"The only source of knowledge is experience"
- Albert Einstein

National Environmental Information Exchange Network

Web Services Implementation: The Beta Phase of EPA Network Nodes

Mark Nobles

Logistics Management Institute

National Emissions Conference

16 April 2002

Why is this project important?

- This is where "the rubber hits the road"

- The Project demonstrated that the Network approach works, and that hands-on experience *during Network development* is crucial to inform and improve the Network approach

- Understanding the Network technical infrastructure is important for understanding the benefits of the Network approach

What is a Network Node?

A simple environmental information web service that initiates or processes requests for information and receives or sends the requested information in the standard XML format

What Is a Network Node?

- A standardized set of tools to exchange information on the Network

- A presence through which Partners exchange information

- A simple environmental web service…

What is a Web Service?

- Software that exposes very simple functionality of business applications through a standard Internet interface

- Web services communicate with other web services via standards-based technologies that can be accessed by trading partners independent of hardware, operating system, or programming environment

Network Node Pilot Project: Background

January 2001

The IMWG, together with the Co-chairs of other Action Teams and the EDSC, decided to produce near-term visible results that demonstrate the advantages of the Network approach

Pilot Project: Initial Goals

- Demonstrate XML technology
- Demonstrate states' capacity to develop Nodes
- Establish node-to-node connections and exchange data between Nodes
- Establish full-production Nodes which concur with TPAs
- Validate Node transactions against registered DETs
- Develop recommendations for Node coordination

Alpha Phase (Spring 2001)

Interested States (Delaware, Nebraska, New Hampshire, Utah) agreed to participate using their own resources

- Project began in February
- "Alpha Nodes" demonstrated in May 2001

Alpha Phase: Successes & Lessons Learned

Produced the first proof-of-concept Node prototypes. Specific technical and management issues recognized, including:

- Versatility of XML and the middleware

- Need to test security, performance, and interoperability (the ability of software and hardware on multiple machines from multiple vendors to communicate)

Beta Phase Objectives

1. Help to identify node design specifications

2. Scope the development of a "node package" or "technology template"

3. Implement a fully normalized (standardized) Facility Data Exchange Templates (XML Schemas)

4. Involve 2-4 more state agencies and EPA (CDX)

5. Test and resolve performance issues

6. Explore security issues and options

Beta Phase Objectives (cont'd)

7. Link State environmental agency node data transfers through EPA's node (CDX)

8. Pilot and/or execute TPAs for additional nodes

9. Automate validation of flows against DET (Schemas)

10. Send production data from nodes to the Facility Registry System (FRS)

11. Test interoperability (a "stretch objective")

Beta Phase Basics

- States: Delaware, Florida, Nebraska, New Hampshire, New Mexico, Utah, Pennsylvania, Mississippi

