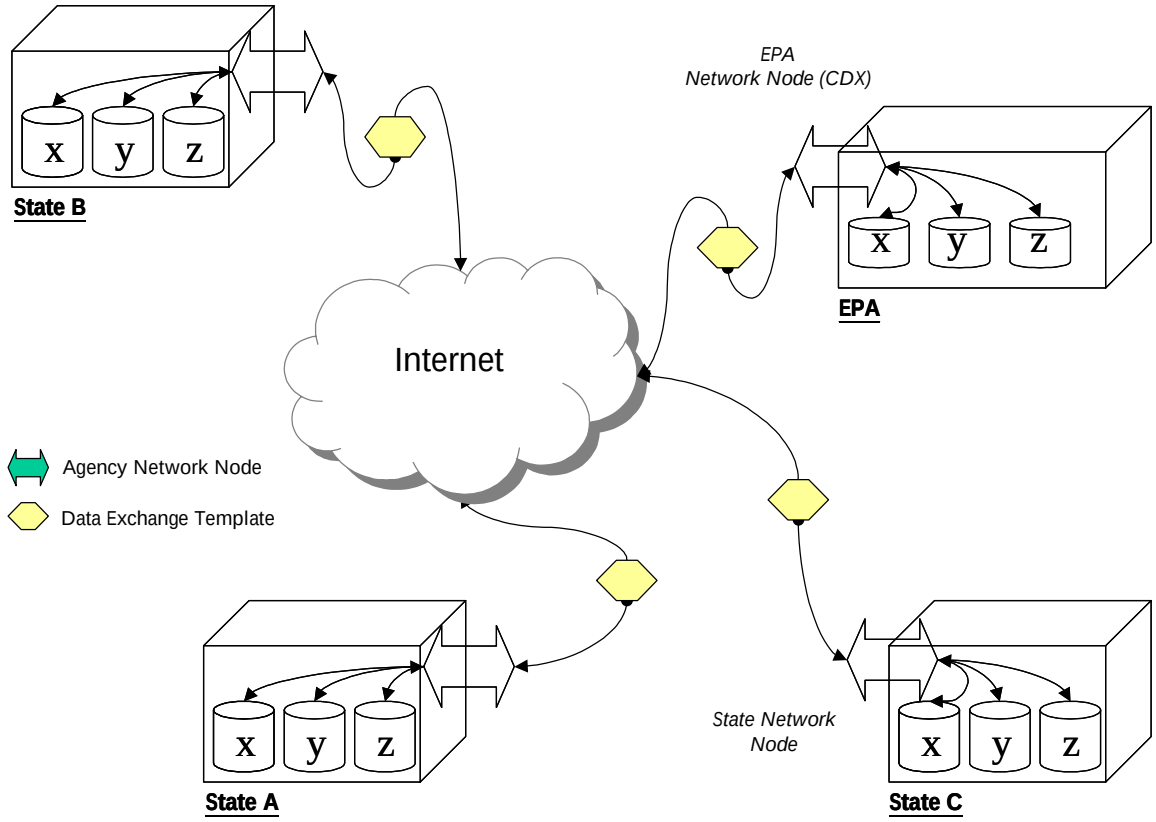
Section 12 presents the recommendations being forwarded in this blueprint.

Section 13 presents a Network example.

3. Overview of Network Design and Design Principles

In overview, the Network facilitates information exchanges between nodes maintained individually by participating partners initially envisioned as state environmental agencies and EPA. As shown in Figure 3, these nodes use the Internet to exchange information via standardized data exchange templates (DETs), using common (Internet-based) protocols. Exchange of this data is governed by trading partner agreements (TPAs), not shown, which document the agreed upon data, exchange format, frequency and related issues.

Figure 3: Conceptual diagram of the Exchange Network



The proposed Network design balances three sometimes-conflicting requirements:

- Desire for an open, dynamic, diverse network of environmental data flows, with an absolute minimum of constraints and overhead on participation.
- Reliable flows of data which is consistent nationally (and/or at other scales) that can be readily accessed and integrated.
- Capacity to fulfill the majority of participants' exchange obligations, whether regulatory, statutory, grant, or otherwise.

The Network proposed in this blueprint involves design requirements and compromises similar to those of the Worldwide Web (WWW), which is diverse and easily joined. Yet underlying it (and mostly invisible) is a strict, technically rigid set of standards for the transmission (TCP/IP) and display (HTML) of information. There is flexibility in some places, but absolutely none in others. For example, the rules of the WWW preclude a non-conforming Internet address (e.g. 207.193.**green**.99.47), or a page that uses a proprietary variant of HTML. It just will not work. These technical design restrictions dramatically constrain what the Web can do, yet are wildly successful.

The proposed Network here manages the conflicts identified above by using the technical infrastructure of the Web to move standardized sets of information in agreed-upon DETs, and where necessary, to officially document the agreement to do so in a TPA. Some DETs will be created and used by only a small number (maybe just two) Network members. Other DETs will likely be used by all members using a network flow to satisfy their obligations to a single member (e.g., states to EPA reporting under a delegation agreement). For EPA's traditional reporting flows, these DETs would function as national standards, but they would be only one part of a diverse and constantly expanding set of standards. They would be superior to the current approach because they would be expressed in a uniform, unambiguous and self-validating formats, rather than through a process of "feeding" a legacy system.

4. What is a Network Node?

A. Defining a Network Node

A Network node is a participant's single, managed point of interaction between trading partners on the Network. The node is the collection of specific technical and policy components that a participating member will manage for providing and receiving information via the Network. Nodes have the following critical features:

- Each Network member has only one node, although that node may handle many kinds and types of data.
- A member's node is the only route for Network delivery and receipt of information.
- The node is the single place for each member to present its standard node catalog of available information and associated network metadata (e.g., their TPAs, description of the information). Data and associated information must be presented on a node to be on the Network.
- The node is the single place where each member implements the minimal but essential transport, security and query protocols described in the blueprint and specified in the TPA.
- The node is the only place where a member's compliance with a TPA can be demonstrated or evaluated.

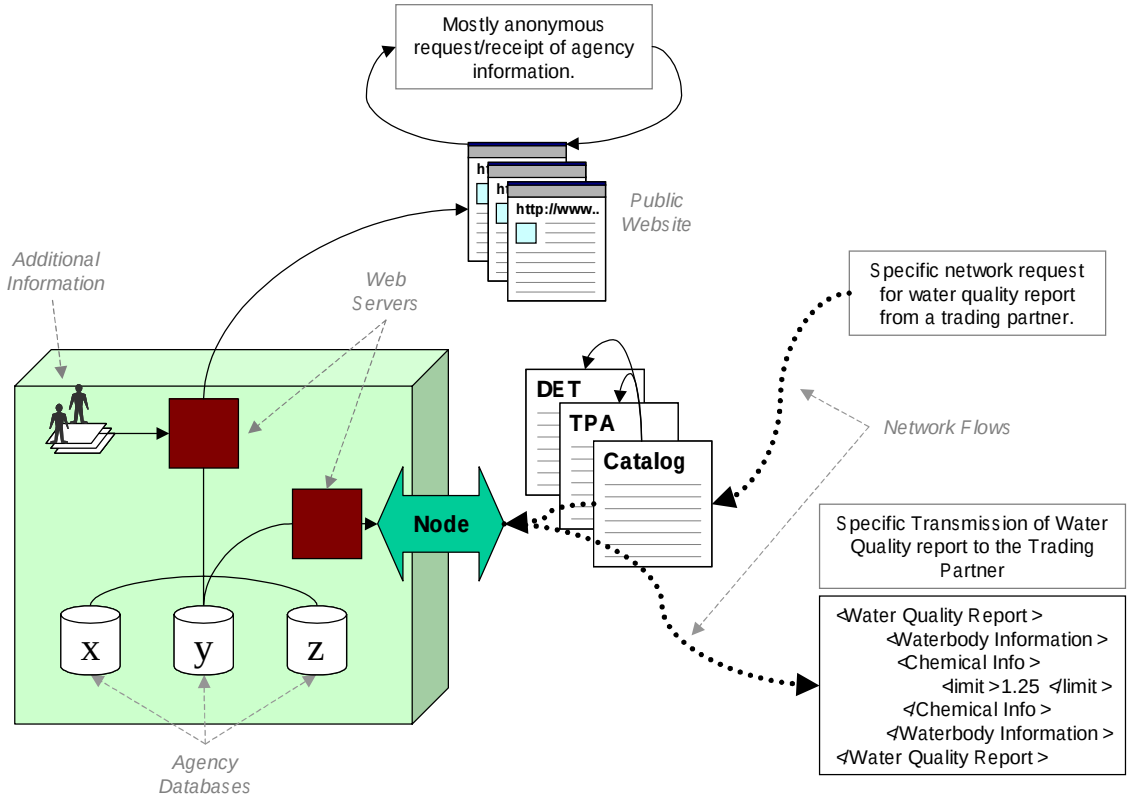
Members may choose to link their nodes with their public access websites, but each performs different functions and Network members will be required to ensure that adequate security is in place to separate the functions. Placing quality information on an attractive, well-designed public access website is a good thing – most agencies are doing this – but a website is not a node. A node presents this information, expressed in an extensible markup language (XML), using a standard DET and governed by a TPA. Figure 4 illustrates the functional differences between an agency's standard website and specific network nodes. Unlike a standard website, flow through the node involves a specific request from a particular trading partner (not anonymous) for information listed or referenced on the node catalog, governed in a TPA, and presented in the correct format specified in a DET.

What Is a Node "Really"?

A node is the central management point for each agency's interaction with the Network. All current flows take a program office-specific, system-specific, state-Region-specific path. This flow is difficult to manage, and the Network concept assumes the following simplification:

- All Network data (e.g., submittal of a quarterly report) flows from the originator's node to its trading partner's node.
- These flows are governed by a TPA signed by a single authorized individual from each partner.

Figure 4: Conceptual diagram of the internal operation of a Network Node.



The transition to this new approach will not be an easy or quick. Existing flows of information and Network flows will co-exist for several years. Managing a state or EPA node will require a new attention to internal roles and responsibilities that have much history. These roles are discussed further in Section 10: Member Organizational Infrastructure. They are introduced here because they touch on nearly every other aspect of the Network.

Today these flows are not from "state" to "EPA" as trading partners, but rather from the state's program, through a separate state system (or double entered), into to a specific program office and system within EPA, with the involvement of a specific person in the EPA Regional Office. This entire flow may be covered by several overlapping PPA, SEA (State/EPA Agreement) and program delegation agreements that are signed by various state and EPA officials to feed program data into an EPA system. These arrangements often vary greatly from state to state, Region to Region and program to program. The locus of authority and responsibility for data quality in this system is unclear. The Network, specifically the TPA, will not attempt to ignore the complexity of current flows; instead, it will simply make explicit the operation and obligations of nodes for a specific flow.