- Nodes based on MS Biztalk, Oracle, and XAware

- Processing Three "input" service requests for facility information

- Used Facility Data Action Team Facility Schemas as Data Exchange Templates

- Used NIST "testbed" Registry to host Schemas for automated validation
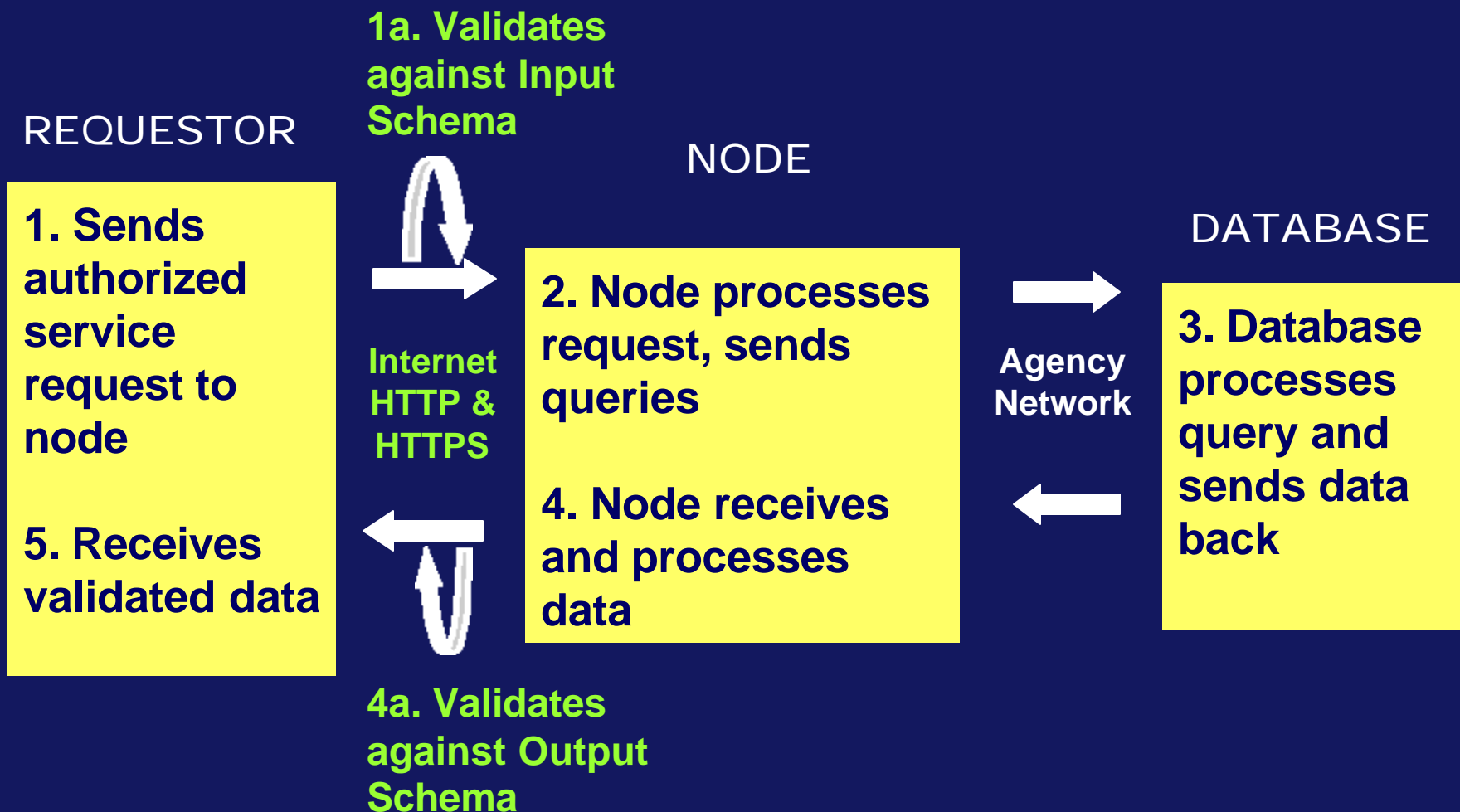
CDX Participation

- Provided subcontractor support to state efforts
- Developed their node using Microsoft SOAP tool kit
- Provided SSL certificates to test security
- Used Florida's Web Service (WSDL) and Oracle tools to retrieve data
- Developing an Oracle based node
- Provided technical expertise

Accomplishments

- Progressed from a proof of concept to utilizing most of the Network components
  - Operational nodes
  - Registry validation
  - Facility Schema usage
  - Security (SSL)
  - Model TPA

Beta Node Information Exchange Process

1a. Validates against Input Schema

REQUESTOR

NODE

DATABASE

1. Sends authorized service request to node

5. Receives validated data

Internet HTTP & HTTPS

2. Node processes request, sends queries

4. Node receives and processes data

Agency Network

3. Database processes query and sends data back

4a. Validates against Output Schema

Beta Phase Results & Lessons Learned

- 10 primary lessons learned
- Most lessons learned relate to overall Network planning, establishing Node specifications, exchange protocol, Template development, etc.
- Need more work on Network