For all Network flows, the point of accountability for performance will be shifted from the program-specific data flow (e.g. manual entry in to a Federal system) to the state node. This clarity is an essential feature of the Network. Trading partners will no longer be left arguing

about issues caused by ad-hoc movements (or manual double entry) of data into the other's system, but can instead focus on the quality and availability of data according as specified in the TPA. It is expected that these arguments will be displaced by discussions about what the data say is happening in the environment.

B. Node Operation, State Nodes and Central Data Exchange (EPA's Node)

Network members will build their nodes as an extension of their existing Web and enterprise architectures. As outlined above, the node has a set of relatively simple technical functions, but its key role is as a management point for data. This role is likely to require additional or new roles and relationships for EPA and state staff. EPA and many states have already begun investments in Web portals that draw information from their official enterprise production systems for public availability. This is very similar to the Network node except that the data would be bound under a TPA and formatted according to the DETs. In addition, each participant would be required to have a formal process for managing the flow of data from the production systems to the portal, since those flows would be official. What was once a person-based flow from one program office in the state to one system at EPA becomes an enterprise flow, both within the originating state and at EPA as data flows through EPA's Central Data Exchange (CDX). See the discussion of roles and responsibilities in Section 10 for more information.

In addition to servicing authorized requests for specific data, each node must be able to provide its catalog to authorized requestors. There are many approaches to formatting and providing the node catalog metadata. As in the case of the Trading Partner Agreement Markup Language (TPAmL) used as the basis for the TPA section of this document, many open source approaches (e.g., the Federal Geographic Data Committee (FGDC) "node" reference format) can be adapted easily to the Network. This "cover sheet" or "lobby" to a node would allow participants to determine what data was available and how, if they are authorized, to access it. At its absolute simplest, this catalog could simply be a single XML file that is always found at the root level of a network node's URL with a common agreed name.

How Will Nodes be Built and Operated?

The Network design is patterned after demonstrated approaches taken in the private and mixed sectors (e.g., healthcare). Some of the base technologies are young (e.g., XML), but as the design team's analysis and independent expert consultations suggest, there is enough experience to support their use. These technologies are described in more detail in Section 9. In overview, participants will build their nodes as an extension of their current enterprise systems. Because the Network will be based on open standards (i.e., not tied to a particular technology or vendor), like XML, participants will be able to build their nodes using a wide and rapidly developing choice of tools. All major software vendors have now embraced these technologies, and many new companies have introduced products that make this market highly competitive and diverse. Perhaps most important, participants are free to implement any tool and any internal architecture for their node–the standards of node function are based purely on performance.

5. Stewardship

The flow of quality data is fundamental to the Network. The concept of stewardship refers to the responsibility for this data quality on the Network. As discussed above, this document seeks to resolve the current ambiguity in many data flows by establishing DETs and TPAs. Effective stewardship is essential for the Network to be successful, which will be achieved as Network data becomes synonymous with "high quality" data. Members will take responsibility for the data they place on the Network and for their interactions with the Network itself.

The concept of stewardship is involved in all of the principles and components of the Network. This section emphasizes some of the most important of these forms of stewardship.

Data Stewardship

By agreeing to host and exchange data and information, each Network trading partner assumes and accepts certain data stewardship responsibilities:

- Assuring that responsibilities for data quality and integrity are clearly defined and understood inside the organization.
- Assuring that that data source, derivation, and accuracy meet specifications.
- Assuring that the data formats and units of measure meet specifications.
- Assuring that any other relevant data or meta-data meet the specification in the TPA.

Node Stewardship

Each partner, whether state, tribal or federal, will be the steward for its own node, making sure it functions properly and that the data available complies with agreed-upon terms.

Network Node

- Assuring that the hardware and software that create, manage, store and provide access to the data work properly.

Transmission/Transaction

- Assuring that the data transmitted and received are complete.
- Assuring that the data transmitted and received comply with agreed-upon formats and time schedules.
- Assuring that data have not been altered.
- Assuring that confidential and sensitive data have not been intercepted.

TPAs will ensure that data quality requirements are built into each data exchange, including quality, format standards, documentation standards, content, sources, accuracy and timeliness, error detection and correction methods, other conditions that affect acceptability of the data, and reconciliation of data quality concerns. The technical infrastructure component describes how the technology supports this stewardship.

Some participants who make data available on the Network will not be the original authors of the data. In these cases, the role would be custodial–to store data for the convenience of access and analysis, with no attempt to govern the data or improve its quality.

Stewardship of Registry Data

Registries, reliable and authoritative sources for commonly used data or code sets made available on the Network, will require shared stewardship across the relevant members. Because of these coordination needs, registries will present special stewardship challenges. One of the first registries to be established on the Network may be the regulated facility registry (FRS), maintained by EPA. Over time, EPA and perhaps other Network participants will expand existing registries and add new registries.

Stewardship of Data Not on a Members Node

The basic Network concept assumes that each trading partner can manage its own data and make this data accessible via its own node on the Network. This capacity will evolve incrementally from state and EPA investments. In some cases, member capacity to steward their data may mature before they have a node operational. For example, EPA's systems are used as the official systems of record for some states, including those with delegated programs. If EPA establishes the technical infrastructure for its node and is technically able to place this data on a "hosted" node for that state (for the state's, EPA's and other members benefit), that state might choose to execute its stewardship through that national system. In this case, states would take on data stewardship and Node stewardship would be shared.

6. Component 1: Data Standards

A. Background

Data standards support the efficient and accurate exchange of data and assist secondary users to understand, interpret and use data appropriately. Note that these standards will apply to the "data" itself and to the "metadata".

States, EPA and tribes recently established the National Environmental Data Standards Council to promote the identification, development and adoption of data standards. The Network will promote and acknowledge the use of all standards developed or endorsed by the Data Standards Council, where they are available. No other mechanism for creating or recognizing data standards is envisioned. The Environmental Data Standards Council has prioritized the standards that need to be developed and chartered workgroups (made up of additional state, EPA, and tribal members) to begin this work. Final standards will be posted on a website, available to all environmental agencies and trading partners. Most importantly, these standards will be used by participants to build DETs.

B. Definition

As defined by the Data Standards Council, data standards are "documented agreements on formats and definitions of common data." Data standards are established to bring better consistency and quality to the information that organizations maintain.

Data standards provide the definitions and formats of the individual data elements (or "words"). Data elements alone are usually meaningful only when placed in data groups (or "sentences"). For example, the data element "mailing address line 1" is grouped with several other data elements, such as city name, state and zip code, to create the data group "mailing address". Some data standards also provide information about the interrelationships of its data groups.

The traditional components of a data standard are defined below.

- Data element – one particular piece of data; for each data element the following information is traditionally provided.
  - Name (e.g., Mailing Address Line 1)
  - Format (e.g., string, integer, date)
  - Definition
- Data group – logical grouping of data elements (e.g., the "Individual" data group in the Facility Identification Data Standard is made up of the elements "First Name, Last Name, Middle Initial, and Title Text")
- Relationships – the relationships between data groups (e.g. the "Facility Site" data group in the Facility Identification Data Standard can be associated with one or more instances of the "Geographic Coordinates" data group.)

Figure 5, below, describes a simple data standard example (the State/EPA agreed-upon Date Standard), which only describes the definition and format for one data element. An example of a complex data standard (the proposed Facility Identification Data Standard), which describes a number of data groups and their relationships to each other, has been provided in Appendix E.

Figure 5: Simple Data Standard Example of the State/EPA Date Data Standard

FINAL DATE DATA STANDARD AS POSTED ON THE ENVIRONMENTAL DATA REGISTRY

DOCUMENT DETAIL

Title: Date Data Standard and Business Rules for Representation of Calendar
Date EPA Document Number: Not Available
Abstract: This data standard and business rules support the implementation and maintenance of the Agency standard for representation of calendar date. This standard provides for consistent numeric representation of calendar date to facilitate interchange of date data among Agency information systems.
Purpose: To layout a data standard and business rules for representation of calendar date.
Organizational Author: Alvin M. Pesachowitz
Version: 1.0
Document Date: 19990120 (YYYYMMDD)
Access Constraints:
Coverage:
Coverage Period:
Cataloging Source:
Create Date: 19990223 (YYYYMMDD)
Change Date: 19990616 (YYYYMMDD)
Program Component:
Expiration Date: (YYYYMMDD)

DATA ELEMENT INFORMATION

Registry Name: Date
Identifier: 5432
Version: 1
Definition: A particular day of a calendar year.
Example: 19961011

VALUE DOMAIN INFORMATION

Datatype: Date
Maximum Character: 8

C. Business Case and Critical Features

Implementation of commonly used data standards on the Network where appropriate will help improve data consistency and quality. Wherever possible, DETs will incorporate data standards to bring consistency to the information being shared. Standardization is especially important for information (like facility or location) likely to be integrated with other users' data. If successful, use on the Network of these cross-program standards in DETs may be one of the most significant contributions the Network itself makes in supporting the integration of what have historically been program specific flows.

D. Government Issues

The Data Standards Council cannot bind an agency to using a standard. Individual agencies will determine if, when and how they might use a standard developed under the auspices of the Data Standards Council.

The Data Standards Council will monitor and act as liaison to other parties creating relevant data standards. Some of the standards currently in use were developed by unrelated government agencies. For example, the standard industrial classification (SIC) codes originally developed by the Department of Commerce are widely utilized by many government agencies and are being updated by a group of federal agencies. Various standards are also being developed by industry groups, the American Chemical Society, American Biological Society, the U.S. Fish and Wildlife Service, and interagency groups such as the Federal Geographic Data Committee. Coordinated development of data standards through the Data Standards Council will prevent agencies from developing standards that already exist. State environmental agencies that have already developed data standards are encouraged to bring these to the attention of the Data Standards Council and appropriate workgroups to expedite their recognition and use in Network DETs.

Data standards will only prove useful if they are widely accepted and used by the trading partners on the Network. EPA, in approving the use of a DET in fulfillment of a delegation agreement, will likely only approve those DETs compliant with the relevant standards. In establishing DETs for trading partners (e.g., other state or local governments), states may apply similar requirements. While no formal mechanism for enforcing the use of data standards is envisioned, the Network (and participants) should promote and encourage the use of these standards whenever possible.

7. Component 2: Data Exchange Templates

A. Background

Data exchange between environmental regulators to date has been characterized by a series of negotiated agreements to use a specific file format or a specific computer program. The vision for Network exchanges relies on agreed-upon, open, neutral, standards-based data exchange templates for defining and describing the information that is exchanged and secure Internet transaction protocols for actually moving the information between trading partners. This foundation will allow for adaptability in the shared information, independence for the partners involved in the exchange, and resilience for the specific flow as new technologies emerge.

The IMWG recognized the many benefits associated with information accessibility, including elimination of the requirement for states to load data into national EPA systems (e.g., PCS, AIRS, RCRIS). Use of data exchange templates that are standards-based and technology-neutral will encourage broad Network participation by states, and preserve existing trading partners' internal mechanisms (database software and structure) for storing and managing their information.