Lessons 1: Need Node Functional Specification

- Node functional specifications and "how tos" are needed, therefore…

    Network-wide strategic planning and communication needed

Lesson 2: Plan Ahead

- Node planning can take far longer than node implementation

- Without a lot of careful forethought and functional specs, time will be spent "undoing" and "redoing"

- Therefore, Node functional specs, DET harmonization, data standards, etc. are needed (for entire Network)

Lesson 3:
Technical and Institutional Capacity are needed

- Technical: Understanding of XML, SOAP, DETs, security, middleware, and existing information systems

- Institutional: Need support from decision-makers, permission to move ahead, permission to use data, etc. (Some of these will be taken care of as the Network progresses)

Lessons 4: We Know What Nodes Are and Do

- Most questions about what Nodes are and what Nodes do have been answered

- There are many ways to establish a Node

  Different physical and software configurations

Lesson 5: Interoperability is Achievable, But Will Require Additional Effort

- Node functional specifications are needed

- Customization, testing, and troubleshooting will be needed

Lesson 6: Connecting to Existing Systems Takes Time

- Relationship of a Partner's Node to their existing information system is complex and unique

    Substantial time spent by States mapping to information systems

    Network-wide support to States desirable, but hard to generalize

Lesson 7: External Support Will be Needed

- Partners may need to and most likely will rely on external technical expertise while more internal technical expertise is developed

  Finding external expertise is not easy, may be costly, but in the long run cheaper while you build your own expertise

Lessons 8: Network Technology is Evolving Quickly

- Node technology will continue to evolve rapidly. Network will need to monitor closely and adopt when appropriate

  - Emergence of WSDL and SOAP v.1.2 during 2001

- Coordination of Network Activities and Node development between Partners, NSB, EDSC, and Action Teams will be required to avoid conflicts and delays in securely placing Nodes on the Internet.

Lesson 9: DET Template Development, Harmonization, and Utilization is Key

- Careful (and rapid) Template development is important to ensure data consistency and reliability
  - Missing data, null values, data element naming constraints, data definition, and the delivery of empty XML tags needed

Lesson 10: The Network Registry/Repository is Very Important

- It worked for the Beta
- The Network Registry/Repository will play an important role in network information exchange

  Additional documentation and specifications will likely be available at the Registry/Repository

Next Steps:

- Follow-on: preliminary objectives –
    - Develop Node Functional specs V.1 (draft)
    - Develop Network information exchange protocol (draft)
    - Plan Do Check Adapt (develop and test metrics for protocol and functional spec compliance)
    - Develop recommendations on technology templates
    - Further test and develop recommendations for interoperability
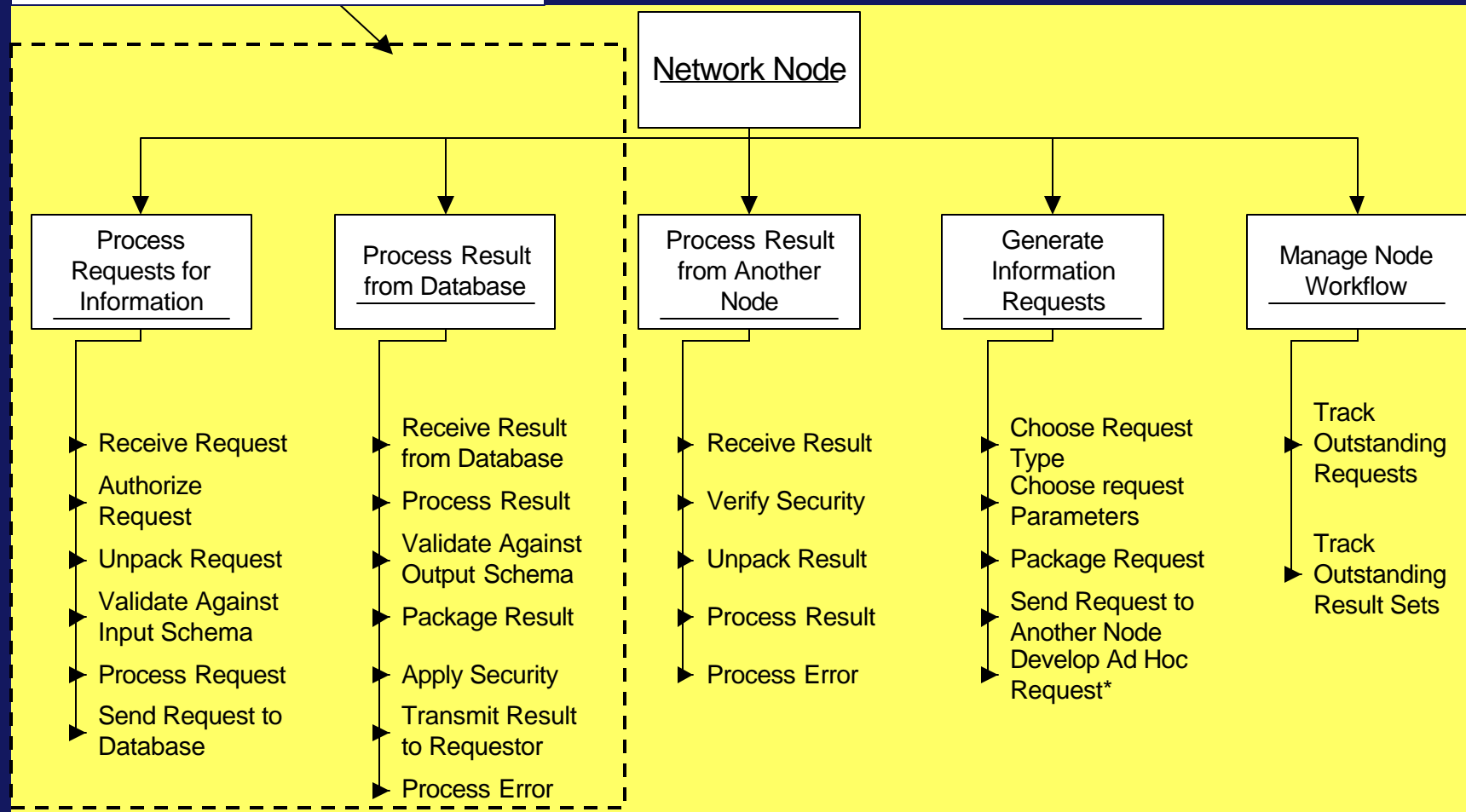
Next Steps (cont'd):

- Develop "How to establish a Node"
- Develop Node Technology "101" training course geared for practitioners
- Working Nodes with a Flow
- Strong candidates for pilot Flow(s):
  - Air Emissions inventory
  - Facility (demonstrate bi-directional flow)
  - Ozone monitoring (could make a simple Schema)

Questions ?

Node Functions

Beta Node Functions

Network Node

Process Requests for Information	Process Result from Database	Process Result from Another Node	Generate Information Requests	Manage Node Workflow
▸ Receive Request	▸ Receive Result from Database	▸ Receive Result	▸ Choose Request Type Choose request Parameters	▸ Track Outstanding Requests
▸ Authorize Request	▸ Process Result	▸ Verify Security	▸ Package Request	▸ Track Outstanding Result Sets
▸ Unpack Request	▸ Validate Against Output Schema	▸ Unpack Result	▸ Send Request to Another Node Develop Ad Hoc Request*	
▸ Validate Against Input Schema	▸ Package Result	▸ Process Result		
▸ Process Request	▸ Apply Security	▸ Process Error		
▸ Send Request to Database	▸ Transmit Result to Requestor			
	▸ Process Error			

Basic Node Architecture

(Common to all Nodes)

Internet

Firewall

Node (Web Server)

SOAP Listener
SOAP Processor

Object Handler
Data Mapping
XML Processor

Database Connectivity

Existing Information System