Wide agreement is nonetheless necessary on what constitutes acceptable DETs. To understand the definition of DETs in the context of the Network, it is important to distinguish between DETs and transactions (templates containing data.)

B. Definitions

Figure 6 presents the hierarchy of components relevant to DETs. Each major component is described in the following sections. (Data elements and data groups are defined above, in Component 1: Data Standards.)

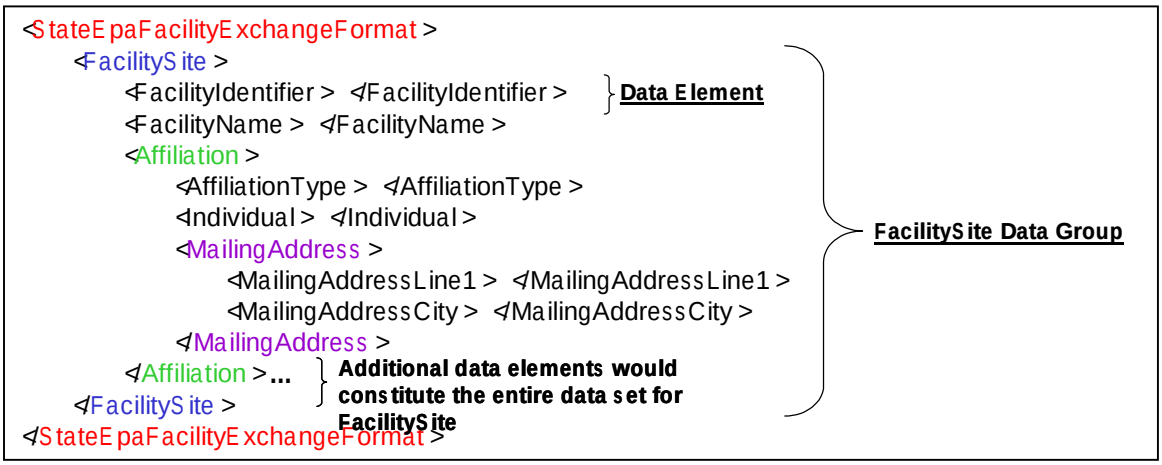
Data Exchange Templates

Data exchange templates identify types of information (data elements and data groups) required or allowable for a particular type of data set according to predefined standards. DETs are empty and contain no data. They simply define the format data must take prior to exchange. DETs will rely on existing data standards where appropriate to increase data quality and consistency. A complete template contains the data groups necessary to describe a specific business event (e.g., issue a permit, initiate an enforcement action.) Figure 7 presents a simplified example of a DET for regulated facility information expressed in extensible markup language (XML).

Figure 6: Data Exchange Template Definitions and Examples

Components:	Examples:
Data Element	First Name, Facility Identifier, Mailing Address Line 1
Data Group	Mailing Address, Facility Site, Affiliation, Enforcement Action
Data Exchange Template	List of enforcement actions taken (conforming to the facility identification and enforcement data standards as developed and adopted.)
Transaction	State's enforcement action records in the State/EPA Data Exchange Template format
Transmission	State's enforcement action records in the State/EPA Data Exchange Template format submitted to EPA on September 8th, 2000

Figure 7: Blank Data Exchange Template



Transactions

Transactions are defined as a specific set of data exchange templates containing data. Figure 8 represents a simplified example of a transaction containing a sample of state environmental

agency regulated facility information. Transactions may consist of multiple instances of a specific data exchange template, each containing data. Information flows over the Network when transactions are exchanged with a trading partner. Transactions will be converted from their electronic format to a human-readable or a different machine-readable format via no- or low-cost commercially available tools (i.e., a browser). Ancillary documents, such as maps, text documents, reference documents and images, may be carried in their native formats or referenced via URL Web links. Existing XML-based formats are available for all these types of data.

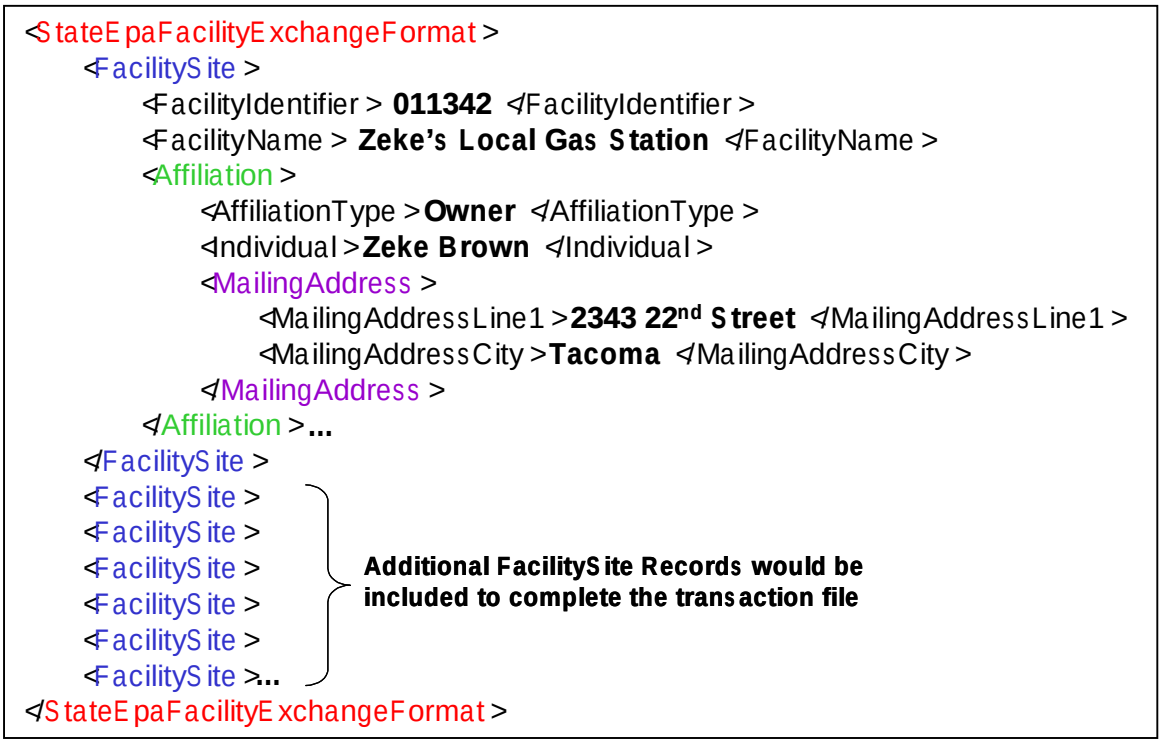


Figure 8: Example Transaction

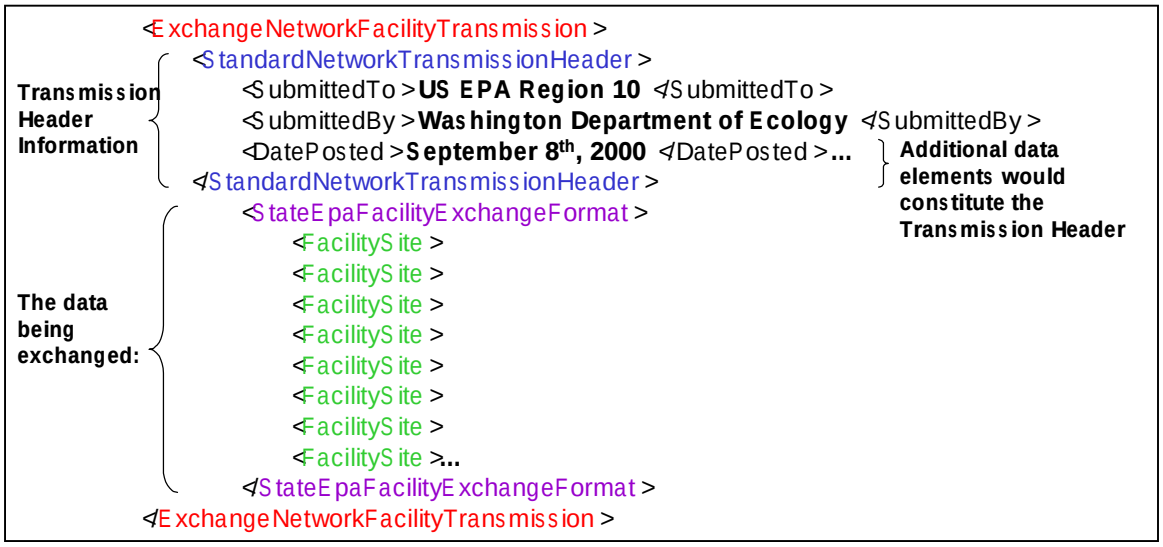
Transmissions

One or more transactions moved across the Network between trading partners constitute a transmission. Figure 9 represents a simplified example of a transmission of regulated facility information from a state agency to EPA.

C. Business Case and Critical Features

Data exchange templates define data available on the Network. It is assumed that the first series of DETs will support traditional state-to-EPA data flows for the major regulatory activities, such as hazardous waste management, air permitting and water quality monitoring.

Figure 9: An Example Transmission



The DETs define not only data groups and elements, but also their logical interrelationships. For example, an appropriate DET can make it clear that a facility has one or more permits; each of which has permit conditions. Such a mechanism allows efficient exchange of data without needing agreement on specific data structures.

Necessary Element of Network Exchange

The existence of an agreed-upon and published DET is a major element in distinguishing Network exchanges from other Internet publication of data. All Network exchanges must conform to a specific DET.

Maintenance of Data Exchange Templates

Designation and maintenance of a template registry is required. Data exchange requires that all parties have access to identical exchange templates. Effective data exchange requires that all trading partners have access to all the DETs being used on the Network. The management of this registry is identified as a core feature of the Network Administration function described in the next section. The choice to use XML as the sole DET language on the Network brings with it a host of tools for the management of these repositories, perhaps most significantly the capability to have the repository referenced in real time each time a DET is used. This means that much of the logistical management of "versions" of DETs becomes automated. When trading partners use a DET, they simply reference the repository and the template is served up. This feature also provides a powerful way to encourage and monitor the use of DETs and the use of standards in DETs.

Extension of the Business Process

Individual DETs mirror specific business events. At first, one or more specific templates are anticipated for each business process involved in data exchange. Over time, these templates may be merged into a smaller and better-integrated set.

Format for Data Exchange Templates

The design team considered a wide range of options for DET languages, including making no specific recommendation. After much deliberation, including the council of outside technical experts, the team elected to focus solely on XML as the language for DETs. XML is the best tool for trading partners to unambiguously express and then validate the data they wish to exchange. The team made this decision with the full knowledge that the technology is still immature, and that few existing XML flows exist now.

D. Government Issues

Data exchange templates can and will be established in a number of ways for use on the Network. For the existing state-to-EPA data flows, DETs will likely be developed by workgroups of state and EPA staff members familiar with those individual flows; a mechanism for joint development, adoption, and sharing of DETs may be desirable. As the Network grows (both in number of trading partners and in the amount of information available) DETs will be created as needed by the trading partners. This flexibility will allow the Network to evolve and meet the needs of a much wider set of trading partners.

8. Component 3: Trading Partner Agreements

A. Background

The electronic commerce most familiar to users of the Internet is the business-to-consumer variety. Typically, a consumer accesses a Web page and is guided by the rules embedded in the application - be it Amazon's shopping cart or some other mechanism. The user interacts live with the application, and may back out if an application imposes unacceptable conditions. For example, if specific personal financial information is needed to complete a transaction, the user may simply decline to submit it, canceling the transaction. Effectively, the website and the user impose conditions and reach agreement through completion or termination of actions.

Increasingly, business and government are seeing the value of electronic transactions that go a step further–electronic transactions initiated by a system owned by one party and negotiated with a system owned by another, without intervention of a user. For example, a business enters a purchase order into an automated system. That system contacts a specified set of vendors, places the order, negotiates details of payment, delivery and terms, and electronically executes the transaction. Prior to implementing such systems, businesses have ensured that such transactions protect the rights of all parties, and that the systems truly reach a common understanding of terms, conditions and other details. At the lowest technical level, considerable agreement is needed simply to begin negotiation of a transaction. In the business-to-business world of electronic commerce, the agreements needed to enable commerce are called trading partner agreements (TPAs). IBM Corporation has developed a standard for such agreements, available in the public domain (Sachs et. al., *Executable Trading-Partner Agreements in Electronic Commerce*). Several of the core elements that IBM includes in its standard and in the electronic language used to express the agreement (TPAml) for the private sector were adapted to the public sector for this component of the Network.

Similar unattended electronic exchanges of information will be needed for data exchanges over the Network. Much of the methodology emerging from the business-to-business e-commerce is directly applicable to any such transactions. Statutory oversight requirements, negotiated agreements between states and EPA, and mandatory reporting requirements introduce additional conditions unique to government or specific to the environmental information Network, which are discussed below.

B. Definition

Trading partner agreements are documents formally adopted by two or more partners for the purpose of defining the responsibilities of each party, the legal standing (if any) of the proposed exchange and the technical details necessary to initiate and conduct electronic information exchange. TPAs may apply to exchanges initiated by the sender ("push" systems) or those initiated at the request of the receiver ("pull" systems). TPAs are necessary when automated exchanges are to take place without operator intervention if the exchange is intended to meet or

replace any mandatory reporting requirement. They are advisable between any parties (e.g., states) who wish to establish an ongoing business process involving automated electronic exchange of information. Specific agreements regarding electronic data exchange between EPA and the states, as currently included in PPA and SEA documents, exist as the current implementation of TPAs. Future TPAs may take the same form, be drafted to complement a PPA and SEA or stand alone. TPAs do not apply to one party's access of data provided through a public access site. Such access may be negotiated when both parties agree they wish to exchange data.

C. Business Case and Critical Features

In practice, a TPA may be lengthy and highly technical, or relatively simple, based on the needs of the specific data flow and the existence of other governing documents. In the more simple case, a flow or exchange may have already been defined as part of a PPA, SEA, Memorandum of Understanding (MOU), or other agreement specifying timetables, data requirements, terms of governance, technical specification and details of flow mechanisms. The TPA for a data flow may then contain only the basic elements missing from the original agreement and reference to these other documents. Such robust technical or governmental frameworks may not exist in other types of flows and state systems. The TPA for a data flow would then need to be more comprehensive and detailed.

The following items should be addressed in the TPA, either directly or by reference to another document:

- **Parties**. This section identifies the organizations involved in the TPA and describes the general purpose of the agreement.

- **Legal Framework.** This section includes governance, standing and applicability issues that apply to the partners. The TPA should address the effect of the agreement on other interparty obligations. For example, it needs to address any reporting requirements met by the agreement. The TPA should also address applicability to all levels of participating organizations.[1]

- **Security.** This section identifies the level of Network security to be used and the specific parameters such as certificates used for authentication, non-repudiation and digital envelope, and other security issues.

- **Data Definition.** This section describes the specific format and structure to be used for exchange and the URL of record for the format.

- **Communication.** This section specifies the transport protocols and electronic addresses of the parties.

[1] If executed by a Region and a state, the relationship to EPA Headquarters requirements must be addressed.

- **Message Exchanges.** This section discusses rules for submitting and responding to requests for data and the timing of data exchange. It includes a list describing the requests that parties can issue to each other. These actions are the independent units of work. The action definitions reflect the associated message flows between the invoker and the service provider, responsiveness, failure handling and other attributes. This section should address the expected update cycle for data of record (e.g., the steward agency will enter data within five business days).

- **Definition of Roles and Responsibilities.** This section outlines specific roles and requirements of parties related to performance, reliability and use of data.

  - Internal Systems Requirements. The TPA does not address partners' internal computer systems unless the electronic exchange is predicated on maintenance of specific internal requirements (e.g., EPA's proposed electronic reporting rule). In such cases, they should be specified.

  - Performance and Reliability. The expected availability of participating systems is specified here.[2] For high-volume systems, the TPA should also identify system performance expectations (e.g., transfer speed, response times).

  - Exchange Failure. Because some exchanges may be mandatory (once voluntarily included in the TPA), the TPA should identify actions required by each party should the exchange fail.[3]

  - System Failure. When the exchange is intended to duplicate data locally, the TPA should address initial synchronization of participating databases and recovery following system failure.

  - Quality and Stewardship. The TPA should specify the definitive source for shared data. The TPA should outline expectations regarding timeliness of data entry, error detection and correction, and other conditions upon which acceptability of the data is predicated. [4]

[2] Situations are already arising where external data is included in public access products. Linking of significant portions of another's web products may be reason to execute a TPA indicating that there is some agreement to maintain specific content at a specific location.

[3] System error response procedures are a part of communications protocols. This item is intended to address business continuation in the event of failure.

[4] "On demand" data exchange introduces these factors. Periodic reports are expected to be complete for the period covered. Where these are replaced by ad hoc sharing of data, the trading partners need an understanding about the condition of the data on an ongoing basis.

- Use of Data. Intended routine uses of the data are specifically addressed to the extent needed in order to understand the responsibilities of the parties. Generally, the allowable uses of data need not be included in a TPA, as the data would be reported by some means in any case. Once delivered, the receiving party is still bound by such considerations as confidential business information (CBI) or enforcement-sensitive data, as if the data had been exchanged in the traditional manner. The TPA may need to address how such data, if mixed with other data, will be identified. If one party wishes to exclude a specific use that would otherwise be enabled by the exchange, it should be addressed. For example, in providing non-mandatory data, states have indicated in a PPA that EPA may not use the data for program evaluation.

- Dispute Resolution. The agreement describes procedures for settling disputes related to the terms of the agreement.

- **Parallel Paper Transactions.** Any expectations for exchange of documents on paper in addition to electronic format for a portion of or the entire duration of the TPA are outlined in this section.

- **Record Retention.** This section addresses issues surrounding transmission logs and requests for historical data.

- **Duration.** This section identifies the period of time for which the agreement will remain in effect.

- **Termination.** This section specifies conditions for termination of the TPA as a whole, including written notice and the effect of termination on other rights and obligations.

- **Addenda.** This section describes if and how addenda may be added to the agreement.

The TPAml Schema noted above, as well as other TPA templates, provide a structure and format for expressing many of these conditions as the Network begins. Other initiatives, such as ebXML (e-business XML), are basing their efforts on the TPAml work. It is likely that initial Network flows will employ a variety of TPA formats. As best practices emerge, they can be codified by the Network administrator into TPA templates. In addition, EPA or other major trading partners may establish templates as a starting point for TPA development.

D. Government Issues

A very important feature of many data exchange (especially e-commerce) networks is that they are bilateral (or peer-to-peer) and therefore self-enforcing. For example, the Internet itself, at any point in time, is simply the collection of computers that have agreed to route TCP/IP among

each other. When users sign onto their Internet service provider (ISP), or when that ISP links to its ISP upstream, they agree to use TCP/IP and abide by a basic user agreement. If users are not willing to use the TCP/IP standard, they cannot connect. If they violate their user agreements, their ISP will turn them off. E-commerce networks operate in a similar way. Sony and IBM execute a TPA and begin exchanging messages. If Sony sends the wrong part, or misrepresents a catalog entry, IBM deals with Sony; the e-commerce "administrator" (i.e., RosettaNet) does not become involved. Thus the larger network polices itself without the involvement of a central authority.

It is envisioned that the Network will be governed by bilateral TPAs and supported by a basic "Network User Agreement" agreed to by all partners when they join the Network. The Network User Agreement will define basic terms and conditions for participation in the Network. The agreement will be common to all Network data flows and will not need to be negotiated separately for each set of trading partners. Any special terms not included in the overall agreement will be included in separately negotiated trading partner agreements when determined necessary.

When a party attempts to provide data that either does not comply with the agreed-upon exchange format or does not meet some other term of the TPA, its partner is in a position to respond using its available authority. That data should not become part of the Network. If the data meets the requirements of agreement, it becomes part of the Network in good standing. By making the TPA explicit about data quality, the Network attempts to establish some baseline for the reliability and trustworthiness of its data. The use of XML provides data originators with significant ability to "self-validate" their own transmissions and recipients with the capability to assess the conformance to the DET.

Unlike engaging in commerce or running the Internet, the purpose of the Network is to support the flow of high quality environmental data. Not all of this data is, or will be, covered under a bilateral TPA. Once the Network is established, members may wish to make a form of quality declaration for given data on their node. For example a state may wish unilaterally to declare a given data source as its "official source of record for the state field burning program." Such a declaration would explicitly document the pledge of the participant to establish and maintain a specific data source as if there were a vigilant trading partner. This declaration would be similar in format and content to the standard bilateral template. It might even include a "complaints" section where data users could contest or otherwise comment on the data. Eventually, the Network administrator or others could fulfill some kind of "audit" function for these data sources, perhaps codified in a TPA, as service to members who choose to offer this type of information. (This external audit function is a feature of some e-commerce networks.) This function would be analogous to a Certified Public Accountant (CPA) auditing a firm's financial statement as accurate. As in the case of the bilateral TPA, the objective of this formal statement would be to provide a baseline of reliability and credibility to Network data.

9. Component 4: Technical Infrastructure and Network Administration

A. Background

The technical infrastructure of the data exchange Network will use the Internet in the same way as many private e-commerce initiatives. Open standards (i.e., standards that are not tied to a specific technology or vendor) will be utilized whenever possible to encourage information sharing. The proposed infrastructure is a "front door-to-front door" framework. The only technology decisions that are being discussed operate on the actual exchange of information between partners and do not deal with the internal workings of how an agency manages and stores its information. These decisions will focus on transfer mechanisms and data exchange formats, which are the two key technical areas that relate to actually exchanging information between trading partners. Because of this, there will be no significant impact on the technologies that an agency chooses to use for database design or application development. The investments and decisions that agencies have made and continue to make concerning internal storage and management of information will not be affected. Also, because the technology infrastructure of the Network will be based on open standards, participating agencies will have tremendous flexibility in choosing hardware, software and service providers to implement the Network-specific technologies that will be needed to fully participate.

B. Definition

The technical infrastructure of the Network is the software, hardware and protocols used to make it function. This blueprint identifies the following elements of this infrastructure:

Basic Network Protocols

All information exchange on the Network will occur utilizing the following protocols:

- Transmission Control Protocol/Internet Protocol (TCP/IP) – communications protocol used to connect hosts on the Internet. TCP/IP is the de facto standard for transmitting data over networks.
- HyperText Transfer Protocol (HTTP) – protocol used to define how messages are formatted and transmitted and what actions servers and browsers should take in response to commands.

Languages for Expression and Construction of Data Exchange Formats

All of the data and all DETs on the Network will be expressed in Extensible Markup Language. (XML). XML is a language for the creation of Web documents and forms. It facilitates the definition, validation and interpretation of data between applications and organizations.

Request, Transmission and Query Protocols

Initial Network flows may use only the simplest possible request/acknowledgements for transport between Nodes. In some cases this may be a simple "get" or "post" command in HTTP. The ability of a node to respond to pre-defined queries, constructed on the basis of DETs, is a powerful but more advanced capacity that will develop over time. Many competing protocols are in development for these kinds of functions, they include SOAP (Simple Object Access Protocol) and XQL (Extensible Query Language). First generation network exchanges may be able to use much simpler sub-sets of these tools as a common starting point. In addition, several broader proposals, such as ebXML (www.ebxml.org) may address both the DET and request/transmission protocols as well as other Network components. As experience is gained in implementing these approaches, and as the approaches themselves mature, they can be standardized and coordinated by the Network administrator.

Limiting queries to those prescribed with the DET allow node managers to ensure that they can be easily serviced.

Security (see table and section below) (sHTTP, SSL and PKI)

- Security– techniques for ensuring that data being transmitted or stored in a computer cannot be read, altered or compromised by those not intended to do so. This will include technology such as Public Key Infrastructure (PKI) that verifies and authenticates the validity of any information network partners involved in an information exchange.

- Secure Socket Layer (SSL) – the connection over which a protocol that uses a private key to encrypt data is transferred. SSL is supported by both Netscape Navigator and Internet Explorer and can be used to transmit any amount of data securely. URLs for Web pages that require a SSL connection start with a "https".

- Secure HTTP (S-HTTP) – protocol for transmitting individual message securely.

C. Business Case and Critical Features

The technical infrastructure of the Network will be based on the small set of core technologies identified above. As in the example of the WWW itself, some technologies will be required while others will present an evolving menu of specific options. It is anticipated that the Network will define several levels of security (described in the table below), available to trading partners as needed. The specific level of security to be used for a given flow would be documented in the TPA, although the tools to implement the agreed-upon security level would not.

Table 1: Description of the four Network Security Levels		
Security Level	**Characteristics**	**Approach**
Level 1	Public information that requires no authentication or certification of integrity. Like all Network information, this information is protected from unauthorized modification at its Node.	This information will be available through the Internet on a public, non-secure web site. Information can be transmitted without encryption or special security measures.
Level 2	Information is approved for public distribution, but is sensitive in that the information requires data integrity.	This information will be available through the Internet on website that is secured using Secure Socket Layer (SSL). The use of SSL allows the users to authenticate that the site being accessed is an approved environmental agency web site, and provides privacy by encrypting all data in transit. SSL also provides data integrity protection.
Level 3	All data submitted by users to environmental agencies is to be treated at this level or higher. This data is of a highly sensitive nature passed between agencies but does not require digital signature. This level can apply to person-to-person and server-to-server transactions.	Access to this information is protected by SSL at the server level, and by the requirement for users' digital identity credentials. These credentials will be in the form of X.509 version 3 digital certificates issued by a Public Key Infrastructure (PKI) that the environmental agency determines meets a sufficient level of assurance in identity proofing and credential protection. Once users have been authenticated, they will be permitted to access only that data to which they are allowed.
Level 4	Information protection that requires non-repudiation in addition to privacy, authentication and data integrity. Generally, this information is the electronic version of current paper processes that require an ink signature. This information may be in the form of data coming from the agency to external users, or may be reports, applications or other information coming from external users to the environmental agency.	This information will be protected by requiring a digital signature "affixed" to the data that can be validated at the time of acceptance of the information by the environmental agency or the external user. Digital certificates issued by an approved PKI will be used for digital signature.

Table 2: Technological Characteristics of the four Network Security Levels				
Security Level	**Standard Internet Firewall**	**Secure Socket Layer (SSL)/Authenticate Originator (Digital Certificate)**	**Authenticate both Trading Partners (Digital Certificate)**	**Digital Signature Affixed**
Level 1	Yes	--	--	--
Level 2	Yes	Yes	--	--

Level 3	Yes	Yes	Yes	--
Level 4	Yes	Yes	Yes	Yes

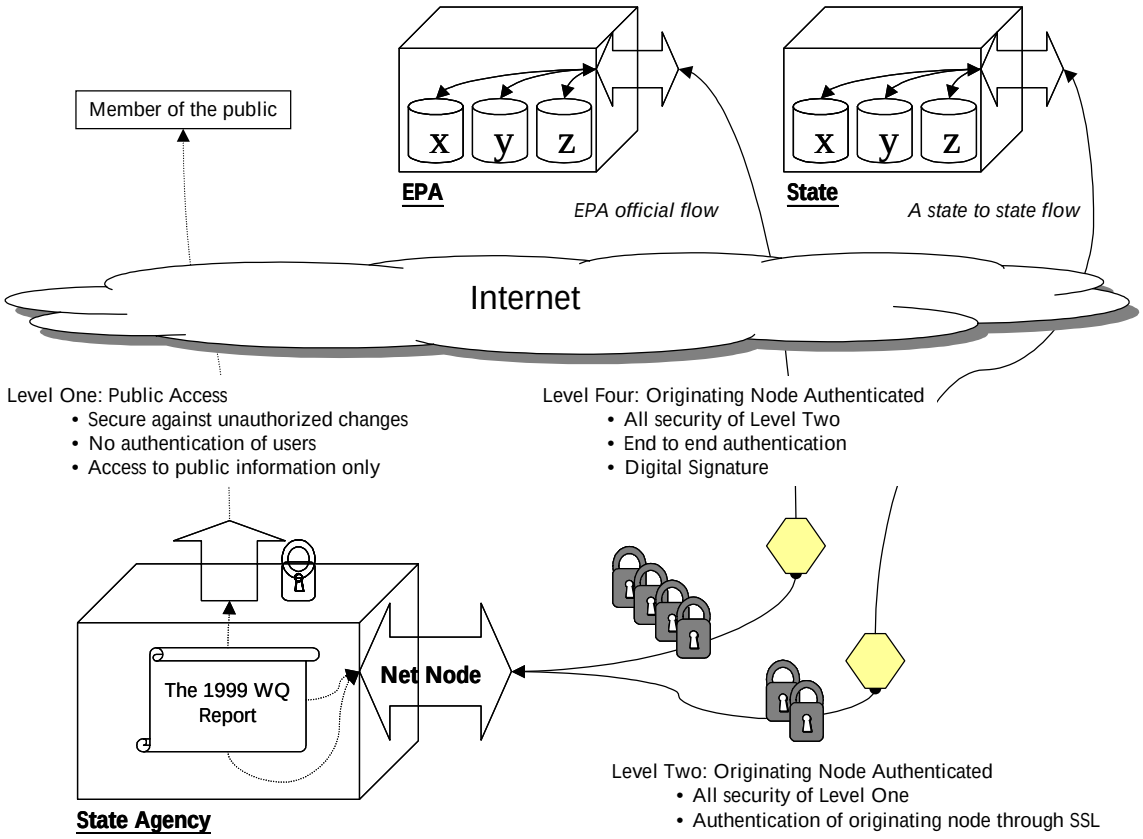
Table 3: Public Sector and Private Sector Examples of the Four Network Security Levels

Security Level	Commercial Examples	Environmental Agency Examples
Level 1	CNN.COM	Public viewing of ambient environmental conditions
Level 2	AMAZON.COM Ordering Process	List of certified state laboratories
Level 3	Transmission of supply and order information between trading partners.	State submission of a formal report required by EPA
Level 4	Contractual binding documents and e-mails	State submission of formal report to EPA which requires an official signature

These levels were developed on the basis of technologies states and EPA are already implementing. EPA and states provide significant information on their websites under Level 1 security. Many states have already established Level 2 security for their commerce functions. Levels 3 and 4 represent combinations of these and will be piloted as part of the Central Data Exchange Action Team chartered by the IMWG. Because they are based on open standards, it is likely that members will use a variety of technical architectures to establish the security of their nodes behind various levels of firewall. XML data travels over the same portion of this infrastructure as web pages, and with the explosion of XML use, security measures anticipating these architectures are readily available. In many cases, these features are built into the server/e-commerce software currently in place.

Most of the information on the Network is anticipated to be public. Certain transmissions of this data (i.e., those constituting official intergovernmental flows) will require a given security level , but the same data may also be available via the data originator's Network node (and perhaps their public access website) at a different security level (e.g., Level 1). As depicted in Figure 10, the identical information may flow from the node under different security levels depending on the partner. The ability to manage these relationships will be a significant portion of the administrative and technical costs of running a Network node. E-commerce software (e.g., Microsoft's Biz-talk server, WebMethods, or Mercator) fulfills this function.

Figure 10: A given information request may flow over the Network under various security levels



This approach is based on the following additional observations/findings:

- Internet security is an issue agencies will increasingly confront whether this or any other Network evolves.
- Agencies will have to face enterprise (i.e., Network node) security issues as they move to the conduct of business and protection of their websites.
- All agencies have to manage the traditional more intrusive relationships they have with trading partners. Many agencies are attempting to minimize these types of interactions to reduce the burden on staff and resources.
- Many of the Network security features discussed here represent significant investments, but they are investments that will be required by any agency wishing to realize the benefits of moving into the Internet age and participating in any form of e-commerce. They offer a great opportunity for synergy and cost savings by allowing Network members to focus on securing a single enterprise Network port rather than an ad-hoc collection of individual feeds and services. Implementation of the Network could reduce the burden placed on state and EPA information technology (IT)

personnel by reducing the number of systems required to communicate with the various EPA programs.

It is even possible that the Network approach could reduce the security risks associated with some data flows by establishing standard protocols and technologies. The more diverse and non-standard the data flows, the greater the security exposure. For example, EPA faces a challenge of providing direct client access to state staff to its national systems. This usually involves EPA developing a piece of software that a state agency uses for the specific purpose of access to and uploading of information to EPA. These states are "clients" to EPA servers. This approach also requires a separate, secure transport mechanism between EPA and each state using the software. This is usually accomplished by setting up a file transfer protocol site for each trading partner – increasing the need to manage multiple security relationships. While many some states will require this level of access for the foreseeable future, many of the flows that currently require this type of access might be migrated to Network flows where EPA initiates data requests from secure state servers and/or states access information on EPA servers via the EPA node, thereby reducing the need for direct access of EPA systems. This offers the potential for dramatic simplification of EPA's security predicament by limiting the number of external clients it allows access. This scenario is also consistent with EPA's decision to focus its enterprise e-commerce flows through the CDX facility.

D. Government Issues

It is critical that the Network remain vendor neutral and flexible. The goal of the Network is to encourage information sharing and to reduce the burden on participating organizations. Use of a particular software or hardware technology cannot be required to participate in the Network.

E. Introduction to Network Administration

The technical infrastructure section above outlines the specific tools and technical standards (e.g., XML, HTTP, SSL) proposed for the Network. Like a local area network (LAN) in agency offices or the Internet itself, the Network will also require a minimal (but critical) administrative capability. A Network administrator would not take the place of lower level technical standards bodies (like the Internet Engineering Task Force (IETF)), the TPA or high level intergovernmental agreements. Like a LAN administrator, a Network administrator would establish recommendations for how to use the Network, not what data to access or what to do with that data. For example, it is expected that states and EPA will use the Network to replace system-dependent flows of data under delegated programs. EPA will use this data as part of its oversight of national programs. The Network administrator will simply support the flow of this data, not its use in oversight. From the administrator, one could learn how to get the status of a facility's permit from a state node, but not whether that status is appropriate or timely.

Parallels In the Public and Private Sector for Network Administration

A Network administrator would undertake whatever functions are needed to support the Network that are not best done by the individual participants acting alone or with their individual trading partners.

- Provision of basic reference information about the Network, its participants and their data.
- Maintenance of a repository for DETs, transaction protocols and trading partner agreement templates, registered on a voluntary basis for participants' reference and use. This registration may be a requirement for a given TPA.
- Maintenance of a repository for TPAs registered on a voluntary basis that result in new data sources at a member node. This registration may be a requirement for a given TPA.
- Provision of minimal “steering group” guidance.

Given the difficulty, expense, and slow pace of wide-scale collaborative change, only the absolute minimum required to initiate the Network is proposed. Many participants will identify scores of other functions for a Network administrator (e.g., maintain a well-indexed search engine, build a value-added portal that links the participants' sites, rate the quality of the data on member websites, provide technical assistance to members). Such ideas can be considered by the IMWG after the basic infrastructure of the Network is established. Furthermore, many of these activities can be done by EPA working with states, by groups of states working together, or by the IMWG itself.

Specifically, the following broader functions that might be performed by a Network administration were considered but deferred:

- Identification, prioritization and sponsorship of DET creation.
- Active promotion and expansion of the Network membership.
- Development of readiness assessment guidelines for potential trading partners.
- Development and distribution of a “quick start” kit that allows partners to participate.
- Shared use and support of an expert team to conduct readiness assessments, and set up a partner site with "quick start" kit.
- Maintenance of a list of lessons learned and frequently asked questions (FAQs) (including node security).
- Establishment of a “test bed” facility to be accessed and used by all partners while developing new transmissions.
- Development of Network performance metrics.

The IMWG may elect to begin work immediately on these functions, but such efforts should be independent of those supporting the minimal Network administrative capability identified above.

Some of these functions may be appropriate for immediate EPA sponsorship. They are discussed further in the Member Organizational Infrastructure Section.

By Whom Would This Function Be Fulfilled?

No single entity governs the Internet or the WWW. The Internet *is* agreements to use common technologies and standards; that is all. The closest things to governance are groups that perform very specific and limited registration functions (often private sector firms that compete with each other) and groups like the W3C, a non-profit consortium that develops "standards" for infrastructure like HTML or XML. None of these groups has legal authority to force people to follow rules.

When people use the Web to conduct business, the technology and the governance of the Internet itself is transparent. This is the target for the Network as well–that participants simply grab, adapt and use the tools offered by the Network to create flows between trading partners. Like an Internet e-mail, individuals do not depend up on their ISP or the WC3 to tell them what they can or should put in their business correspondence, or how to handle any aspect of a debate that might arise as a result of the message itself.

How Would This Function Be Fulfilled?

Models for various aspects of Network governance and administration are listed in the following table:

Table 4: Models of Network Governance

Area	Governance Bodies	Functions
Internet Infrastructure	IETF (Internet Engineering Task Force)	❑ Create technical recommendations for underlying technology (like the format of email messages, and internet addresses like 207.18.19.166 and their domains "www.ibm.com"
	Registrars	❑ Private firms that are authorized to register domain names and addresses.
WWW Infrastructure	W3C	❑ A non-profit consortium that develops underlying technologies of the web like HTML and XML. Issues formal recommendations.
Visa	"Visa" association	❑ A membership association of banks and merchants who agree to abide by "Visa standards" for transactions. A steering group sets technical and performance policies (e.g., you must accept/process Visa charges using a standard transaction set or you don't get to use the Visa network). There is no one Visa corporation - just a holding company whose shares are all owned by members.
RosettaNet	RosettaNet–RosettaNet is a non-profit consortium of companies dedicated to e-commerce tools and standards.	❑ Develop RosettaNet technical and process standards for use by members.
OASIS	OASIS- (Organization for the Advancement of Structure Information Standards)	❑ a non-profit organization that supports members in development of standards ❑ is host to both tpaML and ebXML in partnership with the UN

After review of these and other models of administration, and much debate, the design team proposes the following summary principles for the creation of the Network administrative function:

- It will be kept to the absolute minimum needed to start the Network, and expanded to provide more functionality as it becomes credible to do so.
- It will focus on the core tasks of voluntary registration of members, member node catalogs, TPAs, DETs and query/request protocols. It will host a simple website with reference information about the Network.
- It will have some independence from individual members. This means it will not be solely administered by any state, by EPA or by ECOS. It may have some third party standing.
- It will remain independent from the efforts of participants to promote and expand the Network.

This version of the blueprint does not offer a final recommendation on the specific administrative structure that should be implemented. Instead, the design team requests the IMWG authorize an extension of its charter for a short additional period, after the IMWG meeting to prepare a specific recommendation. **However, this is by no means a reason to delay any aspect of establishing Network flows. We stand to learn much by doing so immediately.**

Government Issues

Aside from the issue of how the Network administrator function is structured, there are few government-specific issues with this component. In most cases these technologies can be used as is because their function is mostly mechanical. By definition, Network administration will NOT include inherently governmental functions.

10. Component 5: Member Organizational Infrastructure

A. Definition

Member organizational infrastructure defines the roles and responsibilities required for Network participants. The term infrastructure is used because these roles and responsibilities will require investment to build, and when effective, should be relatively transparent. Because states and EPA will take the first steps towards implementation together, this section focuses specifically on their near-term roles and responsibilities. As the Network is expanded to other participants (such as tribal governments) their roles and responsibilities will need to be defined as well.

B. Background

Purpose of This Section

The preceding components of the Network Blueprint provide the "plumbing" and "electrical specifications" for moving data and administering the Network. Aside from the discussion "What is a Node Really?" above, the components have addressed what the Network looks like from the front door (or node door) *out*.

This Network component focuses on the infrastructure needed to get EPA, states, and eventually other partners "interested, authorized and able" to participate in the Network for their business. It suggests members' internal roles and responsibilities for operating their nodes and supporting (not administering) the Network itself. This section is among the most important in this document; it is also the most preliminary. Because many of these concepts apply both to EPA and states and because they are all interrelated, there is significant redundancy in the current draft. After the workgroup has debated and clarified some of these issues, and as Network flows begin, the details of the roles and responsibilities outlined here will be further refined and documented. Specifically, this section is offered to frame the IMWG's consideration of the Network and its role in supporting the Network. However, none of the issues debated here should preclude two parties from immediately using other concepts in this document to create Network flows between them.

In addition to the basic organizational infrastructure needed, this section discusses what can be done to increase the ***capacity*** of states and EPA to fulfill these responsibilities. It provides a framework for what states can do for themselves, other states and for EPA; it also describes the complex but critical opportunities EPA can take to increase the capacity of states to build and participate in this Network. As this document makes clear, this Network is fundamentally decentralized; yet EPA plays a critical role. While many blueprint team members believe this Network (or something like it) will arise with or without EPA's participation, all believe that the important things will happen better and faster if EPA is in at the ground floor. An extensive set of specific options and actions were originally developed as part of EPA's Information Integration Initiative (I-3). Key milestones from EPA's I-3 have been included here. These milestones clearly reflect EPA's public and specific commitment to the Network.

State and EPA roles and Responsibilities

State and EPA roles and responsibilities for data exchanges are embedded in a complex, historical web of formal and informal agreements. These include program delegations, annual SEAs, PPAs, PPG and program- or Region-specific agreements. These agreements often overlap, involve different levels of each organization and in some cases conflict. Worse, in many cases roles and responsibilities are ambiguous, with no one accountable for end-to-end data quality. In other cases, stable program-specific arrangements have developed that include agreed-upon metrics for performance and data quality. This wide variety of experiences and problems makes it easy for participants from different programs and states to hold different opinions on the effectiveness of the existing data exchange system. For parties with stable, negotiated formats and expectations, the Network offers an economy of scale and a refined set of technical tools; for partners mired in ambiguous, conflicting agreements, it presents the challenge of making their obligations and metrics explicit, but also offers the tools (especially the TPA) to do so.

As described above in the TPA component, most of these flows are currently described in terms of obligations of states to feed one or more program-specific EPA information systems. The Network will simplify and clarify data exchange roles and responsibilities through the use of TPAs. Each TPA will identify the trading partners and respective node addresses; define the purpose and content of the data exchange; and define expectations for data and transaction quality, security, integrity and frequency. Network participants will need to consider their own requirements for populating internal business applications when developing data exchange templates and TPAs. However, TPAs will not be used to specify how this integration is to be accomplished. The Blueprint Team expressed a strong desire to focus TPAs on business events and processes and on the necessary supporting data and not to constrain the design of DETs and TPAs with the idiosyncrasies of existing internal business applications.

C. Discussion

By agreeing to host and exchange data on the Network, each trading partner, as a Network partner, assumes and accepts certain roles and responsibilities. These roles and responsibilities will include the following:

Role – Node Administrator

The Node Administrator, similar to a Web or systems administrator, will be responsible for:

- Software development and implementation (e.g. security, XML)
- System documentation
- Hardware and software maintenance
- Policies and procedures (e.g., security. documentation, change management, problem management)
- Backup and recovery

Role – Data Steward

The Data Steward, similar to a data administrator, will be responsible for:

- Documenting data and data relationships
- Developing data definitions and data naming standards
- Developing standard calculations and derivations
- Defining data security and retention requirements
- Developing DETs
- Mapping data sources (e.g., business applications) to DETs
- Monitoring data quality

Role – TPA Administrator

The TPA Administrator, similar to a contract administrator, will be responsible for:

- Developing and approving TPAs
- Monitoring compliance with TPAs

As stated in the Introduction above, effective stewardship of the Network is considered fundamental to the idea of the Network and to its success. The roles and responsibilities described above are considered essential for effective shared stewardship of the Network

D. Business Case and Critical Features

The following table outlines some key roles of states, EPA and the IMWG in five distinct areas:

- Supporting the Network Administrator and other shared infrastructure
- Establishing EPA's capacity to build and manage its node
- Establishing EPA's capacity to establish and manage flows with states
- Supporting individual states' capacity to build and manage its node
- Supporting individual states' capacity to establish and manage flows with EPA

These areas are considered from the perspective of EPA, states and the IMWG.

EPA and State Support of the Network Administrator and Other Shared Infrastructure

As indicated in Component 5: Technical Infrastructure and Network Administration, this blueprint does not propose a specific structure and seat for the Network administrator. These arrangements will be developed once the IMWG has endorsed the concept of the Network and considered the blueprint recommendations. Nonetheless, the following roles and responsibilities in supporting this function are clear:

- EPA and states will need to support the IMWG in identifying and establishing the Network administrator.

- State support of this function will likely consist mostly of cooperation and encouragement. States may also be able to contribute direct technical and management resources (staff or expertise) needed to launch this function.
- EPA will also have a special role in presenting and supporting its priority data flows for DET creation, to the extent that this involves the Network administrator.

As the division of responsibilities becomes clear between the Network administrator, IMWG, EPA and states, several additional capacity-building steps could be taken:

- Development of a readiness assessment guideline for potential trading partners.
- Development of a "quick start" kit that allows partners to participate.
- Shared use and support of an expert team to conduct readiness assessments, and set up a partner node with "quick start" kit.
- Maintenance of "lessons learned", FAQs, etc. (including node security).
- Establishment of a "test bed" facility to be accessed and used by all partners while developing new transmissions.

Other support activities were mentioned during development of the Blueprint but have been omitted here for clarity and because their consideration may be premature before the IMWG has discussed the broader blueprint design.

EPA Organizational Infrastructure

Early on, the most important opportunity and challenge for the Network will probably be the credible participation by individual EPA staff at various levels in creating flows with their state counterparts. This task will be more difficult for EPA than for states because of EPA's broader, more diverse and more complex data needs and its multiple state clients. Support from states and the IMWG will be needed. EPA's CDX staffs have a direct link to these in-reach efforts through the IMWG Action Team (CDX Action Team), but most EPA data exchanges remain system- and program-specific. Program offices working on current CDX pilots have already begun the Network oriented data exchange process; but what of the Regional staff person who first hears of these ideas from an eager and aggressive state Chief Information Officer (CIO) who wishes to negotiate something called a TPA? How will *that* person be supported, or at least not stymied? As EPA begins to develop policy and infrastructure to support the Network, it must also ensure that smaller projects succeed. Early Network flow projects (those sponsored by the IMWG and those that arise spontaneously from individual state-EPA initiatives) will form the foundation for later growth.

Under the Network, states and other partners would make their information accessible to EPA's Central Data Exchange facility. EPA would manage its copies of this data (i.e., in the near term, loading data into the existing national systems). While reengineering its systems in concert with EPA's ongoing integration effort, each program will need to help develop exchange formats for its business subject matter area, coordinate with CDX to receive newly retooled transactions based upon these formats, and have the capacity to exchange data in its own system with CDX. These are significant but tractable technical tasks; the real challenge is to manage the following types of change in internal roles and responsibilities:

- Existing programs, policies, processes - Existing delegation agreements that specify information requirements, certain National Environmental Performance Partnership System (NEPPS) agreements, electronic reporting trading partner agreements and informal ad-hoc data acquisition arrangements will all need to converge into documented Network trading partner agreements. These agreements will require the coordination of many people.

- Commitments to conduct business through the Network - Having committed to conducting business through the Network, EPA will need to ensure that its individual programs and regions are able to do so (e.g., have adequate funding and other resources).

- Coordination with internal integration - In addition to retooling information exchanges, and thus system capabilities, EPA is also establishing an enterprise architecture basis for its internal integration efforts. A coordinated, balanced approach may constrict EPA's capacity to retool existing incoming information flows towards the Network vision.

- Central Data Exchange - An operational node on EPA's CDX is required to receive Network data and handle different transmission and exchange formats (transaction sets). Priorities, implementation and resources for CDX development must be established and aligned. CDX and program systems must have the capacity to exchange data. Programs must understand their roles in Network participation, and have the expertise to redevelop their existing information exchanges.

- IT/IRM Policy –EPA's standing IT/IRM policies must be reexamined to determine what is needed to support the Network, concurrent with the reassessment of policies for internal integration and architectural realignment purposes.

- New programs, policies, procedure - A proactive means of handling new laws affecting the Network (e.g., Cross-Media Electronic Reporting and Records Rule (CROMERR)) must be developed. As with other information management concerns, getting in front of the regulatory development process will help create reform that can adapt to future changes.

- Role and responsibilities of Regions – Much of the burden of establishing the Network will fall on the EPA Regional offices. Processes and procedures will have to be harmonized to ensure national consistency. Regions play a central role in the management and organization of their states' TPAs and relationships to NEPPS and other negotiated agreements. Regions can participate with their states to build capacity and extend the Network.

Much of the preliminary planning for I-3 was conducted in parallel with the Network Blueprint work; however, EPA's investment plan for I-3 was due prior to the completion of this document. Because of their direct relevance to the Blueprint, the Exchange Network Infrastructure and Partner Assistance milestones have been included here for discussion purposes only.

Table 5: EPA Milestones for the Network		
EPA Internal Integration	**Exchange Network Infrastructure**	**Partner Assistance**
2000 1. EPA establishes the Information Integration Initiative in support of this vision. 2. EPA makes commitment to internal information integration and begins to realign internal structures and resources is support of I-3. 3. The utility and expanded opportunities in the use of integrated information to environmental protection programs is clarified via FY2000 demonstration projects.	**2000** 1. EPA and state environmental agencies commit to developing a national environmental information exchange network with other partners in environmental protection. A vision of the Exchange Network is documented and supported by ECOS and participating states. 2. The State/EPA Information Management Workgroup takes the active lead in developing this vision and Exchange Network.	**2000** 1. Utilization of the One Stop Network of state officials in defining the vision and determining partner needs.
2001 1. Initial scope of the I-3 project is refined, well-defined and operational. EPA has made a stated commitment to coordinating its internal integration efforts with the evolution of the Exchange Network partnership. 2. A target Enterprise Architecture is in place for EPA's mission functions and is the guiding principle for IT investment decisions and framework for systems development and modernizations efforts. 3. EPA Programs and Regions have launched Information Strategic Planning (ISP) exercises and have realigned systems development plans to include utilization of corporate data services and functions, and/or planning to redeploy business modules as corporate modules. 4. EPA's internal vision for integration of information beyond the regulatory/ambient information realms is clarified.	**2001** 1. Exchange Network Governance and interagency roles are established and designated people are in place. 2. Scope of the first Phase of the Exchange Network is fully defined and operational for a limited subset of shared environmental business functions between EPA and a few prototype states. 3. States and EPA are actively engaged in defining subject matter area 'business model' neutral exchange formats, and retooling existing information exchanges towards the adoption of these formats. 4. State talent and motivation is capitalized on to create as many transaction sets for the Exchange Network as possible. 5. The path towards expansion of the Exchange Network beyond EPA and state environmental agencies is well understood. A clearer vision of the Exchange Network's second phase of development and use is established. 6. The Exchange Network is trusted, and all security concerns are reviewed, well understood and properly	**2001** 1. Initial prototype pilots have identified readiness factors (technical, policy, and organizational) for Exchange Network participants. 2. How best for EPA to assist its partners become ready to be Exchange Network participants is clear and well understood. From these readiness factors an Action Plan for assistance to Exchange Network participants is fully defined. 3. A state/EPA Action Team is actively assisting states evaluate their readiness to participate in the Exchange Network

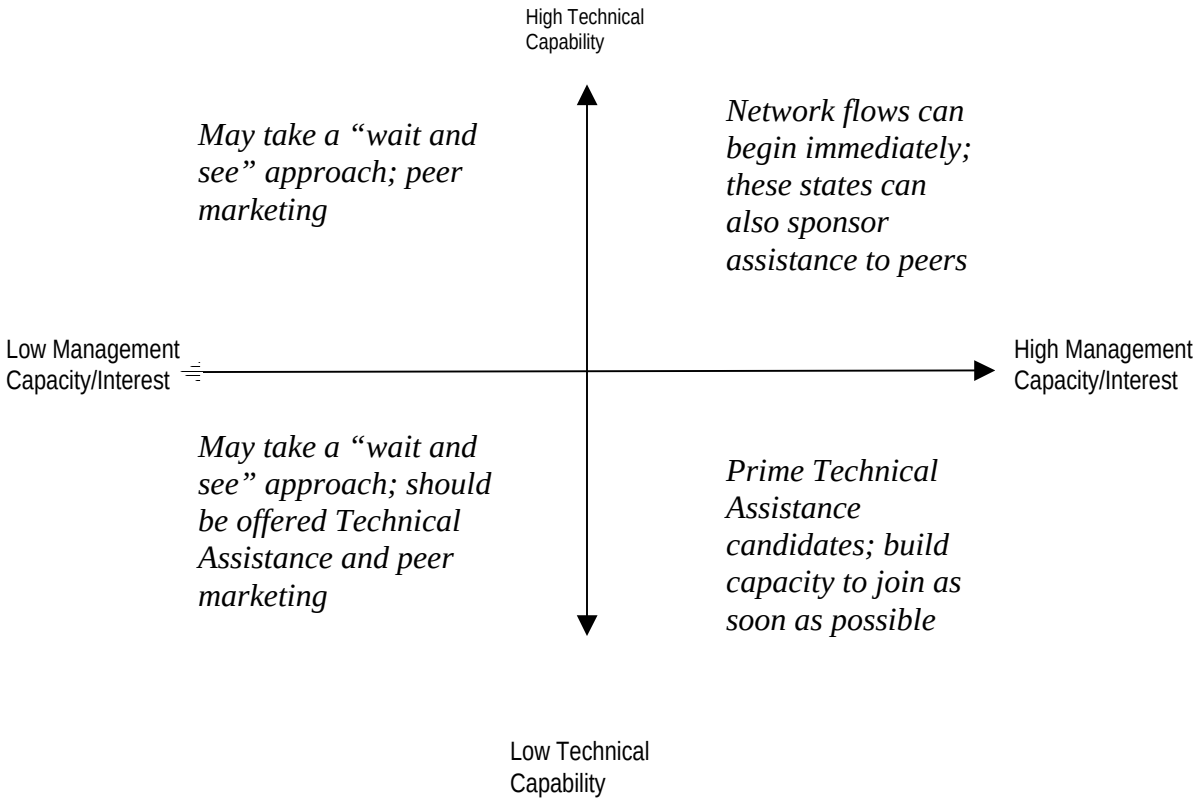
	addressed. 7. EPA and States have determined how best to address the management of the Exchange Network.	
2002 1. The Enterprise Architecture continues to serve as a strategic framework upon which programmatic IT investment decisions are made, with appropriate revisions included. 2. Transition plans in place and migration to the Enterprise Architecture is underway for EPA's major program systems and second tier systems. 3. An expanded set of Foundation components (system of registries, business modules, and functions) are serving as the authoritative source of key agency data and functions. 4. Third tier of foundation components are under development. 5. EPA Programs and Regions have migrated all major and some minor systems to utilize the core infrastructure (where appropriate) and have demonstrated benefits in terms of increased access and analytical capacity and increased efficient utilization of resources. 6. I~3's progress and the approach in its management plan is reviewed.	**2002** 1. Governance and stewardship of the Exchange Network are routine operations 2. Exchange Network is operational. 3. The Exchange Network has reached out beyond EPA and state environmental agencies and is operational with other parties. 4. States, EPA, and the new partners are continuing to retooling existing information exchanges towards the adoption of these formats. Work is underway in identifying new partners with whom information exchanges need transformation. 5. The Exchange Network is trusted, and all security concerns are reviewed, well understood and properly addressed	**2002** 1. A mechanism for assisting partners to assess their readiness to function as Exchange Network portals continue to operate. 2. A mechanism for technical assistance to trading partners to implement and secure their Exchange Network portals continues to operate. 3. An assistance mechanism for states/partners to participate in developing exchange formats in operational.

State Organizational Infrastructure

Each state environmental agency will need to assess its own information management status and level of readiness to join the Network. Three levels of overall technical and management readiness can be examined. We have borrowed this concept of "readiness" from the e-commerce network vocabulary. Large firms (e.g. IBM or Intel) have begun to formally assess the readiness for e-commerce partnerships with their suppliers and distributors. An excellent technical overview of this process as it applies to e-commerce is included in the "RosettaNet" paper included in the Blueprint reference materials.

As indicated in the table below, technical capacity can be thought of as the ability to build a node and the internal systems feeding that node. This is a relatively traditional software/Web

development task. Management capacity and interest are different and more complex. It includes the internal discipline and coordination to ensure that high quality data is available to the node manager, and that TPAs covering that data can be negotiated and implemented. "Interest" is included in this category because the Network is voluntary and the first flows will require pro-active involvement on both sides.



This chart is useful because it depicts a wide range of possible starting points for any given state (and for EPA as an agency). State participants in the Blueprint team span the spectrum of capabilities and interests identified in the table.

- At the highest level of overall readiness are some states (top right of chart) with robust technological and management data exchange capabilities. Several of these states have participated in Network pilots and other projects and could begin Network flows within months. These states are also in an excellent position to partner with EPA and use the IMWG's Knowledge Transfer Action Team to share their experience (and perhaps specific tools and approaches) with other states.

- States in the center of this chart enjoy some of the infrastructure needed, but need further development of some technological or programmatic components in order to join the Network; they are ideal targets for Knowledge Transfer since they are almost ready to go.

- The Network design emphasizes open and flexible tools for partners to use; therefore, few states need be left in the lower right hand portion of the chart. With sufficient management commitment and some support, most states should be able to participate. It may be appropriate for these states to initially rely on EPA for some portions of their technical infrastructure. This would be similar to a state's decision to use EPA's CDX as its electronic reporting infrastructure, while maintaining stewardship/ownership of that data.

- The few states on the left of this chart pose a different challenge. In the upper left (which we think is nearly empty) states have the technical capacity to participate in the Network (perhaps because they are already building portals of their own, and are familiar with XML technologies) but do not have the management capacity or interest. These states should be the targets of "marketing" outreach efforts. EPA will need to ensure that it is offering these states the ability to transition flows to the Network.

- Finally, states in the lower left corner should be offered technical assistance and "peer" outreach, especially to ensure that what appears to be a lack of management commitment is not actually concern that the technical threshold for participation is just too high to merit management investment.

- A special focus similar to the Knowledge Transfer Action Team's "Small States" working group may be an excellent way to document and share the experience of states who have rapidly moved up and into Network participation.

IMWG Organizational Infrastructure

The IMWG is the core forum for state and EPA collective action. As such, the IMWG will play a crucial role in creating Network flows. The IMWG chartered this Blueprint development effort. However this Blueprint does not recommend that the workgroup *be* the Network, nor its administrator. Instead, this blueprint proposes that the IMWG be the venue through which the institutional home and capacity for these functions be identified and launched. In addition, the workgroup is the only body that can provide some high-level coordination and support to ensure that its own sponsored activities are advancing the Network:

- Data Standards Council
  - Encourage DET developers to use the DCS
- Central Data Exchange Action Team
  - Forum for data exchange issues as they relate to EPA's CDX
  - Two Network flow projects (DMR and STORET) launched by the team
  - Security/E-commerce interoperability project
- PCS/IDEF Action Team
  - Encourage/support use of Network concepts as final design for IDEF is established
- Facility Action Team
  - Establish flows of facility data
  - Evaluate of TPAs for facility data

In addition, the IMWG is the only organization positioned to support some of the broader intergovernmental commitments and expectations, such as the following:

- Commitment on behalf of enough participants to conduct business in the new manner to make the endeavor worthwhile.
- Commitment on behalf of participants to take on the implied data stewardship responsibilities[5].
- EPA commitment to retool existing state reporting relationships to accommodate Network principles.
- Commitment on behalf of participants to financially support the Network and work towards establishing self-sustainability for this function.
- EPA commitment to investments for accelerating the Network and support for the DET development process so that DETs are, or can be, available for those who want to use them.
- EPA commitment to maintaining multiple (old and new) business practices for receiving data from partners so that the Network is truly voluntary.

[5] Each participating agency, as a Network partner, in agreeing to host their information, assumes data management responsibility for their portion of the Network. Data quality, timeliness, error correction, meta data expectations, and standard operating procedures will all need to be developed, built into transaction set requirements, and incorporated into TPAs. (The degree of oversight and specificity would vary depending on nature and granularity of the exchange)

11. Relationship of Network Components

The matrix on this and the following page describes the relationship of each of the Network components to the other components.

		Data Standards	Data Exchange Templates
	Data Exchange Templates:	– Data standards will be incorporated into Data Exchange Templates – Cross program data standards implemented in DETs will improve integration.	
	Trading Partner Agreements:	– Trading Partner Agreements will identify which Data standards are being used.	– Trading Partner Agreements will identify which Data Exchange Templates are being used.
	Technical Infrastructure:	– The Technical Infrastructure (e.g. XML Schema)will validate that a data standard is being used. – The Technical Infrastructure will provide easy/open access to all official data standards.	– The Technical Infrastructure will validate that a given transmission is compliant with its Data Exchange Template.
Organizational Infrastructure	**Network Governance:**	– Network Governance for Data standards will be through the Environmental Data Standards Council.	– Coordination/governance of Data Exchange Template development, especially for the traditional State/EPA data flows.
	EPA:	– EPA will develop policy enforcing the use of Data standards in all internal information management activities.	– National programs identify priority areas for Data Exchange Templates development between states and EPA
	State Environmental Agencies:	– State Environmental Agencies will develop policy around the use of data standards developed by the Environmental Data Standards Council in all information management activities.	– National programs identify priority areas for Data Exchange Templates development between states and EPA – State-to-State flows use Data Exchange Templates between state programs.
	State/EPA Information Management Workgroup:	– The State/EPA Information Management Workgroup will continue to provide support to the Environmental Data Standards Council.	– The State/EPA Information Management Workgroup will provide guidance for how to develop Data Exchange Templates.

		Trading Partner Agreements	Technical Infrastructure
Data Exchange Templates:			
Trading Partner Agreements:			
Technical Infrastructure:		– The Technical Infrastructure will include a neutral repository where Trading Partner Agreements will be posted.	
Organizational Infrastructure	**Network Governance:**	– Governance of the Trading Partner Agreement format and development mechanism. – Oversight of network expansion to additional data partners.	– Policies will be established to define general security processes used on the Network.
	EPA:	– Trading Partner Agreements document official flows for regulatory reporting requirements from states to EPA program offices. – Regional role in the governance of Trading Partner Agreements.	– The availability of EPA funding will affect the ability to assist states in developing technical capacity.
	State Environmental Agencies:	– Trading Partner Agreements document official flows for regulatory reporting requirements from states to EPA program offices. – States will coordinate the management of Trading Partner Agreements with their EPA Region.	– Coordination/leveraging of state technical investments via Knowledge Transfer. – States will have some ability to influence EPA's technical decisions/investments.
	State/EPA Information Management Workgroup:	– Oversight of and coordination of the Trading Partner Agreement framework for State/EPA data flows. – Coordination of network expansion to additional data partners.	– The State/EPA Information Management Workgroup comments on technical standards that influence technical infrastructure.

12. Recommendations to the Workgroup

Based on the discussion and analysis documented above, the Blueprint design team makes the following recommendations to the IWMG:

1. The IMWG should approve the Blueprint.
2. The Network Blueprint team should stay intact to develop a specific proposal on Network administration that includes financing options.
3. The IMWG should identify its next steps in advancing the Network, including a plan for outreach.
4. The process of using the Network components to build Network flows should begin immediately.

13. Network Examples

A. Multi-state Watershed Project

Appendix F presents an example of a voluntary agreement among trading partners for the purpose of exchanging data. Not every Network TPA would follow this format, just as many other types of state/EPA operating agreements look quite different from case to case. The particular circumstances of the parties involved and the data being exchanged will influence which elements are included in the agreement and how these issues are described.

References

State/EPA Information Management Workgroup

Shared Expectations of the State/EPA Information Management Workgroup for a National Environmental Information Exchange Network. Working Version, June 12, 2000.

EPA White Papers

The Exchange Network . Draft August 1, 2000.

Shared Network Governance and Stewardship of Data and the Exchange of Data. Draft June 21, 2000.

Industry White Papers

Sachs, et al., *Executable Trading-Partner Agreements in Electronic Commerce*. IBM T.J. Watson Research Center, 2000.

O'Sullivan, Patricia J. and Don S. Whitecar, *Implementing an Industry e-Business Initiative: Getting to RosettaNet*, Intel Technology Journal Q1, 2000.